

VOUCHER REVIEW METHOD AND SYSTEM

TECHNICAL FIELD

[0001] This specification relates to the field of artificial intelligence technologies, and in particular, to a voucher review method and system.

5

BACKGROUND

[0002] When conducting international transactions, enterprises/individuals may engage with enterprises/individuals from different countries. During handling of the international transactions, it is often necessary to upload vouchers related to the transactions and review the vouchers. Different transaction scenarios may correspond to different vouchers and review rules. Currently, a manual review is usually performed. However, as a transaction amount increases and a scenario varies, a manual review has a heavy workload and high costs.

[0003] Therefore, it is necessary to provide a method that can be applied to automated voucher review in a plurality of scenarios.

[0004] Content of the background part is merely information learned of by the inventor, and neither means that the information has entered a public domain before an application date of this specification, nor means that the information can become a conventional technology of this specification.

SUMMARY

[0005] This specification provides a voucher review method and system, which can be applied to automated voucher review in a plurality of scenarios.

[0006] According to a first aspect, this specification provides a voucher review method, including: obtaining at least one to-be-processed image voucher in a target scenario; extracting to-be-reviewed information in the at least one to-be-processed image voucher; determining a key review point in the target scenario, where the voucher review system pre-stores review points in a plurality of scenarios, and the plurality of scenarios include the target scenario; and inputting the to-be-reviewed information and the key review point in the target scenario to a large language model to obtain a review result of the large language model for the at least one to-be-processed image voucher.

[0007] According to a second aspect, this specification further provides a voucher review system, including at least one storage medium and at least one processor. The at least one storage medium stores at least one instruction set used to perform a voucher review method. The at least one processor is communicatively connected to the at least one storage medium. When running, the at least one processor reads the at least one instruction set, and performs the following operations based on an instruction of the at least one instruction set: obtaining at least one to-be-processed image voucher in a target scenario; extracting to-be-reviewed information in the at least one to-be-processed image voucher; determining a key review point in the target scenario, where the voucher review system pre-stores review points in a plurality of scenarios, and the plurality of scenarios include the target scenario; and inputting the to-be-reviewed information and the key review point in the target scenario to a large language model to obtain a review result of the large language model for the at least one to-be-processed image voucher.

[0008] According to a third aspect, this specification further provides a computer-readable nonvolatile storage medium. The computer-readable nonvolatile storage medium stores at least one instruction set, and the at least one instruction set is executed by at least one processor to implement the voucher review method according to the first aspect.

[0009] It can be learned from the above-mentioned technical solutions that, according to the voucher review method and the system provided in this specification, after the at least one to-be-processed image voucher in the target scenario is obtained, the to-be-reviewed information in the at least one to-be-processed image voucher can be extracted, and then based on the key review point that is in the target scenario and that is pre-stored in the voucher review system, the to-be-reviewed information and the key review point are input to the large language model, and the large language model reviews the to-be-reviewed information based on the key review point. According to the voucher review method provided in this specification, review points in different scenarios are pre-configured, and vouchers in the different scenarios are automatically reviewed based on the key review points by using a common capability and a powerful context understanding capability of the large language model. The voucher review method provided in this specification can be applied to automated voucher review in a plurality of scenarios, thereby improving voucher review efficiency, reducing manpower costs, avoiding error-prone problems in a manual review, and improving user experience.

[0010] Other functions of the voucher review method and system provided in this specification are partially listed in the following descriptions. Creative aspects of the voucher review method and system provided in this specification can be fully explained by practice or using the methods, apparatuses, and combinations described in the following detailed examples.

BRIEF DESCRIPTION OF DRAWINGS

[0011] To describe the technical solutions in the embodiments of this specification more clearly, the following briefly describes the accompanying drawings needed for describing the embodiments. Clearly, the accompanying drawings in the following description show merely some
5 embodiments of this specification, and a person of ordinary skill in the art can still derive other drawings from these accompanying drawings without creative efforts.

