CLAIMS

What is claimed is:

1. A method for detectingn abnormality ~~detection method~~ for unmanned checkout,~~, wherein a photographing area for unmanned checkout comprises at least a depth camera of a detection area, the detection area is an area for detecting a to be paid product, and~~ the method comprisinges~~:~~

obtaining, by a ~~image data in the~~ depth camera in a detection area for unmanned checkout, ~~wherein the~~ image data that comprises a first depth image and a~~n~~ corresponding RGB image;

detecting, ~~by~~ using gesture detection and based on the first depth image and the corresponding RGB image~~ data~~, one or more ~~people~~gestures of customers that enter the detection area for unmanned checkout, wherein determining the one or more gestures of the customers comprises:~~the detection comprises one or more of headcount detection and gesture detection; and~~

identifying a first image region within the first depth image, wherein the first image region i) corresponds to hands of the customers and ii) comprises one or more image portions within the first depth image;

in response to identifying the first image region within the first depth image,

mapping the first image region in the first depth image to a second image region in the RGB image, the second image region corresponding to the hands of the customers; and

~~when it is~~ determining~~ed~~, based on the second image region in the RGB image, the one or more gestures of the customers;

determining that the one or more gestures comprise an abnormal gesture; and

in response to determining that the one or more gestures comprise an abnormal gesture, ~~detection result, that a predetermined abnormal condition is satisfied,~~ notifying a central control system to stop the unmanned checkout and trigger an alarm.~~product detection and triggering an alarm, wherein the abnormal condition comprises one or more of the following: the quantity of one or more people that enter the detection area is greater than a predetermined quantity threshold, and a gesture of the one or more people that enter the detection area is a predetermined abnormal gesture.~~

2. The method according to claim 1, further comprising~~wherein the step of headcount detection comprises~~:

determining~~obtaining~~, based on the first~~from the currently obtained~~ depth image, a quantity of customers entering the detection area for the unmanned checkout, wherein determining the quantity of customers comprises:

comparing, the first depth image~~ based on~~ with a background model,~~ and the depth image,~~ to obtain a foreground image, wherein the foreground image is associated with one or more~~ for representing a~~ moving objects, ~~wherein~~and the background model is obtained based on a plurality of second~~ background modeling using a~~ depth images including stationary objects;~~obtained when no moving object passes through the detection area; and~~

identifying, ~~f~~in the foreground image, ~~is used to determine that the moving object is~~ one or more regions representing the one or more moving objects,~~ people,~~ wherein each of the one or more regions is identified based on identifying a cluster of pixel values within a predetermined range of one another; and

determining the quantity of customers based on the one or more regions,

wherein determining the quantity of customers or the gesture detection comprises:

determining that the one or more moving object are the customers; and

performing connected component analysis on the foreground image based on one or more depth values of the first depth image to obtain connected components representing the one or more regions.

~~performing connected component analysis on the foreground image with reference to a depth value in the depth image, and obtaining the quantity of the one or more people in the detection area based on the analysis result.~~

3. The method according to claim 2~~1~~, wherein determining the quantity of customers further~~the depth image and the RGB image are obtained by the depth camera by collecting the same scenario at the same moment, and the step of gesture detection~~ comprises:

obtaining a headcount of the customers in the detection area based on results of the connected component analysis.~~, from the currently obtained depth image based on a background model and the depth image, a foreground image for representing a~~

~~moving object, wherein the background model is obtained based on background modeling using a depth image obtained when no moving object passes through the detection area;~~

~~if the foreground image is used to determine that the moving object is one or more people, performing connected component analysis on the foreground image with reference to a depth value in the depth image, and segmenting a people area in the foreground image based on the analysis result to obtain a connected component that represents an independent people; and~~

~~determining the gesture of the one or more people with reference to the connected component in the depth image and an area in the RGB image and corresponding to the connected component.~~

4. The method according to claim 2~~3~~, wherein the first depth image and the corresponding RGB image are obtained by the depth camera at substantially a same time of a same scenario, and the gesture detection comprises:

detecting~~rmining~~ each of the customers in the foreground image based on results of the ~~gesture of the one or more people with reference to the~~ connected component analysis to obtain the~~in the depth image and an area in the RGB image and corresponding to the~~ connected components that represent the customers; and~~comprises:~~

determining the one or more gestures of the customers based on the connected components in the first depth image and one or more regions corresponding to the connected components in the RGB image.~~locating hands of the people, in the depth image, that represent an independent people;~~

