

CHARACTER SEQUENCE RECOGNITION METHOD AND APPARATUS BASED ON COMPUTER VISION, AND DEVICE AND MEDIUM

CROSS-REFERENCE TO RELATED APPLICATIONS

5 **[0001]** The present application claims priority to Chinese Patent Application No.202010630553.0, titled "METHOD AND APPARATUS FOR RECOGNIZING CHARACTER SEQUENCE BASED ON COMPUTER VISION, DEVICE, AND MEDIUM", filed on July 03, 2020 with the China National Intellectual Property Administration, which is incorporated herein by reference in its entirety.

10

FIELD

[0002] The present disclosure relates to a method and an apparatus for recognizing a character sequence based on computer vision, a computer device, and a storage medium.

BACKGROUND

15 **[0003]** Character sequence recognition has acquired practical application in daily life with development of the computer vision technology. For example, a product serial number, a production date, and an embossing stamp are recognized in industrial scenarios. A general processing of recognizing a character sequence includes detecting a position of the character
20 sequence, retaining the character sequence at the detected position through cropping, and then performs angle determination and recognition on the character sequence in the cropped image to acquire corresponding text content. Alternatively, the character sequence serves as a target of special detection, which is performed by a classifier, and is aggregated into a word through a model based on image structure. Alternatively, a mapping relationship between an
25 image feature, and a position and content of the character sequence, is established based on a neural network algorithm, so as to recognize the character sequence.

SUMMARY

[0004] In a first aspect, a method for recognizing a character sequence based on computer
30 vision is provided. The method includes: acquiring an image including a character sequence

which is to be recognized; acquiring an sub-image of a target region, in which the character sequence is located, from the image based on a position detection model which is pre-constructed; performing horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region; acquiring a disposal state of the character sequence in the horizontally aligned sub-image based on an angle determination model which is pre-constructed; and inputting the horizontally aligned sub-image into a content recognition model which is pre-constructed, in a case that the disposal state of the character sequence is being right side up, to acquire content corresponding to the character sequence.

[0005] In a second aspect, an apparatus for recognizing a character sequence based on computer vision is provided. The apparatus includes: an image acquisition module, configured to acquire an image including a character sequence which is to be recognized; a position detection module, configured to acquire an sub-image of a target region, in which the character sequence is located, from the image based on a position detection model which is pre-constructed; a horizontal calibration module, configured to perform horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region; an angle determination module, configured to acquire a disposal state of the character sequence in the horizontally aligned sub-image based on an angle determination model which is pre-constructed; and a content recognition module, configured to input the horizontally aligned sub-image into a content recognition model which is pre-constructed, in a case that the disposal state of the character sequence is being right side up, to acquire content corresponding to the character sequence.

[0006] In a third aspect, a computer device is provided. The computer device includes a memory and a processor. The memory stores a computer program, and the processor when executing the program implements the foregoing method.

[0007] In a fourth aspect, a computer-readable storage medium is provided. The computer-readable storage medium stores a computer program, and the computer program when executed by a processor implements the foregoing method.

[0008] Hereinafter details of one or more embodiments of the present disclosure are provided in drawings and description. Other features and advantages of the present disclosure would be clarified in the specification, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] For clearer illustration of the technical solutions according to embodiments of the present disclosure or conventional techniques, hereinafter briefly described are the drawings to be applied in embodiments of the present disclosure or conventional techniques. Apparently, the drawings in the following descriptions are only some embodiments of the present disclosure, and other drawings may be obtained by those skilled in the art based on the provided drawings without creative efforts.

[0010] Figure 1 is a schematic flow chart of a method for recognizing a character sequence based on computer vision according to an embodiment of the present disclosure.

[0011] Figure 2 is a schematic flow chart of acquiring a disposal state of a character sequence in a horizontally aligned sub-image of a target region based on a pre-constructed angle determination model according to an embodiment of the present disclosure.

[0012] Figure 3 is a schematic flow chart of acquiring a sub-image of the target region in which a to-be-recognized character sequence is located from an image based on a pre-constructed position detection model according to an embodiment of the present disclosure.

[0013] Figure 4 is a schematic flow chart of inputting a horizontally aligned sub-image of a target region is inputted into a pre-constructed content recognition model to acquire content corresponding to a to-be-recognized character sequence according to an embodiment of the present disclosure.

[0014] Figure 5 is a schematic flow chart of a method for recognizing a character sequence based on computer vision according to another embodiment of the present disclosure.

[0015] Figure 6 is a schematic flow chart of algorithm training and prediction processing according to an embodiment of the present disclosure.

[0016] Figure 7 is a schematic structural diagram of an image pyramid according to an embodiment of the present disclosure.

