

VIDEO GENERATION METHOD AND APPARATUS, AND DEVICE AND STORAGE MEDIUM

[0001] The present disclosure claims a priority right of the Chinese patent filed on October 22, 2021, entitled as “VIDEO GENERATION METHOD, APPARATUS, DEVICE AND STORAGE MEDIUM” and having the application No. 202111234267.3. The full disclosure of this Chinese patent application is combined in the present disclosure by reference.

FIELD

[0002] The present disclosure relates to the field of data processing, and more specifically, to a video generation method, apparatus, device and storage medium.

BACKGROUND

[0003] With the phenomenal growth of multimedia short videos, people are increasingly interested in taking images and shooting videos and the functions of the photographic page are also more diversified. In such case, the technical problem to be solved at present is how to enrich the functions of video shooting to enhance people’s experience in video creation.

SUMMARY

[0004] To fully or at least partly address the above technical problem, the embodiments of the present disclosure provides a video generation method that can automatically apply a corresponding target effect resource to a target object on the photographed picture. As such, the functions of video shooting are enriched, so as to enhance the users’ experience in video creation.

[0005] In a first aspect, the present disclosure provides a video generation method, wherein the method is applied to a client and comprises:

[0006] upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a target identification image based on the photographed picture on the photographic page;

[0007] transmitting the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from

the target identification image;

[0008] receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture;

[0009] generating a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied.

[0010] In a second aspect, the present disclosure provides a video generation method, wherein the method is applied to a server and comprises:

[0011] receiving a target identification image from a client; wherein the target identification image is obtained based on the photographed picture on the photographic page of the client;

[0012] determining whether the target identification image contains any one preset object;

[0013] in response to determining that the target identification image contains any one preset object, determining the preset object as target object and determining a target effect resource corresponding to the target object;

[0014] returning the target effect resource to the client; wherein the target effect resource is used for generating a target video at the client.

[0015] In a third aspect, the present disclosure provides a video generation system, wherein the system includes a client and a server;

[0016] the client is used for upon receiving an identification operation triggered for the photographed picture on a photographic page, obtaining a target identification image based on the photographed picture on the photographic page, and transmitting the target identification image to a server;

[0017] the server is used for receiving a target identification image from a client and determining whether the target identification image contains any one preset object; determining the preset object as a target object and determining a target effect resource corresponding to the target object in response to determining that the target identification image contains any one preset object; and returning the target effect resource to the client;

[0018] the client is further used for receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture; and generating a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied.

[0019] In a fourth aspect, the present disclosure provides a video generation apparatus, wherein the apparatus is applied to client and comprises:

[0020] an acquisition module for upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a target identification image based on the photographed picture on the photographic page;

[0021] a transmission module for transmitting the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image;

[0022] an application module for receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture;

[0023] a generation module for generating a target video based on a photographed picture corresponding to the target object to which the target effect resource is applied.

[0024] In a fifth aspect, the present disclosure proves a video generation apparatus, wherein the apparatus is applied to a server and comprises:

[0025] a first receiving module for receiving a target identification image from a client; wherein the target identification image is obtained based on the photographed picture on the photographic page of the client;

[0026] a first determination module for determining whether the target identification image contains any one preset object;

[0027] a second determination module for in response to determining that the target identification image contains any one preset object, determining the preset object as target object and determining a target effect resource corresponding to the target object;

[0028] a return module for returning the target effect resource to the client; wherein the target effect resource is used for generating a target video at the client.

[0029] In a sixth aspect, the present disclosure provides a computer-readable storage medium stored with instructions therein which, when running on a terminal device, cause the terminal device to perform the above method.

[0030] In a seventh aspect, the present disclosure provides a device, comprising: a memory, a processor and a computer program stored on the memory and executable on the processor, wherein the processor, when executing the computer program, performs the above method.

[0031] In an eighth aspect, the present disclosure provides a program product comprising a computer program/instruction which, when executed, performs the above method.

[0032] In comparison to the prior art, the technical solution provided by the embodiments of the present disclosure at least has following advantages:

5 [0033] The video generation method is provided by the embodiments of the present disclosure. First of all, upon receiving an identification operation triggered for a photographed picture on a photographic page, the client obtains a target identification image based on the photographed picture on the photographic page and transmits the target identification image to a server; wherein the server is used for determining a target effect
10 resource for a target object after identifying the target object from the target identification image. Then, the client receives the target effect resource from the server, applies the target effect resource to the target object on the photographed picture and generates a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied. In the embodiments of the present disclosure, the matching target effect
15 resource can be automatically applied to the target object on the photographed picture to generate a target video during the video shooting. As such, the functions of video shooting are enriched by the embodiments of the present disclosure, so as to enhance the users' experience in video creation.

BRIEF DESCRIPTION OF THE DRAWINGS

20 [0034] The drawings here are incorporated into the description as a part of it, illustrate the embodiments conforming with the present disclosure, and are combined with the description to explain the principles of the present disclosure.

[0035] To more clearly explain the technical solution in the embodiments of the present disclosure or in the prior art, the drawings needed in the description of the embodiments or
25 the prior art are to be introduced simply below. Apparently, those skilled in the art may also obtain other drawings from those illustrated without any exercises of inventive work.

[0036] Fig. 1 illustrates a flowchart of a video generation method provided by embodiments of the present disclosure;

[0037] Fig. 2 illustrates a flowchart of a further video generation method provided by
30 embodiments of the present disclosure;

[0038] Fig. 3 illustrates a schematic diagram of data interaction of a video generation method provided by the embodiments of the present disclosure;

[0039] Fig. 4 illustrates a structural diagram of a video generation system provided by the embodiments of the present disclosure;

5 [0040] Fig. 5 illustrates a structural diagram of a video generation apparatus provided by the embodiments of the present disclosure;

[0041] Fig. 6 illustrates a structural diagram of a further video generation apparatus provide by the embodiments of the present disclosure;

10 [0042] Fig. 7 illustrates a structural diagram of a video generation device provided by the embodiments of the present disclosure.

