

CHARACTER PROCESSING METHOD AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

CROSS-REFERENCE

[0001] The present application claims priority to Chinese Patent Application No. 202311049036.4, filed on August 18, 2023, which is incorporated herein by reference in its entirety as a part of the present application.

TECHNICAL FIELD

[0002] The present disclosure relates to a text processing method and apparatus, a computer device, and a storage medium.

BACKGROUND

[0003] In many scenarios, such as image editing, poster making, etc., it is necessary to add some text to improve the visual effect and make it richer. Generally, when adding text, the added text is often art text, and the art text is generally designed by professional designers, which results in only a few types of art text that can be provided to users, and the types are relatively few; moreover, it takes a lot of time for designers to design art text, and the efficiency is low.

SUMMARY

[0004] Embodiments of the present disclosure at least provide a text processing method and apparatus, a computer device, and a storage medium.

[0005] Embodiments of the present disclosure provide a text processing method, including:

[0006] obtaining a target text and determining font information corresponding to the target text;

[0007] generating a text image corresponding to the target text according to the font information; and

[0008] determining a mask image corresponding to the target text in the text image, and inputting the mask image and style prompt information corresponding to the target text into a stylization model and determining a stylized text image corresponding to the target text.

[0009] In a possible implementation, the obtaining a target text includes:

[0010] obtaining the target text added to an image to be processed; and

[0011] after the determining the stylized text image corresponding to the target text, the method further includes:

- [0012] fusing and displaying the stylized text image with the image to be processed.
- [0013] In a possible implementation, the method further includes: determining the style prompt information corresponding to the target text by using the following method:
- [0014] receiving the style prompt information corresponding to the target text entered by a user; or
- [0015] performing style detection on the image to be processed and determining the style prompt information corresponding to the target text based on a style detection result; or
- [0016] performing content detection on the target text and determining the style prompt information corresponding to the text based on a content detection result.
- [0017] In a possible implementation, in response to the style prompt information including a name of a main object, text in the stylized text image corresponding to the target text is text composed based on the main object.
- [0018] In a possible implementation, in response to the style prompt information not including the name of the main object, the text in the stylized text image corresponding to the target text is text composed based on an associated object associated with the style prompt information.
- [0019] In a possible implementation, in response to the target text including a plurality of texts, different texts correspond to different style prompt information; and
- [0020] the inputting the mask image and style prompt information corresponding to the target text into a stylization model and determining a stylized text image corresponding to the target text includes:
- [0021] inputting the mask image and style prompt information corresponding to each target text into the stylization model and determining the stylized text image corresponding to the target text, in the stylized text image, stylized texts corresponding to different target texts having different style types.
- [0022] In a possible implementation, the obtaining a target text and determining font information corresponding to the target text includes:
- [0023] obtaining an image to be processed, performing content recognition on the image to be processed, and determining, based on a content recognition result, the target text matching the image to be processed and the font information corresponding to the target text; and
- [0024] after the determining a stylized text image corresponding to the target text, the method further includes:
- [0025] fusing and displaying the stylized text image with the image to be processed.
- [0026] Embodiments of the present disclosure further provide a text processing apparatus, including:
- [0027] an obtaining module configured to obtain a target text and determine font information corresponding to the target text;
- [0028] a generation module configured to generate a text image corresponding to the target text according to the font information; and

- [0029] a stylization processing module configured to determine a mask image corresponding to the target text in the text image, and input the mask image and style prompt information corresponding to the target text into a stylization model and determine a stylized text image corresponding to the target text.
- [0030] In a possible implementation, the obtaining module, when obtaining a target text, is configured to:
- [0031] obtain the target text added to an image to be processed; and
- [0032] after determining the stylized text image corresponding to the target text, the stylization processing module is configured to:
- [0033] fuse and display the stylized text image with the image to be processed.
- [0034] In a possible implementation, the stylization processing module is further configured to: determine the style prompt information corresponding to the target text by using the following method:
- [0035] receiving the style prompt information corresponding to the target text entered by a user; or
- [0036] performing style detection on the image to be processed and determining the style prompt information corresponding to the target text based on a style detection result; or
- [0037] performing content detection on the target text and determining the style prompt information corresponding to the text based on a content detection result.
- [0038] In a possible implementation, in response to the style prompt information including a name of a main object, text in the stylized text image corresponding to the target text is text composed based on the main object.
- [0039] In a possible implementation, in response to the style prompt information not including the name of the main object, the text in the stylized text image corresponding to the target text is text composed based on an associated object associated with the style prompt information.
- [0040] In a possible implementation, in response to the target text including a plurality of texts, different texts correspond to different style prompt information; and
- [0041] the stylization processing module, when inputting the mask image and style prompt information corresponding to the target text into a stylization model and determining a stylized text image corresponding to the target text, is configured to:
- [0042] input the mask image and style prompt information corresponding to each target text into the stylization model and determine the stylized text image corresponding to the target text, in the stylized text image, stylized texts corresponding to different target texts having different style types.
- [0043] In a possible implementation, the obtaining module, when obtaining a target text and determining font information corresponding to the target text, is configured to:
- [0044] obtain an image to be processed, perform content recognition on the image to be processed, and determine, based on a content recognition result, the target text matching the image to be processed and the font information corresponding to the target text; and
- [0045] after the determining a stylized text image corresponding to the target text, the method further includes:

