

SPECIFICATION

TARGET OBJECT MATCHING METHOD AND APPARATUS, ELECTRIC DEVICE AND STORAGE MEDIUM

5

The present disclosure claims the benefit of priority of Chinese Patent Application No. 201910882691.5, filed with the Chinese Patent Office on September 18, 2019 and titled “TARGET OBJECT MATCHING METHOD AND DEVICE, ELECTRIC APPARATUS AND STORAGE MEDIUM”, which is incorporated herein by reference in its entity.

10

TECHNICAL FIELD

The present disclosure relates to the technical field of computer vision, and particularly to a target object matching method and device, an electronic apparatus, and a storage medium.

15

BACKGROUND

To match a face with a human body or a hand with a human body is for purpose of determining whether the human body is matched with the face or the hand in a single picture. Sometimes there are many people in an image, each of which may show different actions and have varied sizes, or even these people are shielded by one another. For various reasons, it is very challenging to match a human body with a face and a human body with a hand.

20

SUMMARY

The present disclosure proposes a technical solution for matching target objects.

25

According to one aspect of the present disclosure, there is provided a target object matching method, comprising: acquiring a first target object and a second target object to be matched from an input image, wherein the first target object includes a human body, and the second target object includes at least one of a face and a hand; performing feature processing on a first image corresponding to the first target object and a second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image; establishing, based on the matching degree between the first target object in the first image and the second target object in the second image, a bipartite graph between the first target object and the second target object; and determining, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other. The above-mentioned configuration can improve the

30

35

matching accuracy between target objects and is applicable to a scenario where a plurality of people are shielded by each other, thereby providing better applicability.

In some possible implementations, performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image comprises: performing feature extraction processing on the first image and the second image to obtain a first feature of the first image and a second feature of the second image, respectively; and performing classification processing on a concatenated feature of the first feature and the second feature to obtain the matching degree between the first target object in the first image and the second target object in the second image. The above-mentioned configuration enables conveniently obtaining a matching degree between two target objects, thereby gaining high-precision features and accurate matching degrees.

In some possible implementations, performing classification processing on the concatenated feature of the first feature and the second feature to obtain the matching degree between the first target object in the first image and the second target object in the second image comprises: performing feature fusion processing on a concatenated feature of the first feature and the second feature to obtain a fusion feature; and obtaining the matching degree between the first target object in the first image and the second target object in the second image by providing the fusion feature as an input to a fully connected layer to execute the classification processing on the fusion feature. The above-mentioned configuration can improve the classification efficiency and accuracy with the fusion processing.

In some possible implementations, establishing, based on the matching degree between the first target object in the first image and the second target object in the second image, the bipartite graph between the first target object and the second target object comprises: in response to the second target object including only the face, establishing the bipartite graph between the human body and the face based on the matching degree between the first target object in the first image and the second target object in the second image; in response to the second target object including only the hand, establishing the bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image; and in response to the second target object including the face and the hand, establishing the bipartite graph between the human body and the face and the bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image, wherein the matching degree between the human body and the face

is used as a concatenation weight between the human body and the face in the bipartite graph between the human body and the face, and the matching degree between the human body and the hand is used as a concatenation weight between the human body and the hand in the bipartite graph between the human body and the hand. The above-mentioned configuration can easily construct a relationship between target objects in a manner of establishing a bipartite graph between the target objects.

In some possible implementations, establishing, based on the matching degree between the first target object in the first image and the second target object in the second image, the bipartite graph between the first target object and the second target object comprises: establishing, based on a first target object and a second target object between which a matching degree is greater than a first threshold, the bipartite graph between the first target object and the second target object. The above-mentioned configuration can simplify the structure of the bipartite graph, thereby improving the matching efficiency.

In some possible implementations, determining, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other comprises: selecting, based on the bipartite graph between the first target object and the second target object with a greedy algorithm in a descending order of the matching degree between the first target object and the second target object, a preset number of the second target objects best matched with the first target object as second target objects matched with the first target object. The above-mentioned configuration can conveniently and accurately determine the matched target objects.

In some possible implementations, determining, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other further comprises: in response to the bipartite graph between the first target object and the second target object being a bipartite graph between the human body and the hand, selecting, with a greedy algorithm, at most two second target objects with a type of a hand and best matched with the first target object; and in response to the bipartite graph between the first target object and the second target object being a bipartite graph between the human body and the face, selecting, with a greedy algorithm, a second target object with a type of a face and best matched with the first target object. Based on the configuration above, it is possible to adaptively set different matching values for the second target objects of different types, thereby providing better adaptability.

In some possible implementations, determining, based on the bipartite graph between the first target object and the second target object, the first target object and a second target object that are matched with each other further comprises: in response to the preset number of second

target objects having been determined to be matched with a first target object, ceasing to match other second target objects with the first target object; and in response to a matched first target object has been determined for any second target object, ceasing to match other first target objects with the second target object. The above-mentioned configuration can reduce the probability of matching the same target object with a plurality of target objects, thereby improving the matching accuracy.

In some possible implementations, the acquiring the first target object and the second target object to be matched from the input image comprises at least one of: determining the first target object and the second target object in the input image in response to detecting a box selection of the first target object and the second target object in the input image; detecting the first target object and the second target object in the input image with a target detection neural network; and receiving position information of the first target object and the second target object in the input image to determine the first target object and the second target object in the input image based on the position information. Based on the above-mentioned configuration, a target object to be matched may be determined in different ways, thereby providing better user experience.

In some possible implementations, before performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, the method further comprises: adjusting the first image and the second image to a preset specification respectively; and performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image comprises: performing feature processing on the first and second images adjusted to the preset specification, to obtain the matching degree between the first target object in the first image and the second target object in the second image. The above-mentioned configuration can be applicable to images of different specifications.

In some possible implementations, the method further comprises: displaying the first target object and the second target object that are matched with each other in the input image. The above-mentioned configuration can display a matching result intuitively, thereby providing better user experience.

In some possible implementations, the method further comprises: performing, by a Siamese neural network, the feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object respectively, to obtain the matching degree between the first target object in the first image and the second

target object in the second image. The above-mentioned configuration can improve the accuracy of the feature processing, thereby further improving the matching degree.

In some possible implementations, the method further comprises a step of training the Siamese neural network comprising: obtaining a training sample including a plurality of first training images that are human body images and a plurality of second training images that are face images or hand images; providing the first training image and the second training image as an input to the Siamese neural network to obtain a predicted matching result between the first training image and the second training image; and determining a network loss based on the predicted matching result between the first training image and the second training image to adjust a network parameter of the Siamese neural network according to the network loss until a training requirement is met. The above-mentioned configuration can optimize the Siamese neural network, thereby improving the matching accuracy.

According to a second aspect of the present disclosure, there is provided a target object matching device, comprising:

an acquisition module configured to acquire a first target object and a second target object to be matched from an input image, wherein the first target object includes a human body, and the second target object includes at least one of a face and a hand;

a feature processing module configured to perform feature processing on a first image corresponding to the first target object and a second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image;

a bipartite module configured to establish, based on the matching degree between the first target object in the first image and the second target object in the second image, a bipartite graph between the first target object and the second target object; and

a matching module configured to determine, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other.

In some possible implementations, the feature processing module is further configured to: perform feature extraction processing on the first image and the second image to obtain a first feature of the first image and a second feature of the second image, respectively; and

perform classification processing on a concatenated feature of the first feature and the second feature to obtain the matching degree between the first target object in the first image and the second target object in the second image.

In some possible implementations, the feature processing module is further configured to:

perform feature fusion processing on a concatenated feature of the first feature and the second feature to obtain a fusion feature; and

obtain the matching degree between the first target object in the first image and the second target object in the second image by providing the fusion feature as an input to a fully
5 connected layer to execute the classification processing on the fusion feature.

In some possible implementations, the bipartite module is further configured to:

if the second target object includes only the face, establish a bipartite graph between the human body and the face based on the matching degree between the first target object in the first image and the second target object in the second image;

10 if the second target object includes only the hand, establish a bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image; and

if the second target object includes the face and the hand, establish the bipartite graph between the human body and the face and the bipartite graph between the human body and the
15 hand based on the matching degree between the first target object in the first image and the second target object in the second image,

wherein a matching degree between the human body and the face is used as a concatenation weight between the human body and the face in the bipartite graph between the human body and the face, and a matching degree between the human body and the hand is used
20 as a concatenation weight between the human body and the hand in the bipartite graph between the human body and the hand.

In some possible implementations, the bipartite module is further configured to: establish, based on a first target object and a second target object between which a matching degree is greater than a first threshold, the bipartite graph between the first target object and the second
25 target object.

In some possible implementations, the matching module is further configured to: select, based on the bipartite graph between the first target object and the second target object with a greedy algorithm in a descending order of the matching degree between the first target object and the second target object, a preset number of the second target object best matched with the
30 first target object as second target objects matched with the first target object.