[0012] FIG. 1 is a schematic diagram of an application scenario of a voucher review system according to an embodiment of this specification;

[0013] FIG. 2 is a schematic diagram of a hardware structure of a computing device according
10 to some embodiments of this specification;

[0014] FIG. 3 is a flowchart of a voucher review method according to an embodiment of this specification;

[0015] FIG. 4 is a flowchart of a step of extracting to-be-reviewed information in a voucher review method according to an embodiment of this specification;

[0016] FIG. 5 is a flowchart of a voucher review method according to another embodiment of
15 this specification; and

[0017] FIG. 6 is an example flowchart of a voucher recognition method according to some embodiments of this specification.

DESCRIPTION OF EMBODIMENTS

[0018] Specific application scenarios and requirements of this specification are provided in
20 the following descriptions, so that a person skilled in the art can manufacture and use content of this specification. For the person skilled in the art, various local modifications to the disclosed embodiments are apparent, and the general principles defined herein can be applied to other embodiments and applications without departing from the spirit and scope of this specification.
25 Therefore, this specification is not limited to the shown embodiments, but has a widest scope consistent with that of the claims.

[0019] The terms used herein are only intended to describe a particular example embodiment, but do not impose a limitation. For example, unless otherwise explicitly stated in the context, the singular forms "one", "an", and "the" used herein can include plural forms. When used in this
30 specification, the terms "include", "comprise", and/or "have" indicate/indicates existence of an associated integer, step, operation, element, and/or component, but does not exclude existence of one or more other features, integers, steps, operations, elements, components, and/or groups, or another feature, integer, step, operation, element, component, and/or group can be added to the

system/method.

[0020] In consideration of the following descriptions, these and other features of this specification, operations and functions of related components of the structure, and economy of combination and manufacturing of components can be significantly improved. With reference to the accompanying drawings, all of these form a part of this specification. However, it should be clearly understood that the accompanying drawings are merely used for the purpose of illustration and description, and are not intended to limit the scope of this specification. It should be further understood that the accompanying drawings are not drawn to scale.

[0021] The flowchart used in this specification illustrates operations implemented by a system according to some embodiments of this specification. It should be clearly understood that operations in the flowchart can be implemented out of sequence. On the contrary, operations can be implemented reversely or simultaneously. In addition, one or more other operations can be added to the flowchart. One or more operations can be removed from the flowchart.

[0022] For ease of description, the terms that appear in the following descriptions of this specification are first explained.

[0023] A large language model is usually used in the artificial intelligence field of natural language processing (NLP), and particularly refers to a large machine learning model with a large quantity of parameters and computing resources. The large language model is named for having a huge quantity of parameters and a complex network structure, and has a powerful feature representation capability and feature understanding capability, and can better capture patterns and rules in data when processing complex tasks. Large language models are designed and trained to better understand and generate a natural language.

[0024] Optical character recognition (OCR) refers to a process of performing analysis and recognition processing on an image file of a text material to obtain text and layout information. In this process, various types of images such as handwritten and printed images are converted into machine-coded text. The core of the OCR technology is to recognize text information in an image and convert the text information into a text format that can be edited and processed by a computer.

[0025] In a voucher review scenario, a to-be-reviewed voucher is usually reviewed manually to obtain a review result for the to-be-reviewed voucher. However, with an increasing quantity of review scenarios, a quantity of online reviews increases, a quantity of manual reviews continues to increase, and manpower is nearly saturated. In addition, a review error is prone to occur in a manual review due to negligence of personnel, resulting in low review efficiency. Moreover, a magnitude of a voucher review varies greatly in a peak period or an off-peak period of transaction processing. For example, a quantity of voucher reviews in the peak period can be twice or more of that in the off-peak period. Increasing a quantity of review personnel in the peak period may lead

to a waste of human resources in the off-peak period and a shortage of human resources in the peak period. In other words, a problem of a review error or low review efficiency is prone to occur in a manner of manually performing a voucher review a full scale, and the above-mentioned problem cannot be resolved by simply increasing a quantity of review personnel.