~~if a hand area is obtained, performing skeleton detection of the hands on an area in the RGB image and corresponding to the hand area, and determining the gesture of the one or more people based on the detection result; and~~

~~if a hand area is not obtained, obtaining an independent people area in the corresponding RGB image based on the connected component that represents an independent people in the depth image, performing skeleton detection of the hands on the independent people area in the RGB image, and determining the gesture of the one or more people based on the detection result.~~

5. The method according to ~~any one of~~ claim~~s 1 to~~ 4, wherein determining the one or more gestures further comprises:~~the method further comprises one or more of the following:~~

locating the first image region corresponding to the hands of the customers in the connected components in the first depth image; and~~when it is detected, by using the depth image, that one or more people enter the detection area with no people, notifying the central control system to start product detection; and~~

determining whether the first image region corresponding to the hands are located.~~when it is detected, by using the depth image, that the current one or more people leave and no other people enters the detection area, notifying the central control system to stop product detection.~~

6. The method according to claim 5, further comprising:

performing skeleton detection of the hands on the second image region in the RGB image; and

determining the one or more gestures of the customers based on results of the skeleton detection.~~An abnormality detection apparatus for unmanned checkout, wherein a photographing area for unmanned checkout comprises at least a depth camera of a detection area, the detection area is an area for detecting a to be paid product, and the apparatus comprises:~~

~~a data acquisition module, configured to obtain image data in the depth camera, wherein the image data comprises a depth image and an RGB image;~~

~~an abnormality detection module, configured to detect, by using the image data, one or more people that enter the detection area, wherein the detection comprises one or more of headcount detection and gesture detection; and~~

~~an abnormality processing module, configured to: when it is determined, based on the detection result, that a predetermined abnormal condition is satisfied, notify a central control system to stop product detection and trigger an alarm, wherein the abnormal condition comprises one or more of the following: the quantity of one or more people that enter the detection area is greater than a predetermined quantity threshold, and a gesture of the one or more people that enter the detection area is a predetermined abnormal gesture.~~

7. The method~~apparatus~~ according to claim 5~~6~~, further comprising~~wherein the abnormality detection module is configured to~~:

in response to determining that the first image region corresponding to the hands are not located, identifying regions in the RGB image corresponding to the connected

components in the first depth image;~~obtain, from the currently obtained depth image based on a background model and the depth image, a foreground image for representing a moving object, wherein the background model is obtained based on background modeling using a depth image obtained when no moving object passes through the detection area;~~

~~if the foreground image is used to determine that the moving object is one or more people,~~ performing skeleton detection of the hands on the regions in the RGB image; and

determining the one or more gestures of the customers based on results of the skeleton detection.~~connected component analysis on the foreground image with reference to a depth value in the depth image, and obtain the quantity of the one or more people in the detection area based on the analysis result.~~

8. The method~~apparatus~~ according to claim 1~~6~~, further comprising~~wherein the depth image and the RGB image are obtained by the depth camera by collecting the same scenario at the same moment, and the abnormality detection module is configured to~~:

detecting, based on the first depth image, that the customers enter the detection area that is unoccupied; and

notifying the central control system to start unmanned checkout.~~obtain, from the currently obtained depth image based on a background model and the depth image, a foreground image for representing a moving object, wherein the background model is obtained based on background modeling using a depth image obtained when no moving object passes through the detection area;~~

~~if the foreground image is used to determine that the moving object is one or more people, perform connected component analysis on the foreground image with reference to a depth value in the depth image, and segment a people area in the foreground image based on the analysis result to obtain a connected component that represents an independent people; and~~

~~determine the gesture of the one or more people with reference to the connected component in the depth image and an area in the RGB image and corresponding to the connected component.~~

9. The methodapparatus according to claim 8, further comprisingwherein the abnormality detection module is specifically configured to:

detecting, based on the first depth image, that the customers exit the detection area and the detection area becomes unoccupied; and

notifying the central control system to stop unmanned checkout.locate hands of the people, in the depth image, that represent an independent people;

if a hand area is obtained, perform skeleton detection of the hands on an area in the RGB image and corresponding to the hand area, and determine the gesture of the one or more people based on the detection result; and

if a hand area is not obtained, obtain an independent people area in the corresponding RGB image based on the connected component that represents an independent people in the depth image, perform skeleton detection of the hands on the independent people area in the RGB image, and determine the gesture of the one or more people based on the detection result.