DETAILED DESCRIPTION OF EMBODIMENTS

15 [0043] For a clearer understanding of the above objectives, features and advantages of the present disclosure, the solution of the present disclosure is to be further described below. It is to be appreciated that embodiments of the present disclosure and features within the embodiments may be combined with one another without causing conflicts.

[0044] Many details are elaborated in the following description to provide a more comprehensive understanding of the present disclosure. However, the present disclosure also may be implemented in different ways than those described here. Apparently, the embodiments disclosed in the description is only a part of the embodiments of the present
20 disclosure, rather than all of them.

[0045] To enrich the functions of video shooting and enhance users' experience in video creation, the embodiments of the present disclosure propose a video generation method. First of all, upon receiving an identification operation triggered for a photographed picture on a photographic page, the client obtains a target identification image based on the
25 photographed picture on the photographic page and transmits the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image. Then, the client receives the target effect resource from the server, applies the target effect resource to the target object on the photographed picture and generates a target video based on the
30 photographed picture corresponding to the target object to which the target effect resource is

applied. In the embodiments of the present disclosure, the matching target effect resource can be automatically applied to the target object on the photographed picture to generate a target video during the video shooting. As such, the functions of video shooting are enriched by the embodiments of the present disclosure, so as to enhance the users' experience in video creation.

[0046] On this basis, the embodiments of the present disclosure provide a video generation method. Fig. 1 illustrates a flowchart of a video generation method provided by the embodiments of the present disclosure. As shown in Fig. 1, the method is applied to the client and comprises:

[0047] S101: upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a target identification image based on the photographed picture on the photographic page.

[0048] In the embodiments of the present disclosure, there may be many ways to trigger an identification operation for the photographed picture on the photographic page. For example, one way is to trigger a "long hold" operation for the photographed picture on the photographic page, wherein the long hold operation may be triggered for any position on the photographed picture. A further way is to trigger an operation for identification control configured on the photographic page, wherein the identification control may be disposed at any position on the photographed picture (e.g., right side or underside of the photographic page etc.).

[0049] In the embodiments of the present disclosure, upon receiving an identification operation triggered for photographed picture on a photographic page, a target identification image obtained based on the photographed picture on the photographic page may be one frame of image captured by a current camera, or any one frame of image extracted from a plurality of continuous frames of images, wherein the plurality of continuous frames of images include a plurality of continuous frames of images in which an ending frame is the current image corresponding to the photographed picture. For example, when the current camera captures five continuous frames of images in which an ending frame is the current image corresponding to the photographed picture, the second frame of image is extracted as the target identification image.

[0050] It is to be explained that a plurality of continuous frames of images in which an ending frame is the current image corresponding to the photographed picture included in the

above a plurality of continuous frames of images may be five continuous frames of images or seven continuous frames of images and the like. Besides, the second frame of image or the third frame of image in the above continuous frames of images may be extracted as the target identification image. The embodiments of the present disclosure do not restrict the specific frame number of the plurality of continuous frames of images, nor do define which frame of image should be extracted as the target identification image.

[0051] The target identification image obtained in the above embodiments may be vague or in darker color. It is possible that an invalid target identification image may be subsequently sent to the server for identification. Accordingly, it may be first determined, for the target identification image obtained based on the photographed picture on the photographic page, whether its image definition conforms with a preset definition condition, to reduce invalid input situations.

[0052] In some implementations, the target identification image that the image definition conforms with a preset definition condition is determined according to the following steps A1-A3.

[0053] Step A1: upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a current image corresponding to the photographed picture as an image with definition to be identified.

[0054] In the embodiments of the present disclosure, a current image corresponding to the photographed picture is obtained upon receiving an identification operation triggered for a photographed picture on a photographic page, where the current image corresponding to the photographed picture may be one frame of image, e.g., image 1; the current image corresponding to the photographed picture may also be any one frame of image extracted from a plurality of continuous frames of images, wherein the current image corresponding to the photographed picture is the ending frame of the plurality of continuous frames of images. Therefore, the obtained current image is used as the image with definition to be identified.

[0055] Step A2: determining whether a definition of the image with definition to be identified conforms with a preset definition condition.

[0056] In some implementations, determining whether a definition of the image with definition to be identified conforms with a preset definition condition includes: processing the image with definition to be identified based on a preconfigured fuzzy filtering model, to

obtain an image definition value of the image with definition to be identified; if an image definition value of the image with definition to be identified is greater than or equal to a value threshold indicating conformity with a preset definition condition, it is determined that the image definition of the image with definition to be identified conforms with the preset definition condition; if the image definition value of the image with definition to be identified is smaller than a value threshold indicating conformity with a preset definition condition, it is determined that the image definition of the image with definition to be identified fails to conform with the preset definition condition.

[0057] In the embodiments of the present disclosure, there may be many ways to determine whether a definition of the image with definition to be identified conforms with a preset definition condition. For example, one of them is to compare the image definition of the image with definition to be identified with an image definition pre-stored in the client indicating conformity with a preset definition condition, so as to determine whether the definition of the image with definition to be identified conforms with a preset definition condition. Another way is to configure a fuzzy filtering model in the client in advance, wherein the fuzzy filtering model may be obtained based on image training samples with various definition. The image with definition to be identified is processed by the fuzzy filtering model. Then, an image definition value of the image with definition to be identified may be output. Assuming that a value threshold indicating conformity with a preset definition condition is set to 80, the image conforming with the preset definition condition should have a corresponding image definition value greater than or equal to 80, and so on.

[0058] It is to be explained that the above set value threshold indicating conformity with a preset definition condition may be provided for determining whether an image definition of the image conforms with the preset definition condition. The value threshold may be set to 80, 82, 85 and the like as required. The embodiments of the present disclosure do not restrict the specific value of the value threshold indicating conformity with a preset definition condition.

[0059] Step A3: if it is determined that an image definition of the image with definition to be identified conforms with a preset definition condition, determining the image with definition to be identified as a target identification image.

