

FACE IDENTIFICATION METHODS AND APPARATUSES

TECHNICAL FIELD

[0001] The embodiments of the present disclosure relate to the field of image processing technology, and in particular, to a face identification method and apparatus.

BACKGROUND

[0002] Face identification is the most basic and important part of intelligent video analysis. The target object in the video is needed to be tracked for a long time, and when receiving an identification request issued by the upper-layer application, the face identification and identity determination are performed on the tracked target object.

SUMMARY

[0003] In view of this, the embodiments of the present disclosure provide at least a method and apparatus for face identification.

[0004] In a first aspect, a face identification method is provided, and the method including:

[0005] determining a facial image sequence of a target face, where the facial image sequence includes multiple facial region images of the target face;

[0006] performing face identification on at least one facial region image in the target facial image sequence, and determining, based on a confidence level of a face identification result of the at least one facial region image, a target facial region image from the target facial image sequence;

[0007] determining, based on a face identification result of the target facial region image, identity information of the target face;

[0008] In a second aspect, a face identification apparatus is provided, the apparatus including:

[0009] a target facial image sequence determination module, configured to determining a facial image sequence of a target face, where the facial image sequence includes multiple facial region images of the target face;

[0010] a target facial region image determination module, configured to perform face identification on at least one facial region image in the target facial image sequence, and

determining, based on a confidence of a face identification result of the at least one facial region image, a target facial region image from the target facial image sequence;

[0011] an identity information determination module, configured to determine, based on a face identification result of the target facial region image, identity information of the target face.

[0012] In a third aspect, an electronic device is provided, including a memory and a processor, where the memory is configured to store computer instructions capable of being run on the processor, and the processor implements the face identification method according to the first aspect when executing the computer instructions.

[0013] In a fourth aspect, a computer-readable storage medium is provided, storing a computer program, where steps of the face identification method according to the first aspect are implemented when the program is executed by a processor.

[0014] In the face identification method provided by the technical solution of embodiments in the present disclosure, by performing face identification on a facial region image in a target facial image sequence for a target face, a target facial region image is determined. Further, identify information of the target face is determined. A target facial image sequence is working as a quality selection model. In a case that one of facial region images is failed to be identified, the face identification can be continued to perform on other facial region images in the target facial image sequence, which significantly improves identification success rate for facial region images, that is, a recall rate. Compared with a method that only perform face identification on a high-quality facial region image selected by a face quality model, the method described in this present disclosure does not need to conduct a high-cost data collection to get a large amount of data to train the face quality model for covering all scenarios about low quality facial images. In this way, the cost of face identification is reduced.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] In order to more clearly describe the technical solutions in one or more embodiments of the present disclosure or related technologies, a brief description of the appended drawings involved in the embodiments or related technical descriptions is provided below. Obviously, the drawings in the following description are only concerned with one or more embodiments recorded in this present disclosure. For those of ordinary

skill in the art, other drawings can be acquired according to these drawings without any creative labor.

[0016] FIG. 1 is a flowchart of a face identification method illustrated in an embodiment of the present disclosure.

[0017] FIG. 1A is a flowchart of a method for determining a target facial image sequence of a target face illustrated in an embodiment of the present disclosure;

[0018] FIG. 2 is a flowchart of another face identification method illustrated in an embodiment of the present disclosure;

[0019] FIG. 3 is a block diagram of a face identification apparatus illustrated in an embodiment of the present disclosure;

[0020] FIG. 4 is a block diagram of another face identification apparatus illustrated in an embodiment of the present disclosure;

[0021] FIG. 5 is a schematic diagram of a hardware structure of an electronic device illustrated in embodiments of the present disclosure.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0022] Examples will be described in detail herein, with the illustrations thereof represented in the drawings. When the following descriptions involve the drawings, like numerals in different drawings refer to like or similar elements unless otherwise indicated. Implementations described in the following explanatory embodiments do not represent all implementations consistent with the present disclosure. Rather, they are merely examples of apparatuses and methods consistent with some aspects of the present disclosure as detailed in the appended claims.

[0023] Terms used in the present disclosure are for the purpose of describing particular embodiments only, and are not intended to limit the present disclosure. Terms “a”, “the” and “said” in their singular forms in the description and the appended claims are also intended to include plurality, unless clearly indicated otherwise in the context. It should also be understood that the term “and/or” as used herein refers to and includes any or all possible combinations of one or more of the associated listed items.

[0024] It is to be understood that although different information may be described using terms such as first, second, third, etc. in the present disclosure, the information should not

be limited to these terms. These terms are used only 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 disclosure, and 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 “when” or “as” or “determining in response to”.

[0025] For face identification in video tracking, the key step is how to select a high-quality facial image for face identification from a large number of facial image sequences in the video. If the quality of the to-be-identified facial image is too poor, the to-be-identified facial image cannot be matched with any identity information in the face database, so that an error result that the person cannot be found is obtained, that is, the facial image cannot be recalled.