- [0046] fusing and displaying the stylized text image with the image to be processed.
- [0047] Embodiments of the present disclosure further provide a computer device, including: a processor, a memory, and a bus, the memory stores machine-readable instructions executable by the processor, the processor communicates with the memory through the bus when the computer device is running, and the machine-readable instructions, when executed by the processor, cause the steps of the text processing method provided by any one of the embodiments of the present disclosure to be performed.
- [0048] Embodiments of the present disclosure further provide a computer-readable storage medium, the computer-readable storage medium stores a computer program that, when run by a processor, causes the steps of the text processing method provided by any one of the embodiments of the present disclosure to be performed.
- [0049] In order to make the above objects, features, and advantages of the present disclosure more comprehensible, some embodiments are described in detail below with reference to the drawings.

BRIEF DESCRIPTION OF DRAWINGS

- [0050] In order to more clearly illustrate the technical solutions in the embodiments of the present disclosure, the following briefly introduces the drawings used in the embodiments. The drawings herein are incorporated into and constitute a part of the specification, illustrate embodiments consistent with the present disclosure, and together with the description, serve to explain the technical solutions of the present disclosure. It should be understood that the following drawings show only some embodiments of the present disclosure, and therefore should not be considered as limiting the scope. For those of ordinary skill in the art, other related drawings may also be obtained from these drawings without creative efforts.
- [0051] Fig. 1 shows a flowchart of a text processing method provided by an embodiment of the present disclosure;
- [0052] Fig. 2 shows a schematic diagram of a text image in the text processing method provided by the embodiment of the present disclosure;
- [0053] Fig. 3 shows a schematic diagram of a mask image in the text processing method provided by the embodiment of the present disclosure;
- [0054] Fig. 4a shows a schematic diagram of text in a stylized text image in the text processing method provided by the embodiment of the present disclosure;
- [0055] Fig. 4b shows a schematic diagram of comparison between a stylized text image and a mask image in the text processing method provided by the embodiment of the present disclosure;
- [0056] Fig. 5 shows a schematic architectural diagram of a text processing apparatus provided by an embodiment of the present disclosure; and
- [0057] Fig. 6 shows a schematic structural diagram of a computer device provided by an embodiment of the present disclosure.

DETAILED DESCRIPTION

[0058] In order to make the purpose, technical solutions, and advantages of the embodiments of the present disclosure clearer, the technical solutions in the embodiments of the present disclosure will be described clearly and comprehensively below with reference to the drawings in the embodiments of the present disclosure. Clearly, the described embodiments are only some embodiments of the present disclosure, rather than all the embodiments. The components of the embodiments of the present disclosure, which are generally described and illustrated in the drawings herein, may be arranged and designed in various different configurations. Therefore, the following detailed description of the embodiments of the present disclosure provided in the drawings is not intended to limit the scope of the claimed present disclosure, but merely represents selected embodiments of the present disclosure. All other embodiments obtained by those skilled in the art based on the embodiments of the present disclosure without creative efforts belong to the protection scope of the present disclosure.

[0059] It has been found through research that in many scenarios, such as image editing, poster making, etc., it is necessary to add some text to improve the visual effect and make it richer. Generally, when adding text, in order to make the text match the picture better, the added text is often art text, and the art text is generally designed by professional designers, which results in only a few types of art text that can be provided to users, and the types are relatively few; moreover, it takes a lot of time for designers to design art text, and the efficiency is low.