[0017] Figure 8 is a schematic flow chart of an algorithm for determining an alignment angle of a character sequence according to an embodiment of the present disclosure.

[0018] Figure 9 is a schematic flow chart of an algorithm for recognizing content of a character sequence according to an embodiment of the present disclosure.

5 **[0019]** Figure 10 is a structural block diagram of an apparatus for recognizing a character sequence based on computer vision according to an embodiment of the present disclosure.

[0020] Figure 11 is a schematic diagram of an internal structure of a computer device according to an embodiment of the present disclosure.

10 **DETAILED DESCRIPTION**

[0021] At present, a conventional method for recognizing a character sequence is based on a low-dimensional manual feature, and lacks a capability of adaptive processing for a change in an angle of capturing images in industrial scenarios. Therefore, such conventional method has low accuracy when recognizing character sequences.

15 **[0022]** Hereinafter the present disclosure is described in detail in conjunction with the drawings and embodiments, in order to clarify objectives, technical solutions, and advantages of the present disclosure. It should be understood that the specific embodiments described herein are merely intended for explaining the present disclosure, not for limiting the present disclosure.

20 **[0023]** Reference is made to Figure 1. In an embodiment, a method for recognizing a character sequence based on computer vision is provided. Herein it is taken as an example that the method is applied to a terminal. It is appreciated that the method may be applied to a server, or may be applied to a system including a terminal and a server and implemented through interaction between the terminal and the server. In this embodiment, the method
25 includes following steps S101 to S105.

[0024] In step S101, a terminal acquires an image including a character sequence that is to be recognized.

[0025] The to-be-recognized character sequence refers to a character sequence which a user need to acquire from the image. The image may be captured in an industrial scenario. The

user may record the image including the to-be-recognized character sequence in different scenarios via, for example, a mobile camera or a video collector, and store the image in the terminal. Thereby, the terminal can acquire the image carrying the to-be-recognized character sequence.

5 **[0026]** In step S102, the terminal acquires a sub-image of a target region, in which the character sequence is located, from the image based on a position detection model which is pre-constructed.

10 **[0027]** The position detection model is configured to detect a region, in which the to-be-recognized character sequence is located, in the image. The sub-image of the target region refers to an image of such region. The terminal may detect a position of the character sequence in the image, which includes the to-be-recognized character sequence, through the pre-constructed position detection model. Thereby, the terminal determines the sub-image of the target region in which the to-be-recognized character sequence is located.

15 **[0028]** In step S103, the terminal performs horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region.

20 **[0029]** Generally, the user may capture the image including the to-be-recognized character sequence from variously angles. Hence, the to-be-recognized character sequence is usually not horizontally aligned in the image, but slants from a horizontal position by a certain angle. After acquiring the sub-image of the target region such sub-image in the step S102, the terminal is required to perform horizontal calibration on such sub-image and thereby acquires the horizontally aligned sub-image of the target region, in order to improve accuracy of recognizing the character sequence. The to-be-recognized character sequence is aligned along a horizontal direction in the horizontally aligned sub-image. The terminal may perform affine transformation on the sub-image of the target region to implement the horizontal calibration. Thereby, the terminal acquires the horizontally aligned sub-image of the target region.

25 **[0030]** In step S104, the terminal acquires a disposal state of the character sequence in the horizontally aligned sub-image based on an angle determination model which is pre-constructed.

30 **[0031]** After the terminal performs the horizontal calibration on the sub-image of the target

region in the step S103, the character sequence may be disposed right side up state or upside down in the horizontally aligned sub-image, which depends on an angle of the user capturing the original image. In a case that the character sequence is disposed upside down, such disposal state of the character sequence would affect a final result of recognizing the character sequence. Therefore, after acquiring the horizontally aligned sub-image of the target region, the terminal is required to first determine the disposal state of the character sequence in the acquired horizontally aligned sub-image. The terminal may input the horizontally aligned sub-image of the target region into the pre-constructed angle determination model, so as to determine the disposal state of the character sequence in the horizontally aligned sub-image.

[0032] In step S105, in a case that the disposal state of the character sequence is being right side up, the terminal inputs the horizontally aligned sub-image into a content recognition model that is pre-constructed, so as to acquire content corresponding to the character sequence.

[0033] In a case that the terminal determines that the character sequence is currently disposed right side up, the terminal may input the horizontally aligned sub-image of the target region directly into the pre-constructed content recognition model. The content recognition model is configured recognizing the content of the character sequence in the sub-image of the target region. Therefore, the terminal can acquire the content corresponding to the to-be-recognized character sequence through the content recognition model.