[0026] The traditional method may train a face quality model by collecting large amounts of data to complete the selection of a high-quality facial image from the facial image sequence, and then use the selected high-quality facial image to perform the face identification. However, in practical applications, the environment around the target object is complex, and the target object can have many movements, such as turning the head, bowing the head, covering the face by a hand, looking in the mirror or wearing a mask, etc. It is difficult to cover all scenarios by data collection to train the face quality model.

[0027] During the whole video tracking process, there are sufficient high-quality facial images, in addition to low-quality facial images in some time periods caused by object’s movements, such as turning head, covering head by a hand, or caused by a mirrored face in a reflective surface, such as glass, mirror, ceramic tile, etc. The manner to select a high-quality facial image by the face quality model, specifically the manner to select single facial image, may not guarantee that the high-quality facial image for face identification can be selected, such as, a facial image satisfied with requirements like the face in the image is clear, uncovered and not rotated. If there is no identification result for the selected image, the face identification failure will be declared directly, which cannot be remedied. Therefore, in order to improve the success rate of face identification, a large amount of data needs to be collected to train the face quality model. However, the collection of a large amount of data spends high cost.

[0028] Therefore, the embodiments of the present disclosure provide a face identification method. The method can reduce the impact of low-quality facial image without collecting a

large amount of data, significantly improve the accuracy of face identification, and ensure the recall rate of facial image.

[0029] As shown in FIG. 1, FIG. 1 is a flowchart of face identification method shown in an embodiment of the present disclosure, and the method includes the following steps:

[0030] In step 102, a target facial image sequence for a target face is determined, where the target facial image sequence includes multiple facial region images for the target face.

[0031] In this step, a facial region image can be an image involving the target face of the target object detected from the tracking video when the target object is being tracked, or it can be an image involving the target face of the target object obtained by shooting the target object. Here, the target face is a face of a designated or undesignated person to be identified.

[0032] The face identification method of this embodiment can be executed by a face identification apparatus, for example, can be executed by a terminal device, a server or other processing equipment, where the terminal device can be a user device, a mobile device, a user terminal, a terminal, a cellular phone, a cordless phone, personal digital processing, a handheld device, a computing device, an on-board device, a wearable device, etc. In some possible implementations, the face identification method can be implemented by a processor invoking computer-readable instructions stored in a memory.

[0033] In this embodiment, a specific method of determining the target facial image sequence is not limited. For example, the target facial image sequence can be selected from multiple facial image sequences for multiple different faces maintained by the device in advance, or also can be directly acquired from other devices.

[0034] In addition, in this step, determining the facial image sequence for the target face can be executed at any time. For example, the method in this embodiment may be executed when receiving a face identification request that includes information about the target face; or, the method can also be executed when receiving a track video or continuously shooting images.

[0035] In step 104, face identification is performed on at least one facial region image in the target facial image sequence, and based on a confidence level of a face identification result of the at least one facial region image, a target facial region image is determined from the target facial image sequence.

[0036] For face identification of facial region image, one facial region image can be taken to compare with images of multiple faces respectively in a face database to get comparison results. Each of the comparison results includes a confidence level. The confidence level represents the probability that the face in the facial region image and a face in a compared image of the face database are of an identical object. A comparison result with the highest confidence level is determined as the face identification result of the facial region image.

[0037] If the confidence level of the face identification result of the facial region image reaches a preset confidence level threshold, that is, a face matched with the facial region image is found from the face database, so that face identification succeeds, then determine the facial region image is a target facial region image. If the confidence level of the face identification result of the facial region image doesn't reach a preset confidence level threshold, that is, a face matched with the facial region image is not found from the face database, so that face identification fails, then continue select another facial image which hasn't been selected before in the target facial image sequence to perform face identification until face identification succeeds.

[0038] After face identifications are performed for all facial region images in the target facial image sequence, if there is no face is found in the face database which is corresponding to any one of the facial region images, a facial region image with a face identification result closest to the successful identification result can be selected as the target facial region image in all face identification results. That is, a facial region image with a highest confidence of face identification result as the target facial region image, or face identification can be ended.

[0039] In this step, the order of face identification in the facial region image of the facial image sequence can be arbitrary, such as randomly selecting an image for identification from the target facial image sequence sequentially; or can also be in accordance with the preset order, such as in accordance with image quality from high-quality to low-quality, a facial region image is sequentially selected in the facial image sequence for face identification.

[0040] In the embodiment, a specific face identification method is not limited. For example, the identification may be performed through a neural network, or be performed through other methods.

[0041] In step 106, based on a face identification result of the target facial region image, identity information of the target face is determined.

[0042] The face identification result of the target facial region image includes a facial image corresponding to the target facial region image in the face database, and the identity information associated with the facial image is determined as the identity information of the target face. The identity information can include the ID number, name, registered account number, etc. pre-stored in the face database.

[0043] In the face identification method provided by the embodiment, by performing face identification on multiple facial region images in a target facial image sequence, a target facial region image is determined. Further, identify information of the target face is determined. It is possible to select a low-quality image for identification from the multiple facial region images, but when one of the multiple facial region images cannot be identified, a next facial region image in the target facial image sequence can be used to continue the identification. It is equivalent to that a target facial image sequence is working as a quality selection model, which significantly improves identification success rate for facial region images, that is, a recall rate. Compared with a method that only perform face identification on a high-quality facial region image selected by a face quality model, the method does not need to conduct a high-cost data collection to get a large amount of data to train the face quality model for covering all scenarios about low quality facial images. In this way, the cost of face identification is reduced.