[0060] In the embodiments of the present disclosure, based on the above step A2, if it is determined that an image definition of the image with definition to be identified (e.g., image 1)

conforms with a preset definition condition, the image 1 is identified as a target identification image.

[0061] For example, taking the configuration of the fuzzy filtering model in advance in the client in the above step A2 as an example, it is assumed that the value threshold indicating conformity with a preset definition condition is set to 80. Upon receiving an identification operation triggered for the photographed picture on a photographic page, a current image corresponding to the photographed picture (e.g., image 1) is obtained. The image 1 is input to the fuzzy filtering model for processing, and the resulting image definition value of the image 1 (such as 90) is obtained based on the processing. The resulting image definition value is compared with the preset threshold (e.g., 80). According to the comparison, it is determined that the image definition of the image 1 conforms with the preset definition condition (e.g., 90 greater than 80) and the image 1 is determined as the target identification image.

[0062] In some implementations, following the step A2 in the above embodiments, if it is determined that an image definition of the image with definition to be identified does not conform with a preset definition condition, an image is obtained from the photographed picture on the photographic page based on preset frame extraction conditions and the image with definition to be identified is updated with the image. Furthermore, a step of determining whether a definition of the image with definition to be identified conforms with a preset definition condition continues to be executed, until it is determined that a preset definition identification termination condition is met.

[0063] In some embodiments, obtaining an image from the photographed picture on the photographic page based on preset frame extraction conditions includes: extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined time; or, extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined frame number.

[0064] In the embodiments of the present disclosure, the preset frame extraction conditions may include extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined time; or, extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined frame number. The above interval of a predetermined time may be 20 ms or 30 ms, and the interval of a predetermined frame number may be 20 frames or 30 frames. The embodiments of the

present disclosure are not restricted in the specific values for the interval of a predetermined time and the interval of a predetermined frame number.

[0065] In some embodiments, determining that a preset definition identification termination condition is met includes: if it is determined that a number of image frames extracted from images corresponding to the photographed picture reaches a preset image extraction number, identification for image definition is terminated; or, if it is determined that an image definition value of the image with definition to be identified is smaller than a preset minimum value, identification for image definition is terminated.

[0066] In the embodiments of the present disclosure, the preset definition identification termination condition may include: terminating the image definition identification if the number of image frames extracted reaches a preset image extraction number; alternatively terminating the image definition identification if an image definition value of the image with definition to be identified is smaller than a preset minimum value, wherein the above preset image extraction number may be, 3, 5 or 7. The embodiments of the present disclosure are not limited in the specific value for the preset image extraction number.

[0067] For example, in a scenario where the fuzzy filtering model is configured in the client in advance and the value threshold indicating conformity with a preset definition condition is set to 80, it is assumed that the preset image extraction number in the preset definition identification termination condition is 5. If it is determined that an image definition of the image with definition to be identified (e.g., image 1) does not conform with the preset definition condition, one frame of image (such as image 2) is extracted from the images corresponding to the photographed picture by an interval of a predetermined time (e.g., 30 ms). The image with definition to be identified is updated with the image 2, i.e., image 2 is the new image with definition to be identified. The process continues to determine whether the image definition of the image 2 conforms with the preset definition condition. The image 2 is input to the fuzzy filtering model and the resulting image definition value of the image 2 is smaller than the preset threshold (e.g., 80). In such case, it is determined that the image definition of the image 2 also fails to conform with the preset definition condition. It is further determined whether the obtained image 2 meets a preset definition identification termination condition. If not, the process continues to extract an image (e.g., image 3) from the photographed picture on the photographic page based on the preset image extraction conditions, and so on, until an image 5 is obtained from the photographed picture on the

photographic page. It is then determined whether the image definition of the image 5 conforms with the preset definition condition. If yes, the image 5 is determined as the target identification image. If not and it is determined that the image 5 has met the preset definition identification termination condition (such as the number of the extracted image frame reaching 5), the image definition identification is terminated and the users are prompted with failure of this identification.

[0068] Therefore, in the embodiments of the present disclosure, the client can first determine the target identification image with an image definition conforming with the preset definition condition and then sent it to the server. This can reduce invalid inputs, so as to lower the number of the invalid identification performed at the server while decreasing wastes of network resources between the client and the server.

[0069] S102: transmitting the target identification image to a server.

[0070] wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image.

[0071] In the embodiments of the present disclosure, the target identification image obtained from the above S101 is sent to the server, which server can perform the identification process on the target identification image. If a corresponding target object is identified for the target identification image, a corresponding target effect resource is determined for the target object, wherein the target object may include any classes of objects whose corresponding feature information can be collected, e.g., objects in the classes including pictures, three-dimensional objects and iconic buildings etc., like a give T-shirt pattern, drink can and high-rise building. After the server identifies the target object from the target identification image, the specific implementations for determining the target effect resource for the target object are to be introduced in the following text and will not be repeated here.

[0072] S103: receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture.

[0073] In the embodiments of the present disclosure, according to the above S102, if the server identifies the target object from the target identification image, the target effect resource is determined for the target object and transmitted to the client. The client can receive the target effect resource and apply them to the target object on the photographed picture. Specifically, SDK (Software Development Kit) parses the target effect resource to

obtain props and templates contained in the target effect resource. Subsequent to the initialization of the SDK, the effects are rendered on the photographed picture of the client, such that the users can pre-view on a preview picture of the photographic page.

[0074] Moreover, since the client can apply the target effect resource to the target object on the photographed picture, the client also may receive the position information of the target object on the photographed picture provided by the server in addition to the target effect resource from the server. This may subsequently facilitate the client to apply the target effect resource to the target object on the photographed picture.

[0075] In some implementations, the position information of the target object on the photographed picture is received from the server and the target effect resource is further applied to the target object on the photographed picture based on the position information.

[0076] In the embodiments of the present disclosure, the client also can receive the position information of the target object on the photographed picture from the server in addition to the target effect resource from the server. This accordingly facilitates the client to apply the target effect resource to the target object on the photographed picture based on the position information, wherein the position information of the target object on the photographed picture may be characterized by coordinates.