5 **[0026]** In some application scenarios, for example, in a study-abroad scenario, a related art further provides an automated review link implemented by hard code to review a study-abroad voucher via the automated review link. In this review manner, a review rule needs to be set for each review point to set up a review link. Currently, there are many scenarios that need to be reviewed, such as a remittance scenario, a study-abroad scenario, a salary transfer-in scenario, a trade scenario, and a residence-related scenario (for example, renting a house, buying a house, or applying for a house). A relatively large quantity of types of vouchers are involved in different scenarios, and the types of vouchers involved in the different scenarios may be inconsistent. One type of voucher may involve a plurality of review points. Review points of different types of vouchers may also be different. For voucher reviews in a plurality of scenarios, a review rule needs to be set for each review point of each voucher in each scenario. As a result, code development costs are high, a development period is long, maintenance costs are relatively high, a cold start effect is poor, and scalability or a learning upgrade capability is not strong, that is, reusability is not strong. Consequently, voucher review requirements in the plurality of scenarios cannot be satisfied. Therefore, the manner of setting up the review link by setting the review rule for each review point of each voucher can be applied to only a single scenario, and cannot be applied to a plurality of scenarios.

15 **[0027]** The voucher review method provided in this specification can be applied to a voucher review system. The voucher review system pre-stores review points of a plurality of scenarios. The voucher review system can perform an automated review based on to-be-reviewed information in a to-be-reviewed voucher in a target scenario by using a large language model. According to the voucher review method provided in this specification, the to-be-reviewed information is automatically reviewed by invoking a general capability and a powerful context understanding capability of the large language model. The voucher review method provided in this specification can be applied to voucher reviews in a plurality of scenarios, and code development does not need to be performed for each review rule of each review point, which can greatly reduce a code development workload and shorten a development period.

25 **[0028]** The voucher review method and system provided in this specification have a wide range of application scenarios. For example, the voucher review method and system can be applied to voucher review scenarios of various cross-border transactions, for example, a transaction in which a foreign enterprise engages with a domestic enterprise/individual (such as a salary

transaction, a labor transfer transaction, a payment transaction, or a remittance transaction), or a transaction in which a foreign individual engages with a domestic individual (such as a family support transaction or a remittance transaction). Certainly, the voucher review method and system can also be applied to a transaction scenario in which a domestic enterprise/individual engages with a foreign enterprise/individual, or can be applied to any transaction scenario in which voucher review is required, for example, a scenario in which a plurality of vouchers for house purchase qualification are reviewed, a scenario in which a plurality of vouchers for loan qualification are reviewed, or a scenario in which a plurality of vouchers for admission qualification are reviewed. It should be understood that the above-mentioned embodiment is merely an example for description. A specific transaction scenario to which a voucher review is applied can be flexibly adjusted based on a user requirement, and is not limited to that provided in the above-mentioned embodiment.

[0029] It should be noted that the above-mentioned related application scenario is merely one of a plurality of use scenarios provided in this specification. A person skilled in the art should understand that, when the voucher review method and system provided in this specification are applied to another use scenario, an implementation and technical effects thereof are similar.

[0030] FIG. 1 is a schematic diagram of an application scenario 100 of a voucher review system 130 according to an embodiment of this specification. The voucher review method provided in this specification can be applied to the application scenario 100 shown in FIG. 1. As shown in FIG. 1, the application scenario 100 can include a target user 110, a terminal 120, and the voucher review system 130 (referred to as a system 130 for short below).

[0031] The user 110 can be any user 110 who inputs a to-be-processed image voucher to the system 130, and expects to obtain a review result for the to-be-processed image voucher by using the system 130. The user 110 can submit a review request in a related scenario by using the terminal 120. The review request can include the to-be-processed image voucher. The terminal 120 can send the obtained to-be-processed voucher to the system 130. The system 130 can review the to-be-processed image voucher based on the to-be-processed image voucher and a key review point in a target scenario, and outputs a corresponding review result. For example, the user 110 can be a user who needs to pay fees for studying abroad, can be an overseas user who needs to send remittances to support domestic family members, or can be an overseas user who needs to send remittances to domestic employees. These users 110 need to upload to-be-processed image vouchers of target transactions in target scenarios, and obtain review results output by the system 130.