[0044] In an embodiment, the method can be used to identify a target face in a video stream. In step 102, determining the target facial image sequence for the target face may include that multiple facial region images of the target face are extracted from multi-frame images of the video stream.

[0045] The video stream can be a recorded video or a real-time video. The multi-frame images of the video stream involve the target face of a target object. In this embodiment, the video stream may be a tracking video obtained by tracking the target object's face.

[0046] In an example, before step 102, the video stream can be acquired in advanced. Based on each of frames of images in an acquired video stream, faces in the video stream are detected. For each of multiple faces detected in the video stream, the detected face is tracked in the video stream to determine facial region images for the face in multi-frame

images of the video stream; based on the facial region images for the face in multi-frame images of the video stream, a facial image sequence for the face is generated. Therefore, multiple facial image sequences for different faces presented in the video stream can be acquired in advance, which is beneficial to quickly determine the target facial image sequence for the target face from the facial image sequences for different faces in following processes.

[0047] In this example, for each of multiple faces detected in the video stream, after determining facial region images for the face in multi-frame images of the video stream, a face identifier of the face is generated; according to the face identifier, the facial image sequence composed of facial region images for the face can be quickly determined. In this step, the target facial image sequence can be determined when receiving the face identification request. The face identification request can be issued by upper-layer application. When the face identification request message is received, an acquired face identifier from the face identification request is referred as a face identifier of the target face. According to the face identifier of the target face, the target facial image sequence is determined from the multiple facial image sequences.

[0048] In this example, which face tracking method is used is not limited. For example, the face tracking can be performed by Kalman filtering, or can be performed by neural network.

[0049] In an example, in step 102, determining the target facial image sequence for the target face, as shown in FIG. 1A, which may specifically include the following steps:

[0050] In step 1021, facial region images for the target face in multi-frame images of the video stream are determined as multiple candidate facial region images for the target face.

[0051] For example, in a case that the video stream is a tracking video, the candidate facial region images in the tracking video can be obtained by face detection. One or more candidate facial images can be detected in each frame of image of the tracking video, or for some frames of images, it is possible that no candidate facial region image is detected. Through face tracking, multiple candidate facial region images corresponding to the same face can be obtained, which means that multiple candidate facial region images for the target face are determined. In the case that multiple candidate facial region images are detected in

one frame image, the multiple candidate facial region images can include a physical facial region image and one or more mirrored facial region images.

[0052] In step 1022, image qualities of the multiple candidate facial region images are determined.

[0053] For example, for each of the multiple candidate facial region images, quality of the candidate facial region image is evaluated by a pre-trained facial image quality evaluation model, to determine an evaluated result. For example, the candidate facial region image can be input into a pre-trained facial image quality evaluation model to obtain the evaluated result of image quality. The image quality can be evaluated by integrating any of factors such as image intelligibility, brightness, clarity, facial symmetry and noise, and the evaluated result of image quality can be expressed by grade, score, etc. or other methods.

[0054] A sequence can be used to represent evaluated results of image quality of multiple candidate facial region images for the target face, such as sequence $A = \{a_n, n = 1, 2, 3, \dots, N\}$, a_n represents the n th evaluated result of image quality of a candidate facial region image, N represents the total number of frames of the candidate facial region images of the target face.

[0055] It should be noted that the facial image quality evaluation model used in this embodiment may be a model with a predetermined accuracy obtained by conventional training. For example, a binary classifier based on deep learning may be used to train the model by general face quality data. It is not necessary to spend high-cost to collect a large amount of data to train a high-precision face quality model.

[0056] In step 1023, according to the evaluated results of image qualities, the multiple candidate facial region images are ranked to obtain a first sequence.

[0057] For example, the multiple candidate facial region images can be ranked from high-quality to low-quality or from low-quality to high-quality to obtain a first sequence. Following the above example, sequence A is ranked according to a descending order of the image quality scores to obtain sequence B . In this example, sequence B is used to represent the first sequence.

[0058] In step 1024, based on the first sequence, the target facial image sequence is determined.

[0059] The first sequence can be directly determined as the target facial image sequence; or a sub-sequence of the first sequence may be determined as the target facial image sequence, where the first sub-sequence includes a preset number of candidate facial region images satisfied with preset image quality requirements. By further filtering images in the first sequence based on image quality, and selecting only a sub-sequence including a preset number of candidate facial region images as the target facial image sequence, image quality can be further improved and the image quantity can be effectively controlled of the target facial image sequence, which is improve the accuracy of identification result and efficiency of identification for the target face.

[0060] In some embodiments, based on a corresponding timing sequence of the multiple candidate facial region images in the first sequence, a sparse process is performed on the first sequence to obtain a second sequence. According to the second sequence, the target facial region sequence is determined.