[0077] S104: generating a target video based on photographed picture corresponding to the target object to which the target effect resource is applied.

[0078] In the embodiments of the present disclosure, the target effect resource may include props, effects and templates recommended based on the corresponding target object. For example, the target effect resource may include at least one of virtual overlay effect, occlusion effect, amplification effect and special effect processing effect. A video recording operation is triggered based on the photographed picture corresponding to the target object to which the target effect resource is applied, so as to generate a target video, wherein the target video may include any one or more effects in the target effect resource corresponding to the target object. Accordingly, the video effects are enriched and the users' experience in creation is also enhanced.

[0079] For example, when the target object is a drink can, the client receives the corresponding target effect resource of the drink can determined by the server in accordance with the above embodiments. Assuming that the target effect resource includes the virtual

overlay effect, a combined effect of the virtual world information and the target object may appears on the preview picture of the photographic page after the target effect resource is applied to the target object on the photographed picture. Assuming that the target effect resource includes the occlusion effect, an effect where animation resources partly occlude the target object may appear on the preview picture of the photographic page after the target effect resource is applied to the target object on the photographed picture. Assuming that the target effect resource includes the amplification effect, an effect where the target object is fully or partially amplified may appear on the preview picture of the photographic page after the target effect resource is applied to the target object on the photographed picture. Assuming that the target effect resource includes the special effect processing, an effect where texts resources rotate about the target object may appear on the preview picture of the photographic page after the target effect resource is applied to the target object on the photographed picture, and so on.

[0080] In the video generation method provided by the embodiments of the present disclosure, first of all, upon receiving an identification operation triggered for a photographed picture on a photographic page, the client obtains a target identification image based on the photographed picture on the photographic page and transmits the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image. Then, the client receives the target effect resource from the server, applies the target effect resource to the target object on the photographed picture and generates a target video based on photographed picture corresponding to the target object to which the target effect resource is applied. In the embodiments of the present disclosure, the matching target effect resource can be automatically applied to the target object on the photographed picture to generate a target video during the video shooting. As such, the functions of video shooting are enriched by the embodiments of the present disclosure, so as to enhance the users' experience in video creation.

[0081] Based on the above video generation system, the embodiments of the present disclosure provide a further video generation method. Fig. 2 illustrates a flowchart of a further video generation method provided by the embodiments of the present disclosure. As shown in Fig. 2, the method is applied to the server and comprises:

[0082] S201: receiving a target identification image from a client.

[0083] Wherein the target identification image is obtained from the photographed picture on the photographic page of the client.

[0084] In the embodiments of the present disclosure, the client receives the target identification image from the client, wherein the target identification image is one frame of image obtained by the client based on the photographed picture on the photographic page upon receiving an identification operation triggered for the photographed picture on a photographic page. In some embodiments, the target identification image whose image definition is determined to conform with the preset definition condition at the client may serve as the target identification image received by the server.

[0085] S202: determining whether the target identification image contains any one preset object.

[0086] In the embodiments of the present disclosure, the preset object may include any classes of objects whose corresponding feature information can be collected, e.g., objects in the class including pictures, three-dimensional objects and iconic buildings etc., like a give T-shirt pattern, drink can and high-rise building. On the basis of the target identification image obtained from the above S201, it is determined whether the target identification image contains any one preset object. For example, if the photographed picture on the photographic page is a table with one drink can placed thereon, the preset object contained in the target identification image may be the drink can; if the photographed picture on the photographic page is a plurality of high buildings, the preset object contained in the target identification image may be high-rise buildings, and so on.

[0087] In some implementations, in order to determine whether the target identification image contains any one preset object, image features of the target identification image are first extracted, then the image features match with feature information corresponding to the preset objects.

[0088] In the embodiments of the present disclosure, a corresponding relation between the preset object and its feature information is pre-configured at the server. After receiving the target identification image from the client, the feature information of the target identification image is extracted and further matches with the preset feature information corresponding to the preset object.

[0089] For example, an information comparison model is configured at the server, wherein

the information comparison model may be obtained from training on the basis of a sample set of corresponding relations between the different preset objects and feature information corresponding to preset objects. The server enters the extracted feature information of the target identification image into the information comparison model to obtain a match result
5 corresponding to the feature information of the target identification image.

[0090] In the embodiments of the present disclosure, in view of the match result in the above embodiments, if image features of the target identification image match feature information corresponding to any one preset object, the preset object is determined as target object.

[0091] For example, in a scenario where the information comparison model is configured at the server the sample data of the information comparison model includes a corresponding relation between the drink can and its feature information (e.g., data X (length, width and height) of the drink can), and the extracted feature information of the target identification image includes the data X of length, width and height. By inputting the feature information
15 of the target identification image into the information comparison model, a result indicating a successful match between the feature information (e.g., length, width and height data X) and the feature information corresponding to the preset object (e.g., drink can) may be obtained. Accordingly, the preset object (e.g., drink can) is determined as the target object and the target object identified from the target identification image is the drink can.

[0092] S203: If it is determined that the target identification image contains any one preset object, determining the preset object as target object and determining a target effect resource corresponding to the target object.

[0093] In the embodiments of the present disclosure, it is determined in the above S202 whether the target identification image contains any one preset object. If it is determined
25 that the target identification image contains any one preset object, the preset object is determined as target object and a target effect resource corresponding to the target object is also decided, wherein the target effect resource may include props, effects and templates recommended based on the corresponding target object.

[0094] In the embodiments of the present disclosure, if it is determined that the target identification image does not include any one preset object, it indicates failure of
30 identification and the server may determine common effect resources for the client.

[0095] S204: returning the target effect resource to the client.

[0096] Wherein the target effect resource is used for generating a target video at the client.

[0097] In some embodiments, after determining the preset object as target object, the method also comprises: returning position information of the target object on the photographed picture to the client, such that the client applies the target effect resource to the target object on the photographed picture based on the position information.