In some possible implementations, the matching module is further configured to: select, with a greedy algorithm, a second target object with a type of a face and best matched with the first target object if the bipartite graph between the first target object and the second target object is a bipartite graph between a human body and the face.

In some possible implementations, the matching module is further configured to: if the preset number of second target objects have been determined to be matched with a first target object, ceasing to match other second target objects with the first target object; and

if a matched first target object has been determined for any second target object, ceasing to match other first target objects with the second target object.

In some possible implementations, the acquisition module is configured to acquire the first target object and the second target object to be matched from an input image by at least one of:

determining the first target object and the second target object in the input image in response to detecting a box selection of the first target object and the second target object in the input image;

detecting the first target object and the second target object in the input image with a target detection neural network; and

receiving position information of the first target object and the second target object in the input image to determine the first target object and the second target object in the input image based on the position information.

In some possible implementations, the feature processing module is further configured to: before performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, adjust the first and second images to a preset specification respectively; and

performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image comprises:

performing feature processing on the first and second images adjusted to the preset specification, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

In some possible implementations, the device further comprises a display module configured to display, in the input image, the first target object and the second target object that are matched with each other.

In some possible implementations, the feature processing module is further configured to: perform, by a Siamese neural network, the feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

In some possible implementations, the device further comprises a training module, configured to train the Siamese neural network, wherein a step of training the Siamese neural network comprises: obtaining a training sample including a plurality of first training images that are human body images and a plurality of second training images that are face images or
5 hand images;

obtaining a predicted matching result between the first training image and the second training image by providing the first training image and the second training image as an input to the Siamese neural network; and

determining a network loss based on the predicted matching result between the first
10 training image and the second training image to adjust a network parameter of the Siamese neural network according to the network loss until a training requirement is met.

According to a third aspect of the present disclosure, there is provided an electronic apparatus, comprising:

a processor; and

15 a memory configured to store processor-executable instructions;

wherein the processor is configured to invoke instructions stored in the memory to execute the method according to any implementation of the first aspect.

According to a fourth aspect of the present disclosure, there is provided a computer readable storage medium having computer program instructions stored thereon, wherein the
20 computer program instructions, when executed by a processor, implement the method according to any implementation of the first aspect.

According to a fifth aspect of the present disclosure, there is provided a computer readable code, wherein when the computer readable code is executed in an electronic apparatus, a processor of the electronic apparatus executes the method according to any implementation of
25 the first aspect.

In an embodiment of the present disclosure, it is possible to firstly acquire a first image of a first target object to be matched and a second image of a second target object to be matched, wherein the first target object may be a human body, and the second target object may be a face and/or a hand; then to obtain a matching degree between the first target object in the first image
30 and the second target object in the second image by performing feature processing on the first image and the second image; and to determine a matching result between the first target object in the first image and the second target object in the second image in a manner of establishing a bipartite graph between the first target object and the second target object. The embodiment of the present disclosure firstly detects a matching degree between each of first target objects and
35 each of second target objects, then constrains the detected matching degree in a manner of

establishing a bipartite graph between the first and second target objects, and finally determines the second target object matched with the first target object, such that the final matching result is more accurate.

It should be understandable that the general description above and the following detailed description are merely exemplary and explanatory, instead of restricting the present disclosure.

Additional features and aspects of the present disclosure will become apparent from the following detailed description of exemplary embodiments with reference to the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The drawings herein, which are incorporated in and constitute part of the specification, may illustrate embodiments in line with the present disclosure and be used to explain the technical solutions of the present disclosure together with the specification.

FIG. 1 shows a flowchart of a target object matching method according to an embodiment of the present disclosure.

FIG. 2 shows a schematic diagram for position areas of target objects in an input image according to an embodiment of the present disclosure.

FIG. 3 shows a flowchart of obtaining a matching degree between a first target object and a second target object by a neural network according to an embodiment of the present disclosure.

FIG. 4 shows a schematic structural diagram for a Siamese neural network according to an embodiment of the present disclosure.

FIG. 5 shows a schematic diagram for a bipartite graph established between a human body and a hand and a matching result thereof according to an embodiment of the present disclosure.

FIG. 6 shows a flowchart of training a Siamese neural network according to an embodiment of the present disclosure.

FIG. 7 shows a block diagram for a target object matching device according to an embodiment of the present disclosure.

FIG. 8 shows a block diagram for an electronic apparatus according to an embodiment of the present disclosure.

FIG. 9 is a block diagram showing another electronic apparatus according to an embodiment of the present disclosure.

DETAILED DESCRIPTION

Various exemplary embodiments, features and aspects of the present disclosure will be described in detail with reference to the drawings. The same reference numerals in the drawings represent parts having the same or similar functions. Although various aspects of the
5 embodiments are shown in the drawings, it is unnecessary to proportionally draw the drawings unless otherwise specified.

Herein the specific term “exemplary” means “used as an example, or embodiment, or explanatory”. An “exemplary” embodiment given here is not necessarily construed as being superior to or better than other embodiments.

10 The term “and/or” used herein represents only an association relationship for describing associated objects, and represents three possible relationships. For example, A and/or B may represent the following three cases: A exists alone, both A and B exist, and B exists alone. In addition, the term “at least one” used herein indicates any one of multiple listed items or any combination of at least two of multiple listed items. For example, including at least one of A, B,
15 and C may indicate including any one or more elements selected from the group consisting of A, B, and C.

In addition, numerous details are given in the following specific embodiments for the purpose of better explaining the present disclosure. It should be understood by a person skilled in the art that the present disclosure can still be realized even without some of those details. In
20 some of the examples, methods, means, units and circuits that are well known to a person skilled in the art are not described in detail so that the principle of the present disclosure becomes apparent.

In an embodiment of the present disclosure, there is provided a target object matching method, which can conveniently determine whether objects in two images are matched with
25 each other or not, for example, detect whether a face is matched with a human body, or whether a hand is matched with a human body. This method may be applicable to an arbitrary image processing apparatus, such as an electronic apparatus or a server. The electronic apparatus may include terminal apparatuses such as a cellphone, a laptop, and PAD, or wearable apparatuses such as a smart bracelet and a smart watch, or other handheld apparatuses, or the like. The
30 server may include a cloud server or a local server, etc. As long as an apparatus is capable of executing image processing, it can be used to execute the target object matching method according to an embodiment of the present disclosure.

FIG. 1 shows a flowchart of a target object matching method according to an embodiment of the present disclosure. As shown in FIG. 1, the target object matching method may comprise
35 the following steps:

Step S10 of acquiring a first target object and a second target object to be matched from an input image, wherein the first target object includes a human body, and the second target object includes at least one of a face and a hand.

In some possible implementations, an embodiment of the present disclosure can realize a matching between a face and a human body as well as a matching between a hand and a human body, namely, determine whether the face and the human body in an input image belong to the same person, and whether the hand and the human body belong to the same person, so as to realize a matching among the face, the hand, and the human body of each person. An image of the target object to be matched in the input image may be acquired first. The target object may include a human body, and at least one of a face and a hand. For example, a target detection processing can be performed on the input image to detect target objects in the input image. That is, a first target object and a second target object to be matched in the input image are obtained firstly, e.g., obtaining positions of the first target object and the second target image in the input image. Furthermore, an image area corresponding to the first target object and an image area corresponding to the second target object can be determined. The first target object includes a human body, and the second target object includes at least one of a face and a hand.

Step S20 of performing feature processing on a first image corresponding to the first target object and a second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image.

In some possible implementations, in the case of obtaining the first target object and the second target object to be matched in the input image, namely, in the case of acquiring the positions where the first target object and the second target object to be matched are in the input image respectively, image areas corresponding to the first target object and the second target object in the input image can be determined, namely, it is possible to determine a first image corresponding to the position of the first target object in the input image, and a second image corresponding to the position of the second target object in the input image, wherein the first image and the second image take up a portion of image area in the input image, respectively.

In the case of obtaining the first image and the second image, feature processing may be performed on the first image and the second image respectively to detect the matching between the first target object in the first image and the second target object in the second image, thereby obtaining a matching degree between the first target object and the second target object.

In some possible implementations, the above-mentioned matching degree between the first target object and the second target object can be acquired by a neural network. Image features of the first image and the second image may be obtained respectively, to further determine the matching degree between the first target object and the second target object according to the image features. In an example, the neural network may include a feature extraction module, a feature fusion module, and a fully connected module. The feature extraction module can perform feature extraction processing on inputted first and second images. The feature fusion module can realize a feature fusion of feature information of the first and second images. The fully connected module can obtain a binary classification result of the first target object and the second target object, that is, the matching degree between the first target object and the second target object. The matching degree may be a value equal to or greater than 0 and less than or equal to 1. The greater the matching degree, the higher the probability that the first target object and the second target object belong to the same person.

In an example, the neural network may be a Siamese neural network, in which the feature extraction module may include two feature extraction branches, for which processing operations and parameters are all the same. The feature information of the first and second images can be extracted, respectively, with the two feature extraction branches. The Siamese neural network enables the detection of the matching degree, which can improve the accuracy of the detected matching degree.

