METHOD AND APPARATUS FOR FILTRATING IMAGES AND ELECTRONIC DEVICE

Technical Field

The present disclosure relates to the field of computer vision technologies, and in particular, to a method and apparatus for filtrating images and an electronic device.

Background

In recent years, with the continuous development of artificial intelligence technology, artificial intelligence technology has achieved relatively good results in computer vision, speech recognition and other aspects. In some relatively special scenarios, such as board game scenarios, there is a need to identify board objects. Due to the problem of mutual occlusion between objects in a large number of collected images, the identification accuracy and identification efficiency are low.

Summary

The present disclosure provides a method solution for filtrating images.

Specifically, the present disclosure is implemented through the following technical solutions.

According to an aspect of embodiments of the present disclosure, provided is a method for filtrating images. The method includes: obtaining a first image, where the first image is an image frame in a video stream obtained by collecting images of a target area; detecting the first image to obtain a first detection result of a target object in the first image; determining a state of a target object which state is to be determined according to the first detection result of the target object in the first image and a second detection result of the target object which state is to be determined, where the target object which state is to be determined is one target object in the first image, the second detection result of the target object which state is to be determined is a detection result of the target object which state is to be determined in at least one second image frame in N image frames, adjacent to the first image, in the video stream, obtained by detecting the second image, and N is a positive integer; and determining the quality level of an image in a bounding box of the target object which state is to be determined according to the state of the target object which state is to be determined, where the bounding box of the target object which state is to be determined is determined according to the first detection result of the target object which state is to be determined.

With reference to any implementation provided by the present disclosure, the state of the target object which state is to be determined includes an occlusion state and a motion state, and

the determining a state of a target object which state is to be determined according to the first detection result of the target object in the first image and a second detection result of the target object which state is to be determined includes: determining a motion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined; determining whether the motion state of the target object which state is to be determined satisfies a preset motion state condition; and in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determining an occlusion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined.

With reference to any implementation provided by the present disclosure, the first detection result of the target object in the first image includes a bounding box of the target object in the first image, and the in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determining an occlusion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined includes: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determining an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined.

With reference to any implementation provided by the present disclosure, the target object which state is to be determined is a first-category target object, and the video stream is collected at an overhead view of the target area; and the in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determining an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined includes: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined

and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, determining that the target object which state is to be determined is in an unoccluded state; and in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, determining that the target object which state is to be determined is in an occluded state.

With reference to any implementation provided by the present disclosure, the target object which state is to be determined is a second-category target object, and the video stream is collected at a side view of the target area; and the in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determining an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined includes: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, determining that the target object which state is to be determined is in an unoccluded state.

With reference to any implementation provided by the present disclosure, the target object which state is to be determined is a second-category target object, and the video stream is collected at a side view of the target area; and the in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determining an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined includes: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, determining, according to the position of the target object which state is to be determined in a synchronous image, the position of a

target object with the intersection-over-union of a bounding box and the bounding box of the target object which state is to be determined greater than zero in the synchronous image, and the position of an image collection device for collecting the video stream, whether the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between each target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, where the synchronous image is collected synchronously with the first image from the overhead view of the target area; in response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between any target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, determining that the target object which state is to be determined is in an unoccluded state; and in response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is greater than the distance between one target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, determining that the target object which state is to be determined is in an occluded state.

With reference to any implementation provided by the present disclosure, the determining a motion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined includes: determining a first position of the target object which state is to be determined in the first image according to the first detection result of the target object which state is to be determined; determining a second position of the target object which state is to be determined in the second image according to the second detection result of the target object which state is to be determined; determining a motion speed of the target object which state is to be determined according to the first position, the second position, the time when the first image is collected, and the time when the second image is collected; and determining the motion state of the target object which state is to be determined according to the motion speed of the target object which state is to be determined; and the determining whether the motion state of the target object which state is to be determined satisfies a preset motion state condition includes: determining whether the motion state of the target object which state is to be determined satisfies the preset motion state condition according to the motion speed of the target

object which state is to be determined and an image collection frame rate of the image collection device for collecting the video stream.

With reference to any implementation provided by the present disclosure, the state of the target object which state is to be determined includes an occlusion state and a motion state, the occlusion state of the target object which state is to be determined includes an unoccluded state and an occluded state, and the motion state of the target object which state is to be determined includes satisfying the preset motion state condition and dissatisfying the preset motion state condition; the determining the quality level of an image in a bounding box of the target object which state is to be determined according to the state of the target object which state is to be determined includes: in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the unoccluded state, determining that the image in the bounding box of the target object which state is to be determined is a first quality image; in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the occluded state, determining that the image in the bounding box of the target object which state is to be determined is a second quality image; and in the case that the motion state of the target object which state is to be determined dissatisfies the preset motion state condition, determining that the image in the bounding box of the target object which state is to be determined is a third quality image.

With reference to any implementation provided by the present disclosure, the method further includes: determining a quality classification result of the image in the bounding box of the target object which state is to be determined in the first image via a neural network, where the neural network is trained with sample images annotated with quality levels, and one sample image includes at least one target object which state is to be determined; and in response to that the quality classification result of the image in the bounding box of the target object which state is to be determined determined by the neural network is consistent with the quality level of the image in the bounding box of the target object which state is to be determined determined according to the state of the target object which state is to be determined, taking the quality level of the image in the bounding box of the target object which state is to be determined as the target quality level of the image in the bounding box of the target object which state is to be determined.

According to an aspect of embodiments of the present disclosure, provided is an apparatus for filtrating images, including: an image obtaining unit, configured to obtain a first image,

where the first image is an image frame in a video stream obtained by collecting images of a target area; a detection result obtaining unit, configured to detect the first image to obtain a first detection result of a target object in the first image; a state determining unit, configured to determine a state of a target object which state is to be determined according to the first detection result of the target object in the first image and a second detection result of the target object which state is to be determined, where the target object which state is to be determined is one target object in the first image, the second detection result of the target object which state is to be determined is a detection result of the target object which state is to be determined in at least one second image frame in N image frames, adjacent to the first image, in the video stream, obtained by detecting the second image, and N is a positive integer; and a quality determining unit, configured to determine the quality level of an image in a bounding box of the target object which state is to be determined according to the state of the target object which state is to be determined, where the bounding box of the target object which state is to be determined is determined according to the first detection result of the target object which state is to be determined.

With reference to any implementation provided by the present disclosure, the state determining unit is specifically configured to: determine a motion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined; determine whether the motion state of the target object which state is to be determined satisfies a preset motion state condition; and in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determine an occlusion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined.

With reference to any implementation provided by the present disclosure, the first detection result of the target object in the first image includes a bounding box of the target object in the first image, and the state determining unit configured to determine an occlusion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is

to be determined satisfies the preset motion state condition, determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined.

With reference to any implementation provided by the present disclosure, the target object which state is to be determined is a first-category target object, and the video stream is collected at an overhead view of the target area; and the state determining unit configured to determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, determine that the target object which state is to be determined is in an unoccluded state; and in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, determine that the target object which state is to be determined is in an occluded state.

With reference to any implementation provided by the present disclosure, the target object which state is to be determined is a second-category target object, and the video stream is collected at a side view of the target area; and the state determining unit configured to determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to

zero, determine that the target object which state is to be determined is in an unoccluded state.