[0098] In the embodiments of the present disclosure, the target effect resource corresponding to the target object is determined from the above S203. The server returns the target effect resource to the client and also may return to the client position information of the target object on the photographed picture, to facilitate the client to generate a target video based on the target effect resource and the position information of the target object on the photographed picture.

[0099] In the video generation method provided by the embodiments of the present disclosure, first of all, the server receives a target identification image from a client; wherein the target identification image is obtained based on the photographed picture on the photographic page of the client. Then, the server determines whether the target identification image contains any one preset object. If it is determined that the target identification image contains any one preset object, the preset object is determined as target object and a target effect resource corresponding to the target object is also decided. Further, the target effect resource is returned to the client, wherein the target effect resource is used for generating a target video at the client. In the embodiments of the present disclosure, the matching target effect resource can be automatically applied to the target object on the photographed picture to generate a target video during the video shooting. As such, the functions of video shooting are enriched by the embodiments of the present disclosure, so as to enhance the users' experience in video creation.

[00100] In accordance with the above video generation method, the embodiments of the present disclosure also provide a video generation method. Fig. 3 illustrates a schematic diagram of data interaction of the video generation method provided by the embodiments of the present disclosure. As shown in Fig. 3, the method comprises:

[00101] S301: the client, upon receiving an identification operation triggered for a photographed picture on a photographic page, obtains a current image corresponding to the

photographed picture as an image with definition to be identified.

[00102] In the embodiments of the present disclosure, upon receiving an identification operation triggered for the photographed picture on a photographic page, the client obtains one frame of current image corresponding to the photographed picture, e.g., image 1, as an image with definition to be identified.

[00103] S302: the client determines whether an image definition of the image with definition to be identified conforms with a preset definition condition.

[00104] In the embodiments of the present disclosure, the image 1 is input to the fuzzy filtering model preset in the client, wherein it is assumed that a value threshold indicating conformity with a preset definition condition corresponding to the fuzzy filtering model is set to 80. The value (such as 90) of the image 1 is output based on the processing of the fuzzy filtering model.

[00105] S303: in response to determining that an image definition of the image with definition to be identified conforms with a preset definition condition, the image with definition to be identified is determined as a target identification image.

[00106] In the embodiments of the present disclosure, the value of the image 1 output from the above S302 is compared with the preset threshold (e.g., 80). If it is determined that an image definition of the image 1 has conforms with a preset definition condition (e.g., 90 greater than the threshold of 80), the image 1 is determined as a target identification image.

[00107] S304: the client transmits the target identification image to a server.

[00108] In the embodiments of the present disclosure, the client transmits the target identification image determined (i.e., image 1) in the above S303 to the server.

[00109] S305: the client extracts image features of the target identification image.

[00110] In the embodiments of the present disclosure, an information comparison model is pre-configured in the server and the target identification image (such as image 1) is input to the information comparison model to extract image features of the image 1.

[00111] S306: the server matches the image features with feature information corresponding to preset objects and determines whether the image features match with the feature information corresponding to any one preset object. If yes, S307 is performed; if not, S310 is executed.

[00112] In the embodiments of the present disclosure, the information comparison model in the server may be obtained from training based on a sample set of corresponding relations between the different preset objects and their corresponding feature information. The image 1 is input to the information comparison model, and the image features of the image 1 are matched with the feature information corresponding to the preset object to produce a corresponding match result for the feature information of the image 1.

[00113] S307: if image features of the target identification image match feature information corresponding to any one preset object successfully, the successfully matched preset object is determined as target object and the corresponding target effect resource of the target object is decided.

[00114] In the embodiments of the present disclosure, based on the match result in the above S306, if image features (e.g., length, width and height data X) of the image 1 match feature information (such as length, width and height data Y of the drink can) corresponding to any one preset object successfully, the preset object (e.g., drink can) is determined as target object and the corresponding target effect resource of the target object is decided, wherein the target effect resource may include props, effects and templates recommended on the basis of the corresponding target object. For example, the target effect resource may include at least one of virtual overlay effect, occlusion effect, amplification effect and special effect processing effect.

[00115] S308: the server transmits to the client the target effect resource and the position information of the target object on the photographed picture.

[00116] In the embodiments of the present disclosure, if the client determines the preset object as the target object, the target effect resource corresponding to the target object and the position information of the target object on the photographed picture can be determined. The server transmits to the client the target effect resource and the position information of the target object on the photographed picture, to facilitate the client to subsequently apply the target effect resource to the target object on the photographed picture based on the position information.

[00117] S309: the client applies the target effect resource to the target object on the photographed picture based on the position information to generate a target video.

[00118] In the embodiments of the present disclosure, after receiving the target effect

resource and the position information of the target object on the photographed picture, the client parses the target effect resource to obtain configuration, such as props and templates et al., contained in the target effect resource. Subsequent to the initialization of the SDK, the effects are rendered on the photographed picture of the client, such that the users can pre-view on a preview picture of the photographic page.

[00119] In the embodiments of the present disclosure, when the server receives an operation that triggers the recording, the recording is performed based on the preview picture to which the target effect resource is applied, so as to generate a target video.

[00120] S310: if the image features of the target identification image do not match with the feature information corresponding to any one preset object, the server determines common effect resources for the client.

[00121] In the embodiments of the present disclosure, according to the match result in the above S306, if the image features (e.g., length, width and height data X) of the image 1 do not match with the feature information corresponding to any one preset object, it indicates failure of identification for the target object and the server may determine common effect resources for the client.

[00122] S311: the server transmits to the client the common effect resources.

[00123] S312: the client applies the common effect resources to the target object on the photographed picture to generate a target video.

[00124] In the embodiments of the present disclosure, after receiving the common effect resources, the client parses the common effect resources to obtain configurations, such as props and templates, contained in the common effect resources. Subsequent to the initialization of the SDK, the effects are rendered on the photographed picture of the client, such that the users can pre-view on a preview picture of the photographic page. When the server receives an operation that triggers the recording, the recording is performed based on the preview picture to which the common effect resources are applied, so as to generate a target video.