[0061] Where, an interval distance between any two adjacent candidate facial region images in the second sequence in the corresponding timing sequence of the video stream is greater than the preset interval threshold.

[0062] The following sparse process method can be used to remove some of the candidate facial region images in sequence B to obtain sequence C, so that the interval distance between any two adjacent candidate facial region images in sequence C is greater than the preset interval threshold, which is noted as *step*:

[0063] *a)* Traverse values in sequence B, and note the serial number of the current candidate facial region image as i .

[0064] *b)* For each i , from i to start traverse values of sequence B, note the current candidate facial region image number as j .

[0065] *c)* If the interval distance between B_i and B_j is less than the preset interval threshold, which is noted as *step*, then delete B_j .

[0066] *d)* When finish traversing for i , all remaining candidate facial region images in sequence B are assigned to sequence C.

[0067] *e)* End.

[0068] In an example, assuming that the sequence B includes the following six candidate facial region images which are B_1 , B_2 , B_3 , B_4 , B_5 , B_6 . And the preset interval threshold is 0.05ms. When i is equal to 1, sequentially calculate intervals between B_1 and B_2 , B_3 , B_4 , B_5 , B_6 respectively. The interval between B_1 and B_2 is 0.02ms, the interval between B_1 and B_3 is 0.06ms, the interval between B_1 and B_4 is 0.15ms, the interval between B_1 and B_5 is 0.04ms, the interval between B_1 and B_6 is 0.10ms. In this case, B_2 and B_5 can be deleted from sequence B. Currently, the remaining elements in sequence B are B_1 , B_3 , B_4 , B_6 . When i is equal to 3, sequentially calculate intervals between B_3 and B_4 , B_6 respectively. The interval between B_3 and B_4 is 0.09ms, the interval between B_3 and B_6 is 0.04ms. In this case, B_6 can be deleted from sequence B. Currently, the remaining elements in sequence B are B_1 , B_3 , B_4 . Now, traversing for i is finished, and elements in sequence C are B_1 , B_3 , B_4 .

[0069] In implements, the timing sequence can be represented by time labels. The candidate facial region image is a part of the image frames of the video stream, and the time label of the candidate facial region image may be a time label of the corresponding image frame in the video stream.

[0070] For the target object in the tracking video, when the target object has some movements such as turning head, covering head by a hand, or drinking water, an image captured these movements by the tracking video will not be suitable for face identification. Generally, these movements will be lasted in a certain period of time. Sequential continuous image frames in the tracking video are usually similar, and image qualities are generally similar. Repeated face identification on similar images is not worthwhile, so the time sequential sparse process can be performed on collected candidate facial region images to obtain the second sequence, so as to avoid face identification on low-quality images for many times and improve the efficiency of face identification.

[0071] In the embodiment, a specific method of the time sequential sparse process is not limited. Any method, which can realize that an interval distance between any two adjacent candidate facial region images in the second sequence is greater than the preset interval threshold, can be adopted. For example, multiple time labels can be calculated according to the preset interval threshold, and the candidate facial region images in the first sequence within the difference range of the multiple time labels can be combined into a second sequence; for another example, elements in the first sequence can be traversed in order, and

the adjacent candidate facial region images whose interval distance is less than the preset interval threshold are removed to obtain the second sequence.

[0072] After the second sequence is obtained, according to the second sequence, the second sequence can be directly determined as the target facial image sequence, or the second sub-sequence can be extracted from the second sequence and then the second sub-sequence is determined as the target facial image sequence. Where, the second sub-sequence includes a preset number of candidate facial region images satisfied with preset image quality requirements.

[0073] The image quality requirements can require for ranking of image quality. For example, when the second sequence includes the candidate facial region images ranked in descending order of image quality from high to low, the sub-sequence composed of candidate facial region images with a preset number from head of the second sequence may be determined as the target facial image sequence; or, when the second sequence includes the candidate facial region images ranked in an ascending order of image quality from low to high, the sub-sequence composed of candidate facial region images with a preset number from tail of the second sequence may be determined as the target facial image sequence.

[0074] The preset number can be represented by K , which can be set by those skilled in the art according to actual needs. Take K candidate facial region images from head of the sequence C as the sequence D , which is the final target facial image sequence. In another example, when ranking in the first sequence is the ascending order, K candidate facial region images from tail of the sequence C can be selected as the sequence D .

[0075] The image quality requirement may also require for the level or score of the image quality, that is, the image quality is required to meet the preset level or the preset score. This embodiment does not limit the image quality requirement. According to the face quality requirement, a preset number of candidate facial region images are further filtered from the second sequence as the target facial image sequence, which can further simplify the target facial region images for the target face, so that the final target facial image sequence is discrete on timing and is high-quality, which improves the efficiency of face identification. After obtaining the target facial image sequence, face identification can be performed on the facial region images in order of image quality from high to low.

[0076] In another example, determining the target facial image sequence may be acquiring the target facial image sequence maintained by and from other devices. Other devices may be terminal devices, servers, or other processing devices, and other devices may execute steps 1021 to 1024 shown in FIG.1A to obtain the target facial image sequence.