[0034] Hereinabove illustrated is the method for recognizing the character sequence based on computer vision. The terminal acquires the image including the to-be-recognized character sequence, acquires the sub-image of the target region in which the to-be-recognized character sequence is located from the image based on the pre-constructed position detection model, performs the horizontal calibration on the sub-image of the target region to acquire the horizontally aligned sub-image of the target region, acquires the disposal state of the character sequence in the horizontally aligned sub-image based on the pre-constructed angle determination model, and inputs the horizontally aligned sub-image into the pre-constructed content recognition model in a case that the disposal state of the character sequence is being right side up, so as to acquire the content corresponding to the to-be-recognized character sequence. Herein the terminal performs the horizontal calibration on the sub-image of the target region, which implements adaptive processing for a change in an angle of capturing

images in the industrial scenes. Thereby, accuracy of recognizing character sequences is improved.

[0035] Reference is made to Figure 2. In an embodiment, the step S104 includes following steps S201 and S202.

5 **[0036]** In step S201, the terminal acquires a disposal angle of the horizontally aligned sub-image of the target region based on the angle determination model.

[0037] The angle determination model is configured to determine an angle of the horizontally aligned sub-image of the target region. The disposal state of the character sequence depends mainly on the angle of the user capturing the original image captured.

10 Hence, the terminal may determine the disposal angle of the horizontally aligned sub-image of the target region through the angle determination model, and determine the disposal state of the character sequence according to the disposal angle.

[0038] In step S202, the terminal determines the disposal state of the character sequence according to an angle range in which the disposal angle is located.

15 **[0039]** There may be a small deviation between the disposal angle, of the horizontally aligned sub-image of the target region, acquired by the terminal and an angle corresponding to standard horizontal alignment. Hence, after determining the disposal angle of the horizontally aligned sub-image through the angle determination model in step S201, the terminal may further select an angle range corresponding to the disposal angle from a preset
20 list of angle ranges, as the angle range in which the disposal angle is located. The terminal determines the disposal state of the character sequence by using such angle range.

[0040] Further, the angle ranges may include a first angle range and a second angle range, and the disposal state of the character sequence may include being right side up and being upside down. The step S202 may further include following sub-steps. In a case that the
25 angle range in which the disposal angle is located is the first angle range, the terminal determines that the disposal state of the character sequence is being right side up. In a case that the angle range in which the disposal angle is located is the second angle range, the terminal determines that the disposal state of the character sequence is being upside down.

[0041] The first angle range is different from the second angle range. The first angle
30 range and the second angle range characterizes two disposal states, respectively, of the

character sequence. The terminal may determine that the horizontally aligned sub-image of the target region is currently disposed right side up, in a case that the obtained disposal angle of the horizontally aligned sub-image is located in the first angle range. The terminal may determine that the horizontally aligned sub-image of the target region is currently disposed
5 upside down, in a case that the obtained disposal angle of the horizontally aligned sub-image is located in the second angle range.

[0042] Moreover, in a case that the disposal state of the character sequence is being upside down, the horizontally aligned sub-image of the target region is first rotated to the right-side-up state and then inputted into the content recognition model, so as to acquire the
10 content of the character sequence.

[0043] The content of the character sequence, which is acquired through the content recognition model, may deviate from the actual content of the character sequence, in a case that the terminal inputs the horizontally aligned sub-image, which is disposed upside down, directly into the content recognition model. Therefore, before being inputted into the
15 content recognition model, the horizontally aligned sub-image of the target region should be rotated first such that it is disposed right side up. For example, the horizontally aligned sub-image of the target region is rotated by 180° with respect to its center, and the rotated horizontally aligned sub-image of the target region can be inputted into the content recognition model, so as to acquire the content corresponding to the to-be-recognized
20 character sequence.

[0044] Herein the terminal can acquire the disposal angle of the horizontally aligned sub-image of the target region through the angle determination model, and thereby determine the disposal state of the character sequence. In a case that the disposal state of the character sequence being upside down, the terminal may rotate the horizontally aligned sub-image to
25 the right-side-up state, and input the horizontally aligned sub-image which is disposed right side up into the content recognition model to acquire the content of the character sequence. Thereby, the acquired content of the character sequence is more accurate.

[0045] Reference is made to Figure 3. In an embodiment, the step S102 includes following steps S301 to S303.

[0046] In step S301, the terminal extracts an image feature concerning a character region
30 from the image through the position detection model.

[0047] The image feature concerning a character region is configured to determine a position of the character sequence. The terminal may extract such image feature from the acquired image, which includes the to-be-recognized character sequence, through the position detection model.

5 **[0048]** In step S302, the terminal acquires a prediction mask for the sub-image of the target region according to the image feature.

[0049] A mask refers to occluding a to-be-processed image totally or partially with a selected image, a selected graphic, or a selected object, such that the processing itself or an image region subject to the processing can be controlled. The terminal may acquire the
10 prediction mask corresponding to the image feature concerning the character region according to the image feature itself.