Step S30: establishing, based on the matching degree between the first target object in the first image and the second target object in the second image, a bipartite graph between the first target object and the second target object.

In some possible implementations, in the case of obtaining the matching degree between the first target object and the second target object, a bipartite graph between the first target object and the second target object may be established. The input image may include at least one person, and include at least one first target object and at least one second target object. Based on the matching degree between each of the first target objects and each of the second target objects, a bipartite graph between the first target object and the second target object may be established. The first target objects and the second target objects may be used as two sets of points in the bipartite graphs respectively, and the matching degree between the first target object and the second target object is used as a concatenation weight between the two sets of points.

For example, different bipartite graphs may be established depending upon the type of the second target object. If the type of the second target object is a face, a resulting bipartite graph is a bipartite graph between a human body and the face. If the type of the second target object

is a hand, a resulting bipartite graph is a bipartite graph between a human body and the hand. If the second target object includes a face and a hand, resulting bipartite graphs includes a bipartite graph between a human body and the face and a bipartite graph between a human body and the hand.

5 Step S40 of determining, based on the bipartite graph between the first target object and the second target object, a first target object and a second target object that are matched with each other.

In some possible implementations, in the case of obtaining the bipartite graph between the first target object and the second target object, it is possible to determine a second target object
10 matched with the first target object according to the bipartite graph, namely, determine a second target object belonging to the same person as the first target object.

As described above, the concatenation weight between the first target object and the second target object in the bipartite graph is the matching degree therebetween. In an embodiment of the present disclosure, the second target objects matched with the first target
15 object may be determined in a descending order of the matching degree.

In an example, if the bipartite graph is the bipartite graph between a human body and a face, it is possible to determine a face (i.e. a second target object) best matched with each human body (i.e. a first target object) in the descending order of the matching degree. If the bipartite graph is the bipartite graph between a human body and a human body, it is possible to
20 determine at most two hands (i.e. second target objects) best matched with each human body (i.e. a first target object) in the descending order of the matching degree.

In an embodiment of the present disclosure, the second target object matched with the first target object as described above may be obtained with a greedy algorithm. If a first target object has been matched with a corresponding second target object, the matching of other
25 objects with the first target object and the second target object will be ceased.

Based on the above-mentioned configuration, it is possible in an embodiment of the present disclosure to predict the matching degree between each of the first target objects and each of the second target objects in the input image, and then determine a matching result between the first target object and the second target object in a manner of establishing the
30 bipartite graph between the first target object and the second target object, so as to obtain a matching result with higher accuracy.

The embodiments of the present disclosure will be illustrated in detail below with reference to the drawings. In the embodiments of the present disclosure, input images may be obtained firstly, wherein the input images may be arbitrary images including a person. The
35 input image may be obtained by at least one of following actions: capturing an input image by

an image capture apparatus, receiving an input image transmitted by other apparatuses, and reading an input image from a memory. The image capture apparatus may be any apparatus with an image capture function, such as a camera, a video camera, a cellphone, or a computer, but the present disclosure is not specifically limited thereto. In addition, the memory may be a local memory or a cloud memory. The foregoing is only an exemplary description of how to obtain the input image. In other embodiments, the input image may also be obtained by other methods, but the present disclosure is not specifically limited thereto.

When an input image is obtained, it is possible to further obtain a first target object and a second target object to be matched in the input image, e.g., obtain the position areas where the first target object and the second target object are located. In an embodiment of the present disclosure, an input image may be provided as an input to a neural network capable of realizing detection of a target object that may include a human body, a face, and a hand. For example, an input image may be provided as an input to a neural network capable of detecting a target object, to obtain a position area where the first target object is located and a position area where the second target object is located in the input image with the detection of the target object by the neural network, wherein the position areas of the first target objects and the second target objects may be indicated with detection boxes in the input image. In addition, information of types (a human body, a face, or a hand) of the target objects corresponding to detection boxes may be included. The position areas where the first target object and the second target object are located can be determined based on the positions corresponding to the above-mentioned detection boxes. Types of the first target object and the second target object can be determined from identifiers. For example, the neural network used for detecting the target object in an embodiment of the present disclosure may be either a region proposal network (RPN) or a target recognition convolutional neural network (RCNN), but the present disclosure is not limited thereto. In this way, all of the first target objects and the second target objects in the input image can be easily and accurately identified.

In some possible implementations, the first target object and the second target object in the input image may also be determined according to the performance of a box selection on the input image. That is, the box selection is performed by a user in an embodiment of the present disclosure to select a first target object and a second target object to be matched from the input image with a box, namely, select the position areas corresponding to the first target object and the second target object with a box. The position areas determined by the box selection may have a rectangular shape or other shapes, but the present disclosure is not limited thereto. A type of an object (such as a human body, a face, or a hand) can be identified in each selected position area while the box selection is performed. In this way, the first target object and the

second target object to be matched may be determined based on the selection of the target objects by the user. For instance, at least one of the first target objects and at least one of the second target objects in the input image may be used as the first target object and the second target object to be matched, which provides better flexibility and applicability.

5 In some possible implementations, the position areas corresponding to the first target object and the second target object can be determined by directly receiving the position information of the first target object and the second target object, such as vertex coordinates of corresponding position areas of the first target object and the second target object and their height values. Alternatively, the position areas of the first target object and the second target
10 object in the input image can be determined by receiving coordinates of two vertex angles corresponding to the respective position areas of the first target object and the second target object, thereby obtaining the first target object and the second target object in the input image. The foregoing description is only an exemplary explanation. In other embodiments, the position information of the position areas may also be indicated in other ways. In this manner,
15 the first target object and the second target object to be matched can be determined based on the position information sent by the user. For example, at least one of the first target objects and at least one of the second target objects in the input image may be used as the first target object and the second target object to be matched, which provides better flexibility and applicability.

20 According to the above-mentioned configuration, it is possible to determine the position areas where the target objects are located in the input image to obtain a first image of each of the first target objects and a second image of each of the second target objects in the input image based on the position areas. FIG. 2 shows a schematic diagram for position area of each of target objects in an input image according to an embodiment of the present disclosure. In
25 FIG. 2, A1 and B1 represent the position areas of the first target objects A and B respectively, wherein the first target objects are human bodies. A2 and B2 represent the position areas of the second target objects with a type of a face respectively, and A3 and A4 represent the position areas of the second target objects with a type of a hand respectively. In FIG. 2, all of the bodies, faces and hands may be used as the first target object and the second target object to be
30 matched. In an embodiment of the present disclosure, it is also possible to use only a portion of the first and second target objects in the input image as the first target object and the second target object to be matched. Particular examples are not illustrated here.

In the case of obtaining the first target object and the second target object to be matched, it is possible to perform feature processing on the image areas corresponding to the first target
35 object and the second target object to predict a matching degree between the first target object

and the second target object. In an embodiment of the present disclosure, it is possible to perform the aforesaid feature processing by a neural network to obtain the corresponding matching degree between the first target object and the second target object. FIG. 3 shows a flowchart of obtaining a matching degree between a first target object and a second target object by a neural network according to an embodiment of the present disclosure.

As shown in FIG. 3, in an embodiment of the present disclosure, performing feature processing on a first image corresponding to the first target object and a second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image may comprise the following steps:

Step S21 of performing feature extraction processing on the first image and the second image to obtain a first feature of the first image and a second feature of the second image, respectively.

In some possible implementations, the feature extraction processing may be performed on the image areas of the first target object and the second target object in the input image, wherein the first image refers to the image area corresponding to the position of the first target object, and the second image refers to the image area corresponding to the position of the second target object. In the case of determining the first image and the second image, feature extraction processing may be performed on the first image and the second image. The feature extraction processing may be performed by a feature extraction module of the neural network. The feature extraction module may include a feature extraction branch, which can be used to perform the feature extraction processing on the first image and the second image, respectively. In the case where there are a plurality of first target objects and a plurality of second target objects, it is possible to perform the feature extraction processing on a plurality of first images and a plurality of second images. In addition, the feature extraction module may alternatively include two feature extraction branches. The two feature extraction branches may or may not share the same network structure, and can be applied in the embodiments of the present disclosure as long as they are capable of performing feature extraction. In the case where there are two feature extraction branches, the first image and the second image may be inputted in one-to-one correspondence into two feature extraction branches, respectively. For example, one feature extraction branch is used to perform the feature extraction processing on the first image to obtain a first feature corresponding to the first image, and the other feature extraction branch is used to perform the feature extraction processing on the second image to obtain a second feature corresponding to the second image. In other embodiments, it is also possible to include at least three feature extraction branches configured to perform the feature extraction

processing on the first images and the second images, but the present disclosure is not limited thereto. According to the above-mentioned embodiment, it is possible to precisely implement feature processing and determine the matching degree.