[0077] FIG. 2 provides a face identification method according to another embodiment of the present disclosure. The method may include the following processing, where the steps that are the same as the procedures of the foregoing embodiment will not be described in detail.

[0078] In step 202, the target facial image sequence of the target face is determined, and image qualities of multiple facial region images are determined.

[0079] The target facial image sequence includes multiple facial region images for the target face.

[0080] A pre-trained facial image quality evaluation model with a predetermined accuracy can be used to evaluate image qualities for facial region images to get evaluated results.

[0081] For example, after determining the target facial image sequence, image qualities of the facial region images in the target facial image sequence can be evaluated to obtain the evaluated results.

[0082] Or before determining target facial image sequence, image qualities of facial region images can be determined. For example, multiple candidate facial region images can be obtained from the multi-frame images of the video stream first, then image qualities of the candidate facial region images are evaluated to obtain evaluated results. Then, facial region images of the target facial image sequence are determined from the candidate facial region images according to the evaluated results.

[0083] In step 204, in the case the target facial region image is not determined, face identification is performed on a first facial region image that has not been identified and with highest image quality in the target facial image sequence.

[0084] In the case the target facial region image is not determined, facial region images in the target facial images sequence can be identified successively according to the image quality order to obtain face identification results.

[0085] For example, the facial region images in the target facial image sequence D can be as the first facial region images and are sequentially input into the face identification model one by one, so as to extract face features in the facial region images, and compare them with a facial image of the face database based on the face features. Confidence levels can be used to represent the comparison results. For each of facial region images in facial image sequence D, multiple confidence levels can be obtained when the image is compared in the face database. Where, the comparison result with the highest confidence level is determined as the face identification result of the facial region image.

[0086] In step 206, in response to determining that the confidence level of the face identification result of the first facial region image is greater than a preset threshold, the first facial region image is determined as the target facial region image.

[0087] When a confidence level of the face identification result of a first facial region image exceeds the preset threshold, the face identification succeeds. The first facial image is determined as the target facial region image, and the face identification is no longer carried out for other facial region images in the target facial image sequence.

[0088] When the confidence level of the face identification result of the first facial region image does not exceed the preset threshold, and the first facial region image is not the last facial region image in the image sequence, then continue to perform face identification on a first facial region image that has not been identified and with highest image quality in the target facial image sequence.

[0089] In step 208, in response to determining that confidence levels of face identification results of all facial region images in the target facial image sequence are less than a preset threshold, a second facial region image, whose face identification result with a maximum confidence level, is determined as the target facial region image.

[0090] When face identification is performed on all facial region images in the target facial image sequence, confidence levels of face identification results of the facial region images face identification are all less than the preset threshold, then the second facial region image, whose face identification result with a maximum confidence level in all face identification results of facial region images, is taken as the target facial region image.

[0091] In step 210, based on a face identification result of the target facial region image, identity information of the target face is determined.

[0092] The identity information associated with the face identification result of the target facial region image in the face database is taken as the identity information of the target face. Therefore, the identity information of the target object is determined by face identification.

[0093] In the face identification method provided by the embodiment, during performing face identification on multiple facial region images in a target facial image sequence, a facial region image with the highest image quality is identified firstly. When the identification fails, continue to use a facial region image with the highest quality of remaining images in the target facial image sequence for identification. It is equivalent to that the target facial image sequence is working as a quality selection model, which significantly improves an identification success rate for facial region images and a recall rate for high quality facial region images. Compared with a method that only performs face identification on a high-quality facial region image selected by a face quality model, the method does not need to conduct a high-cost data collection to get a large amount of data to train the face quality model for covering all scenarios about low quality facial images. In this way, the cost of face identification is reduced.

[0094] In an embodiment, the face identification method provided by this embodiment in the present disclosure can be applied to game place environment. In a scene of intelligent play place, players seated-in needs to be tracking for a long time. And when receiving the identification request issued by the upper-layer application, the face identification is performed on the tracked player to verify identity. However, the game place environment is really complex, such as, lightings change greatly, players keep many postures. It is difficult to cover all problem scenarios by data collection to train the face quality model to select a high-quality facial image for face identification.

[0095] The following is an illustration of the application of a face identification method in the game place provided by this embodiment of the present disclosure.

[0096] First of all, a face detection model, a facial image quality evaluation model and a face identification model need to be trained in advance.

[0097] For the face detection model, common face detection models such as Retina Net, YoloV3 or PCN can be used to complete training with general face data, or general face data and game-place-specific face data can be used to improve the accuracy of the model.

[0098] For the facial image quality evaluation model, a binary classifier based on deep learning may be used to complete training with general face quality data, or general face data and game-place-specific face data can be used to improve the accuracy of the model. For example, binary cross entropy loss can be used as a loss function during training.

[0099] For the face identification model, common neural networks such as resenet50 or SqueezeNet can be used to complete training with general face identification data. Common face identification loss functions such as ArcFace can be used for training.

[0100] Based on the models prepared above, the tracking video captured in real time by the camera in the game place can be processed as follows to maintain facial image sequences with time dispersion and high quality. In other examples, a tracking video pre-captured by the camera in the game place can also be processed.