[0050] In step S303, the terminal computes a connected domain and a minimum bounding rectangle based on the prediction mask, so as to acquire the sub-image of the target region.

[0051] After acquiring the prediction mask for the sub-image of the target region in the step
15 S302, the terminal may compute the connected domain based on the prediction mask and accordingly compute the minimum bounding rectangle, so as to acquire the sub-image of the target region.

[0052] Moreover, the image including the to-be-recognized character sequence, which is acquired by the terminal, may be unclear or under poor illumination, which reduces accuracy
20 of recognizing the character sequence. In order to address such issue, the step S301 may further include following steps in one embodiment. The terminal pre-processes the image and extracts a high-dimensional image feature from the pre-processed image. The terminal performing first feature enhancement on the high-dimensional image feature through an image pyramid. The terminal determines the enhanced high-dimensional image feature as
25 the image feature concerning the character region.

[0053] The pre-processing refers to that the terminal filters the image including the to-be-recognized character sequence to remove a region in which a small or unidentifiable character sequence is located, so that the terminal is capable to extract the high-dimensional image feature from the image including the to-be-recognized character sequence. Moreover,
30 the terminal may further perform the first feature enhancement on the extracted

high-dimensional image feature through the image pyramid, which facilitates an expressing capability of the image feature concerning the character region. Thereby, the prediction mask for the sub-image of the target region can be accurately generated in an environment without a clear feature.

5 **[0054]** Herein the terminal may extract the image feature concerning the character region from the image to generate the prediction mask corresponding to such image feature, and computes the connected domain and the minimum bounding rectangle based on the prediction mask to acquire the sub-image of the target region accurately. In addition, in order to avoid the character sequence being missed or erroneously recognized due to an unclear feature, the
10 terminal may perform the first feature enhancement on the extracted image feature through the image pyramid, so that the expressing capability of the image feature concerning the character region is improved. Hence, the accuracy of recognizing the character sequence can be further improved.

15 **[0055]** Reference is made to Figure 4. In an embodiment, the step S105 includes following steps S401 to S403.

[0056] In step S401, the terminal extracts a global image feature from the horizontally aligned sub-image of the target region through the content recognition model, so as to acquire an image feature which concerns the character sequence and corresponds to the horizontally aligned sub-image.

20 **[0057]** The content recognition model is configured to recognizing character content in the character sequence which is included in the horizontally aligned sub-image of the target region. The terminal may extract the global image feature from the horizontally aligned sub-image of the target region through the content recognition model, so as to acquire the image feature, which concerns the character sequence, corresponding to the horizontally
25 aligned sub-image of the target region.

[0058] In step S402, the terminal performs second feature enhancement on the image feature concerning the character sequence along a horizontal direction by using a convolution kernel which is a row vector.

30 **[0059]** The second feature enhancement refers to enhancing the image feature concerning the character sequence. After the image feature concerning the character sequence is

acquired in the step S401, the second feature enhancement is performed on such image feature along the horizontal direction, that is, along an alignment direction of the character sequence, by using the row-vector convolution kernel.

[0060] In step S403, the terminal performs parallel prediction on the to-be-recognized character sequence based on the image feature concerning the character sequence, which is after the second feature enhancement, so as to acquire the content of the character sequence.

[0061] The terminal may recognize the content of the character sequence based on the image feature concerning the character sequence, which has been subject to the second feature enhancement, in order to further improve the efficiency of recognizing the character sequence. Such recognition is the parallel prediction, which can predict multiple character sequences simultaneously. Thereby, the content of the character sequences can be predicted with high efficiency.

[0062] Herein the terminal may recognize the content of the character sequence accurately through the content recognition model, and perform the second feature enhancement on the image feature concerning the character sequence, which can improve the expression capability of the image feature, and thereby improve accuracy of recognizing the content of the character sequence. Moreover, all character sequences are predicted in parallel, which can further improve the efficiency of recognizing the content of the character sequence.

[0063] Reference is made to Figure 5. In an embodiment, a method for recognizing a character sequence based on computer vision is provided. In this embodiment, it is taken as an example that the method is applied to a terminal. The method includes following steps S501 to S510.

[0064] In step S501, a terminal acquires an image including a to-be-recognized character sequence.

[0065] In step S502, the terminal pre-processes the image, extracts a high-dimensional image feature from the pre-processed image, performs first feature enhancement on the high-dimensional image feature through an image pyramid, and determines the enhanced high-dimensional image feature as an image feature concerning a character region.

[0066] In step S503, the terminal acquires a prediction mask for a sub-image of a target region according to the image feature concerning the character region, and computes a

connected domain and a minimum bounding rectangle based on the prediction mask to acquire the sub-image of the target region.