[0032] The terminal 120 can be a device interacting with the user 110. The user 110 can input various image vouchers to the terminal 120 by using an input/output device of the terminal 120,

or can input various filled information to the terminal 120, or can receive review results for the image vouchers from the terminal 120. In some embodiments, the terminal 120 can include a hardware device with a data information processing function and a necessary application program needed for driving the hardware device to work. The application program can provide a capability and an interface for interacting with the outside through a network for the user 110. For example, the terminal 120 can have an application program that provides the above-mentioned transaction processing, for example, an application program that provides a payment service for studying abroad, or an application program that provides an overseas remittance. The terminal 120 can receive, by using the interaction interface of the application program, an image voucher input by the user 110, and output a review result for the image voucher to the user 110. For example, the terminal 120 can receive an image voucher input by the user 110 for a family support transaction application, and output a review result for the image voucher to the user 110. For example, the terminal 120 can receive an image voucher input by the user 110 for a common transaction application, and output a review result for the image voucher to the user 110.

[0033] In some embodiments, the terminal 120 can include a mobile device, a tablet computer, a notebook computer, a built-in device in a motor vehicle or similar content, or any combination thereof. In some embodiments, the mobile device can include a smart home device, a smart mobile device, a virtual reality device, an augmented reality device or a similar device, or any combination thereof. In some embodiments, the smart home apparatus can include a smart television, a desktop computer, or any combination. In some embodiments, the smart mobile device can include a smartphone, a personal digital assistance, a game device, a navigation device, etc., or any combination thereof. In some embodiments, the built-in apparatus in the motor vehicle can include a vehicle-mounted computer, a vehicle-mounted television, etc.

[0034] The system 130 can be a background server that provides a service to the terminal 120. For example, the system 130 can be a background server for voucher review. The system 130 can be communicatively connected to the terminal 120. The system 130 can store data or instructions for performing the voucher review method described in this specification, and can execute or be configured to execute the data or the instructions. For example, when the user 110 submits a to-be-processed image voucher and a voucher review request by using the terminal 120, the system 130 receives the to-be-processed image voucher and responds to the request to perform the voucher review method provided in this specification. In some embodiments, a large language model can be deployed in the system 130. In some embodiments, the system 130 can invoke the large language model. The large language model can output a review result for a to-be-processed image voucher based on the to-be-processed image voucher.

[0035] The system 130 can be a single computing device, or can be a cluster system including

a plurality of computing devices. The data or the instructions that are stored in the system 130 and that are used to perform the voucher review method described in this specification can be in any form of system architecture, for example, a layered architecture, an event-driven architecture, a micro kernel architecture, a microservice architecture, or a cloud architecture.

5 **[0036]** It should be understood that the application scenario shown in FIG. 1 is merely one of a plurality of application scenarios provided in this specification. In some embodiments, the terminal 120 may not exist in the application scenario 100. For example, the user 110 can directly operate the system 130.

10 **[0037]** It should be further understood that quantities of users 110, terminals 120, and systems 130 shown in FIG. 1 are merely examples. There can be any quantity based on an implementation requirement.

15 **[0038]** The system 130 provided in this specification can output a review result based on the large language model. The large language model can review an input to-be-processed image voucher for the to-be-processed image voucher and review points that are in a plurality of scenarios and that are pre-stored in the system 130, and output a review result for the to-be-processed image voucher. Content of the review result output by the system 130 provided in this specification is complete, accurate, and smooth. The voucher review method and system provided in this specification can be applied to a wide range of application scenarios in various industries and various tasks.

20 **[0039]** FIG. 2 is a schematic diagram of a hardware structure of a computing device 200 according to some embodiments of this specification. As shown in FIG. 2, the computing device 200 includes at least one storage medium 230 and at least one processor 220. In some embodiments, the computing device 200 can further include an internal communication bus 210. In some embodiments, the computing device 200 can further include a communication port 250. In some
25 embodiments, the computing device 200 can further include an I/O component 260.