[00125] In the embodiments of the present disclosure, the server can determine target effect resource based on the target object identified on the photographed picture, such that the client can apply the target effect resource determined by the server to the object image on the photographed picture, to generate a video. As such, the functions of video shooting are

enriched, thereby enhancing the users' experience in video creation.

[00126] In accordance with the above video generation method, the embodiments of the present disclosure propose a video generation system. Fig. 4 is a structural diagram of a video generation system provided by the embodiments of the present disclosure. As shown in Fig. 4, the video generation system may include a client 41 and a server 42. The client 41 is communicatively linked to the server 42.

[00127] The client 41 is used for obtaining a target identification image based on a photographed picture on the photographic page upon receiving an identification operation triggered for the photographed picture on a photographic page, and transmitting the target identification image to the server 42.

[00128] The server 42 is used for receiving a target identification image from the client 41 and determining whether the target identification image contains any one preset object; determining the preset object as target objects and determining target effect resource corresponding to the target object if it is determined that the target identification image contains any one preset object; and returning the target effect resource to the client 41.

[00129] The client 41 is also used for receiving the target effect resource from the server 42 and applying the target effect resource to the target object on the photographed picture; and generating a target video on based on the photographed picture corresponding to the target object to which the target effect resource is applied.

[00130] In some implementations, the client 41 also obtains a current image corresponding to the photographed picture as an image with definition to be identified upon receiving an identification operation triggered for the photographed picture on a photographic page; determines whether an image definition of the image with definition to be identified conforms with a preset definition condition; and if it is determined that an image definition of the image with definition to be identified conforms with a preset definition condition, the image with definition to be identified is determined as a target identification image.

[00131] In some implementations, the client 41 also obtains an image from the photographed picture on the photographic page based on preset frame extraction conditions if it is determined that the image definition of the image with definition to be identified does not conform with a preset definition condition, and updates the image with definition to be identified with the image; and continues to execute a step of the determining whether a

definition of the image with definition to be identified conforms with a preset definition condition, until it is determined that a preset definition identification termination condition is met.

[00132] In some implementations, the client 41 further processes the image with definition to be identified based on a preconfigured fuzzy filtering model, to obtain an image definition value of the image with definition to be identified; if an image definition value of the image with definition to be identified is greater than or equal to a value threshold indicating conformity with a preset definition condition, it is determined that image definition of the image with definition to be identified conforms with the preset definition condition; and if the image definition value of the image with definition to be identified is smaller than a value threshold indicating conformity with a preset definition condition, it is determined that image definition of the image with definition to be identified fails to conform with the preset definition condition.

[00133] In some implementations, the client 41 also extracts one frame of image from images corresponding to the photographed picture by an interval of a predetermined time; or, extracts one frame of image from images corresponding to the photographed picture by an interval of a predetermined frame number.

[00134] In some implementations, the client 41 also terminates identification for image definition if it is determined that a number of image frames extracted from images corresponding to the photographed picture reaches a preset image extraction number; or, terminates identification for image definition if it is determined that the image definition value of the image with definition to be identified is smaller than a preset minimum value.

[00135] In some implementations, the client 41 also receives from the server 42 position information of the target object on the photographed picture; and applies the target effect resource to the target object on the photographed picture based on the position information.

[00136] In some implementations, the server 42 also extracts image features of the target identification image; and matches the image features with feature information corresponding to preset objects; if image features of the target identification image successfully match feature information corresponding to any one preset object, the server determines the preset object as target object.

[00137] In some implementations, the server 42 also returns position information of the target

object on the photographed picture to the client, such that the client 41 applies the target effect resource to the target object on the photographed picture based on the position information.

[00138] In the video generation system provided by the embodiments of the present disclosure, the server 42 can determine a target effect resource based on the target object identified on the photographed picture, such that the client 41 can apply the target effect resource determined by the server 42 to the object image on the photographed picture, to generate a video. As such, the functions of video shooting are enriched, thereby enhancing the users' experience in video creation.

[00139] In accordance with the above method embodiments, the present disclosure also proposes a video generation apparatus. Fig. 5 is a structural diagram of a video generation apparatus provided by the embodiments of the present disclosure. The apparatus is applied to the client and the video generation apparatus 500 includes:

[00140] an acquisition module 501 for upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a target identification image based on the photographed picture on the photographic page;

[00141] a transmission module 502 for transmitting the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image;

[00142] an application module 503 for receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture;

[00143] a generation module 504 for generating a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied.

[00144] In some implementations, the acquisition module 501 includes:

[00145] an acquisition sub-module for obtaining a current image corresponding to the photographed picture as an image with definition to be identified upon receiving an identification operation triggered for the photographed picture on a photographic page;

[00146] a first determination sub-module for determining whether an image definition of the image with definition to be identified conforms with a preset definition condition;

[00147] a second determination sub-module for determining the image with definition to be identified as a target identification image if it is determined that an image definition of the

image with definition to be identified conforms with a preset definition condition.

[00148] In some implementations, the apparatus also includes:

[00149] an updating module for obtaining an image from the photographed picture on the photographic page based on preset frame extraction conditions if it is determined that an image definition the image with definition to be identified does not conform with a preset definition condition, and updating the image with definition to be identified with the image.

[00150] an identification module for continuing to execute a step of the determining whether an image definition of the image with definition to be identified conforms with a preset definition condition, until it is determined that a preset definition identification termination condition is met.

[00151] In some implementations, the first determination sub-module is specifically used for:

[00152] processing the image with definition to be identified based on a preconfigured fuzzy filtering model, to obtain an image definition value of the image with definition to be identified; if an image definition value of the image with definition to be identified is greater than or equal to a value threshold indicating conformity with a preset definition condition, determining that image definition of the image with definition to be identified conforms with the preset definition condition; and if the image definition value of the image with definition to be identified is smaller than a value threshold indicating conformity with a preset definition condition, determining that image definition of the image with definition to be identified fails to conform with the preset definition condition.