With reference to any implementation provided by the present disclosure, the target object which state is to be determined is a second-category target object, and the video stream is collected at a side view of the target area; and the state determining unit configured to determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, determine, according to the position of the target object which state is to be determined in a synchronous image, the position of a target object with the intersection-over-union of a bounding box and the bounding box of the target object which state is to be determined greater than zero in the synchronous image, and the position of an image collection device for collecting the video stream, whether the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between each target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, where the synchronous image is collected synchronously with the first image from the overhead view of the target area; in response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between any target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, determine that the target object which state is to be determined is in an unoccluded state; and in response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is greater than the distance between one target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, determine that the target object which state is to be determined is in an occluded state.

With reference to any implementation provided by the present disclosure, when determining a motion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined, the state determining unit is specifically configured to: determine a first position of the target object which state is to be determined in the first image according to the first detection result of the target object which state is to be determined; determine a second position of the target object which state is to be determined in the second image according to the second detection result of the target object which state is to be determined; determine a motion speed of the target object which state is to be determined according to the first position, the second position, the time when the first image is collected, and the time when the second image is collected; and determine the motion state of the target object which state is to be determined according to the motion speed of the target object which state is to be determined; and when determining whether the motion state of the target object which state is to be determined satisfies a preset motion state condition, the state determining unit is specifically configured to: determine whether the motion state of the target object which state is to be determined satisfies the preset motion state condition according to the motion speed of the target object which state is to be determined and an image collection frame rate of the image collection device for collecting the video stream.

With reference to any implementation provided by the present disclosure, the state of the target object which state is to be determined includes an occlusion state and a motion state, the occlusion state of the target object which state is to be determined includes an unoccluded state and an occluded state, and the motion state of the target object which state is to be determined includes satisfying the preset motion state condition and dissatisfying the preset motion state condition; the quality determining unit is specifically configured to: in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the unoccluded state, determine that the image in the bounding box of the target object which state is to be determined is a first quality image; in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the occluded state, determine that the image in the bounding box of the target object which state is to be determined is a second quality image; and in the case that the motion state of the target object which state is to be determined dissatisfies the preset motion state condition, determine that the image in the bounding box of the target object which state is to be determined is a third quality image.

With reference to any implementation provided by the present disclosure, the apparatus further includes: a classification unit, configured to determine a quality classification result of the image in the bounding box of the target object which state is to be determined in the first image via a neural network, where the neural network is trained with sample images annotated with quality levels, and one sample image includes at least one target object which state is to be determined; and in response to that the quality classification result of the image in the bounding box of the target object which state is to be determined determined by the neural network is consistent with the quality level of the image in the bounding box of the target object which state is to be determined determined according to the state of the target object which state is to be determined, take the quality level of the image in the bounding box of the target object which state is to be determined as the target quality level of the image in the bounding box of the target object which state is to be determined.

The embodiments of the present disclosure also provide an electronic device, including: a memory and a processor, where the memory is configured to store computer instructions run on the processor, and the processor is configured to implement the method for filtrating images according to any embodiment of the present disclosure when the computer instructions are executed.

The embodiments of the present disclosure also provide a computer-readable storage medium having a computer program stored thereon. When the program is executed by a processor, the method for filtrating images in the embodiments of the present disclosure is implemented.

In the embodiments of the present disclosure, the state of the target object of the to-be-detected state in the first image is determined according to the first detection result of a target object in a first image in a video stream obtained by collecting images of a target area, and a second detection result of the target object which state is to be determined in at least one second image frame of adjacent multiple image frames, so as to determine the quality level of the image in the bounding box of the target object of the to-be-detected state, thereby filtrating high-quality images for the target objects of the to-be-determined state to improve the identification efficiency.

It should be understood that the above general description and the following detailed description are merely exemplary and explanatory, and are not intended to limit the present disclosure.

Brief Description of the Drawings

The accompanying drawings here incorporated in the specification and constituting a part of the specification illustrate the embodiments consistent with the present disclosure and are intended to explain the technical solutions of the present disclosure together with the specification.

FIG. 1 is a flowchart of a method for filtrating images provided by at least one embodiment of the present disclosure;

FIG. 2 is a schematic diagram of an application scenario provided by at least one embodiment of the present disclosure;

FIG. 3A is a schematic diagram of a target object provided by at least one embodiment of the present disclosure;

FIG. 3B is a schematic diagram of another target object provided by at least one embodiment of the present disclosure;

FIG. 4 is a flowchart of a method for determining a motion state of a target object which state is to be determined provided by at least one embodiment of the present disclosure;

FIG. 5 is a schematic diagram of an apparatus for filtrating images provided by at least one embodiment of the present disclosure; and

FIG. 6 is a schematic structural diagram of an electronic device provided by at least one embodiment of the present disclosure.

Detailed Description

Exemplary embodiments will be described in detail herein, examples of which are illustrated in the accompanying drawings. If the drawings are involved in the following description, the same figure in different drawings refers to the same or similar elements, unless otherwise indicated. The implementations described in the following exemplary embodiments do not represent all implementations consistent with the present application. Instead, the implementations are merely examples of apparatuses and methods described in detail in the appended claims and consistent with some aspects of the present application.

The terms used in the present application are merely intended to describe particular embodiments, and are not intended to limit the present application. The singular forms "a" and "the" used in the present application and the appended claims are also intended to include most forms, unless other meanings are clearly indicated in the context. It should also be understood that the term "and/or" as used herein refers to and encompasses any or all possible combinations of one or more associated listed items. In addition, the term "at least one" herein means any one

of a plurality of items or any combination of at least two of a plurality of items.

It should be understood that although the terms such as first, second, and third may be used in the present application to describe various information, such information should not be limited to these terms. These terms are only used to distinguish the same type of information from each other. For example, the first information may also be referred to as the second information without departing from the scope of the present application. Similarly, the second information may also be referred to as the first information. Depending on the context, the word "if" as used herein may be interpreted as "at the time of" or "when" or "in response to determination".

To make a person skilled in the art well understand the technical solutions in the embodiments of the present disclosure, and make the foregoing objects, features, and advantages in the embodiments of the present disclosure apparent, the technical solutions in the embodiments of the present disclosure are further described below in detail with reference to the accompanying drawings.

In one exemplary board game scenario of the present disclosure, multiple people sit around a game table which may include multiple game areas, and different game areas may have different game meanings. Moreover, in a multiplayer game, users may play the game by means of redeemed items (such as game coins).

For example, the user may exchange the redeemed item with some items belonging to the user, and the redeemed items are placed in different game areas of the game table to play the game. For example, a first user may exchange multiple self-owned watercolor pens for chess pieces for the game, and play the game between different game areas on the game table according to the game rules with the chess pieces. If a second user wins the first user in the game, the chess pieces of the first user may belong to the second user. For example, the game described above is suitable for entertainment activities among family members during leisure time such as holidays.

With the continuous development of artificial intelligence technology, many places are trying to perform intelligent construction. For example, one of the topics is the construction of intelligent game places. Then, one of the requirements of the intelligent game place is to automatically identify objects on the board in the game, for example, to automatically identify the number of redeemed items.

FIG. 1 is a flowchart of a method for filtrating images provided by at least one embodiment of the present disclosure. As shown in FIG. 1, the method may include steps 101 to 104.

At step 101, a first image is obtained, where the first image is an image frame in a video

stream obtained by collecting images of a target area.