[0060] Based on the above research, the present disclosure provides a text processing method and apparatus, a computer device, and a storage medium. In the method, after a target text is obtained, a text image corresponding to the target text may be generated according to font information corresponding to the target text, then a mask image corresponding to the text image may be determined, and the mask image is processed by a stylization model according to style prompt information to obtain a stylized text image corresponding to the target text. Through this method, a personalized stylized text image may be generated in real time, and the processing speed and processing quality of the stylized text are improved.

[0061] It should be noted that similar reference numerals and letters represent similar items in the following drawings, and therefore, once an item is defined in one drawing, it does not need to be further defined and explained in subsequent drawings.

[0062] The term "and/or" used herein only describes an association relationship, which means that three relationships may exist, for example, A and/or B may mean: only A exists, both A and B exist, and only B exists. In addition, the term "at least one" used herein represents any one of a plurality of items or any combination of at least two of the plurality of items, for example, including at least one of A, B, and C may mean including any one or more elements selected from a set composed of A, B, and C.

[0063] It may be understood that before the use of the technical solutions disclosed in the embodiments of the present disclosure, users shall be informed of the type, range of use, use scenarios, etc., of personal

information involved in the present disclosure in an appropriate manner and the authorization of the users shall be obtained in accordance with relevant laws and regulations.

[0064] For example, in response to reception of an active request from a user, prompt information is sent to the user to clearly inform the user that the requested operation will require access to and use of personal information of the user. In this way, the user may independently choose, according to the prompt information, whether to provide the personal information to software or hardware, such as an electronic device, an application, a server, or a storage medium, that performs the operations of the technical solutions of the present disclosure.

[0065] As an optional but non-limiting implementation, in response to the reception of the active request from the user, the prompt information may be sent to the user in the form of, for example, a pop-up window, and the prompt information may be presented in text in the pop-up window. In addition, the pop-up window may also include a selection control for the user to choose whether to "agree" or "disagree" to provide the personal information to the electronic device.

[0066] It may be understood that the above process of notifying and obtaining user authorization is only illustrative and does not constitute a limitation on the implementations of the present disclosure, and other manners that satisfy the relevant laws and regulations may also be applied to the implementations of the present disclosure.

[0067] In order to facilitate understanding of the present embodiment, a text processing method disclosed in an embodiment of the present disclosure is first introduced in detail below. An execution body of the text processing method provided by the embodiment of the present disclosure is generally a computer device with a certain computing power, and the computer device includes, for example, a server, a terminal device, etc.

[0068] Fig. 1 is a flowchart of a text processing method provided by an embodiment of the present disclosure. The method includes steps 101 to 103.

[0069] Step 101: obtaining a target text and determining font information corresponding to the target text.

[0070] Step 102: generating a text image corresponding to the target text according to the font information.

[0071] Step 103: determining a mask image corresponding to the target text in the text image, and inputting the mask image and style prompt information corresponding to the target text into a stylization model and determining a stylized text image corresponding to the target text.

[0072] The following is a detailed description of the above steps.

[0073] The target text may be a text to be stylized. In a possible implementation, when the target text is obtained, it may be text entered by a user, or optical character recognition (OCR) may be performed on an image, and then the recognized text is used as the target text.

[0074] In a possible application scenario, the target text may be text for adding to an image to be processed, for example, in a poster design scenario, the target text may be text that needs to be added to a poster, therefore, the obtaining the target text may be obtaining the target text added to the image to be processed.

- [0075] The font information corresponding to the target text may include, for example, font types such as Song typeface, Kai typeface, and Hei typeface. When the font information corresponding to the target text is determined, it may be in response to a setting instruction from the user, and the font information carried in the setting instruction is determined as the font information corresponding to the target text.
- [0076] Optionally, the font information may further include font size, font weight, etc.
- [0077] In another possible implementation, the obtaining a target text and determining font information corresponding to the target text can be performed after obtaining an image to be processed, and includes performing content recognition on the image to be processed, and determining, based on a content recognition result, the target text matching the image to be processed and the font information corresponding to the target text.
- [0078] Specifically, in a poster making scenario, after a picture part of the poster is completed, text may be added to the poster by means of automatic copywriting, and the target text may be text automatically generated during automatic copywriting for the image.
- [0079] Exemplarily, a copywriting model may be used for automatic copywriting, and after the image to be processed is input to the copywriting model, the copywriting model may automatically output the target text matching the image to be processed and the font information corresponding to the target text.
- [0080] For example, if the image to be processed is an ink and wash landscape painting, the target text matching the image to be processed may be a certain ancient poem.
- [0081] The font information corresponding to the target text may match an image style of the image to be processed, for example, if the image style of the image to be processed is an ink and wash landscape painting, the font information corresponding to the target text may be Kai typeface, and if the image style of the image to be processed is a cartoon, the font information corresponding to the target text may be Hei typeface, etc.
- [0082] The generating the text image corresponding to the target text according to the font information may refer to displaying an image of a target background color according to a preset size, and then drawing, on the image of the target background color, the target text according to the font information.
- [0083] The target background color may be, for example, white; when the target text is drawn on the image, a color different from the target background color may be used for drawing, and for example, the default color black may be used for drawing. Exemplarily, if the target text is "ABC", the generated text image is as shown in Fig. 2.
- [0084] It should be noted that here, after the text image is generated, it may not be displayed, but only used for a subsequent stylization processing process.
- [0085] The mask image corresponding to the target text in the text image may be a binary image. In the mask image, a pixel value of a text area where the target text is located is 1, and pixel values of other areas other than the text area are 0, exemplarily as shown in Fig. 3.