[0040] The internal communication bus 210 can be connected to different system components, including the storage medium 230 and the processor 220. The I/O component 260 supports input/output between the computing device 200 and another component.

30 **[0041]** The storage medium 230 can include a data storage apparatus. The data storage apparatus can be a non-transitory storage medium, or can be a transitory storage medium. For example, the data storage apparatus can include one or more of a disk 232, a read-only storage medium (ROM) 234, or a random access storage medium (RAM) 236. The storage medium 230 further includes at least one instruction set stored in the data storage apparatus. The instruction set is computer program code, and the computer program code can include a program, a routine, an
35 object, a component, a data structure, a process, a module, etc. for performing the voucher review

method provided in this specification.

[0042] The communication port 250 is used for data communication between the computing device 200 and the outside. For example, the computing device 200 can be connected to a network through the communication port 250. The at least one processor 220 is communicatively connected to the at least one storage medium 230 through the internal communication bus 210. The at least one processor 220 is configured to execute the at least one instruction set. When the system 130 runs, the at least one processor 220 reads the at least one instruction set, and performs, based on an instruction of the at least one instruction set, the voucher review method provided in this specification.

[0043] The processor 220 can perform all steps included in the voucher review method. The processor 220 can be in a form of one or more processors. The processor 220 can send an execution instruction. The processor 220 can include one or more hardware processors, for example, a microcontroller, a microprocessor, a reduced instruction set computer (RISC), an application-specific integrated circuit (ASIC), an application-specific instruction set processor (ASIP), a central processing unit (CPU), a graphics processing unit (GPU), a physical processing unit (PPU), a microcontroller unit, a digital signal processor (DSP), a field programmable gate array (FPGA), an advanced RISC machine (ARM), a programmable logic device (PLD), any circuit or processor that can perform one or more functions, or any combination thereof.

[0044] For only the purpose of description, only one processor 220 is described in the computing device 200 in this specification. However, it should be noted that the computing device 200 in this specification can further include a plurality of processors. Therefore, the operations and/or method steps disclosed in this specification can be performed by one processor or can be performed jointly by a plurality of processors, as described in this specification. For example, if the processor 220 in the computing device 200 in this specification performs step A and step B, it should be understood as that step A and step B can be jointly or separately performed by two different processors 220 (for example, a first processor performs step A, and a second processor performs step B, or a first processor and a second processor jointly perform steps A and B).

[0045] FIG. 3 is a flowchart of a voucher review method according to an embodiment of this specification. The voucher review method P100 is applied to the voucher review system 130. As described above, the user 110 can submit a to-be-processed image voucher and a voucher review request by using the terminal 120. The system 130 can receive the to-be-processed image voucher and the voucher review request, and perform the voucher review method P100 provided in this specification in response to the voucher review request. As described above, the computing device 200 can perform the voucher review method P100 in this specification. Specifically, the processor 220 in the computing device 200 can read an instruction set stored in a local storage medium of

the computing device, and then perform the voucher review method P100 in this specification based on a stipulation of the instruction set. As shown in FIG. 3, the method P100 provided in this specification can include the following steps.

[0046] S110: Obtain at least one to-be-processed image voucher in a target scenario.

5 **[0047]** The system 130 can receive at least one to-be-processed image voucher input by the user 110. For example, a manner of inputting the at least one to-be-processed image voucher by the user 110 can be as follows: After logging in to an application program for performing voucher review on the terminal 120, the user 110 enters the target scenario through a preset entry of the application program, and then enters a voucher uploading interface through a voucher uploading
10 entry in the target scenario, and uploads the at least one to-be-processed image voucher through the voucher uploading interface. In other words, the user uploads the at least one to-be-processed image voucher online. For example, a manner of uploading the at least one to-be-processed image voucher by the user 110 can be as follows: The user 110 selects at least one pre-captured image from an album of the terminal 120 as a to-be-processed image voucher to be uploaded; or the user
15 110 grants a camera to capture at least one to-be-processed image voucher by using the camera, and uploads the at least one to-be-processed image voucher; or the user 110 can select one or more to-be-processed image vouchers from at least one to-be-processed image voucher uploaded in history as a to-be-processed image voucher to be uploaded this time. It should be understood that the above-mentioned embodiment is merely an example for description. A specific manner of
20 uploading a to-be-processed image voucher can be flexibly adjusted based on a user requirement, and is not limited herein in this specification.