In the embodiments of the present disclosure, the target area is an area on which a target object is placed. For example, the target area may be a plane (e.g., a board), a container (e.g., a box), or the like. The target object may be one or more objects. In some relatively common situations, the target object is a sheet-shaped object of various shapes, such as game coins, banknotes, and cards. FIG. 2 shows a partial schematic diagram of a game table in a board game scenario. The game table includes multiple target areas, where each closed area represents one target area. The target object in this scenario is, for example, game coins on the game table.

At step 102, the first image is detected to obtain a first detection result of a target object in the first image.

In some embodiments, the first image may be input to a pre-trained target detection network to obtain a first detection result of the target object in the first image. The target detection network may be trained by using a sample image annotated with the category of the target object. The first detection result includes a bounding box of each target object, the position of the bounding box, and a classification result of the target objects.

At step 103, a state of a target object which state is to be determined is determined according to the first detection result of the target object in the first image and a second detection result of the target object which state is to be determined.

In the embodiments of the present disclosure, the target object which state is to be determined is one target object in the first image, the second detection result of the target object which state is to be determined is a detection result of the target object which state is to be determined in at least one second image frame in N image frames, adjacent to the first image, in the video stream, obtained by detecting the second image, and N is a positive integer.

In some embodiments, the state of the target object which state is to be determined includes an occlusion state and a motion state. The occlusion state is used for representing whether the target object which state is to be determined is occluded by other target objects, and the motion state is used for representing whether the target object which state is to be determined satisfies a preset motion state condition. Persons skilled in the art should understand that the state of the target object which state is to be determined may also include other states, and is not limited to the states described above.

In the case that the first image is a first image frame in the video stream, detection may be performed according to at least one second image frame in N image frames located behind the first image, to obtain a detection result of the target object which state is to be determined in the

second image, thereby determining the state of the target object which state is to be determined. In the case that the first image is not the first image frame in the video stream, detection may be performed according to at least one second image frame in N image frames located in front of the first image, to obtain a detection result of the target object which state is to be determined in the second image, thereby determining the state of the target object which state is to be determined.

At step 104, the quality level of an image in a bounding box of the target object which state is to be determined is determined according to the state of the target object which state is to be determined.

In the embodiments of the present disclosure, the bounding box of the target object which state is to be determined is determined according to the first detection result of the target object which state is to be determined.

In one example, for the bounding box of the target object which state is to be determined in the first detection result, the image in the bounding box may be cropped, and the quality level of the cropped image is determined according to the state of the target object which state is to be determined. The quality level of the image in the bounding box of the target object which state is to be determined in the first image may be determined according to the state of the target object which state is to be determined.

In the embodiments of the present disclosure, the state of the target object of the to-be-detected state in the first image is determined according to the first detection result of a target object in a first image in a video stream obtained by collecting images of a target area, and a second detection result of the target object which state is to be determined in at least one second image frame of adjacent multiple image frames, so as to determine the quality level of the image in the bounding box of the target object of the to-be-detected state, thereby filtrating high-quality images for the target objects of the to-be-determined state to improve the identification efficiency.

In some embodiments, the state of the target object which state is to be determined includes an occlusion state and a motion state. The state of the target object which state is to be determined may be determined in the following ways.

First, a motion state of the target object which state is to be determined is determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined. That is, the change in position of the target object which state is to be determined is determined according to the first

detection result of the target object which state is to be determined in the first image (also called a current image frame), and a second detection result of the target object which state is to be determined in the second image (an image frame in front of the first image or an image frame behind the first image), and the motion state of the target object which state is to be determined may be determined by combining a time interval between collections of the first image and the second image.

Next, whether the motion state of the target object which state is to be determined satisfies a preset motion state condition is determined.

In one example, the preset motion state condition may be set as: the motion speed is less than a set speed threshold.

The motion speed of the target object may be determined according to the change in position of the target object which state is to be determined in the first image and the second image, and the time interval. In response to that the motion speed is zero, it may be determined that the target object is in a still state, and then it may be determined that the motion state satisfies the preset motion state condition. In response to that the motion speed is less than the set speed threshold, it may also be determined that the motion state satisfies the preset motion state condition. Persons skilled in the art should understand that the set speed threshold may be specifically set according to requirements to the image quality, which is not limited in the embodiments of the present disclosure.

In response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, an occlusion state of the target object which state is to be determined is determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined.

In the case that the motion state of the target object which state is to be determined dissatisfies the set state condition, for example, when the motion speed is greater than or equal to a set motion threshold, it is indicated that the motion speed of the target object is relatively high. For an object on the game table, it is generally occluded. For example, when moved by a hand, the object is occluded by the hand. Moreover, the identification accuracy of the target object with a relatively high motion speed is reduced. Therefore, in the embodiments of the present disclosure, only the occlusion state of the target object which state is to be determined with the motion state satisfying the preset motion state condition is determined. That is, the occlusion state of the target object which state is to be determined is determined according to the first

detection result of the target object which state is to be determined with the motion state satisfying the preset motion state condition in the first image and the first detection result of other target objects in the first image, so as to determine the occlusion state of the target object which state is to be determined.

In some embodiments, the first detection result of the target object in the first image includes a bounding box of the target object in the first image. In response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, an occlusion state of the target object which state is to be determined is determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined.

In one example, in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is obtained. In response to that the intersection-over-union of the bounding box of each of other target objects and the bounding box of the target object which state is to be determined is less than or equal to a set threshold, e.g., zero, it is determined that the target object which state is to be determined is in an unoccluded state. In response to that the intersection-over-union of the bounding box of at least one of other target objects and the bounding box of the target object which state is to be determined is greater than the set threshold, e.g., zero, it is determined that the target object which state is to be determined is in an occluded state. There are two cases here, one is that the target object which state is to be determined occludes other target objects, and the other is that the target object which state is to be determined is occluded by other target objects.

In the embodiments of the present disclosure, the occlusion state of the target object is determined by means of the intersection-over-union of the bounding box of other target objects in the first image and the bounding box of the target object which state is to be determined, and the quality level of the image in the bounding box of the target object which state is to be determined is determined according to the occlusion state, so as to filtrate high-quality images for the target object which state is to be determined, thereby improving the identification efficiency.

In the embodiments of the present disclosure, an image collection device may be disposed at the periphery of the target area for collecting a video stream of the target area. Exemplarily, an image collection device (i.e., a top image collection image) may be disposed above the target

area, so that the image collection device collects the video stream of the target area in an overhead view. An image collection device (i.e., a side image collection image) may be disposed at the left and right sides (or multiple sides) of the target area, so that the image collection device collects the video stream of the target area in a side view. An image collection device may also be disposed above and the left and right sides (or multiple sides) of the target area, so that the image collection synchronous collects the video stream of the target area at the overhead view and the side view.

The classification of the target object which state is to be determined may be determined according to the first detection result and/or the second detection result of the target object which state is to be determined. Regarding the first-category target object, the video stream is collected at the overhead view of the target area. That is, the video stream of the target area is collected by the image collection device disposed above the target area at the overhead view. The first-category target object includes currency, cards, etc., and may also include game coins stacked in a horizontal direction and the like. FIG. 3A shows a schematic diagram of the game coins stacked in the horizontal direction, and the stacking mode may be referred to as a float stack. Persons skilled in the art should understand that the first-category target object may also include other forms of items, or items placed in other forms, and is not limited to the above.