- [0086] In a possible implementation, when the mask image corresponding to the target text in the text image is determined, edge detection may be performed on the text image, and the mask image may be determined based on an edge detection result.
- [0087] Here, a size of the mask image is the same as a size of the text image, and both may be the preset size, and the preset size may be a preset image size for stylization model processing.
- [0088] The stylization model may be an artificial intelligence generated content (AIGC) model; the style prompt information, prompt, corresponding to the target text is used for prompting the stylization model what style of image the text image is to be processed into. Optionally, when the style prompt information corresponding to the target text is determined, any of the following methods may be used:
- [0089] Method 1: receiving the style prompt information corresponding to the target text entered by a user.
- [0090] Specifically, the user may simultaneously enter the style prompt information corresponding to the target text when entering the target text.
- [0091] Method 2: performing style detection on the image to be processed and determining the style prompt information corresponding to the target text based on a style detection result.
- [0092] Here, the image to be processed may refer to an image where the target text is located when the target text is obtained. Because the target text needs to be added to the image to be processed, the final stylized text needs to be consistent with the style of the image to be processed.
- [0093] In a specific implementation, when style detection is performed on the image to be processed, a pre-trained style detection model may be used, the image to be processed is input to the style detection model, and the style detection model may output the style detection result, the style detection result includes a style to which the image to be processed belongs.
- [0094] After the style detection result is determined, a text style corresponding to an image style of the image to be processed may be determined according to a pre-established correspondence between image styles and text styles, and the determined text style is used as the style prompt information.
- [0095] Method 3: performing content detection on the target text and determining the style prompt information corresponding to the text based on a content detection result.
- [0096] Specifically, the performing content detection on the target text may be performing word segmentation processing on the target text to determine keywords included in the target text, then detecting whether the keywords included in the target text include a target keyword, and using the target keyword as the style prompt information. The word segmentation processing may be performed using, for example, an N-gram model.
- [0097] The target keyword may be a word used for representing style information, such as a cartoon, ancient style, cool, etc. The target keyword may be a preset keyword.
- [0098] In another possible implementation, the stylization model may be a model that may only generate a fixed style. In this case, the style prompt information of the stylization model may be fixed, that is, a preset default value.

- [0099] Similarly, if each stylization model may only generate a stylized text of one style, in this case, the method may further include a step of selecting a stylization model.
- [0100] Specifically, in response to a selection operation of the user, the stylization model selected by the selection operation may be used as the model for performing step 103, or the model for performing step 103 may be a model matching the target text.
- [0101] Exemplarily, semantic detection may be performed on the target text, and a model matching a semantic detection result is determined as the model for performing step 103 (for example, a matching relationship between each semantic detection result and a model may be preset).
- [0102] In a possible implementation, the target text may include a plurality of texts, and different texts may correspond to different style prompt information, or not all texts correspond to the same style prompt information, and style prompt information corresponding to some texts may be the same, and style prompt information corresponding to some texts may be different.
- [0103] In this way, when the mask image and the style prompt information corresponding to the target text are input into the stylization model to determine the stylized text image corresponding to the target text, the mask image and the style prompt information corresponding to each target text may be input into the stylization model to determine the stylized text image corresponding to the target text, in the stylized text image, stylized texts corresponding to different target texts have different style types.
- [0104] Exemplarily, if the target text includes three texts, the style prompt information may include three parts of information, which respectively correspond to the three characters of the target text, and the generated stylized text image includes stylized text respectively corresponding to the three target texts, and the three stylized texts have different style types.
- [0105] In a possible implementation, in order to improve similarity between the text in the stylized text image and the target text, during a stylization process of the stylization model, a feature constraint may further be added to the stylization model.
- [0106] Specifically, the mask image may be input to a control model to determine a constraint feature corresponding to the mask image, and then the style prompt information, the mask image, and the constraint feature may be input to the stylization model to determine the stylized text image corresponding to the target text.
- [0107] The control model may be, for example, a controlnet model, and after the mask image is input to the control model, the control model may output constraint features of different dimensions. The constraint feature may be understood as a feature that may best represent the target text, for example, a structural feature of the target text.
- [0108] The inputting the style prompt information, the mask image, and the constraint feature into the stylization model may be understood as that the style prompt information and the mask image are input into the stylization model at the same time, and the constraint feature may act on an intermediate feature map of