[0101] Face detecting: Each frame image in the tracking video is input into the face detection model sequentially, and a face detection box in each frame image is obtained. An image in the face detection box is a candidate facial region image.

[0102] Image quality evaluating: Candidate facial region images are input into the facial image quality evaluation model frame by frame in real time, to obtain image quality scores.

[0103] Face tracking: Through a fast face tracking method, such as Kalman filter, the quality scores of all candidate facial region images corresponding to the same player's face can be obtained. These scores can be represented by a quality score sequence.

[0104] High-quality facial image sequence selecting: the sequence can be ranked according to a descending order of quality scores, then sparse process is performed on timing sequence, to make an interval distance between any two adjacent elements in the sequence is greater than the preset interval threshold; Then, K candidate facial region images are taken from head of the sequence as the final facial image sequence; the final facial image sequence includes K facial region images that ranked from high-quality to low-quality.

[0105] It should be noted that if the number of players in the tracking video is N , N facial image sequences also need to be maintained. When the tracking video keeps going over time, the maintained facial image sequences will also be updated. When receiving a face identification request issued by the upper-layer application, face identification is performed on a specified player in the face identification request.

[0106] The facial region images in the facial image sequence for the player are input into the face identification model one by one for feature extraction and identity searching. If the searching confidence level of the current facial region image is greater than the preset threshold, that is, the searching is successful, identity information in the face database associated with the searching result with the searching confidence level is output. If the searching confidence level of the current facial region image is less than the preset threshold, this searching confidence is recorded and the next facial region image is input into the face identification model. If the current facial region image is already the last one in the facial image sequence, identity information in the face database associated with the searching result with the highest searching confidence level in recorded searching confidence levels is output.

[0107] In a scene of intelligent play place, an identity of a player captured by a camera needs to be identified. The solution herein can be used to finish this goal, which can guarantee relative high identification accuracy. Based on the recall rates of high-quality facial region images and faces of players, subsequent works such as game table monitoring and user analysis can be better performed.

[0108] Those skilled in the art can understand that in the above methods of detailed description, the writing order of each step does not mean a strict execution order to limit any implementation process, and the specific execution order of each step should be determined according to its function and possible internal logic.

[0109] As shown in FIG. 3, FIG. 3 is the block diagram of a face identification apparatus shown in this embodiment of the present disclosure. The apparatus includes target facial image sequence determination module 31, target facial region image determination module 32, and identity information determination module 33.

[0110] The target facial image sequence determination module 31 is configured to determine a target facial image sequence for a target face, where the target facial image sequence comprises multiple facial region images for the target face.

[0111] The target facial region image determination module 32 is configured to perform face identification on at least one facial region image in the target facial image sequence, and determine, based on a confidence level of a face identification result of the at least one facial region image, a target facial region image from the target facial image sequence.

[0112] The identity information determination module 33 is configured to determine, based on a face identification result of the target facial region image, identity information of the target face.

[0113] In an example, the target facial image sequence includes the multiple facial region images for the target face extracted from multi-frame images in a video stream.

[0114] In an example, as shown in FIG. 4, based on foregoing embodiment of the apparatus, the apparatus also includes: facial image sequence generation module 30.

[0115] Facial image sequence generation module 30 is configured to: perform face detection based on each frame of image in an acquired video stream; for each of multiple faces detected in the video stream: track, in the video stream, the detected face to determine facial region images for the face in multi-frame images of the video stream; generate, based on the facial region images for the face in multi-frame images of the video stream, a facial image sequence for the face.

[0116] In an example, the target facial image sequence determination module 31 is configured to: determine that facial region images for the target face in multi-frame images of the video stream are multiple candidate facial region images for the target face; determine image qualities of the multiple candidate facial region images; rank, according to the image qualities, the multiple candidate facial region images to obtain a first sequence; determine the target facial image sequence based on the first sequence.

[0117] In an example, the target facial image sequence determination module 31 is configured to determine that a first sub-sequence of the first sequence is the target facial image sequence, the first sub-sequence comprising a preset number of candidate facial region images satisfied with preset image quality requirements.

[0118] In an example, the target facial image sequence determination module 31 is configured to: perform, based on corresponding timing sequence of the multiple candidate facial region images in the first sequence, sparse process for the first sequence to obtain a second sequence, where an interval distance between any two adjacent candidate facial region images in the second sequence in the corresponding timing sequence of the video stream is greater than the preset interval threshold.

[0119] In an example, the target facial image sequence determination module 31 is further configured to determine that a second sub-sequence of the second sequence is the target

facial image sequence, the second sub-sequence comprising a preset number of candidate facial region images satisfied with preset image quality requirements.

[0120] In an example, when determining image qualities of the multiple candidate facial region images, the target facial image sequence determination module 31 is configured to: for each of the multiple candidate facial region images, perform quality evaluation for the candidate facial region image by a pre-trained facial image quality evaluation model, to determine a evaluated result of image quality of the candidate facial region image.