In the case that the target object which state is to be determined is the first-category target object, and the video stream is collected at the overhead view of the target area, the occlusion state of the target object which state is to be determined may be determined in the following ways: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, that is, there is no overlapping area between the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the image collected at the overhead view, it is determined that the target object which state is to be determined is in an unoccluded state. The other target objects are, for example, hands, water glasses, and the like. Persons skilled in the art should understand that the categories of the other target objects may be specifically set according to needs, which is not limited in the present disclosure.

In response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other

target objects in the first image except the target object which state is to be determined is greater than zero, that is, there is an overlapping area between the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the image collected at the side view, it is determined that the target object which state is to be determined the is in an occluded state.

Regarding the second-category target object, the video stream is collected at the side view of the target area. That is, the video stream of the target area is collected by the image collection device disposed at the side (the left and right sides or multiple sides) of the target area at the side view. The second-category target object may include game coins stacked in a vertical direction. FIG. 3B shows a schematic diagram of the redeemed items stacked in the vertical direction, and the stacking mode may be referred to as a stand stack. Persons skilled in the art should understand that the second-category target object may also include other forms of items, or items placed in other forms, and is not limited to the above.

In the case that the target object which state is to be determined is the second-category target object, and the video stream is collected at the side view of the target area, the occlusion state of the target object which state is to be determined may be determined in the following ways: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, that is, there is no overlapping area between the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the image collected at the side view, it is determined that the target object which state is to be determined is in an unoccluded state.

In response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, that is, there is an overlapping area between the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the image collected at the side view, since an overlapping area exists between two bounding boxes in the image collected at the side view, and is related to the relative positions between corresponding target objects in the two bounding boxes, and the relative positions between two target objects and the image collection device, the occlusion state of the target object which state

is to be determined may further be determined according to a synchronous image collected synchronously with the first image from the overhead view of the target area. For ease of description, a target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero in the image collected at the side view is referred to as a side-view occlusion object. There may be one or more side-view occlusion objects.

That is, the relationship of the distance between the target object which state is to be determined and the image collection device for collecting the video stream and the distance between each side-view occlusion object and the image collection device for collecting the video stream is determined according to the position of the target object which state is to be determined in a synchronous image, the position of the side-view occlusion object in the synchronous image, and the position of an image collection device for collecting the video stream. Since the synchronous image is collected by the top image collection device from the overhead view, after the positions of the target object which state is to be determined and the side-view occlusion object in the synchronous image are determined, and by combining the position of the image collection device for collecting the video stream, the distance relationship of the target object which state is to be determined and the side-view occlusion object and the side image collection device in the horizontal direction may be determined.

In response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between any one of the side-view occlusion objects and the image collection device for collecting the video stream, it is determined that the target object which state is to be determined is in an unoccluded state. That is, when the bounding box of the side-view occlusion object overlaps the bounding box of the target object which state is to be determined in the image collected at the side view, and the distance between the target object which state is to be determined and the image collection device is less than the distance between the target object which state is to be determined and the side-view occlusion object, it may be determined that the target object which state is to be determined is not occluded by the side-view occlusion object, and thus, it is determined that the target object which state is to be determined is in an unoccluded state.

In response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is greater than the distance between one of the side-view occlusion objects and the image collection device for collecting the video stream, it is determined that the target object which state is to be determined is in an occluded state. That is, when the bounding box of the side-view occlusion object overlaps the bounding

box of the target object which state is to be determined in the image collected at the side view, and the distance between the target object which state is to be determined and the image collection device is greater than the distance between the target object which state is to be determined and the side-view occlusion object, it may be determined that the target object which state is to be determined is occluded by the side-view occlusion object, and thus, it is determined that the target object which state is to be determined is in an occluded state.

FIG. 4 is a flowchart of a method for determining a motion state of a target object which state is to be determined provided by at least one embodiment of the present disclosure. As shown in FIG. 4, the method includes steps 401 to 404.

At step 401, a first position of the target object which state is to be determined in the first image is determined according to the first detection result of the target object which state is to be determined.

The first position of the target object which state is to be determined in the first image may be determined according to the position of the bounding box of the target object which state is to be determined in the first detection result. For example, the central position of the bounding box may be used as the first position of the corresponding target object.

At step 402, a second position of the target object which state is to be determined in the second image is determined according to the second detection result of the target object which state is to be determined.

Similar to step 401, the second position of the target object which state is to be determined in the second image may be determined according to the position of the bounding box of the target object which state is to be determined in the second detection result.

At step 403, a motion speed of the target object which state is to be determined is determined according to the first position, the second position, the time when the first image is collected, and the time when the second image is collected.

The change in position of the target object which state is to be determined in the first image and the second image may be determined according to the first position and the second position. The time corresponding to the occurrence of change in position may be determined by combining the time when the first image is collected and the time when the second image is collected. Therefore, the motion speed of the target object which state is to be determined in a pixel plane coordinate system (a uv coordinate system) may be determined.

At step 404, the motion state of the target object which state is to be determined is

determined according to the motion speed of the target object which state is to be determined.

After the motion state of the target object which state is to be determined is determined, whether the motion state of the target object which state is to be determined satisfies the preset motion state condition is determined according to the motion speed and an image collection frame rate of the image collection device for collecting the video stream.

A motion speed threshold may be determined according to the image collection frame rate of the image collection device for collecting the video stream. When the motion speed of the target object which state is to be determined in the uv coordinate system is less than the motion speed threshold, the target object captured by the image collection device is a clear state, and the motion state in which the motion speed is less than the motion speed threshold is determined to satisfy the preset motion state condition. When the motion speed of the target object which state is to be determined in the uv coordinate system exceeds the motion speed threshold, the target object captured by the image collection device is a motion blurring state, and the motion state in which the motion speed exceeds the motion speed threshold is determined to dissatisfy the preset motion state condition.

In the embodiments of the present disclosure, the motion state of the target object which state is to be determined, and whether the target object which state is to be determined satisfies the preset motion state condition are determined according to the motion speed of the target object which state is to be determined, so as to filtrate an image of clear target object which state is to be determined, thereby improving the identification efficiency.

In some embodiments, the state of the target object which state is to be determined includes an occlusion state and a motion state. The occlusion state of the target object which state is to be determined includes an unoccluded state and an occluded state. The motion state of the target object which state is to be determined includes satisfying the preset motion state condition and dissatisfying the preset motion state condition.

According to the states above, the quality level of the image in the bounding box of the target object which state is to be determined may be determined in the following ways.

I. In the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the unoccluded state, it is determined that the image in the bounding box of the target object which state is to be determined is a first quality image. That is, the image in the bounding box corresponding to the target object which state is to be determined that is not occluded by other objects and in a non-motion blurring state is determined as the first quality image, i.e., the

high-quality image.

II. In the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the occluded state, it is determined that the image in the bounding box of the target object which state is to be determined is a second quality image. That is, the image in the bounding box corresponding to the target object which state is to be determined that is occluded by other objects and in a non-motion blurring state is determined as the second quality image, i.e., the medium-quality image.

III. In the case that the motion state of the target object which state is to be determined dissatisfies the preset motion state condition, it is determined that the image in the bounding box of the target object which state is to be determined is a third quality image. That is, the image in the bounding box corresponding to the target object which state is to be determined that is in a motion blurring state is determined as the third quality image, i.e., the low-quality image.

In the embodiments of the present disclosure, the quality level of the image in the bounding box of the target object is determined according to the occlusion state of the target object which state is to be determined and whether the motion state satisfies the preset motion state condition, so that the frame images in the video stream are filtrated by means of the determined quality level, to improve the identification accuracy of the target object when the target object is identified by using the filtrated image.