Here the Siamese neural network is taken as an example for illustration. FIG. 4 shows a schematic structure diagram for a Siamese neural network according to an embodiment of the present disclosure. The feature extraction module according to an embodiment of the present disclosure may include two feature extraction branches, and the two feature extraction branches of the Siamese neural network have completely the same structure and parameters. The feature extraction branches may include a residual network. That is, the feature extraction module according to an embodiment of the present disclosure may be composed of the residual network which may be used to perform the feature extraction processing on the first image and the second image to extract feature information from the first and second images. The residual network may be a resnet 18, but the present disclosure is not limited thereto. Besides, the feature extraction module may also be another network module capable of performing feature extraction, which is not specifically limited in the present disclosure. As shown in FIG. 4, the first image I1 may be an image corresponding to a human body area, and the second image I2 may be a second image corresponding to a face area or to a hand area. In the case where there are a plurality of the first images and a plurality of the second images, each of the first images and each of the second images may be inputted into two feature extraction branches to perform the feature extraction processing on the inputted first and second images. Alternatively, an embodiment of the present disclosure may input only one image each time to the feature extraction branches respectively to perform the feature extraction on the inputted two images, and then in the case of obtaining the matching degree of the target objects in the two images, input the first image and the second image required to be detected for the matching degree next time.

In addition, an embodiment of the present disclosure may also assign an identifier to each image, and may identify types of target objects included in the image at the same time. That is, in an embodiment of the present disclosure, each of the first images and the second images may include an image identifier and a type identifier used to discriminate respective images and the types of target objects in the images in subsequent processing.

In addition, in some possible implementations, after the first image of each of the first target objects and the second image of each of the second target objects are obtained, the first images and the second images may be adjusted as images with preset specifications. For example, by zooming out or in, up-sampling, or down-sampling or the like, the first image and the second image may be adjusted to a preset specification, e.g., 224*224 (the present

disclosure is not limited thereto), and thereafter the first and second images adjusted to the preset specification are provided as an input to the neural network to perform feature extraction on the first and second images to obtain a first feature corresponding to the first image and a second feature corresponding to the second image.

5 Step S22 of performing classification processing on a concatenated feature of the first feature and the second feature to obtain the matching degree between the first target object in the first image and the second target object in the second image.

10 In some possible implementations, an embodiment of the present disclosure may perform feature fusion processing on a concatenated feature of the first feature and the second feature to obtain a fusion feature; and provide the fusion feature as an input to a fully connected layer to execute the classification processing on the fusion feature, so as to obtain the matching degree between the first target object in the first image and the second target object in the second image.

15 The first feature and the second feature obtained in an embodiment of the present disclosure may be expressed in the form of a matrix or a vector respectively, and the first feature and the second feature may have the same scale. Then, the first and second features may be concatenated by a concatenated function (such as a concat function), for example, in the channel direction to obtain a concatenated feature. In the case of obtaining the concatenated feature of the first feature and the second feature, feature fusion processing can be performed on the concatenated feature. For example, the feature fusion processing may be implemented by at least one layer of convolution. For instance, in an embodiment of the present disclosure, a resnet_block may be used to perform resnet processing on the concatenated feature and then perform the feature fusion processing to obtain a fusion feature. Thereafter, classification prediction of the matching degree is performed based on the fusion feature, which can obtain the classification result of whether the first target object and the second target object are matched with each other, and the corresponding matching degree between the first target object and the second target object.

25 In an example, the classification prediction of the matching degree may be implemented through a fully connected (FC) layer. That is, the fusion feature may be provided as an input to the fully connected layer, and it is possible to output the above prediction results, i.e., the matching degree between the first target object and the second target object and the matching result determined based on the matching degree based on the processing of the fusion feature by the fully connected layer. In the case where the matching degree is greater than a first threshold, it can be determined that the first target object is matched with the second target object and the matching result may be identified with a first identifier, such as “1”. In the case

where the matching degree is less than the first threshold, however, it can be determined that the first target object and the second target object are not matched with each other and the matching result may be identified with a second identifier, such as “0”. The above-mentioned first identifier and second identifier may be different with each other, and are used to indicate the respective matching results that the first target object and the second target object belong to the same person and do not belong to the same person.

In the case of obtaining the matching degree between each of the first target objects and each of the second target objects to be matched in the input image, a bipartite graph between the first target object and the second target object may be established according to the obtained matching degree.

An expression of $G=(V, E)$ represents an undirected graph, wherein a set of vertices may be divided into two mutually disjoint subsets, and two vertices attached to each edge in the graph belong to these two mutually disjoint subsets. In an embodiment of the present disclosure, the first target objects and the second target objects can be constructed as vertex sets V and E in the bipartite graph, respectively, and the connection between the vertices, i.e., each edge in the bipartite graph, may be used to indicate the matching degree between the first target object and the second target object corresponding to the two vertices respectively.

In some possible implementations, a corresponding bipartite graph may be established based on the type of the second target object to be matched in the input image. For instance, when the second target object to be matched in the input image includes only a face, a bipartite graph between a human body and the face may be established based on the matching degree between the first target object in the first image and the second target object in the second image; when the second target object to be matched in the input image includes only a hand, a bipartite graph between a human body and the hand may be established based on the matching degree between the first target object in the first image and the second target object in the second image; and when the second target object to be matched in the input image includes a face and a hand, a bipartite graph between a human body and a face and a bipartite graph between a human body and a hand may be established based on the matching degree between the first target object in the first image and the second target object in the second image. That is, each of the first target objects and the second target object with a type of a hand may be used to establish a bipartite graph between a human body and a hand, and each of the first target objects and the second target object with a type of a face may be used to establish a bipartite graph between a human body and the face. In each the dipartite graphs, the matching degree between a human body and a face may be used as a concatenation weight between the human body and the face in the bipartite graph between the human body and the face, and the

matching degree between a human body and a hand may be used as a concatenation weight between the human body and the hand in the bipartite graph between the human body and the hand.

To put it differently, in an embodiment of the present disclosure, the first target object and the second target object may be used as point sets of the vertices in the bipartite graph, and the point sets may be classified as three kinds, i.e. a human body, a face and a hand. Furthermore, it is possible to establish a bipartite graph between a human body and a face and a bipartite graph between a human body and a hand, respectively. A weight of an edge between two vertices may be used as the matching degree between the first target object and the second target object corresponding to the two vertices outputted by the neural network.

It should be noted here that in an embodiment of the present disclosure, in the case of obtaining the matching degree between each of the first target objects and each of the second target objects, it is possible to select each of the first target objects and each of the second target objects between which a matching degree is greater than the first threshold, and to determine the bipartite graph between the first target object and the second target object based on the first target object and the second target object between which a matching degree is greater than the first threshold.

As for each of the first target objects, if there is a second target object whose matching degrees with all of the first target objects are less than the first threshold, this second target object is not used to form a bipartite graph. Conversely, if there is a first target object whose matching degrees with all of the second target objects with a type of a face are less than the first threshold, this first target object is not used to form a bipartite graph between a human body and a face; and if there is a first target object whose matching degrees with all of the second target objects with a type of a human body are less than the first threshold, this first target object is not used to form a bipartite graph between a human body and a hand.

A structure of the bipartite graph can be simplified by setting the first threshold, thereby improving the matching efficiency between the first target object and the second target object.

In the case of obtaining the bipartite graph between the first target object and the second target object, it is possible to obtain, based on the bipartite graph between the first target object and the second target object with the greedy algorithm, at most a preset number of the second target objects matched with the first target objects with each type of a human body. The preset number may vary for the second target objects of different types. For example, in the case where the second target object is a hand, the preset number may be 2; and in the case where it is a face, the preset number may be 1. Specifically, different preset numbers may be selected

depending upon different types of a target object, but the present disclosure is not limited thereto.

At most a preset number of the second target objects matched with the first target object may be determined in a descending order of the matching degree between each of the second target objects and the first object. In an embodiment of the present disclosure, a greedy algorithm may be utilized to determine whether a first target object is matched with a second target object or not. That is, a second target object is matched with a first target object in the descending order of the matching degree. If the number of the second target object matched with a first target object reaches the preset number, the process for matching the second target object with this first target object will be terminated, namely, the matching of the first target object with any of the other second target objects will be ceased. In addition, if a matched first target object is determined for any second target object, the process for matching the second target object with any of the other first target objects will be terminated, namely, the matching of the second target object with any of the other first target objects will be ceased.

In some possible implementations, in the process of determining second target objects matched with a first target object in the descending order of the matching degree, if iteration is performed until a matching degree between a first target object and a second target object is less than the first threshold, the matching process can be terminated at this time. A bipartite graph between a human body and a face can be illustrated, for example. It is assumed that the descending order of the matching degrees may be arranged as: 90% of the matching degree between X1 and Y1, 80% of the matching degree between X2 and Y2, 50% of the matching degree between X2 and Y1, and 30% of the matching degree between X1 and Y2, and the first threshold may be 60%, wherein X1 and X2 represent two first target objects respectively, and Y1 and Y2 represent two second target objects respectively. According to the order of the matching degree, the first target object X1 and the second target object Y1 having a matching degree of 90% may be determined to be matched, and the first target object X2 and the second target object Y2 having a matching degree of 80% are determined to be matched, but then the next matching degree is 50% and thus less than the first threshold, so the matching process can be terminated. It can thus be determined that the faces matched with the first target objects X1 and X2 are Y1 and Y2, respectively.