[0048] As described above, the system 130 can perform voucher review in a plurality of scenarios. The plurality of scenarios can be the scenarios listed above, such as a remittance scenario, a family support scenario, a scenario of paying fees for studying abroad, a salary transfer-
25 in scenario, a house purchase scenario, and a trade scenario. The target scenario can be any scenario in the plurality of scenarios. The to-be-processed image voucher can be any type of image voucher that needs to be reviewed in the target scenario. In a same scenario, one or more types of vouchers may need to be submitted. For example, a type of voucher that needs to be submitted in the family support scenario can include an identity voucher, a relationship voucher, etc.; a type of
30 voucher that needs to be submitted in the salary transfer-in scenario can include a labor service certificate, an account entry certificate, etc.; and a type of voucher that needs to be submitted in the scenario of paying fees for studying abroad includes an identity certificate for studying abroad, a payment certificate, an identity certificate, etc. A type of voucher can be in a plurality of different forms. For example, the identity certificate can be in various forms, such as an identity card, a
35 passport, an XXX pass, etc. For another example, the relationship certificate can include a

marriage certificate, an identity verification document, etc. For another example, the labor service certificate can include an employment contract, an employment certificate, etc. For another example, the account entry certificate can include a payroll, a bank flow, etc. For another example, the identity certificate for studying abroad can include an admission letter, a student ID card, etc.

For example, in a family support transaction scenario, the to-be-processed image voucher may be a plurality of image vouchers of an identity verification document of the user 110 or a marriage certificate voucher of the user 110. Alternatively, in a study-abroad transaction scenario, the to-be-processed image voucher may be a plurality of enrollment certificate vouchers of the user 110, an identity certificate of the user 110, an expense list voucher of a corresponding school where the user 110 studies abroad, etc. The to-be-processed image voucher can alternatively be another image voucher.

[0049] As shown in FIG. 3, the method P100 further includes the following step.

[0050] S130: Extract to-be-reviewed information in the at least one to-be-processed image voucher.

[0051] The to-be-reviewed information can be understood as one or more pieces of key information that need to be reviewed in each to-be-processed image voucher, for example, identity information (such as a name, an age, and an identity number) in an identity certificate, identity information of both parties in a relationship certificate, and an employment contract. The to-be-reviewed information is used as a basis for review in a subsequent review process, and how to efficiently and correctly extract the to-be-reviewed information in each to-be-processed image voucher is particularly important.

[0052] FIG. 4 is a flowchart of a step of extracting to-be-reviewed information in a voucher review method according to an embodiment of this specification. As shown in FIG. 4, in step S130, the to-be-reviewed information in each of the at least one to-be-processed image voucher needs to be extracted. To extract at least one piece of to-be-reviewed information, steps S131 to S135 need to be performed on a to-be-processed image voucher. A to-be-processed image voucher can be any one of the at least one to-be-processed image voucher. Extracting the to-be-reviewed information in each to-be-processed image voucher can be performing steps S131 to S135 on the to-be-processed image voucher.

[0053] S131: Perform character recognition on the to-be-processed image voucher.

[0054] Character recognition can be converting, based on the above-mentioned OCR technology, text in the to-be-processed image voucher into a digital text format that can be edited and processed by a computer. In some possible embodiments, performing character recognition on the to-be-processed image voucher can be directly performing the OCR technology on the to-be-processed image to obtain the text in the to-be-processed image voucher.