After the quality level of the image in the bounding box of the target object which state is to be determined is obtained according to the foregoing method, the quality classification result of the image may further be obtained by using the neural network to verify the determined quality level, so as to obtain the final target quality level.

First, the quality classification result of the image in the bounding box of the target object which state is to be determined in the first image is determined by using the neural network.

The neural network may be trained by the sample images annotated with the quality level, and one sample image includes at least one target object which state is to be determined. The sample image may determine the quality level according to the method for filtrating images provided by at least one embodiment of the present disclosure, and is annotated with the determined quality level. For example, in the case that the image of the bounding box of the target object which state is to be determined in an image is determined as the first quality image according to the method for filtrating images provided by one of embodiments of the present disclosure, the image may be annotated as the first quality image, and the image is used as a

sample image to train the neural network. Persons skilled in the art should understand that the image of the quality level is determined as the sample image by using other methods, to train the neural network. It should be noted that the quality level annotation of the sample image should be consistent with the image quality level determined according to the method for filtrating images provided by the embodiments of the present disclosure.

In response to that the quality classification result of the image in the bounding box of the target object which state is to be determined determined by the neural network is consistent with the quality level of the image in the bounding box of the target object which state is to be determined determined according to the state of the target object which state is to be determined, the quality level of the image in the bounding box of the target object which state is to be determined is used as the target quality level of the image in the bounding box of the target object which state is to be determined.

For an image frame in the video stream, the quality level of the image in the bounding box corresponding to the image frame is determined according to the state of the target object which state is to be determined in the image by means of the method for filtrating images provided by the embodiments of the present disclosure. Then, the quality classification result in the bounding box of the target object which state is to be determined in the image is obtained according to the neural network. In the case that the quality classification result obtained by the neural network is consistent with the quality level determined according to the method for filtrating images provided by the embodiments of the present disclosure, the quality level may be determined as the target quality level.

For example, in the case that the image of the bounding box of the target object which state is to be determined in an image is determined as the first quality image according to the method for filtrating images provided by one of the embodiments of the present disclosure, if the quality classification result obtained by the network is also the first quality image, it may be determined that the image in the bounding box of the target object which state is to be determined in the image is the first quality image.

In the embodiments of the present disclosure, the quality classification result of the image in the bounding box of the target object which state is to be determined is determined by the neural network, to further verify the quality level of the image, so as to improve the accuracy of the quality level classification of the image.

The target area 200 of the game table shown in FIG. 2 is taken as an example to describe the method for filtrating images according to at least one embodiment of the present disclosure.

Persons skilled in the art should understand that the method for filtrating images may also be applied to other target areas, but is not limited to the target areas of the game table.

The image collection device 201 disposed in a left area 201 of a dotted line A may be regarded as a side image collection device, which collects the image of the target area from a left-side angle of view. The image collection device 202 disposed in a right area 202 of a dotted line B may also be regarded as a side image collection device, which collects the image of the target area from a right-side angle of view. In addition, a top image collection device (not shown in FIG. 2) may be further provided above the target area 200 of the game table to collect an image of the target area from an overhead view.

First, an image frame in a video stream obtained by collecting images of a target area by any of the foregoing image collection devices is obtained, and the image frame may be referred to as a first image. The first image may be an image collected at an overhead view, or an image obtained from a side view.

Next, the first image is detected to obtain a first detection result of a target object in the first image. The target object in the first image may include a target object which state is to be determined, and the target object is a target object for image quality filtrating. In the board game scenario, the target object which state is to be determined includes a first-category target object, e.g., game coins stacked in the horizontal direction (as shown in FIG. 3A), and a second-category target object (as shown in FIG. 3B), e.g., game coins stacked in the vertical direction. Other target objects except the target object which state is to be determined include hands. The obtained first detection result includes the target object which state is to be determined, and the bounding boxes, positions and the classification result of other target objects.

Next, a second detection result of the target object which state is to be determined in at least one second image frames in N image frames adjacent to the first image is obtained. The state of the target object which state is to be determined may be determined according to the first detection result and the second detection result, including an occlusion state and a motion state, where the occlusion state includes an occluded state and an unoccluded state, and the motion state includes satisfying the preset motion state condition and dissatisfying the preset motion state condition.

The method for determining the occlusion state is described below.

For the first-category target object, e.g., the game coins stacked in the horizontal direction, the occlusion state of the first-category target object may be determined by a first image

collected by the top image collection device. For example, in the case that the intersection-over-union of the bounding box of the horizontally stacked game coins in the first image and the bounding box of each hand detected is less than or equal to zero, it is determined that the horizontally stacked game coins are in an unoccluded state. On the contrary, in the case that the intersection-over-union of the bounding box of the horizontally stacked game coins in the first image and the bounding box of one of the hands detected is greater than zero, it is determined that the horizontally stacked game coins are in an occluded state.

For the second-category target object, e.g., the game coins stacked in the vertical direction, the occlusion state of the second-category target object may be determined by a first image collected by the side image collection device. For example, in the case that the intersection-over-union of the bounding box of the vertically stacked game coins in the first image and the bounding box of each hand detected is less than or equal to zero, it is determined that the vertically stacked game coins are in an unoccluded state.

In the case that the intersection-over-union of the bounding box of the vertically stacked game coins in the first image and the bounding box of one of the hands detected is greater than zero, it is determined further according to the position relationship of the vertically stacked game coins, the hand, and the side image collection device that the vertically stacked game coins are in an occluded state. For ease of description, the hand with the intersection-over-union of the bounding boxes greater than zero is called occlusion hand.

In one example, the position relationship of the vertically stacked game coins, the hand, and the side image collection device may be determined by the synchronous image collected by the top image collection device. For example, the distance between the vertically stacked game coins and the side image collection device, and the distance between the occlusion hand and the side image collection device may be determined according to the position of the vertically stacked game coins in the synchronous image, the position of the occlusion hand in the synchronous image, and the position of the side image collection device.

In the case that the distance between the vertically stacked game coins and the side image collection device is less than the distance between the occlusion hand and the side image collection device, it may be determined that the vertically stacked game coins are in an unoccluded state. On the contrary, in the case that the distance between the vertically stacked game coins and the side image collection device is greater than the distance between the occlusion hand and the side image collection device, it may be determined that the vertically stacked game coins are in an occluded state.

The method for determining the motion state is described below.

First, the first position of the target object which state is to be determined in the first image is determined according to the first detection result of the target object which state is to be determined. The target object which state is to be determined includes game coins stacked in the horizontal direction and game coins stacked in the vertical direction, which are collectively called stacked game coins for ease of descriptions. That is, the first position of the stacked game coins in the first image is first determined.

Next, the second position of the target object which state is to be determined in the second image is determined according to the second detection result of the stacked game coins. Taking the second image to be an image frame in N image frames adjacent to the first image as an example, the position of stacked game coins in an image frame in front of the second image is obtained.

The motion speed of the stacked game coins in the uv coordinate system may be determined according to the time when the first image is collected, the time when the second image is collected, the first position, and the second position, so as to determine the motion state of the stacked game coins.

A corresponding motion speed threshold may be obtained according to the image collection frame rate of the image collection device for collecting the video stream. In the case that the motion speed of the stacked game coins in the uv coordinate system is less than or equal to the motion speed threshold, it may be determined that the motion state satisfies the preset motion state condition. In the case that the motion speed of the stacked game coins in the uv coordinate system is greater than the motion speed threshold, it may be determined that the motion state dissatisfies the preset motion state condition.