[0067] In step S504, the terminal performs horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region.

5 **[0068]** In step S505, the terminal acquires a disposal angle of the horizontally aligned sub-image of the target region through an angle determination model.

[0069] In step S506, the terminal determines that a disposal state of the character sequence is being right side up in a case that the disposal angle is located in a first angle range, and determines that a disposal state of the character sequence is being upside down in a case that
10 the disposal angle is located in a second angle range.

[0070] In step S507, the terminal inputs the horizontally aligned sub-image of the target region into a pre-constructed content recognition model in a case that the disposal state of the character sequence is being right side up, and rotates the horizontally aligned sub-image of the target region to be disposed right side up and then inputs the rotated horizontally aligned
15 sub-image into the content recognition model in a case that the disposal state of the character sequence is being upside down.

[0071] In step S508, the terminal extracts a global image feature from the horizontally aligned sub-image of the target region through the content recognition model to acquire an image feature which concerns the character sequence and corresponds to the horizontally
20 aligned sub-image of the target region.

[0072] In step S509, the terminal performs second feature enhancement on the image feature concerning the character sequence along a horizontal direction by using a convolution kernel which is a row vector.

[0073] In step S510, the terminal performs parallel prediction on the to-be-recognized
25 character sequence based on the image feature concerning the character sequence, which is after the second feature enhancement, to acquire character sequence content.

[0074] Herein the terminal performs the horizontal calibration on the sub-image of the target region, which achieves the adaptive processing for a change in an angle of capturing images in industrial scenarios. Hence, accuracy of recognizing the character sequence is
30 improved. In addition, the terminal acquires the disposal angle of the horizontally aligned

sub-image of the target region through the angle determination model, so as to determine the disposal state of the character sequence. In a case that the disposal state of the character sequence being upside down, the terminal may rotate the horizontally aligned sub-image to be the right-side-up state, so as to further improve accuracy of the acquired content of the character sequence. The terminal may further perform the first feature enhancement on the extracted high-dimensional image feature through the image pyramid, and perform the second feature enhancement on the image feature concerning the character sequence, which can improve the expression ability of these features. Hence, accuracy of recognizing the content of the character sequence is further improved. Moreover, all character sequences are predicted in parallel, which can improve efficiency of recognizing the content of the character sequence.

[0075] In an embodiment, an algorithm for recognizing a character sequence under any angle in an industrial scenario is further provided. The algorithm is capable to address an issue that characters are missed or erroneously recognized under dim environment, poor illumination, or varying angles in convention algorithms for recognizing characters, and hence improve recognition accuracy. The present disclosure is applicable to an industrial scenario having a poor environment for camera imaging, while ensuring high efficiency and accuracy of the recognition algorithm and supporting character recognition under various angles even the upside down state. Figure 6 shows a flow chart of training and prediction of the algorithm, where the process includes training and prediction. During the training, it is required to train three different models, which are configured to detect a position of the character sequence, determine an angle of the character sequence, and recognize content of the character sequence, respectively. During the prediction, the trained models perform position detection, angle determination, and content recognition sequentially on the to-be-detected image. Hence, the position and the corresponding content of the character sequence are acquired.

[0076] Hereinafter a processing flow of all modules is illustrated.

[0077] I. Training

[0078] 1.1 Algorithm for detecting a position of a character Sequence

[0079] A training sample refers to a whole sample image including the character sequence. A corresponding label is a position box for the character sequence in the image, and includes

coordinate information of a position of the character sequence, for example, coordinate information of an upper left corner of a head of the character sequence and coordinate information of a lower right corner of a tail of the character sequence. Since different training samples have various dimensions and colors, the samples are required to be normalized, and the position box of the character sequence that is small or unidentifiable is removed through filtering. Data of the pre-processed the image serves as an input of the algorithm for detecting the position of the character sequence, and the feature of the image is enhanced through a deep neural network in conjunction with an image pyramid. As shown in Figure 7, the "conv"s represent different convolution layers, and the "stride"s represent different strides. Each extracted feature under different scales are up-sampled, and then is superimposed with a feature acquired from the previous network(s), to acquire a final image feature containing both spatial information and semantic information. The image feature acquired by the algorithm for detecting the position is configured to predict a mask corresponding to a final region in which the character sequence is located. The position box of the character sequence can be acquired by computing a connected domain and a minimum bounding rectangle based on the mask.