the stylization model, for example, the constraint feature may be fused with the intermediate feature map output from the stylization model to constrain the processing process of the stylization model.

[0109] In practical applications, in response to the style prompt information including a name of a main object, text in the stylized text image corresponding to the target text is text composed based on the main object.

[0110] Here, the style prompt information including the main object refers to that the style prompt information includes a name of the main object. When detecting whether the style prompt information includes the main object, exemplarily, word segmentation processing may be first performed on the style prompt information to determine a word segmentation result corresponding to the style prompt, and then the word segmentation result is matched with names of respective main objects in a main object name library, and if the match is successful, it is determined that the style prompt information includes the name of the main object.

[0111] The text composed based on the main object may refer to text composed based on an image of the main object, and the composed here may be understood as putting together.

[0112] Exemplarily, if the style prompt information is "nut" and the target text is "ABC", the text in the generated stylized text image is "ABC" composed based on nuts, exemplarily as shown in Fig. 4a.

[0113] In response to the style prompt information not including the name of the main object, the text in the stylized text image corresponding to the target text is text composed based on an associated object associated with the style prompt information.

[0114] Here, the associated object associated with the style prompt information may be understood as an object that may reflect the style indicated by the style prompt information, exemplarily, if the style prompt information is a fruit, the associated object may be an apple, a grape, a banana, etc.; and if the style prompt information is a flower, the associated object may be a lily, a sunflower, a rose, etc.

[0115] It should be noted that the main object/associated object that constitutes the text may be one or more types of objects.

[0116] Generally, the stylized text image and the mask image may have the same size. Further, in order to make the text in the stylized text image more vivid, the area where the text in the stylized text image is located may not completely coincide with the mask area in the mask image (the mask area is the area corresponding to the target text in the mask image), that is, the area where the text in the stylized text image is located may be larger or smaller than the mask area in the mask image, and the larger or smaller here does not refer to scaling up or scaling down in proportion, but refers to incomplete coincidence under the same size.

[0117] Exemplarily, as shown in Fig. 4b, in Fig. 4b, the black-and-white image on the left half is the mask image, the target text is "P", the gray image on the right half is the stylized text image, and in the area where the text in the stylized text image is located, the position selected by the dotted frame is to fully display the

flower, and the displayed flower exceeds the original mask area. Other areas other than the dotted frame also have content exceeding the mask area, which will not be introduced as examples here.

[0118] In practical applications, during a process of training the stylization model, a model parameter of the stylization model may be adjusted according to a size of the area where the text in the stylized text image generated during the training process is located and the size of the mask area in the mask image, so as to control a size difference between the two areas.

[0119] In a possible implementation, in response to the target text being obtained based on the image to be processed, after the stylized text image corresponding to the target text is generated, the stylized text image and the image to be processed may be fused and displayed.

[0120] In practical applications, after the stylized text image corresponding to the target text is generated, the stylized text image may be displayed at a target position (such as a center) of the image to be processed, and then a size and a display position of the stylized text image may be set in response to a setting instruction of the user for the stylized text image.

[0121] In another possible implementation, an image shape of the stylized text image may be the target text, that is, the stylized text image is an image in a shape of the target text, as shown in Fig. 4a. The text in the stylized text image is different from ordinary text, and the text in the stylized text image cannot be modified.

[0122] In a possible implementation, the text in the stylized text image may be deleted. Exemplarily, in response to a deletion instruction from the user for any text in the stylized text image, an image area corresponding to the text in the stylized text image may be segmented, and the image area corresponding to the segmented text may be deleted.