The quality level of the image in the bounding box of the stacked game coins may be determined according to the occlusion state and the motion state of the determined stacked game coins.

For example, in the case that the motion state of the stacked game coins satisfies the preset motion state condition, and the stacked game coins are in an unoccluded state, the image in the bounding box of the stacked game coins is the first quality image. In the case that the motion state of the stacked game coins satisfies the preset motion state condition, and the stacked game coins are in an occluded state, the image in the bounding box of the stacked game coins is the second quality image. In the case that the motion state of the stacked game coins dissatisfies the preset motion state condition, the image in the bounding box of the stacked game coins is the

third quality image.

The first image or the images in the bounding box of the stacked game coins in the first image are filtrated according to the quality level of the image in the bounding box of the stacked game coins, so that the identification efficiency and accuracy during the identification of the stacked game coins with the filtrated image may be improved.

As shown in FIG. 5, at least one embodiment of the present disclosure also provides an apparatus for filtrating images, including: an image obtaining unit 501, configured to obtain a first image, where the first image is an image frame in a video stream obtained by collecting images of a target area; a detection result obtaining unit 502, configured to detect the first image to obtain a first detection result of a target object in the first image; a state determining unit 503, configured to determine a state of a target object which state is to be determined according to the first detection result of the target object in the first image and a second detection result of the target object which state is to be determined, where the target object which state is to be determined is one target object in the first image, the second detection result of the target object which state is to be determined is a detection result of the target object which state is to be determined in at least one second image frame in N image frames, adjacent to the first image, in the video stream, obtained by detecting the second image, and N is a positive integer; and a quality determining unit 504, configured to determine the quality level of an image in a bounding box of the target object which state is to be determined according to the state of the target object which state is to be determined, where the bounding box of the target object which state is to be determined is determined according to the first detection result of the target object which state is to be determined.

In some embodiments, the state determining unit 503 is specifically configured to: determine a motion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined; determine whether the motion state of the target object which state is to be determined satisfies a preset motion state condition; and in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determine an occlusion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined.

In some embodiments, the first detection result of the target object in the first image includes a bounding box of the target object in the first image, and the state determining unit 503

configured to determine an occlusion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and a first detection result of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined.

In some embodiments, the target object which state is to be determined is a first-category target object, and the video stream is collected at an overhead view of the target area; and the state determining unit 503 configured to determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, determine that the target object which state is to be determined is in an unoccluded state; and in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, determine that the target object which state is to be determined is in an occluded state.

In some embodiments, the target object which state is to be determined is a second-category target object, and the video stream is collected at a side view of the target area; and the state determining unit 503 configured to determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined in response to that the motion

state of the target object which state is to be determined satisfies the preset motion state condition, is specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of each of other target objects in the first image except the target object which state is to be determined is less than or equal to zero, determine that the target object which state is to be determined is in an unoccluded state.

In some embodiments, the target object which state is to be determined is a second-category target object, and the video stream is collected at a side view of the target area; and the state determining unit 503 configured to determine an occlusion state of the target object which state is to be determined according to an intersection-over-union of the bounding box of the target object which state is to be determined and a bounding box of each of other target objects in the first image except the target object which state is to be determined in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, s specifically configured to: in response to that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object which state is to be determined and the bounding box of at least one of other target objects in the first image except the target object which state is to be determined is greater than zero, determine, according to the position of the target object which state is to be determined in a synchronous image, the position of a target object with the intersection-over-union of a bounding box and the bounding box of the target object which state is to be determined greater than zero in the synchronous image, and the position of an image collection device for collecting the video stream, whether the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between each target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, where the synchronous image is collected synchronously with the first image from the overhead view of the target area; in response to that the distance between the target object which state is to be determined and the image collection device for collecting the video stream is less than the distance between any target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, determine that the target object which state is to be determined is in an unoccluded state; and in response to that the distance between the target object which state is to be determined and the image collection device for

collecting the video stream is greater than the distance between one target object with the intersection-over-union of the bounding box and the bounding box of the target object which state is to be determined greater than zero and the image collection device for collecting the video stream, determine that the target object which state is to be determined is in an occluded state.

In some embodiments, when determining a motion state of the target object which state is to be determined according to the first detection result of the target object which state is to be determined and the second detection result of the target object which state is to be determined, the state determining unit 503 is specifically configured to: determine a first position of the target object which state is to be determined in the first image according to the first detection result of the target object which state is to be determined; determine a second position of the target object which state is to be determined in the second image according to the second detection result of the target object which state is to be determined; determine a motion speed of the target object which state is to be determined according to the first position, the second position, the time when the first image is collected, and the time when the second image is collected; and determine the motion state of the target object which state is to be determined according to the motion speed of the target object which state is to be determined. When determining whether the motion state of the target object which state is to be determined satisfies a preset motion state condition, the state determining unit is specifically configured to: determine whether the motion state of the target object which state is to be determined satisfies the preset motion state condition according to the motion speed of the target object which state is to be determined and an image collection frame rate of the image collection device for collecting the video stream.

In some embodiments, the state of the target object which state is to be determined includes an occlusion state and a motion state, the occlusion state of the target object which state is to be determined includes an unoccluded state and an occluded state, and the motion state of the target object which state is to be determined includes satisfying the preset motion state condition and dissatisfying the preset motion state condition. The quality determining unit 504 is specifically configured to: in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the unoccluded state, determine that the image in the bounding box of the target object which state is to be determined is a first quality image; in the case that the motion state of the target object which state is to be determined satisfies the preset motion state condition, and the target object which state is to be determined is in the occluded state, determine that the image in the bounding box of the target object which state is to be determined is a second quality image; and

in the case that the motion state of the target object which state is to be determined dissatisfies the preset motion state condition, determine that the image in the bounding box of the target object which state is to be determined is a third quality image.

With reference to any implementation provided by the present disclosure, the apparatus further includes: a classification unit, configured to determine a quality classification result of the image in the bounding box of the target object which state is to be determined in the first image via a neural network, where the neural network is trained with sample images annotated with quality levels, and one sample image includes at least one target object which state is to be determined; and in response to that the quality classification result of the image in the bounding box of the target object which state is to be determined determined by the neural network is consistent with the quality level of the image in the bounding box of the target object which state is to be determined determined according to the state of the target object which state is to be determined, take the quality level of the image in the bounding box of the target object which state is to be determined as the target quality level of the image in the bounding box of the target object which state is to be determined.

In some embodiments, the functions provided by or the modules included in the apparatuses provided in the embodiments of the present disclosure may be used to implement the methods described in the foregoing method embodiments. For specific implementations, reference may be made to the description in the method embodiments above. For the purpose of brevity, details are not described herein repeatedly.

The apparatus embodiments described above are merely illustrative, where the modules described as separate components may or may not be physically separated, and the components displayed as modules may or may not be physical modules, that is, may be located a same position, or may also be distributed to multiple network modules. Some or all of the modules may be selected according to actual requirements to achieve the objectives of the solutions in the specification. A person of ordinary skill in the art may understand and implement without involving any inventive effort.