[0080] 1.2 Algorithm for determining an alignment angle of the character sequence

[0081] Reference is made to Figure 8. In a case that the alignment angle of the character sequence is greater than 0° and less than 180° , an image of the horizontally aligned character sequence can be obtained through affine transformation. After the horizontal calibration, it is possible that character sequence is disposed either right side up or upside down, which depends on the angle of capturing the original image. Hence, the angle determination algorithm is applied to determine whether the character sequence is disposed upside down. In a case that the corrected character sequence is disposed upside down, the corrected character sequence is rotated by 180° with respect to its center. In a case that the corrected character sequence is disposed right side up, the character sequence is directly outputted without processing. Hence, it is ensured that the final image of the character sequence is always in the right-side-up state when outputted for content recognition in the next stage.

[0082] 1.3 Algorithm for recognizing content of the character sequence

[0083] Reference is made to Figure 9. When recognizing the content of the character sequence, an image feature concerning the character sequence is learned through the deep

neural network. The extracted image feature is subject to feature enhancement along the alignment direction of the character sequence by using a row-vector convolution kernel, such that a feature concerning the whole sequence is obtained. Thereby, the content of the character sequences can be predicted in parallel with high efficiency.

5 [0084] II. Prediction

[0085] The to-be-detected image is inputted, and a position of the character sequence in the image is detected through the algorithm for detecting the position of the character sequence. Then, a region of the image, which corresponds to the detected position, is cropped and subject to affine transformation. Afterwards, the cropped and transformed region is inputted
10 into the algorithm for determining the alignment angle of the character sequence. The image of the cropped and transformed region is rotated by 180° with respect to its center in a case that such image is determined to be disposed upside down, and is subject to no processing in a case that such image is determined to be disposed right side up. After being processed by the algorithm for detecting the position of the character sequence and the algorithm for
15 determining the alignment angle of the character sequence, the region serves as an input of the network for recognizing the content of the character sequence. Finally, the position and the corresponding content of the character sequence in the image are acquired through the content recognition network.

[0086] Herein the algorithm for recognizing the character sequence, which is stable and
20 efficient for common industrial scenarios having varying imaging sharpness, varying imaging angle, and varying imaging change, is acquired by concatenating algorithms for three stages, i.e., the algorithm for detecting the position of the character sequence, the algorithm for determining the alignment angle of the character sequence, and the algorithm for recognizing the content of the character sequence. Such algorithm provides a good basis for recognizing
25 character sequences in industrial scenarios.

[0087] Although steps in the foregoing flowcharts are illustrated in a sequence indicated by arrows, it is not necessary to execute these steps according to the indicated sequence. Unless clearly specified in this specification, there is no strict limitation on a sequence of executing these steps, and the steps may be performed in another sequence. In addition, at least one
30 step in these flowcharts may include multiple sub-steps or multiple stages. It is not necessary to execute the sub-steps or the stages simultaneously. Rather, their execution may

differ in time. Moreover, it is not necessary to execute the sub-steps or the stages sequentially. Rather, they may be executed in turn or alternately along with at least one sub-step or stage of another step.

[0088] Reference is made to Figure 10. In an embodiment, an apparatus for recognizing a character sequence based on computer vision is provided. The apparatus includes an image acquisition module 1001, a position detection module 1002, a horizontal calibration module 1003, an angle determination module 1004, and a content recognition module 1005.

[0089] The image acquisition module 1001 is configured to acquire an image including a character sequence which is to be recognized.

[0090] The position detection module 1002 is configured to acquire a sub-image of a target region, in which the character sequence is located, from the image based on a position detection model which is pre-constructed.

[0091] The horizontal calibration module 1003 is configured to perform horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region.

[0092] The angle determination module 1004 is configured to acquire a disposal state of the character sequence in the horizontally aligned sub-image based on an angle determination model which is pre-constructed.

[0093] The content recognition module 1005 is configured to input the horizontally aligned sub-image into a content recognition model which is pre-constructed, in a case that the disposal state of the character sequence is being right side up, to acquire content corresponding to the character sequence.

[0094] In an embodiment, the angle determination module 1004 is further configured to: acquire a disposal angle of the horizontally aligned sub-image of the target region based on the angle determination model; and determine the disposal state of the character sequence according to an angle range in which the disposal angle is located.

[0095] In an embodiment, the angle range is a first angle range or a second angle range. The disposal state of the character sequence is being right side up or being upside down. The angle determination module 1004 is further configured to: determine that the disposal state of the character sequence is being right side up, in a case that the angle range is the first

angle range; and determine that the disposal state of the character sequence is being upside down, in a case that the angle range is the second angle range.

[0096] In an embodiment, the content recognition module 1005 is further configured to rotate the horizontally aligned sub-image of the target region to be disposed right side up, and then input the rotated horizontally aligned sub-image of the target region into the content recognition model, in a case that the disposal state of the character sequence is being upside down, to acquire the content corresponding to the character sequence.