[00153] In some implementations, the updating module is used specifically for:

[00154] extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined time; or, extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined frame number.

[00155] In some implementations, the identification module is used for:

[00156] if it is determined that a number of image frames extracted from images corresponding to the photographed picture reaches a preset image extraction number, terminating identification for image definition; or, if it is determined that an image definition value of the image with definition to be identified is smaller than a preset minimum value, terminating identification for image definition.

[00157] In some implementations, the apparatus also comprises:

[00158] a second receiving module for receiving position information of the target object on the photographed picture from the server;

[00159] correspondingly, the application module 503 includes:

- 5 **[00160]** an application sub-module for applying the target effect resource to the target object on the photographed picture based on the position information.

[00161] In the video generation apparatus provided by the embodiments of the present disclosure, first of all, upon receiving an identification operation triggered for a photographed picture on a photographic page, the client obtains a target identification image based on the photographed picture on the photographic page and transmits the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image. Then, the client receives the target effect resource from the server, applies the target effect resource to the target object on the photographed picture and generates a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied. In the embodiments of the present disclosure, the matching target effect resource can be automatically applied to the target object on the photographed picture to generate a target video during the video shooting. As such, the functions of video shooting are enriched by the embodiments of the present disclosure, so as to enhance the users' experience in video creation.

10
15
20

[00162] Based on the above method embodiments, the present disclosure also provides a video generation apparatus. Fig. 6 illustrates a structural diagram of a further video generation apparatus provided by the embodiments of the present disclosure. The apparatus is applied to the server and the video generation apparatus 600 comprises:

- 25 **[00163]** a first receiving module 601 for receiving a target identification image from a client; wherein the target identification image is obtained based on the photographed picture on the photographic page of the client;

[00164] a first determination module 602 for determining whether the target identification image contains any one preset object;

- 30 **[00165]** a second determination module 603 for determining the preset object as target object if it is determined that the target identification image contains any one preset object, and

determining target effect resource corresponding to the target object;

[00166] a return module 604 for returning the target effect resource to the client; wherein the target effect resource is used for generating a target video at the client.

[00167] In some implementations, the first determination module 602 includes:

5 **[00168]** an extraction sub-module for extracting image features of the target identification image;

[00169] a matching sub-module for matching the image features with feature information corresponding to preset objects;

[00170] correspondingly, the second determination module 603 includes:

10 **[00171]** a third determination sub-module for determining the preset object as target object if image features of the target identification image successfully match feature information corresponding to any one preset object.

[00172] In some implementations, the return module 604 is also used for:

15 **[00173]** returning position information of the target object on the photographed picture to the client, such that the client applies the target effect resource to the target object on the photographed picture based on the position information.

20 **[00174]** In the video generation apparatus provided by the embodiments of the present disclosure, first of all, the server receives a target identification image from a client; wherein the target identification image is obtained based on the photographed picture on the photographic page of the client. Then, the server determines whether the target identification image contains any one preset object. If it is determined that the target identification image contains any one preset object, the preset object is determined as target object and a target effect resource corresponding to the target object is also decided. Further, the target effect resource is returned to the client, wherein the target effect resource is used for
25 generating a target video at the client. In the embodiments of the present disclosure, the matching target effect resource can be automatically applied to the target object on the photographed picture to generate a target video during the video shooting. As such, the functions of video shooting are enriched by the embodiments of the present disclosure, so as to enhance the users' experience in video creation.

30 **[00175]** In addition to the above method and apparatus, the embodiments of the present

disclosure also provide a computer-readable storage medium stored thereon with instructions which, when running on a terminal device, cause the terminal device to implement the video generation method according to the embodiments of the present disclosure.

[00176] Embodiments of the present disclosure also provide a computer program product comprising a computer program/instruction which, when executed by a processor, implements the video generation method according to the embodiments of the present disclosure.

[00177] Moreover, embodiments of the present disclosure also provide a video generation device 700. According to Fig. 7, the device may comprise:

[00178] a processor 701, a memory 702, an input unit 703 and an output unit 704. The video generation device may include one or more processors 701 and Fig. 7 only shows one processor as the example. In some embodiments of the present disclosure, the processor 701, the memory 702, the input unit 703 and the output unit 704 may be connected by buses or other components. Fig. 7 illustrates the bus connection as an example.

[00179] The memory 702 may be used for storing software programs and modules and the processor 701 operates the software programs and modules stored in the memory 702 to execute the various functional applications and data processing of the video generation device. The memory 702 mainly includes program storage area and data storage area, wherein the program storage area may store operating systems and applications desired by at least one function and the like. Besides, the memory 702 may include a high random access memory and also a non-volatile memory, such as at least one magnetic disk storage device, flash device or other volatile solid-state storage devices. The input unit 703 may be provided for receiving input digital or character information and generating signal inputs associated with the user setting and the functional control of the video generation device.

[00180] Specifically, in the embodiments, the processor 701 would load into the memory 702 executable files corresponding to the process of one or more applications in accordance with the following instructions. The processor 701 also operates the applications stored in the memory 702 to implement the above various functions of the video generation device.

[00181] It is to be explained that relational terms such as “first” and “second” in the text are disclosed only to distinguish one entity or operation from a further entity or operation without requiring or suggesting any actual relations or sequences between these entities or operations. Besides, the terms “comprise”, “contain” or any other variants are not featured with

exclusiveness. Accordingly, the procedure, method, object or device including a series of factors may also include other factors not clearly listed in addition to those stated. Alternatively, inherent factors of the procedure, method, object or device are also included. The procedure, method or device including factors defined by the expression of “comprising one ...” may also contain other same factors if no more restrictions are provided.

[00182] The above is just the specific implementations of the present disclosure, which are provided for those skilled in the art to understand or implement the present disclosure. Many modifications to the embodiments are obvious for those skilled in the art. General principles defined in the text may be implemented in other embodiments without deviating from the spirit or scope of the present disclosure. Therefore, the present disclosure will not be restricted to the embodiments disclosed herein and instead has a broadest scope consistent with the principles and novel features disclosed here.