The apparatus embodiments of the present disclosure may be applied to computer devices, such as servers or terminal devices. The apparatus embodiments may be implemented by software, or may be implemented by hardware or a combination of hardware and software. Taking the software implementation as an example, a logical apparatus is formed by reading a corresponding computer program instruction in a non-volatile memory into an internal storage by means of a processor of file processing in which the apparatus is located. From the hardware aspect, FIG. 6 shows a hardware structure diagram of an electronic device in which the apparatus

in the specification is located. In addition to the processor, an internal bus, a network interface, and the non-volatile memory as shown in FIG. 6, the server or the electronic device in which the apparatus in the embodiments is located may also include other hardware according to the actual function of the computer device, and details are not described herein.

Accordingly, the embodiments of the present disclosure also provide a computer storage medium having a computer program stored thereon. When the program is executed by a processor, the method for filtrating images according to any embodiment is implemented.

Accordingly, the embodiments of the present disclosure also provide a computer device, including a memory, a processor, and a computer program stored on the memory and running on the processor. When the program is executed by the processor, the method for filtrating images according to any embodiment is implemented.

The present disclosure may take the form of a computer program product implemented on one or more storage media (including but not limited to a disk memory, a CD-ROM, an optical memory, etc.) including program codes. Computer-usable storage media includes permanent and non-permanent, removable and non-removable media, and may implement information storage by means of any method or technology. Information may be computer-readable commands, data structures, modules of programs, or other data. Examples of computer storage media include, but are not limited to: Phase-change Random Access Memory (PRAM), Static Random Access Memory (SRAM), Dynamic Random Access Memory (DRAM), other types of Random Access Memories (RAMs), Read-Only Memory (ROM), Electrically Erasable Programmable Read-Only Memory (EEPROM), flash memory or other memory technologies, Compact Disk Read Only Memory (CD-ROM), Digital Versatile Disc (DVD) or other optical storage, magnetic cassettes, magnetic tape storage or other magnetic storage devices or any other non-transmitting medium that may be used to store information that may be accessed by a computing device.

Persons skilled in the art would easily conceive of other embodiments of the present disclosure after considering the specification and practicing the specification disclosed herein. The present disclosure is intended to cover any variations, uses, or adaptive changes of the present disclosure. These variations, uses, or adaptive changes conform to the general principles of the present disclosure and include the common general knowledge or conventional technical measures in the technical field that are not disclosed in the present disclosure. The specification and embodiments are considered as exemplary only, and the real scope and spirit of the present disclosure are indicated by the following claims.

It should be understood that the present disclosure is not limited to the precise structure that

is described above and illustrated in the drawings, and various modifications and changes may be made without departing from the scope thereof. The scope of the present disclosure is limited only by the following claims.

The above are only preferred embodiments of the present disclosure, and are not intended to limit the present disclosure. Any modification, equivalent substitution, improvement, etc. made within the spirit and principle of the present disclosure should be included in the scope of protection of the present disclosure.

The foregoing descriptions of various embodiments emphasize differences between the embodiments. For a same or similar part, reference may be made to each other. For brevity, details are not described again.

CLAIMS

1. A method for filtrating images, comprising:

obtaining a first image, wherein the first image is an image frame in a video stream obtained by collecting images of a target area;

detecting the first image to obtain a first detection result of a target object in the first image;

determining a state of the target object according to the first detection result of the target object in the first image and a second detection result of the target object, wherein the target object is one target object in the first image, the second detection result of the target object is a detection result of the target object in at least one second image frame in N image frames, adjacent to the first image, in the video stream, obtained by detecting the second image, and N is a positive integer; and

determining a quality level of an image in a bounding box of the target object according to the state of the target object, wherein the bounding box of the target object is determined according to the first detection result of the target object,

wherein the state of the target object comprises an occlusion state and a motion state, and

the determining the state of the target object according to the first detection result of the target object in the first image and the second detection result of the target object comprises:

determining the motion state of the target object according to the first detection result of the target object and the second detection result of the target object;

determining whether the motion state of the target object satisfies a preset motion state condition; and

in response to that the motion state of the target object satisfies the preset motion state condition, determining the occlusion state of the target object according to the first detection result of the target object and a first detection result of other target objects in the first image except the target object.

2. The method according to claim 1, wherein the first detection result of the target object in the first image comprises a bounding box of the target object in the first image, and the in response to that the motion state of the target object satisfies the preset motion state condition, determining an occlusion state of the target object according to the first detection result of the target object and a first detection result of other target objects in the first image except the target object comprises:

in response to that the motion state of the target object satisfies the preset motion state condition, determining an occlusion state of the target object according to an intersection-over-union of the bounding box of the target object and a bounding box of each of other target objects in the first image except the target object.

3. The method according to claim 2, wherein the target object is a first-category target object, and the video stream is collected at an overhead view of the target area; and the in response to that the motion state of the target object satisfies the preset motion state condition, determining an occlusion state of the target object according to an intersection-over-union of the bounding box of the target object and a bounding box of each of other target objects in the first image except the target object comprises:

in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of each of other target objects in the first image except the target object is less than or equal to zero, determining that the target object is in an unoccluded state; and

in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of at least one of other target objects in the first image except the target object is greater than zero, determining that the target object is in an occluded state.

4. The method according to claim 2, wherein the target object is a second-category target object, and the video stream is collected at a side view of the target area; and the in response to that the motion state of the target object satisfies the preset motion state condition, determining an occlusion state of the target object according to an intersection-over-union of the bounding box of the target object and a bounding box of each of other target objects in the first image except the target object comprises:

in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of each of other target objects in the first image except the target object is less than or equal to zero, determining that the target object is in an unoccluded state.

5. The method according to claim 2, wherein the target object is a second-category target object, and the video stream is collected at a side view of the target area; and the in response to that the motion state of the target object satisfies the preset motion state condition, determining an occlusion state of the target object according to an intersection-over-union of the bounding

box of the target object and a bounding box of each of other target objects in the first image except the target object comprises:

in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of at least one of other target objects in the first image except the target object is greater than zero, determining, according to the position of the target object in a synchronous image, the position of the other target object with the intersection-over-union of a bounding box and the bounding box of the target object greater than zero in the synchronous image, and the position of an image collection device for collecting the video stream, whether the distance between the target object and the image collection device for collecting the video stream is less than the distance between the other target object with the intersection-over-union of the bounding box and the bounding box of the target object greater than zero and the image collection device for collecting the video stream, wherein the synchronous image is collected synchronously with the first image at the overhead view of the target area;

in response to that the distance between the target object and the image collection device for collecting the video stream is less than the distance between any target object with the intersection-over-union of the bounding box and the bounding box of the target object greater than zero and the image collection device for collecting the video stream, determining that the target object is in an unoccluded state; and

in response to that the distance between the target object and the image collection device for collecting the video stream is greater than the distance between one target object with the intersection-over-union of the bounding box and the bounding box of the target object greater than zero and the image collection device for collecting the video stream, determining that the target object is in an occluded state.

6. The method according to any one of claims 1 to 5, wherein the determining a motion state of the target object according to the first detection result of the target object and the second detection result of the target object comprises:

determining a first position of the target object in the first image according to the first detection result of the target object;

determining a second position of the target object in the second image according to the second detection result of the target object;

determining a motion speed of the target object according to the first position, the second

position, the time when the first image is collected, and the time when the second image is collected; and

determining the motion state of the target object according to the motion speed of the target object; and

the determining whether the motion state of the target object satisfies a preset motion state condition comprises:

determining whether the motion state of the target object satisfies the preset motion state condition according to the motion speed of the target object and an image collection frame rate of the image collection device for collecting the video stream.