[0123] In another possible implementation, after the stylized text image is generated based on the above method, a style text corresponding to the stylized text image may also be published in response to a publishing instruction from the user. Here, after the style text corresponding to the stylized text image is published, when others use the style text, after entering the text to be stylized, without performing other operations (such as setting style prompt information and other operations), the input text to be stylized may be directly processed based on the same processing logic as the generation method of the published stylized text image.

[0124] In the text processing method provided by the embodiment of the present disclosure, after the target text is obtained, the text image corresponding to the target text may be generated according to the font information corresponding to the target text, then the mask image corresponding to the text image may be determined, and the mask image is processed by the stylization model according to the style prompt information to obtain the stylized text image corresponding to the target text. Through this method, a personalized stylized text image may be generated in real time, and the processing speed and processing quality of the stylized text are improved.

- [0125] Those skilled in the art can understand that, in the methods of the specific implementations, the order in which the steps are written does not imply a strict execution order, and does not constitute any limitation on the implementation process. The specific execution order of the steps should be determined by their functions and possible internal logics.
- [0126] Based on the same inventive concept, embodiments of the present disclosure further provide a text processing apparatus corresponding to the text processing method. Because a problem-solving principle of the apparatus in embodiments of the present disclosure is similar to that of the text processing method in embodiments of the present disclosure, for the implementation of the apparatus, reference may be made to the implementation of the method, and repeated parts are not described herein.
- [0127] Reference is made to Fig. 5, which is a schematic architectural diagram of a text processing apparatus provided by an embodiment of the present disclosure, the apparatus includes: an obtaining module 501, a generation module 502, and a stylization processing module 503.
- [0128] The obtaining module 501 is configured to obtain a target text and determine font information corresponding to the target text.
- [0129] The generation module 502 is configured to generate a text image corresponding to the target text according to the font information.
- [0130] The stylization processing module 503 is configured to determine a mask image corresponding to the target text in the text image, and input the mask image and style prompt information corresponding to the target text into a stylization model and determine a stylized text image corresponding to the target text.
- [0131] In a possible implementation, the obtaining module 501, when obtaining the target text, is configured to:
- [0132] obtain the target text added to an image to be processed; and
- [0133] after determining the stylized text image corresponding to the target text, the stylization processing module 503 is further configured to:
- [0134] fuse and display the stylized text image with the image to be processed.
- [0135] In a possible implementation, the stylization processing module 503 is further configured to determine the style prompt information corresponding to the target text by using the following method:
- [0136] receiving the style prompt information corresponding to the target text entered by a user; or
- [0137] performing style detection on the image to be processed and determining the style prompt information corresponding to the target text based on a style detection result; or
- [0138] performing content detection on the target text and determining the style prompt information corresponding to the text based on a content detection result.
- [0139] In a possible implementation, in response to the style prompt information including a name of a main object, text in the stylized text image corresponding to the target text is text composed based on the main object.

- [0140] In a possible implementation, in response to the style prompt information not including a name of a main object, text in the stylized text image corresponding to the target text is text composed based on an associated object associated with the style prompt information.
- [0141] In a possible implementation, in response to that the target text includes a plurality of texts, different texts correspond to different style prompt information;
- [0142] the stylization processing module 503, when inputting the mask image and the style prompt information corresponding to the target text into the stylization model to determine the stylized text image corresponding to the target text, is configured to:
- [0143] input the mask image and style prompt information corresponding to each target text into the stylization model to determine the stylized text image corresponding to the target text; in the stylized text image, stylized texts corresponding to different target texts having different style types.
- [0144] In a possible implementation, the obtaining module 501, when obtaining the target text and determining the font information corresponding to the target text, is configured to:
- [0145] obtain an image to be processed, perform content recognition on the image to be processed, and determine, based on a content recognition result, the target text matching the image to be processed and the font information corresponding to the target text; and
- [0146] after determining the stylized text image corresponding to the target text, the method further includes:
- [0147] fusing and displaying the stylized text image with the image to be processed.
- [0148] For the description of the processing processes of various modules in the apparatus, and the interaction processes between the modules, reference may be made to the related description in the method embodiments, and details are not described herein again.
- [0149] Based on the same technical concept, an embodiment of the present disclosure further provides a computer device. Reference is made to Fig. 6 which is a schematic structural diagram of a computer device 600 provided by an embodiment of the present disclosure. The computer device in the embodiments of the present disclosure may include, but is not limited to, mobile terminals such as a mobile phone, a notebook computer, a digital broadcasting receiver, a personal digital assistant (PDA), a tablet computer (PAD), a portable multimedia player (PMP), and a vehicle-mounted terminal (e.g., a vehicle-mounted navigation terminal), and fixed terminals such as a digital TV and a desktop computer, or various forms of servers, such as a standalone server or a server cluster. The computer device shown in Fig. 6 is only an example, and should not impose any limitation on the function and scope of use of the embodiments of the present disclosure.
- [0150] As shown in Fig. 6, the computer device 600 may include a processing apparatus (e.g., a central processing unit or a graphics processing unit) 601 that may perform a variety of appropriate actions and processing in accordance with a program stored in a read-only memory (ROM) 602 or a program loaded from a storage apparatus 608 into a random-access memory (RAM) 603. The RAM 603 further stores