The foregoing is only an exemplary explanation of terminating the matching process by setting the first threshold, but the present disclosure is not specifically limited thereto. In other embodiments, it is also possible to match at most a preset number of the second target objects with each of the first target objects based on only the descending order of the matching degrees between the first target objects and the second target objects. The at most preset number of the

second target objects used herein indicates that each person may be matched with two hands when the second target object is a hand, but the first target object is possibly matched with only one second target object with a type of a hand due to the setting of the first threshold and the number of the second target objects in the input image during the matching process.

Hereafter the second target with a type of a hand is taken as an example for illustration. FIG. 5 shows a schematic diagram for a bipartite graph established between a human body and a hand and a matching result therebetween according to an embodiment of the present disclosure. FIG. 5 represents a bipartite graph between a human body and a hand established based on the matching degree between the first target object and the second target object. The human body and the hand may be used as sets of two categories of vertices of the bipartite graph, respectively. P1, P2 and P3 represent three first target objects, namely, three human bodies, respectively. H1, H2, H3, H4 and H5 represent five second target objects with a type of a hand, respectively. A connection line between any of the first target objects and any of the second target objects may indicate the matching degree between the first target object and the second target object.

Based on the bipartite graph between the human body and the hand, it is possible to assign matched second target object(s) to each of the first target objects in the descending order of the matching degree, wherein at most two second target objects are matched with each of the first target objects. If a second target object is determined to be matched with a first target object in the descending order of the matching degree, the matching of the second target object with the other first target objects may be ceased, while determining whether the number of the second target objects matched with the first target object reaches the preset number. If the number of the second target objects matched with the first target object reaches the preset number, the matching of the other second target objects with the first target object may be ceased. If the number of the second target objects matched with the first target object does not reach the preset number, a second target object with the next matching degree in the descending order of the matching degree may be matched with the corresponding first target object, to determine whether the second target object is the one matched with the other first target objects, and whether the number of the second target objects matched with the first target object reaches the preset number. If the second target object is not matched with any of the first target objects, and the number of the second target object matched with the first target object is less than the preset number, it is determined that the first target object and the second target object are matched with each other. Analogically, the above-mentioned process may be iterated repeatedly on the first target object and the second target object corresponding to each matching degree until the termination condition is met. The termination condition may include at least one of: matching

each of the first target objects with a corresponding second target object, performing the above matching process based on the first target object and the second target object between which a matching degree is the minimum matching degree, and a matching degree being less than the first threshold.

5 The process of determining a second target object matched with the first target object based on the bipartite graph between a human body and a face, which will not be repeated here, is similar to the aforementioned process.

10 Furthermore, in an embodiment of the present disclosure, in the case of obtaining a second target object matched with each of first target objects, it is possible to display the position areas of the first target object and the second target object that are matched with each other. For example, in an embodiment of the present disclosure, boundary boxes of the position areas of the first target object and the second target object that are matched with each other may be displayed in the same display status, and the boundary boxes may be detection boxes of the position areas obtained in S10. In an example, the boundary boxes of the position areas of
15 the first target object and the second target object that are matched with each other may be displayed in the same color, but the present disclosure is not limited thereto. As shown in FIG. 2, for each person, the human body boxes, the hand boxes, and the face boxes corresponding to different persons may be discriminated by line widths of the displayed boxes, for example, so as to conveniently discriminate the matching results.

20 Based on the above-mentioned configuration in an embodiment of the present disclosure, the second target object best matched with each of the first target objects may be selected in a manner of establishing a bipartite graph between the second target object and the first target object, to improve the matching accuracy between the target objects.

25 As described above, the embodiments of the present disclosure may be applied to a neural network, for example, a Siamese neural network. For instance, in an embodiment of the present disclosure, it is possible to perform, by the Siamese neural network, feature processing on the first image corresponding to the position area of the first target object and on the second image corresponding to the position area of the second target object respectively, to obtain the matching degree between the first target object in the first image and the second target object in
30 the second image.

FIG. 6 shows a flow chart of training a Siamese neural network according to an embodiment of the present disclosure. The step of training the Siamese neural network may comprise the following steps:

Step S51 of obtaining a training sample including a plurality of first training images that are human body images and a plurality of second training images that are face images or hand images.

In some possible implementations, the first training image and the second training image may be image areas extracted from a plurality of images, or image areas of target objects with the corresponding type identified by target detection from a plurality of images, or any images including a human body, a hand, or a face, but the present disclosure is not limited thereto.

Step S52 of obtaining a predicted matching result between the first training image and the second training image by providing the first training image and the second training image as an input to the Siamese neural network.

In some possible implementations, the Siamese neural network is utilized to perform the feature extraction, feature concatenation, feature fusion and classification processing on the first training image and the second training image, to finally predict the matching degree between the first training image and the second training image, thereby determining the matching result between the first training image and the second training image according to the matching degree. The matching result may be represented by a first identifier and a second identifier. For example, the first identifier may be 1 and the second identifier may be 0, which are used to indicate the matching result of whether the first training image and the second training image are matched with each other or not. Specifically, the matching result may be determined according to the comparison result between the matching degree and the first threshold. If the matching degree is greater than the first threshold, the corresponding first training image and second training image are determined to be matched with each other, and the matching result may be indicated by the first identifier, otherwise the matching result may be indicated by the second identifier.

Step S53 of adjusting, based on the predicted matching result between the first training image and the second training image, a network parameter of the Siamese neural network until a training requirement is met.

In an embodiment of the present disclosure, the real matching result between the first training image and the second training image may be used for supervision, so as to determine a network loss according to the predicted matching result and the real matching result between the first training image and the second training image, i.e., determining the network loss based on a difference between the two matching results.

In the case of obtaining the network loss, parameters of the Siamese neural network such as a convolution parameter may be adjusted according to the network loss. In the case where the obtained network loss is less than a loss threshold, it is determined that the training

requirement is met, and thus the training can be terminated. If the obtained network loss is equal to or greater than the loss threshold, the network parameters are adjusted according to the network loss, and the matching result between each of the first training images and each of the second training images are predicted again until the obtained network loss is less than the loss threshold. The loss threshold may be a preset value, e.g., 1%, or may be another value, but the present disclosure is not limited thereto. Based on the above implementation, it is possible to optimize the Siamese neural network, thereby improving the feature processing and matching accuracy.

To render the embodiments of the present disclosure clearer, the specific processes thereof are illustrated below. A human body picture and a face/hand picture extracted from an input image may be resized to a fixed size, e.g., 224*224, and then the pictures are inputted into two feature extraction branches of the Siamese network, respectively. The two branches of the Siamese network extract features of the human body and the face or the hand, respectively. Finally, the two branches concatenate the extracted feature maps of the human body and the face or the hand, and then perform the binary scoring, wherein the scores range between 0 and 1. If the human body is matched with the face or the hand, the score approximates 1; otherwise, the score approximates 0. For example, in FIG. 4, the two branches of the Siamese network use resnet18 to extract features respectively; the extracted feature maps are merged, then subjected to a resnet_block convolutional layer, and finally to a fully connected layer for classification to obtain a matching degree. Thereafter, the point set is classified into three categories: a human body, a face, and a hand. Fully connected bipartite graphs are established between the human body and the face and between the human body and a hand, respectively, and the weight of the corresponding edge is the score (matching degree) outputted by the Siamese network. The bipartite graph is constrained by some rules, for example, one human body can be matched with at most two hands, and one human body can be matched with at most one face. The scores are sorted. The matching is performed with a greedy algorithm in a descending order of the scores. A plurality of edges which are unnecessary and improper are removed and the iteration processes are constantly performed until the matching process is ended. In an embodiment of the present disclosure, it is possible to learn more association relationships of complicated scenarios with the Siamese network. In addition, in an embodiment of the present disclosure, it is possible to use the bipartite graph to constrain the results outputted by the Siamese network in the final association, so that the accuracy of the final result is higher.

In conclusion, in the embodiments of the present disclosure, it is possible to obtain a first image of a first target object and a second image of a second target object to be matched, wherein the first target object may be a human body, and the second target object may be a face

and/or a hand, then to perform feature processing on the first image and the second image to obtain a matching degree between the first target object in the first image and the second target object in the second image, and to determine a matching result between the first target object in the first image and the second target object in the second image in a manner of establishing a bipartite graph. In the embodiments of the present disclosure, a matching degree between each of the first target objects and each of the second target objects is detected firstly; the above detected matching degree is constrained in a manner of establishing a bipartite graph; and finally the second target object matched with the first target object is determined, so that the accuracy of the finally associated matching result is higher.

It can be understandable that the above-mentioned method embodiments of the present disclosure may be combined with one another to form a combined embodiment without departing from the principle and the logics, which, due to limited space, will not be repeatedly described in the present disclosure.