7. The method according to any one of claims 1 to 5, wherein the occlusion state of the target object comprises an unoccluded state and an occluded state; and

the determining the quality level of an image in a bounding box of the target object according to the state of the target object comprises:

in the case that the motion state of the target object satisfies the preset motion state condition, and the target object is in the unoccluded state, determining that the image in the bounding box of the target object is a first quality image;

in the case that the motion state of the target object satisfies the preset motion state condition, and the target object is in the occluded state, determining that the image in the bounding box of the target object is a second quality image; and

in the case that the motion state of the target object dissatisfies the preset motion state condition, determining that the image in the bounding box of the target object is a third quality image.

8. The method according to any one of claims 1 to 7, further comprising:

determining a quality classification result of the image in the bounding box of the target object in the first image via a neural network, wherein the neural network is trained with sample images annotated with quality levels, and one sample image comprises at least one target object; and

in response to that the quality classification result of the image in the bounding box of the target object determined by the neural network is consistent with the quality level of the image in the bounding box of the target object determined according to the state of the target object, taking the quality level of the image in the bounding box of the target object as the target quality

level of the image in the bounding box of the target object.

9. An apparatus for filtrating images, comprising:

an image obtaining unit, configured to obtain a first image, wherein the first image is an image frame in a video stream obtained by collecting images of a target area;

a detection result obtaining unit, configured to detect the first image to obtain a first detection result of a target object in the first image;

a state determining unit, configured to determine a state of the target object according to the first detection result of the target object in the first image and a second detection result of the target object, wherein the target object is one target object in the first image, the second detection result of the target object is a detection result of the target object in at least one second image frame in N image frames, adjacent to the first image, in the video stream, obtained by detecting the second image, and N is a positive integer; and

a quality determining unit, configured to determine a quality level of an image in a bounding box of the target object according to the state of the target object, wherein the bounding box of the target object is determined according to the first detection result of the target object,

wherein the state of the target object comprises an occlusion state and a motion state, and

the state determining unit is specifically configured to: determine the motion state of the target object according to the first detection result of the target object and the second detection result of the target object; determine whether the motion state of the target object satisfies a preset motion state condition; and in response to that the motion state of the target object satisfies the preset motion state condition, determine the occlusion state of the target object according to the first detection result of the target object and a first detection result of other target objects in the first image except the target object.

10. The apparatus according to claim 9, wherein the first detection result of the target object in the first image comprises a bounding box of the target object in the first image, and

the state determining unit configured to determine an occlusion state of the target object according to the first detection result of the target object and a first detection result of other target objects in the first image except the target object in response to that the motion state of the target object satisfies the preset motion state condition is specifically configured to: in response to that the motion state of the target object satisfies the preset motion state condition, determine an occlusion state of the target object according to an intersection-over-union of the bounding box

of the target object and a bounding box of each of other target objects in the first image except the target object.

11. The apparatus according to claim 10, wherein the target object is a first-category target object, and the video stream is collected at an overhead view of the target area; and

the state determining unit configured to determine an occlusion state of the target object according to an intersection-over-union of the bounding box of the target object and a bounding box of each of other target objects in the first image except the target object in response to that the motion state of the target object satisfies the preset motion state condition is specifically configured to: in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of each of other target objects in the first image except the target object is less than or equal to zero, determine that the target object is in an unoccluded state; and in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of at least one of other target objects in the first image except the target object is greater than zero, determine that the target object is in an occluded state.

12. The apparatus according to claim 10, wherein the target object is a second-category target object, and the video stream is collected at a side view of the target area; and

the state determining unit configured to determine an occlusion state of the target object according to an intersection-over-union of the bounding box of the target object and a bounding box of each of other target objects in the first image except the target object in response to that the motion state of the target object satisfies the preset motion state condition is specifically configured to: in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of each of other target objects in the first image except the target object is less than or equal to zero, determine that the target object is in an unoccluded state.

13. The apparatus according to claim 10, wherein the target object is a second-category target object, and the video stream is collected at a side view of the target area; and

the state determining unit configured to determine an occlusion state of the target object according to an intersection-over-union of the bounding box of the target object and a bounding box of each of other target objects in the first image except the target object in response to that the motion state of the target object satisfies the preset motion state condition is specifically

configured to: in response to that the motion state of the target object satisfies the preset motion state condition, and the intersection-over-union of the bounding box of the target object and the bounding box of at least one of other target objects in the first image except the target object is greater than zero, determine, according to the position of the target object in a synchronous image, the position of the other target object with the intersection-over-union of a bounding box and the bounding box of the target object greater than zero in the synchronous image, and the position of an image collection device for collecting the video stream, whether the distance between the target object and the image collection device for collecting the video stream is less than the distance between the other target object with the intersection-over-union of the bounding box and the bounding box of the target object greater than zero and the image collection device for collecting the video stream, wherein the synchronous image is collected synchronously with the first image at the overhead view of the target area; in response to that the distance between the target object and the image collection device for collecting the video stream is less than the distance between any target object with the intersection-over-union of the bounding box and the bounding box of the target object greater than zero and the image collection device for collecting the video stream, determine that the target object is in an unoccluded state; and in response to that the distance between the target object and the image collection device for collecting the video stream is greater than the distance between one target object with the intersection-over-union of the bounding box and the bounding box of the target object greater than zero and the image collection device for collecting the video stream, determine that the target object is in an occluded state.

14. The apparatus according to any one of claims 9 to 13, wherein the state determining unit configured to determine a motion state of the target object according to the first detection result of the target object and the second detection result of the target object is specifically configured to: determine a first position of the target object in the first image according to the first detection result of the target object; determine a second position of the target object in the second image according to the second detection result of the target object; determine a motion speed of the target object according to the first position, the second position, the time when the first image is collected, and the time when the second image is collected; and determine the motion state of the target object according to the motion speed of the target object; and

the state determining unit configured to determine whether the motion state of the target object satisfies a preset motion state condition is specifically configured to: determine whether the motion state of the target object satisfies the preset motion state condition according to the motion speed of the target object and an image collection frame rate of the image collection

device for collecting the video stream.

15. The apparatus according to any one of claims 9 to 13, wherein the occlusion state of the target object comprises an unoccluded state and an occluded state; and

the quality determining unit is specifically configured to: in the case that the motion state of the target object satisfies the preset motion state condition, and the target object is in the unoccluded state, determine that the image in the bounding box of the target object is a first quality image; in the case that the motion state of the target object satisfies the preset motion state condition, and the target object is in the occluded state, determine that the image in the bounding box of the target object is a second quality image; and in the case that the motion state of the target object dissatisfies the preset motion state condition, determine that the image in the bounding box of the target object is a third quality image.

16. The apparatus according to any one of claims 9 to 15, further comprising:

a classification unit configured to: determine a quality classification result of the image in the bounding box of the target object in the first image via a neural network, wherein the neural network is trained with sample images annotated with quality levels, and one sample image comprises at least one target object; and in response to that the quality classification result of the image in the bounding box of the target object determined by the neural network is consistent with the quality level of the image in the bounding box of the target object determined according to the state of the target object, take the quality level of the image in the bounding box of the target object as the target quality level of the image in the bounding box of the target object.

17. An electronic device, comprising a memory and a processor, wherein the memory is configured to store computer instructions run on the processor, and the processor is configured to implement the method according to any one of claims 1 to 8 when the computer instructions are executed.

18. A computer-readable storage medium having a computer program stored thereon, wherein when the program is executed by a processor, the method according to any one of claims 1 to 8 is implemented.