[0097] In an embodiment, the position detection module 1002 is further configured to: extract an image feature concerning a character region from the image through the position detection model; acquire a prediction mask for the sub-image of the target region according to the image feature; and compute a connected domain and a minimum bounding rectangle based on the prediction mask to acquire the sub-image of the target region.

[0098] In an embodiment, the position detection module 1002 is further configured to: pre-process the image; extract a high-dimensional image feature from the pre-processed image; and perform first feature enhancement on the high-dimensional image feature through an image pyramid to acquire the image feature concerning the character region.

[0099] In an embodiment, the content recognition module 1005 is further configured to extract a global image feature from the horizontally aligned sub-image of the target region through the content recognition model to acquire an image feature which concerns the character sequence and corresponds to the horizontally aligned sub-image; perform second feature enhancement on the image feature concerning the character sequence along a horizontal direction by using a convolution kernel which is a row vector; and perform parallel prediction on the to-be-recognized character sequence based on the image feature concerning the character sequence, which is after the second feature enhancement, to acquire the content of the character sequence.

[0100] Specific limitations on the apparatus for recognizing the character sequence based on computer vision may refer to those on the methods in the foregoing description, and are not repeated herein. Each module in the foregoing apparatus for recognizing the character sequence based on computer vision may be wholly or partially implemented through software, hardware, or a combination of the two. The foregoing modules may be hardware built in, or independent from, a processor of a computer device, or may be software stored in a memory

of the computer device, so as to facilitate the processor invoking the foregoing modules to perform corresponding operations.

[0101] In an embodiment, a computer device is provided. The computer device may be a terminal, and an internal structural diagram thereof may be as shown in Figure 11. The computer device includes a processor, a memory, and a communication interface, a display screen, and an input apparatus that are connected via a system bus. The processor of the computer device is configured to provide computing and control capabilities. The memory of the computer device includes a non-volatile storage medium and an internal memory. The non-volatile storage medium stores an operating system and computer-readable instructions. The internal memory provides an environment for running the operating system and the computer-readable instructions in the non-volatile storage medium. The communication interface of the computer device is configured to communicate with an external terminal in a wired or wireless manner. The wireless communication may be implemented by Wi-Fi, an operator network, near field communication (NFC), or other technologies. The computer-readable instructions when executed by the processor implements the foregoing method for face liveness detection. The display screen of the computer device may be a liquid crystal display screen or an electronic ink display screen. An input apparatus of the computer device may be a touch control layer covering the display screen, may be a button, a trackball, or a touchpad disposed on housing of the computer device, and may be a keyboard, a touchpad, a mouse, or the like that is connected to the computer device.

[0102] The structures as shown in Figure 11 are merely a block diagram of some structures related to the present disclosure, and do not constitute a limitation on the computer devices to which the present disclosure is applicable. A specific example of such computer devices may include components more or less than those as shown in Figure 11, some components as shown may be combined, or the components as shown may have different arrangement.

[0103] In one embodiment, a computer device is further provided. The computer device includes a memory and a processor. The memory stores computer-readable instructions, and the computer-readable instructions when executed by the processor implement the steps in the foregoing method embodiments.

[0104] In one embodiment, a computer-readable storage medium is provided. The computer-readable storage medium stores computer-readable instructions, and the

computer-readable instructions when executed by a processor implement the steps in the foregoing method embodiments.

[0105] Those skilled in the art may understand that all or some of procedures of the method in the foregoing embodiments may be implemented by a computer program instructing relevant hardware. The program may be stored in a non-volatile computer-readable storage medium. The procedures of the foregoing method embodiments may be implemented when the program is executed. All references to the memory, the storage, the database, or other medium used in embodiments of the present disclosure may include a non-volatile and/or a volatile memory. The non-volatile memory may include a read-only memory (ROM), a cassette, a floppy disk, a flash memory, and an optical storage. The volatile memory may include a random access memory (RAM) or an external cache. By way of description rather than limitation, the RAM may be implemented in various forms, such as a static RAM (SRAM) and a dynamic RAM (DRAM).

[0106] Technical features of the foregoing embodiments may be arbitrarily combined. In order to make description concise, not all possible combinations of the technical features in the foregoing embodiments are described. The combinations of these technical features shall be considered as falling within the scope covered by this specification as long as there is no conflict.

[0107] The foregoing embodiments only show some implementations of the present disclosure, and are described in detail. These embodiments should not be construed as a limitation on the scope of the present disclosure. Several transformations and improvements can be made by those skilled in the art without departing from the concept of the present disclosure, and belong to the protection scope of the present disclosure. Therefore, the protection scope of the present disclosure shall be subject to the appended claims.