In addition, the present disclosure further provides a target object device, an electronic apparatus, a computer readable storage medium, and a program, which are all capable of realizing any one of the target object matching methods provided in the present disclosure. For the corresponding technical solution and descriptions which will not be repeated, reference may be made to the corresponding descriptions of the method.

A person skilled in the art may understand that, in the foregoing method according to specific embodiments, the order of describing the steps does not means a strict order of execution that imposes any limitation on the implementation process. Rather, a specific order of execution of the steps should depend on the functions and possible inherent logics of the steps.

FIG. 7 shows a block diagram for a target object matching device according to an embodiment of the present disclosure. As shown in FIG. 7, the target object matching device comprises:

an acquisition module 10 configured to acquire a first target object and a second target object to be matched from an input image, wherein the first target object includes a human body, and the second target object includes at least one of a face and a hand;

a feature processing module 20 configured to perform feature processing on a first image corresponding to the first target object and a second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image;

a bipartite module 30 configured to establish, based on the matching degree between the first target object in the first image and the second target object in the second image, a bipartite graph between the first target object and the second target object; and

a matching module 40 configured to determine, based on the bipartite graph between the first target object and the second target object, a first target object and a second target object that are matched with each other.

In some possible implementations, the feature processing module is further configured to:

perform feature extraction processing on the first image and the second image to obtain a first feature of the first image and a second feature of the second image, respectively; and

perform classification processing on a concatenated feature of the first feature and the second feature to obtain the matching degree between the first target object in the first image and the second target object in the second image.

In some possible implementations, the feature processing module is further configured to:

perform feature fusion processing on a concatenated feature of the first feature and the second feature to obtain a fusion feature; and

obtain the matching degree between the first target object in the first image and the second target object in the second image by providing the fused feature as an input to a fully connected layer to execute the classification processing on the fusion feature.

In some possible implementations, the bipartite module is further configured to:

if the second target object includes only the face, establish a bipartite graph between the human body and the face based on the matching degree between the first target object in the first image and the second target object in the second image;

if the second target object includes only the hand, establish a bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image; and

if the second target object includes the face and the hand, establish the bipartite graph between the human body and the face and the bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image,

wherein a matching degree between the human body and the face is used as a concatenation weight between the human body and the face in the bipartite graph between the human body and the face, and a matching degree between the human body and the hand is used as a concatenation weight between the human body and the hand in the bipartite graph between the human body and the hand.

In some possible implementations, the bipartite module is further configured to: establish, based on a first target object and a second target object between which a matching degree is greater than a first threshold, the bipartite graph between the first target object and the second target object.

5 In some possible implementations, the matching module is further configured to: select, based on the bipartite graph between the first target object and the second target object with a greedy algorithm in a descending order of the matching degree, a preset number of the second target objects best matched with the first target object as second target objects matched with the first target object.

10 In some possible implementations, the matching module is further configured to: select, with a greedy algorithm, a second target object with a type of a face and best matched with the first target object if the bipartite graph between the first target object and the second target object is a bipartite graph between a human body and the face.

In some possible implementations, the matching module is further configured to:

15 if the preset number of second target objects have been determined to be matched with a first target object, cease to match other second target object with the first target object; and

if the second target object have been determined to be matched with a first target object, cease to match other first target objects with the second target object.

In some possible implementations, the acquisition module is configured to acquire the first target object and the second target object to be matched from an input image by at least one of:

determining the first target object and the second target object in the input image in response to detecting a box selection of the first target object and the second target object in the input image;

25 detecting the first target object and the second target object in the input image with a target detection neural network; and

receiving position information of the first target object and the second target object in the input image to determine the first target object and the second target object in the input image based on the position information.

30 In some possible implementations, the feature processing module is further configured to: before performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, adjust the first image and the second image to a preset specification respectively; and

perform feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image comprises:

5 perform feature processing on the first and second images, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

In some possible implementations, the device further comprises a display module, configured to display, in the input image, the first target object and the second target object that
10 are matched with each other.

In some possible implementations, the feature processing module is further configured to: perform, by a Siamese neural network, the feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object respectively, to obtain the matching degree between the first target object in the first image and
15 the second target object in the second image.

In some possible implementations, the device further comprises a training module, configured to train the Siamese neural network, wherein a step of training the Siamese neural network comprises:

obtaining a training sample including a plurality of first training images that are human
20 body images and a plurality of second training images that are face images or hand images;

obtaining a predicted matching result between the first training image and the second training image by providing the first training image and the second training image as an input to the Siamese neural network; and

determining a network loss based on the predicted matching result between the first
25 training image and the second training image to adjust a network parameter of the Siamese neural network according to the network loss until a training requirement is met.

In some embodiments, functions of or modules included in the device provided in the embodiments of the present disclosure may be configured to execute the method described in the foregoing method embodiments. For specific implementation of the functions or modules,
30 reference may be made to descriptions of the foregoing method embodiments. For brevity, details are not described here again.

The embodiments of the present disclosure further propose a computer readable storage medium having computer program instructions stored thereon, wherein the computer program instructions, when executed by a processor, implement the above-described method. The

computer readable storage medium may be a volatile computer readable storage medium or a non-volatile computer readable storage medium.

The embodiments of the present disclosure further propose an electronic apparatus, comprising: a processor; and a memory configured to store processor-executable instructions; wherein the processor is configured to carry out the above-described method.

The electronic apparatus may be provided as a terminal, a server, or an apparatus in other forms.

The embodiments of the present disclosure further provide a computer readable code, wherein when the computer readable code is executed in an electronic apparatus, a processor of the electronic apparatus executes the method provided above.

FIG. 8 shows a block diagram for an electronic apparatus according to an embodiment of the present disclosure. For example, the electronic apparatus 800 may be a terminal, such as a mobile phone, a computer, a digital broadcasting terminal, a message transmitting and receiving apparatus, a game console, a tablet apparatus, a medical equipment, a fitness equipment, and a personal digital assistant.

Referring to FIG. 8, the electronic apparatus 800 may include one or more components of: a processing component 802, a memory 804, a power supply component 806, a multimedia component 808, an audio component 810, an Input/Output (I/O) interface 812, a sensor component 814, and a communication component 816.

The processing component 802 is configured usually to control overall operations of the electronic apparatus 800, such as the operations associated with display, telephone calls, data communications, camera operations, and recording operations. The processing component 802 can include one or more processors 820 configured to execute instructions to perform all or part of the steps included in the above-described methods. In addition, the processing component 802 may include one or more modules configured to facilitate the interaction between the processing component 802 and other components. For example, the processing component 802 may include a multimedia module configured to facilitate the interaction between the multimedia component 808 and the processing component 802.

The memory 804 is configured to store various types of data to support the operation of the electronic apparatus 800. Examples of such data include instructions for any applications or methods operated on the electronic apparatus 800, contact data, phonebook data, messages, pictures, video, etc. The memory 804 may be implemented using any type of volatile or non-volatile memory apparatus, or a combination thereof, such as a static random access memory (SRAM), an electrically erasable programmable read-only memory (EEPROM), an erasable programmable read-only memory (EPROM), a programmable read-only memory

(PROM), a read-only memory (ROM), a magnetic memory, a flash memory, a magnetic disk, or an optical disk.

The power component 806 is configured to provide power to various components of the electronic apparatus 800. The power component 806 may include a power management system,
5 one or more power sources, and any other components associated with the generation, management, and distribution of power in the electronic apparatus 800.

The multimedia component 808 includes a screen providing an output interface between the electronic apparatus 800 and the user. In some embodiments, the screen may include a liquid crystal display (LCD) and a touch panel (TP). If the screen includes the touch panel, the
10 screen may be implemented as a touch screen to receive input signals from the user. The touch panel may include one or more touch sensors configured to sense touches, swipes, and gestures on the touch panel. The touch sensors may sense not only a boundary of a touch or swipe action, but also a period of time and a pressure associated with the touch or swipe action. In some embodiments, the multimedia component 808 may include a front camera and/or a rear
15 camera. The front camera and/or the rear camera may receive an external multimedia datum while the electronic apparatus 800 is in an operation mode, such as a photographing mode or a video mode. Each of the front camera and the rear camera may be a fixed optical lens system or may have focus and/or optical zoom capabilities.

The audio component 810 is configured to output and/or input audio signals. For example,
20 the audio component 810 may include a microphone (MIC) configured to receive an external audio signal when the electronic apparatus 800 is in an operation mode, such as a call mode, a recording mode, and a voice recognition mode. The received audio signal may be further stored in the memory 804 or transmitted via the communication component 816. In some embodiments, the audio component 810 further includes a speaker configured to output audio
25 signals.

The I/O interface 812 is configured to provide an interface between the processing component 802 and peripheral interface modules, such as a keyboard, a click wheel, buttons, and the like. The buttons may include, but are not limited to, a home button, a volume button, a starting button, and a locking button.