various programs and data required for the operation of the computer device 600. The processing apparatus 601, the ROM 602, and the RAM 603 are connected to one another through a bus 604. An input/output (I/O) interface 605 is also connected to the bus 604.

[0151] Generally, the following apparatus may be connected to the I/O interface 605: an input apparatus 606 including, for example, a touchscreen, a touchpad, a keyboard, a mouse, a camera, a microphone, an accelerometer, and a gyroscope; an output apparatus 607 including, for example, a liquid crystal display (LCD), a speaker, and a vibrator; the storage apparatus 608 including, for example, a magnetic tape and a hard disk drive; and a communication apparatus 609. The communication apparatus 609 may allow the computer device 600 to perform wireless or wired communication with other devices to exchange data. Although Fig. 6 shows the computer device 600 having various apparatus, it should be understood that it is not required to implement or have all of the shown apparatus. It may be an alternative to implement or have more or fewer apparatus.

[0152] In particular, according to the embodiments of the present disclosure, the process described above with reference to the flowchart may be implemented as a computer software program. For example, an embodiment of the present disclosure includes a computer program product, which includes a computer program carried on a computer-readable medium, the computer program includes program code for performing the text processing method. In such an embodiment, the computer program may be downloaded and installed from a network through the communication apparatus 609, installed from the storage apparatus 608, or installed from the ROM 602. When the computer program is executed by the processing apparatus 601, the above-mentioned functions defined in the method of the embodiments of the present disclosure are performed.

[0153] An embodiment of the present disclosure further provides a computer-readable storage medium having a computer program stored thereon that, when executed by a processor, causes the steps of the text processing method described in the method embodiments to be performed. The storage medium may be a volatile or non-volatile computer-readable storage medium.

[0154] An embodiment of the present disclosure further provides a computer program product carrying program code, instructions included in the program code can be used to perform the steps of the text processing method described in the above-mentioned method embodiments. For details, reference may be made to the above-mentioned method embodiments, and details are not described herein again.

[0155] The computer program product may be implemented in the form of hardware, software, or a combination thereof. In an optional embodiment, the computer program product is specifically embodied as a computer storage medium. In another optional embodiment, the computer program product is specifically embodied as a software product, such as a software development kit (SDK).

[0156] It can be clearly understood by those skilled in the art that, for convenience and brevity of description, for the specific operation processes of the system and apparatus described above, reference may be made to the corresponding processes in the foregoing method embodiments, and details are not repeated

herein. In the several embodiments provided in the present disclosure, it should be understood that the disclosed system, apparatus, and method may be implemented in other manners. The apparatus embodiment described above is merely an example. For example, the unit division is merely logical function division and there may be other division during actual implementation. For another example, a plurality of units or components may be combined or integrated into another system, or some features may be omitted or not implemented. In addition, the displayed or discussed mutual couplings, direct couplings, or communication connections may be implemented through some communication interfaces. The indirect couplings or communication connections between the apparatuses or units may be implemented in electrical, mechanical, or other forms.

[0157] The units described as separate parts may or may not be physically separate, and parts displayed as units may or may not be physical units, and may be located at one position, or may be distributed on a plurality of network units. Some or all of the units may be selected according to actual needs to achieve the objectives of the solutions of the embodiments.

[0158] In addition, various functional units in the various embodiments of the present disclosure may be integrated into one processing unit, each of the units may exist alone physically, or two or more units may be integrated into one unit.

[0159] If the functions are implemented in the form of a software functional unit and sold or used as an independent product, the functions may be stored in a non-volatile computer-readable storage medium executable by a processor. Based on such an understanding, a part of the technical solutions of the present disclosure that makes a contribution essentially or to the related art, or a part of the technical solutions may be embodied in the form of a software product. The computer software product is stored in a storage medium and includes a plurality of instructions for instructing a computer device (which may be a personal computer, a server, or a network device) to perform all or some of the steps of the methods described in the embodiments of the present disclosure. Moreover, the foregoing storage medium includes a USB flash drive, a removable hard disk, a read-only memory (ROM), a random-access memory (RAM), a magnetic disk, an optical disc, or other various media that can store program code.