10. A non-transitory, computer-readable medium storing one or more instructions executable by a computer system to perform operations comprising:

obtaining, by a depth camera in a detection area for unmanned checkout, image data that comprises a first depth image and a corresponding RGB image;

detecting, using gesture detection and based on the first depth image and the corresponding RGB image, one or more gestures of customers that enter the detection area for unmanned checkout, wherein determining the one or more gestures of the customers comprises:

identifying a first image region within the first depth image, wherein the first image region i) corresponds to hands of the customers and ii) comprises one or more image portions within the first depth image;

in response to identifying the first image region within the first depth image,

mapping the first image region in the first depth image to a second image region in the RGB image, the second image region corresponding to the hands of the customers; and

determining, based on the second image region in the RGB image, the one or more gestures of the customers;

determining that the one or more gestures comprise an abnormal gesture; and

in response to determining that the one or more gestures comprise an abnormal gesture, notifying a central control system to stop the unmanned checkout and trigger an alarm.~~The apparatus according to any one of claims 6 to 9, wherein the abnormality processing module is further configured to perform one or more of the following:~~

~~when it is detected, by using the depth image, that one or more people enter the detection area with no people, notifying the central control system to start product detection; and~~

~~when it is detected, by using the depth image, that the current one or more people leave and no other people enters the detection area, notifying the central control system to stop product detection.~~

11. The non-transitory, computer-readable medium according to claim 10, wherein the operations further comprise:

determining, based on the first depth image, a quantity of customers entering the detection area for the unmanned checkout, wherein determining the quantity of customers comprises:

comparing, the first depth image with a background model, to obtain a foreground image, wherein the foreground image is associated with one or more moving objects, and the background model is obtained based on a plurality of second depth images including stationary objects;

identifying, in the foreground image, one or more regions representing the one or more moving objects, wherein each of the one or more regions is identified based on identifying a cluster of pixel values within a predetermined range of one another; and

determining the quantity of customers based on the one or more regions,

wherein determining the quantity of customers or the gesture detection comprises:

determining that the one or more moving object are the customers; and

performing connected component analysis on the foreground image based on one or more depth values of the first depth image to obtain connected components representing the one or more regions. ~~A computer device, comprising a memory, a processor, and a computer program stored in the memory and running on the processor, wherein a photographing area for~~

~~unmanned checkout comprises at least a depth camera of a detection area, the detection area is an area for detecting a to be paid product, and the processor implements the following method when executing the program:~~

~~obtaining image data in the depth camera, wherein the image data comprises a depth image and an RGB image;~~

~~detecting, by using the image data, one or more people that enter the detection area, wherein the detection comprises one or more of headcount detection and gesture detection; and~~

~~when it is determined, based on the detection result, that a predetermined abnormal condition is satisfied, notifying a central control system to stop product detection and triggering an alarm, wherein the abnormal condition comprises one or more of the following: the quantity of one or more people that enter the detection area is greater than a predetermined quantity threshold, and a gesture of the one or more people that enter the detection area is a predetermined abnormal gesture.~~

12. The non-transitory, computer-readable medium according to claim 11, wherein determining the quantity of customers further comprises:

obtaining a headcount of the customers in the detection area based on results of the connected component analysis.

13. The non-transitory, computer-readable medium according to claim 11, wherein the first depth image and the corresponding RGB image are obtained by the depth camera at substantially a same time of a same scenario, and the gesture detection comprises:

detecting each of the customers in the foreground image based on results of the connected component analysis to obtain the connected components that represent the customers; and

determining the one or more gestures of the customers based on the connected components in the first depth image and one or more regions corresponding to the connected components in the RGB image.

14. The non-transitory, computer-readable medium according to claim 13, wherein determining the one or more gestures further comprises:

locating the first image region corresponding to the hands of the customers in the connected components in the first depth image; and

determining whether the first image region corresponding to the hands are located.

15. The non-transitory, computer-readable medium according to claim 14, the operations further comprising:

performing skeleton detection of the hands on the second image region in the RGB image; and

determining the one or more gestures of the customers based on results of the skeleton detection.