The sensor component 814 includes one or more sensors configured to provide status assessments of various aspects of the electronic apparatus 800. For example, the sensor component 814 may detect at least one of an open/closed status of the electronic apparatus 800, relative positioning of components, e.g., the components being the display and the keypad of the electronic apparatus 800. The sensor component 814 may further detect a change of
30 position of the electronic apparatus 800 or one component of the electronic apparatus 800,
35

presence or absence of contact between the user and the electronic apparatus 800, location or acceleration/deceleration of the electronic apparatus 800, and a change of temperature of the electronic apparatus 800. The sensor component 814 may include a proximity sensor configured to detect the presence of nearby objects without any physical contact. The sensor component 814 may also include a light sensor, such as a CMOS or CCD image sensor, for use in imaging applications. In some embodiments, the sensor component 814 may also include an accelerometer sensor, a gyroscope sensor, a magnetic sensor, a pressure sensor, or a temperature sensor.

The communication component 816 is configured to facilitate wired or wireless communication between the electronic apparatus 800 and other apparatus. The electronic apparatus 800 can access a wireless network based on a communication standard, such as WiFi, 2G, or 3G, or a combination thereof. In an exemplary embodiment, the communication component 816 receives a broadcast signal or broadcast associated information from an external broadcast management system via a broadcast channel. In an exemplary embodiment, the communication component 816 may include a near field communication (NFC) module to facilitate short-range communications. For example, the NFC module may be implemented based on a radio frequency identification (RFID) technology, an infrared data association (IrDA) technology, an ultra-wideband (UWB) technology, a Bluetooth (BT) technology, or any other suitable technologies.

In exemplary embodiments, the electronic apparatus 800 may be implemented with one or more application specific integrated circuits (ASICs), digital signal processors (DSPs), digital signal processing devices (DSPDs), programmable logic devices (PLDs), field programmable gate arrays (FPGAs), controllers, micro-controllers, microprocessors, or other electronic components, for performing the above-described methods.

In exemplary embodiments, there is also provided a non-volatile computer readable storage medium including computer program instructions, such as those included in the memory 804, executable by the processor 820 of the electronic apparatus 800, for completing the above-described methods.

FIG. 9 is another block diagram showing an electronic apparatus according to an embodiment of the present disclosure. For example, the electronic apparatus 1900 may be provided as a server. Referring to FIG. 9, the electronic apparatus 1900 includes a processing component 1922, which further includes one or more processors, and a memory resource represented by a memory 1932 configured to store instructions such as application programs executable for the processing component 1922. The application programs stored in the memory 1932 may include one or more than one module of which each corresponds to a set of

instructions. In addition, the processing component 1922 is configured to execute the instructions to execute the above-mentioned methods.

The electronic apparatus 1900 may further include a power component 1926 configured to execute power management of the electronic apparatus 1900, a wired or wireless network interface 1950 configured to connect the electronic apparatus 1900 to a network, and an Input/Output (I/O) interface 1958. The electronic apparatus 1900 may be operated on the basis of an operating system stored in the memory 1932, such as Windows ServerTM, Mac OS XTM, UnixTM, LinuxTM or FreeBSDTM.

In exemplary embodiments, there is also provided a nonvolatile computer readable storage medium, for example, memory 1932 including computer program instructions, which are executable by the processing component 1922 of the electronic apparatus 1900, to complete the above-described methods.

The present disclosure may be a system, a method, and/or a computer program product. The computer program product may include a computer readable storage medium having computer readable program instructions thereon for causing a processor to carry out aspects of the present disclosure.

The computer readable storage medium can be a tangible device that can retain and store instructions for use by an instruction execution device. The computer readable storage medium may be, for example, but is not limited to, an electronic storage apparatus, a magnetic storage apparatus, an optical storage apparatus, an electromagnetic storage apparatus, a semiconductor storage apparatus, or any suitable combination thereof. A non-exhaustive list of more specific examples of the computer readable storage medium includes: a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), a static random access memory (SRAM), a portable compact disc read-only memory (CD-ROM), a digital versatile disk (DVD), a memory stick, a floppy disk, a mechanically encoded apparatus such as punch-cards or raised structures in a groove having instructions recorded thereon, and any suitable combination thereof. A computer readable storage medium, as used herein, is not to be construed as being transitory signals per se, such as radio waves or other freely propagating electromagnetic waves, electromagnetic waves propagating through a waveguide or other transmission media (e.g., light pulses passing through a fiber-optic cable), or electrical signals transmitted through a wire.

Computer readable program instructions described herein can be downloaded into respective computing/processing apparatuses from a computer readable storage medium or into an external computer or external storage device via a network, for example, the Internet, a local

area network, a wide area network and/or a wireless network. The network may comprise copper transmission cables, optical transmission fibers, wireless transmission, routers, firewalls, switches, gateway computers and/or edge servers. A network adapter card or network interface in each computing/processing apparatus receives computer readable program instructions from the network and forwards the computer readable program instructions for storage in a computer readable storage medium within the respective computing/processing apparatus.

Computer readable program instructions for carrying out operations of the present disclosure may be assembler instructions, instruction-set-architecture (ISA) instructions, machine instructions, machine dependent instructions, microcode, firmware instructions, state-setting data, or either source code or object code written in any combination of one or more programming languages, including an object oriented programming language such as Smalltalk, C++ or the like, and conventional procedural programming languages, such as the “C” programming language or similar programming languages. The computer readable program instructions may execute entirely on the user’s computer, partly on the user’s computer, as a stand-alone software package, partly on the user’s computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user’s computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider). In some embodiments, electronic circuitry including, for example, programmable logic circuitry, field-programmable gate arrays (FPGA), or programmable logic arrays (PLA) may execute the computer readable program instructions by utilizing state information of the computer readable program instructions to personalize the electronic circuitry, in order to perform aspects of the present disclosure.

Aspects of the present disclosure are described herein with reference to flowcharts and/or block diagrams of methods, apparatus (systems), and computer program products according to embodiments of the present disclosure. It will be appreciated that each block of the flowcharts and/or block diagrams, and combinations of blocks in the flowcharts and/or block diagrams, can be implemented by computer readable program instructions.

These computer readable program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing devices to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing devices, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks. These computer readable program instructions may also be stored in a computer readable storage medium that

can direct a computer, a programmable data processing device, and/or other devices to function in a particular manner, such that the computer readable storage medium having instructions stored therein comprises an article of manufacture including instructions which implement aspects of the function/act specified in the flowchart and/or block diagram block or blocks.

5 The computer readable program instructions may also be loaded onto a computer, other programmable data processing devices, or other apparatuses to cause a series of operational steps to be performed on the computer, other programmable devices or other apparatuses to produce a computer implemented process, such that the instructions which execute on the computer, other programmable devices, or other apparatuses implement the functions/acts
10 specified in the flowchart and/or block diagram block or blocks.

 The flowcharts and block diagrams in the drawings illustrate the architecture, functionality, and operation of possible implementations of systems, methods and computer program products according to various embodiments of the present disclosure. In this regard, each block in the flowcharts or block diagrams may represent a module, program segment, or portion of
15 instruction, which comprises one or more executable instructions for implementing the specified logical function(s). In some alternative implementations, the functions noted in the block may occur out of the order noted in the drawings. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be
20 noted that each block of the block diagrams and/or flowcharts, and combinations of blocks in the block diagrams and/or flowcharts, can be implemented by special purpose hardware-based systems that perform the specified functions or acts, or combinations of special purpose hardware and computer instructions.

 Although the embodiments of the present disclosure have been described above, the
25 foregoing descriptions are exemplary but not exhaustive, and the disclosed embodiments are not limiting. For a person skilled in the art, a number of modifications and variations are obvious without departing from the scope and spirit of the described embodiments. The terms used herein are intended to provide the best explanations on the principles of the embodiments, practical applications, or technical improvements to the technologies in the market, or to make
30 the embodiments described herein understandable to other persons skilled in the art.

CLAIMS

1. A target object matching method, comprising:

acquiring a first target object and a second target object to be matched from an input
5 image, wherein the first target object includes a human body, and the second target object
includes at least one of a face and a hand;

performing feature processing on a first image corresponding to the first target object and
a second image corresponding to the second target object in the input image respectively, to
obtain a matching degree between the first target object in the first image and the second target
10 object in the second image;

establishing, based on the matching degree between the first target object in the first
image and the second target object in the second image, a bipartite graph between the first
target object and the second target object; and

determining, based on the bipartite graph between the first target object and the second
15 target object, the first target object and the second target object that are matched with each
other.

2. The method according to claim 1, wherein the performing feature processing on the
first image corresponding to the first target object and the second image corresponding to the
second target object in the input image respectively, to obtain a matching degree between the
20 first target object in the first image and the second target object in the second image comprises:

performing feature extraction processing on the first image and the second image to obtain
a first feature of the first image and a second feature of the second image, respectively; and

performing classification processing on a concatenated feature of the first feature and the
second feature to obtain the matching degree between the first target object in the first image
25 and the second target object in the second image.