[0121] In an example, for the each of multiple faces detected in the video stream, after tracking, in the video stream, the detected face to determine the facial region images for the face in multi-frame images of the video stream, the facial image sequence generation module 30 is further configured to generate a facial identifier of the face. The target facial image sequence determination module 31 is configured to: take an acquired facial identifier in a face identification request as a facial identifier of the target face; determine, based on the facial identifier of the target face, the target facial image sequence for the target face from facial image sequences for the multiple faces.

[0122] In an example, the target facial image sequence determination module 31 is further configured to determine image qualities of the multiple candidate facial region images. The target facial region image determination module 32 is specifically configured to: in response to determining that a confidence level of the face identification result of a first facial region image is greater than a preset threshold, determine that the first facial region image is the target facial region image.

[0123] In an example, the target facial region image determination module is further configured to: in response to determining that confidence levels of face identification results of all facial region images in the target facial image sequence are less than a preset threshold, determine a second facial region image, whose face identification result with a maximum confidence, and take the second facial region image as the target facial region image.

[0124] The functions of each module in the above-mentioned apparatuses is detailed in the implementation of corresponding steps in the above-mentioned methods, which are not described here.

[0125] In the embodiments of the present disclosure, an electronic device is also provided, as shown in FIG. 5. The electronic device includes a memory 51 and a processor 52, where

the memory is configured to store computer instructions capable of being run on the processor, and the processor 52 implements the face identification method of any one of embodiments of the present disclosure when executing the computer instructions.

[0126] In the embodiments of the present disclosure, a computer program product is also provided. The computer program product includes computer programs codes/instructions, when executed in a processor to perform the method of any one of embodiments of the present disclosure.

[0127] In the embodiments of the present disclosure, a computer-readable storage medium is provided, storing a computer program, where steps of the face identification method of any one of embodiments of the present disclosure are implemented when the program is executed by a processor.

[0128] For the embodiments of the apparatus, since the apparatus substantially corresponds to the method embodiments, relevant information can be referred to the description of the method embodiments. The apparatus embodiments described above are merely illustrative, and the modules described as separate components may or may not be physically separate, and the components displayed as modules may or may not be physical modules, that is, may be located in one place, or may be distributed to a plurality of network modules. Some or all of the modules may be selected according to actual needs to achieve the purpose of the solution of the present description. A person of ordinary skill in the art would understand and implement without creative efforts.

[0129] Specific embodiments of the present description are described above. Other embodiments are within the scope of the appended claims. In some cases, the actions or steps described in the claims may be performed in an order different from that in the embodiments and may still achieve the desired results. Moreover, the processes depicted in the figures do not necessarily require the particular order or sequence shown to achieve the desired results. In some embodiments, multitasking and parallel processing are also possible or may be advantageous.

[0130] A person skilled in the art upon consideration of the specification and practice of the present disclosure disclosed herein will readily appreciate other implementations of the present disclosure. The present disclosure is intended to cover any variations, uses, or adaptations of the present disclosure, and the variations, uses, and adaptations follow a

general principle of the present disclosure and include common sense or common technical means in this technical field that are not disclosed in the present disclosure. The specification and the embodiments are considered as merely exemplary, and the real scope and spirit of the present disclosure are pointed out in the following claims.

[0131] It is to be understood that this specification is not limited to the precise construction already described above and illustrated in the accompanying drawings, and that various modifications and changes may be made without departing from its scope. The scope of this specification is limited only by the appended claims.

[0132] The foregoing descriptions are merely preferred embodiments of one or more embodiments of this specification, but are not intended to limit the one or more embodiments of this specification. Any modification, equivalent replacement, or improvement made within the spirit and principle of one or more embodiments of this specification shall fall within the protection scope of the one or more embodiments of this specification.

CLAIMS

1. A face identification method, comprising:

determining a target facial image sequence for a target face, wherein the target facial image sequence comprises multiple facial region images for the target face;

performing face identification on at least one facial region image in the target facial image sequence;

determining, based on a confidence level of a face identification result of the at least one facial region image, a target facial region image from the target facial image sequence; and

determining, based on the face identification result of the target facial region image, identity information of the target face.

2. The method according to claim 1, wherein the target facial image sequence comprises the multiple facial region images for the target face extracted from multi-frame images in a video stream.

3. The method according to claim 1 or 2, further comprising:

performing face detection based on each frame of image in an acquired video stream; and

for each of multiple faces detected in the video stream:

tracking, based on the video stream, the face to determine facial region images for the face in multi-frame images of the video stream; and

generating, based on the facial region images for the face in multi-frame images of the video stream, a facial image sequence for the face.

4. The method according to claim 3, wherein determining the target facial image sequence for the target face comprises:

determining the facial region images for the target face in multi-frame images of the video stream as multiple candidate facial region images for the target face;

determining image qualities of the multiple candidate facial region images;

ranking, according to the image qualities, the multiple candidate facial region images to obtain a first sequence; and

determining the target facial image sequence based on the first sequence.

5. The method according to claim 4, wherein determining the target facial image sequence based on the first sequence comprises:

determining a first sub-sequence of the first sequence as the target facial image sequence, the first sub-sequence comprising a preset number of candidate facial region images satisfied with a preset image quality requirement.