16. The non-transitory, computer-readable medium according to claim 14, the operations further comprising:

in response to determining that the first image region corresponding to the hands are not located, identifying regions in the RGB image corresponding to the connected components in the first depth image;

performing skeleton detection of the hands on the regions in the RGB image; and

determining the one or more gestures of the customers based on results of the skeleton detection.

17. The non-transitory, computer-readable medium according to claim 10, the operations further comprising:

detecting, based on the first depth image, that the customers enter the detection area that is unoccupied; and

notifying the central control system to start unmanned checkout.

18. The non-transitory, computer-readable medium according to claim 17, the operations further comprising:

detecting, based on the first depth image, that the customers exit the detection area and the detection area becomes unoccupied; and

notifying the central control system to stop unmanned checkout.

19. A computer-implemented system, comprising:

one or more computers; and

one or more computer memory devices interoperably coupled with the one or more computers and having tangible, non-transitory, machine-readable media storing one or more instructions that, when executed by the one or more computers, perform operations comprising:

obtaining, by a depth camera in a detection area for unmanned checkout, image data that comprises a first depth image and a corresponding RGB image;

detecting, using gesture detection and based on the first depth image and the corresponding RGB image, one or more gestures of customers that enter the detection area for unmanned checkout using gesture detection, wherein determining the one or more gestures of the customers comprises:

identifying a first image region within the first depth image, wherein the first image region i) corresponds to hands of the customers and ii) comprises one or more image portions within the first depth image;

in response to identifying the first image region within the first depth image,

mapping the first image region in the first depth image to a second image region in the RGB image, the second image region corresponding to the hands of the customers; and

determining, based on the second image region in the RGB image, the one or more gestures of the customers;

determining that the one or more gestures comprise an abnormal gesture; and

in response to determining that the one or more gestures comprise an abnormal gesture, notifying a central control system to stop the unmanned checkout and trigger an alarm.

20. The computer-implemented system according to claim 19, wherein the operations further comprise:

determining, based on the first depth image, a quantity of customers entering the detection area for the unmanned checkout, wherein determining the quantity of customers comprises:

comparing, the first depth image with a background model, to obtain a foreground image, wherein the foreground image is associated with one or more

moving objects, and the background model is obtained based on a plurality of second depth images including stationary objects;

identifying, in the foreground image, one or more regions representing the one or more moving objects, wherein each of the one or more regions is identified based on identifying a cluster of pixel values within a predetermined range of one another; and

determining the quantity of customers based on the one or more regions,

wherein determining the quantity of customers or the gesture detection comprises:

determining that the one or more moving object are the customers; and

performing connected component analysis on the foreground image based on one or more depth values of the first depth image to obtain connected components representing the one or more regions.

21. The computer-implemented system according to claim 20, wherein determining the quantity of customers further comprises:

obtaining a headcount of the customers in the detection area based on results of the connected component analysis.

22. The computer-implemented system according to claim 20, wherein the first depth image and the corresponding RGB image are obtained by the depth camera at substantially a same time of a same scenario, and the gesture detection comprises:

detecting each of the customers in the foreground image based on results of the connected component analysis to obtain the connected components that represent the customers; and

determining the one or more gestures of the customers based on the connected components in the first depth image and one or more regions corresponding to the connected components in the RGB image.

23. The computer-implemented system according to claim 22, wherein determining the one or more gestures further comprises:

locating the first image region corresponding to the hands of the customers in the connected components in the first depth image; and

determining whether the first image region corresponding to the hands are located.

24. The computer-implemented system according to claim 23, the operations further comprising:

performing skeleton detection of the hands on the second image region in the RGB image; and

determining the one or more gestures of the customers based on results of the skeleton detection.

25. The computer-implemented system according to claim 23, the operations further comprising:

in response to determining that the first image region corresponding to the hands are not located, identifying regions in the RGB image corresponding to the connected components in the first depth image;

performing skeleton detection of the hands on the regions in the RGB image; and

determining the one or more gestures of the customers based on results of the skeleton detection.

26. The computer-implemented system according to claim 19, the operations further comprising:

detecting, based on the first depth image, that the customers enter the detection area that is unoccupied; and

notifying the central control system to start unmanned checkout.

27. The computer-implemented system according to claim 26, the operations further comprising:

detecting, based on the first depth image, that the customers exit the detection area and the detection area becomes unoccupied; and

notifying the central control system to stop unmanned checkout.