3. The method according to claim 2, wherein the performing classification processing on
the concatenated feature of the first feature and the second feature to obtain the matching
degree between the first target object in the first image and the second target object in the
second image comprises:

30 performing feature fusion processing on the concatenated feature of the first feature and
the second feature to obtain a fusion feature; and

obtaining the matching degree between the first target object in the first image and the
second target object in the second image by providing the fusion feature as an input to a fully
connected layer to execute the classification processing on the fusion feature.

4. The method according to any one of claims 1 to 3, wherein the establishing, based on the matching degree between the first target object in the first image and the second target object in the second image, the bipartite graph between the first target object and the second target object comprises:

5 in response to the second target object including only the face, establishing the bipartite graph between the human body and the face based on the matching degree between the first target object in the first image and the second target object in the second image;

in response to the second target object including only the hand, establishing the bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image; and

10 in response to the second target object including the face and the hand, establishing the bipartite graph between the human body and the face and the bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image,

15 wherein the matching degree between the human body and the face is used as a concatenation weight between the human body and the face in the bipartite graph between the human body and the face, and the matching degree between the human body and the hand is used as a concatenation weight between the human body and the hand in the bipartite graph between the human body and the hand.

20 5. The method according to any one of claims 1 to 4, wherein the establishing, based on the matching degree between the first target object in the first image and the second target object in the second image, the bipartite graph between the first target object and the second target object comprises:

25 establishing, based on a first target object and a second target object between which a matching degree is greater than a first threshold, the bipartite graph between the first target object and the second target object.

6. The method according to any one of claims 1 to 5, wherein the determining, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other comprises:

30 selecting, based on the bipartite graph between the first target object and the second target object with a greedy algorithm in a descending order of the matching degree between the first target object and the second target object, a preset number of the second target objects best matched with the first target object as second target objects matched with the first target object.

7. The method according to claim 6, wherein the determining, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other further comprises:

in response to the bipartite graph between the first target object and the second target object being a bipartite graph between the human body and the hand, selecting, with a greedy algorithm, at most two second target objects with a type of a hand and best matched with the first target object; and

in response to the bipartite graph between the first target object and the second target object being a bipartite graph between the human body and the face, selecting, with a greedy algorithm, a second target object with a type of a face and best matched with the first target object.

8. The method according to claim 6 or 7, wherein the determining, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other further comprises:

in response to the preset number of second target objects having been determined to be matched with a first target object, ceasing to match other second target objects with the first target object; and

in response to a matched first target object has been determined for any second target object, ceasing to match other first target objects with the second target object.

9. The method according to any one of claims 1 to 8, wherein the acquiring the first target object and the second target object to be matched from the input image comprises at least one of:

determining the first target object and the second target object in the input image in response to detecting a box selection of the first target object and the second target object in the input image;

detecting the first target object and the second target object in the input image with a target detection neural network; and

receiving position information of the first target object and the second target object in the input image to determine the first target object and the second target object in the input image based on the position information.

10. The method according to any one of claims 1 to 9, wherein before performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, the method further comprises:

adjusting the first image and the second image to a preset specification respectively; and

performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image comprises:

5 performing feature processing on the first and second images adjusted to the preset specification, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

11. The method according to any one of claims 1 to 10, wherein the method further comprises:

10 displaying, in the input image, the first target object and the second target object that are matched with each other.

12. The method according to any one of claims 1 to 11, wherein the method further comprises: performing, by a Siamese neural network, the feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

13. The method according to claim 12, wherein the method further comprises a step of training the Siamese neural network comprising:

obtaining a training sample including a plurality of first training images that are human body images and a plurality of second training images that are face images or hand images;

20 providing the first training image and the second training image as an input to the Siamese neural network to obtain a predicted matching result between the first training image and the second training image; and

determining a network loss based on the predicted matching result between the first training image and the second training image to adjust a network parameter of the Siamese neural network according to the network loss until a training requirement is met.

14. A target object matching device, comprising:

an acquisition module configured to acquire a first target object and a second target object to be matched from an input image, wherein the first target object includes a human body, and the second target object includes at least one of a face and a hand;

30 a feature processing module configured to perform feature processing on a first image corresponding to the first target object and a second image corresponding to the second target object in the input image respectively, to obtain a matching degree between the first target object in the first image and the second target object in the second image;

a bipartite module configured to establish, based on the matching degree between the first target object in the first image and the second target object in the second image, a bipartite graph between the first target object and the second target object; and

a matching module configured to determine, based on the bipartite graph between the first target object and the second target object, the first target object and the second target object that are matched with each other.

15. The device according to claim 14, wherein the feature processing module is further configured to:

perform feature extraction processing on the first image and the second image to obtain a first feature of the first image and a second feature of the second image, respectively; and

perform classification processing on a concatenated feature of the first feature and the second feature to obtain the matching degree between the first target object in the first image and the second target object in the second image.

16. The device according to claim 15, wherein the feature processing module is further configured to:

perform feature fusion processing on a concatenated feature of the first feature and the second feature to obtain a fusion feature; and

obtain the matching degree between the first target object in the first image and the second target object in the second image by providing the fusion feature as an input to a fully connected layer to execute the classification processing on the fusion feature.

17. The device according to any one of claims 14 to 16, wherein the bipartite module is further configured to:

if the second target object includes only the face, establish a bipartite graph between the human body and the face based on the matching degree between the first target object in the first image and the second target object in the second image;

if the second target object includes only the hand, establish a bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image; and

if the second target object includes the face and the hand, establish the bipartite graph between the human body and the face and the bipartite graph between the human body and the hand based on the matching degree between the first target object in the first image and the second target object in the second image,

wherein a matching degree between the human body and the face is used as a concatenation weight between the human body and the face in the bipartite graph between the human body and the face, and a matching degree between the human body and the hand is used

as a concatenation weight between the human body and the hand in the bipartite graph between the human body and the hand.

18. The device according to any one of claims 14 to 17, wherein the bipartite module is further configured to establish, based on a first target object and a second target object between which a matching degree is greater than a first threshold, the bipartite graph between the first target object and the second target object.

19. The device according to any one of claims 14 to 18, wherein the matching module is further configured to select, based on the bipartite graph between the first target object and the second target object with a greedy algorithm in a descending order of the matching degree between the first target object and the second target object, a preset number of the second target objects best matched with the first target object as second target objects matched with the first target object.

20. The device according to claim 19, wherein the matching module is further configured to: select, with a greedy algorithm, a second target object with a type of a face and best matched with the first target object if the bipartite graph between the first target object and the second target object is a bipartite graph between a human body and the face.

21. The device according to claim 19 or 20, wherein the matching module is further configured to:

if the preset number of second target objects have been determined to be matched with a first target object, cease to match other second target objects with the first target object; and

if a matched first target object has been determined for any second target object, cease to match other first target objects with the second target object.

22. The device according to any one of claims 14 to 21, wherein the acquisition module is configured to acquire the first target object and the second target object to be matched from the input image by at least one of:

determining the first target object and the second target object in the input image in response to detecting a box selection of the first target object and the second target object in the input image;

detecting the first target object and the second target object in the input image with a target detection neural network; and

receiving position information of the first target object and the second target object in the input image to determine the first target object and the second target object in the input image based on the position information.

23. The device according to any one of claims 14 to 22, wherein the feature processing module is further configured to:

before performing feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, adjust the first and second images to a preset specification respectively; and

perform feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object in the input image respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image comprises:

perform feature processing on the first and second images adjusted to the preset specification, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

24. The device according to any one of claims 14 to 23, wherein the device further comprises a display module configured to display, in the input image, the first target object and the second target object that are matched with each other.

25. The device according to any one of claims 14 to 24, wherein the feature processing module is further configured to: perform, by a Siamese neural network, the feature processing on the first image corresponding to the first target object and the second image corresponding to the second target object respectively, to obtain the matching degree between the first target object in the first image and the second target object in the second image.

26. The device according to claim 25, wherein the device further comprises a training module configured to train the Siamese neural network, wherein a step of training the Siamese neural network comprises:

obtaining a training sample including a plurality of first training images that are human body images and a plurality of second training images that are face images or hand images;

obtaining a predicted matching result between the first training image and the second training image by providing the first training image and the second training image as an input to the Siamese neural network; and

determining a network loss based on the predicted matching result between the first training image and the second training image to adjust a network parameter of the Siamese neural network according to the network loss until a training requirement is met.

27. An electronic apparatus, comprising:

a processor; and

a memory configured to store processor-executable instructions;

wherein the processor is configured to invoke instructions stored in the memory to execute the method according to any one of claims 1 to 13.

28. A computer readable storage medium having computer program instructions stored thereon, wherein the computer program instructions, when executed by a processor, implement the method according to any one of claims 1 to 13.

29. A computer program, comprising a computer readable code, wherein when the
5 computer readable code is executed in an electronic apparatus, a processor of the electronic apparatus executes the method according to any one of claims 1 to 13.