6. The method according to claim 4, wherein determining the target facial image sequence based on the first sequence comprises:

performing, based on a corresponding timing sequence of the multiple candidate facial region images in the first sequence, sparse process for the first sequence to obtain a second sequence, wherein an interval distance between any two adjacent candidate facial region images in the second sequence is greater than a preset interval threshold; and

determining the target facial image sequence based on the second sequence.

7. The method according to claim 6, wherein determining the target facial image sequence based on the second sequence comprises:

determining a second sub-sequence of the second sequence as the target facial image sequence, the second sub-sequence comprising a preset number of candidate facial region images satisfied with a preset image quality requirement.

8. The method according to any one of claims 4 to 7, wherein determining image qualities of the multiple candidate facial region images comprises:

for each of the multiple candidate facial region images, performing quality evaluation for the candidate facial region image by a pre-trained facial image quality evaluation model, to determine an evaluated result of image quality of the candidate facial region image.

9. The method according to any one of claims 3 to 8, wherein, for the each of multiple faces detected in the video stream, after tracking, based on the video stream, the face to determine the facial region images for the face in multi-frame images of the video stream, the method further comprises:

generating a facial identifier of the face; and

determining the target facial image sequence for the target face comprises:

taking an acquired facial identifier in a face identification request as a facial identifier of the target face; and

determining, based on the facial identifier of the target face, the target facial image sequence for the target face from facial image sequences for the multiple faces.

10. The method according to any one of claims 1 to 9, further comprising:

determining image qualities of the multiple facial region images; and

performing face identification on the at least one facial region image in the target facial image sequence, and determining, based on the confidence level of the face identification result of the at least one facial region image, the target facial region image from the target facial image sequence, comprises:

in the case that the target facial region image is not determined, performing face identification on a first facial region image without face identification and with highest image quality in the target facial image sequence; and

in response to determining that a confidence level of the face identification result of the first facial region image is greater than a preset threshold, determining that the first facial region image is the target facial region image.

11. The method according to claim 10, wherein performing face identification on the at least one facial region image in the target facial image sequence, and determining, based on the confidence level of the face identification result of the at least one facial region image, the target facial region image from the target facial image sequence, further comprises:

in response to determining that confidence levels of face identification results of all facial region images in the target facial image sequence are less than the preset threshold, determining a second facial region image with a maximum confidence level; and

taking the second facial region image as the target facial region image.

12. A face identification apparatus, comprising:

a target facial image sequence determination module, configured to determine a target facial image sequence for a target face, wherein the target facial image sequence comprises multiple facial region images for the target face;

a target facial region image determination module, configured to perform face identification on at least one facial region image in the target facial image sequence, and determine, based on a confidence level of a face identification result of the at least one facial region image, a target facial region image from the target facial image sequence; and

an identity information determination module, configured to determine, based on a face identification result of the target facial region image, identity information of the target face.

13. The apparatus according to claim 12, wherein the target facial image sequence comprises the multiple facial region images for the target face extracted from multi-frame images in a video stream.

14. The apparatus according to claim 12 or 13, wherein the apparatus further comprises a facial image sequence generation module, configured to:

perform face detection based on each frame of image in an acquired video stream; and
for each of multiple faces detected in the video stream:

track, based on the video stream, the face to determine facial region images for the face in multi-frame images of the video stream; and

generate, based on the facial region images for the face in multi-frame images of the video stream, a facial image sequence for the face.

15. The apparatus according to claim 14, wherein the target facial image sequence determination module is configured to:

determine the facial region images for the target face in multi-frame images of the video stream as multiple candidate facial region images for the target face;

determine image qualities of the multiple candidate facial region images;

rank, according to the image qualities, the multiple candidate facial region images to obtain a first sequence; and

determine the target facial image sequence based on the first sequence.

16. The apparatus according to claim 15, wherein target the facial image sequence determination module is further configured to:

determine a first sub-sequence of the first sequence as the facial image sequence, the first sub-sequence comprising a preset number of candidate facial region images satisfied with a preset image quality requirement.

17. The apparatus according to claim 15, wherein the target facial image sequence determination module is further configured to:

perform, based on a corresponding timing sequence of the multiple candidate facial region images in the first sequence, sparse process for the first sequence to obtain a second sequence, wherein an interval distance between any two adjacent candidate facial region images in the second sequence is greater than a preset interval threshold; and

determine the target facial image sequence based on the second sequence.

18. The apparatus according to claim 17, wherein the target facial image sequence determination module is further configured to:

determine a second sub-sequence of the second sequence as the target facial image sequence, the second sub-sequence comprising a preset number of candidate facial region images satisfied with a preset image quality requirement.

19. An electronic device, comprising a memory and a processor, wherein the memory is configured to store computer instructions capable of being run on the processor, and the

processor implements the method according to any one of claims 1 to 11 when executing the computer instructions.

20. A computer-readable storage medium, storing a computer program, wherein the method according to any one of claims 1 to 11 is performed when the program is executed by a processor.