[0160] It should be finally noted that the embodiments described above are merely specific implementations of the present disclosure, and are used for illustrating rather than limiting the technical solutions of the present disclosure, and the scope of protection of the present disclosure is not limited thereto. Although the present disclosure has been described in detail with reference to the foregoing embodiments, those of ordinary skill in the art should understand that, within the technical scope disclosed in the present disclosure, any person skilled in the art could still modify the technical solutions described in the foregoing embodiments, or readily figure out any variations thereof, or make equivalent substitutions to some of the technical features thereof. However, these modifications, variations, or substitutions do not make the essence of the corresponding technical solutions depart from the spirit and scope of the technical solutions of the embodiments of the present disclosure, and shall fall within the scope of protection of the present

disclosure. Therefore, the scope of protection of the present disclosure shall be subject to the scope of protection of the claims.

CLAIMS

1. A text processing method, comprising:
obtaining a target text and determining font information corresponding to the target text;
generating a text image corresponding to the target text according to the font information; and
determining a mask image corresponding to the target text in the text image, and inputting the mask image and style prompt information corresponding to the target text into a stylization model and determining a stylized text image corresponding to the target text.
2. The method according to claim 1, wherein the obtaining a target text comprises:
obtaining the target text added to an image to be processed; and
after the determining the stylized text image corresponding to the target text, the method further comprises:
fusing and displaying the stylized text image with the image to be processed.
3. The method according to claim 2, further comprising: determining the style prompt information corresponding to the target text by using the following method:
receiving the style prompt information corresponding to the target text entered by a user; or
performing style detection on the image to be processed and determining the style prompt information corresponding to the target text based on a style detection result; or
performing content detection on the target text and determining the style prompt information corresponding to the text based on a content detection result.
4. The method according to any one of claims 1 to 3, wherein the inputting the mask image and style prompt information corresponding to the target text into a stylization model to determine a stylized text image corresponding to the target text comprises:
inputting the mask image into a control model to determine a constraint feature corresponding to the mask image;
and
inputting the style prompt information, the mask image, and the constraint feature into the stylization model to determine the stylized text image corresponding to the target text.
5. The method according to any one of claims 1 to 4, wherein in response to the style prompt information comprising a name of a main object, text in the stylized text image corresponding to the target text is text composed based on the main object.
6. The method according to claim 5, wherein in response to the style prompt information not comprising the name of the main object, the text in the stylized text image corresponding to the target text is text composed based on an associated object associated with the style prompt information.

7. The method according to any one of claims 1 to 6, wherein in response to the target text comprising a plurality of texts, different texts correspond to different style prompt information; and
the inputting the mask image and style prompt information corresponding to the target text into a stylization model and determining a stylized text image corresponding to the target text comprises:
inputting the mask image and style prompt information corresponding to each target text into the stylization model and determining the stylized text image corresponding to the target text, wherein in the stylized text image, stylized texts corresponding to different target texts have different style types.
8. The method according to claim 1, wherein the obtaining a target text and determining font information corresponding to the target text comprises:
obtaining an image to be processed, performing content recognition on the image to be processed, and
determining, based on a content recognition result, the target text matching the image to be processed and the font information corresponding to the target text; and
after the determining a stylized text image corresponding to the target text, the method further comprises:
fusing and displaying the stylized text image with the image to be processed.
9. A text processing apparatus, comprising:
an obtaining module configured to obtain a target text and determine font information corresponding to the target text;
a generation module configured to generate a text image corresponding to the target text according to the font information; and
a stylization processing module configured to determine a mask image corresponding to the target text in the text image, and input the mask image and style prompt information corresponding to the target text into a stylization model and determine a stylized text image corresponding to the target text.
10. A computer device, comprising: a processor, a memory, and a bus, wherein the memory stores machine-readable instructions executable by the processor, the processor communicates with the memory through the bus when the computer device is running, and the machine-readable instructions, when executed by the processor, cause the steps of the text processing method according to any one of claims 1 to 8 to be performed.
11. A computer-readable storage medium, wherein the computer-readable storage medium stores a computer program that, when run by a processor, causes the steps of the text processing method according to any one of claims 1 to 8 to be performed.