I/We Claim:

1. A video generation method, wherein the method is applied to a client and comprises:
upon receiving an identification operation triggered for a photographed picture on a
photographic page, obtaining a target identification image based on the photographed picture
on the photographic page;

transmitting the target identification image to a server, wherein the server is used for
determining a target effect resource for a target object after identifying the target object from
the target identification image;

receiving the target effect resource from the server and applying the target effect
resource to the target object on the photographed picture;

generating a target video based on the photographed picture corresponding to the
target object to which the target effect resource is applied.

2. The method of claim 1, wherein upon receiving an identification operation triggered
for a photographed picture on a photographic page, obtaining a target identification image
based on the photographed picture on the photographic page comprises:

upon receiving an identification operation triggered for a photographed picture on a
photographic page, obtaining a current image corresponding to the photographed picture as an
image with definition to be identified;

determining whether a definition of the image with definition to be identified
conforms with a preset definition condition;

in response to determining that the definition of the image with definition to be
identified conforms with a preset definition condition, determining the image with definition
to be identified as a target identification image.

3. The method of claim 2, wherein after determining whether a definition of the image
with definition to be identified conforms with a preset definition condition, the method further
comprises:

in response to determining that a definition of the image with definition to be
identified fails to conform with a preset definition condition, obtaining an image from the
photographed picture on the photographic page based on preset frame extraction conditions
and updating the image with definition to be identified with the image;

continuing to execute a step of determining whether a definition of the image with

definition to be identified conforms with a preset definition condition, until it is determined that a preset definition identification termination condition is met.

4. The method of claim 2 or 3, wherein determining whether a definition of the image with definition to be identified conforms with a preset definition condition comprises:

processing the image with definition to be identified based on a preconfigured fuzzy filtering model, to obtain an image definition value of the image with definition to be identified;

if an image definition value of the image with definition to be identified is greater than or equal to a value threshold indicating conformity with a preset definition condition, determining that image definition of the image with definition to be identified conforms with the preset definition condition;

if an image definition value of the image with definition to be identified is smaller than a value threshold indicating conformity with a preset definition condition, determining that image definition of the image with definition to be identified fails to conform with the preset definition condition.

5. The method of claim 3, wherein obtaining an image from a photographed picture on the photographic page based on preset frame extraction conditions comprises:

extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined time;

or,

extracting one frame of image from images corresponding to the photographed picture by an interval of a predetermined frame number.

6. The method of claim 3, wherein determining that a preset definition identification termination condition is met comprises:

in response to determining that a number of image frames extracted from images corresponding to the photographed picture reaches a preset image extraction number, terminating identification for image definition;

or,

in response to determining that an image definition value of the image with definition to be identified is smaller than a preset minimum value, terminating identification for image

definition.

7. The method of claim 1, wherein before applying the target effect resource to the target object on the photographed picture, the method further comprises:

5 receiving position information of the target object on the photographed picture from the server;

wherein applying the target effect resource to the target object on the photographed picture comprises:

10 applying the target effect resource to the target object on the photographed picture based on the position information.

8. A video generation method, wherein the method is applied to a server and comprises:

15 receiving a target identification image from a client; wherein the target identification image is obtained based on a photographed picture on a photographic page of the client;

determining whether the target identification image contains any one preset object;

in response to determining that the target identification image contains any one preset object, determining the preset object as a target object and determining a target effect resource corresponding to the target object;

20 returning the target effect resource to the client; wherein the target effect resource is used for generating a target video at the client.

9. The method of claim 8, wherein determining whether the target identification image contains any one preset object comprises:

25 extracting image features of the target identification image;

matching the image features with feature information corresponding to preset objects;

correspondingly, wherein in response to determining that the target identification image contains any one preset object, determining the preset object as a target object comprises:

30 if image features of the target identification image successfully match feature information corresponding to any one preset object, determining the preset object as the target object.

10. The method of claim 8, wherein after determining the preset object as a target object, the method further comprises:

returning position information of the target object on the photographed picture to the client, such that the client applies the target effect resource to the target object on the photographed picture based on the position information.

11. A video generation system, wherein the system comprises a client and a server;

the client is used for upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a target identification image based on the photographed picture on the photographic page, and transmitting the target identification image to the server;

the server is used for receiving a target identification image from the client and determining whether the target identification image contains any one preset object; in response to determining that the target identification image contains any one preset object, determining the preset object as a target object and determining a target effect resource corresponding to the target object; and returning the target effect resource to the client;

the client is further used for receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture; and generating a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied.

12. A video generation apparatus, wherein the apparatus is applied to client and comprises:

an acquisition module for upon receiving an identification operation triggered for a photographed picture on a photographic page, obtaining a target identification image based on photographed picture on the photographic page;

a transmission module for transmitting the target identification image to a server; wherein the server is used for determining a target effect resource for a target object after identifying the target object from the target identification image;

an application module for receiving the target effect resource from the server and applying the target effect resource to the target object on the photographed picture;

a generation module for generating a target video based on the photographed picture corresponding to the target object to which the target effect resource is applied.

13. A video generation apparatus, wherein the apparatus is applied to a server and comprises:

a first receiving module for receiving a target identification image from a client;

5 wherein the target identification image is obtained based on a photographed picture on a photographic page of the client;

a first determination module for determining whether the target identification image contains any one preset object;

10 a second determination module for in response to determining that the target identification image contains any one preset object, determining the preset object as a target object and determining target effect resource corresponding to the target object;

a return module for returning the target effect resource to the client; wherein the target effect resource is used for generating a target video at the client.

15 14. A computer-readable storage medium stored therein with instructions which, when running on a terminal device, cause the terminal device to perform the method according to any of claims 1-10.

20 15. A device comprising: a memory, a processor and computer programs stored on the memory and running on the processor, wherein the processor, when executing the computer programs, performs the method according to any of claims 1-10.

16. A computer program product comprising a computer program/instruction which, when executed by a processor, performs the method according to any of claims 1-10.

25