CLAIMS

1. A method for recognizing a character sequence based on computer vision, comprising:

acquiring an image including a character sequence which is to be recognized;

acquiring an sub-image of a target region, in which the character sequence is located, from the image based on a position detection model which is pre-constructed;

performing horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region;

acquiring a disposal state of the character sequence in the horizontally aligned sub-image based on an angle determination model which is pre-constructed; and

inputting the horizontally aligned sub-image into a content recognition model which is pre-constructed, in a case that the disposal state of the character sequence is being right side up, to acquire content corresponding to the character sequence,

wherein inputting the horizontally aligned sub-image into the content recognition model to acquire the content corresponding to the character sequence comprises:

extracting a global image feature from the horizontally aligned sub-image of the target region through the content recognition model to acquire an image feature which concerns the character sequence and corresponds to the horizontally aligned sub-image;

performing second feature enhancement on the image feature concerning the character sequence along a horizontal direction by using a convolution kernel which is a row vector; and

performing parallel prediction on the to-be-recognized character sequence based on the image feature concerning the character sequence, which is after the second feature enhancement, to acquire the content of the character sequence.

2. The method according to claim 1, wherein acquiring the disposal state of the character sequence in the horizontally aligned sub-image based on the angle determination model which is pre-constructed comprises:

acquiring a disposal angle of the horizontally aligned sub-image of the target region

based on the angle determination model; and

determining the disposal state of the character sequence according to an angle range in which the disposal angle is located.

3. The method according to claim 2, wherein:

the angle range is a first angle range or a second angle range;

the disposal state of the character sequence is being right side up or being upside down;
and

determining the disposal state of the character sequence according to the angle range in which the disposal angle is located comprises:

determining that the disposal state of the character sequence is being right side up,
in a case that the angle range is the first angle range.

4. The method according to claim 2, wherein:

the angle range is a first angle range or a second angle range;

the disposal state of the character sequence is being right side up or being upside down;
and

determining the disposal state of the character sequence according to the angle range in which the disposal angle is located comprises:

determining that the disposal state of the character sequence is being upside down,
in a case that the angle range is the second angle range.

5. The method according to claim 3 or 4, further comprising:

rotating the horizontally aligned sub-image of the target region to be disposed right side up, and then inputting the rotated horizontally aligned sub-image of the target region into the content recognition model, in a case that the disposal state of the character sequence is being upside down, to acquire the content corresponding to the character sequence.

6. The method according to claim 1, wherein acquiring the sub-image of the target region, in which the character sequence is located, from the image based on the position detection model which is pre-constructed comprises:

extracting an image feature concerning a character region from the image through the

position detection model;

acquire the prediction mask for the sub-image of the target region according to the image feature; and

computing a connected domain and a minimum bounding rectangle based on the prediction mask to acquire the sub-image of the target region.

7. The method according to claim 6, wherein extracting the image feature concerning the character region from the image comprises:

pre-processing the image;

extracting a high-dimensional image feature from the pre-processed image; and

performing first feature enhancement on the high-dimensional image feature through an image pyramid to acquire the image feature concerning the character region.

8. The method according to claim 7, wherein the pre-processing comprises:

filtering the image including the to-be-recognized character sequence to remove a region, in which a small or unidentifiable character sequence is located, such that the high-dimensional image feature is capable to be extracted from the filtered image.

9. An apparatus for recognizing a character sequence based on computer vision, comprising:

an image acquisition module, configured to acquire an image including a character sequence which is to be recognized;

a position detection module, configured to acquire an sub-image of a target region, in which the character sequence is located, from the image based on a position detection model which is pre-constructed;

a horizontal calibration module, configured to perform horizontal calibration on the sub-image of the target region to acquire a horizontally aligned sub-image of the target region;

an angle determination module, configured to acquire a disposal state of the character sequence in the horizontally aligned sub-image based on an angle determination model which is pre-constructed; and

a content recognition module, configured to input the horizontally aligned sub-image

into a content recognition model which is pre-constructed, in a case that the disposal state of the character sequence is being right side up, to acquire content corresponding to the character sequence;

wherein the content recognition module is configured to:

extract a global image feature from the horizontally aligned sub-image of the target region through the content recognition model to acquire an image feature which concerns the character sequence and corresponds to the horizontally aligned sub-image;

perform second feature enhancement on the image feature concerning the character sequence along a horizontal direction by using a convolution kernel which is a row vector; and

perform parallel prediction on the to-be-recognized character sequence based on the image feature concerning the character sequence, which is after the second feature enhancement, to acquire the content of the character sequence..

10. A computer device, comprising;

a memory, storing a computer program; and

a processor;

wherein the processor when executing the computer program implements the method according to any one of claims 1 to 8.

11. A computer-readable storage medium, storing a computer program, wherein:

the computer program when executed by a processor implements the method according to any one of claims 1 to 8.