[0055] In some possible embodiments, performing character recognition on the to-be-processed image voucher data can be as follows: The user 110 selects the to-be-reviewed information in the to-be-processed image voucher, and then character recognition is performed on a selected area to obtain text in the selected area. For example, performing character recognition on the to-be-processed image voucher data can include: determining, based on a selection operation performed by the user on the to-be-reviewed information in the to-be-processed image voucher, a character recognition area corresponding to the to-be-reviewed information in the to-be-processed image voucher; and performing character recognition on the character recognition area to obtain a character recognition result. To be specific, the system 130 sends a selection indication corresponding to the to-be-reviewed information to the terminal 120; the terminal 120 displays the selection indication to the user 110; and the user 110 selects a corresponding position in the to-be-processed image voucher based on the selection indication, to select a position of the to-be-reviewed information in the to-be-processed image voucher. When one to-be-processed image voucher includes a plurality of pieces of to-be-reviewed information, the system 130 can sequentially generate corresponding selection indications for all pieces of to-be-reviewed information, and prompt, by using the terminal 120, the user 110 to sequentially select different to-be-reviewed information in the to-be-processed image voucher based on the selection indications. The system 130 can alternatively simultaneously display, by using the terminal 120, selection indications for a plurality of pieces of to-be-reviewed information to the user 110, and prompt the user 110 to select the plurality of pieces of to-be-reviewed information in the to-be-processed image voucher based on the selection indications. The system 130 can alternatively simultaneously display, by using the terminal 120, selection indications for a plurality of pieces of to-be-reviewed information to the user 110, display an arrangement relationship between the plurality of pieces of to-be-reviewed information, and prompt the user 110 to sequentially select the plurality of pieces of to-be-reviewed information in the to-be-processed image voucher based on the arrangement relationship. It should be understood that the above-mentioned embodiment is merely an example for description. A specific manner of selecting the character recognition area can be flexibly adjusted based on a user requirement, and is not limited to that provided in the above-mentioned embodiment.

[0056] For example, in an image voucher of an identity card, the system 130 sends a selection indication for a "name" to the terminal 120; the terminal 120 displays the selection indication for the "name" to the user 110, where the selection indication for the "name" displayed by the terminal 120 can be, for example, "please select a name area in the image"; the user 110 performs a selection operation at a position corresponding to the "name" in the image voucher of the identity card; and the system 130 uses an area selected by the user 110 as a character recognition area of the "name",

performs character recognition on the character recognition area of the "name", and uses a character recognition result as a character recognition result for the "name". For example, in an image voucher of an admission letter, the system 130 sends a selection instruction for a "school name" to the terminal 120; the terminal 120 displays the selection indication for the "school name" to the user 110, where the selection indication for the "school name" displayed by the terminal 120 can be, for example, "please select a school name area in the image"; the user 110 performs a selection operation at a position corresponding to the "school name" in the image voucher of the admission letter; and the system 130 uses an area selected by the user 110 as a character recognition area of the "school name", performs character recognition on the character recognition area of the "school name", and uses a character recognition result as a character recognition result for the "school name". The system 130 sends a selection instruction for a "payment year" to the terminal 120; the terminal 120 displays the selection indication for the "payment year" to the user 110, where the selection indication for the "payment year" displayed by the terminal 120 can be, for example, "please select a payment year area in the image"; the user 110 performs a selection operation at a position corresponding to the "payment year" in the image voucher of the admission letter; and the system 130 uses as an area selected by the user 110 as a character recognition area of the "payment year", performs character recognition on the character recognition area of the "payment year", and uses a character recognition result as a character recognition result for the "payment year". It should be understood that the above-mentioned embodiment is merely an example for description, and indication content displayed when each selection indication is specifically displayed can be flexibly adjusted based on a preset scenario requirement, and is not limited to that provided in the above-mentioned embodiment.

[0057] The system 130 sends a selection indication to the terminal 120, and the terminal 120 displays, in a graphical user interface of the terminal 120, the selection indication corresponding to a selection instruction to prompt the user 110 to select the to-be-reviewed information, to reduce a recognition range and a recognition difficulty of the OCR, improve recognition precision, and reduce a recognition range of a subsequent review model, thereby improving subsequent review precision.

[0058] In some possible embodiments, the system 130 can combine the manner of directly performing character recognition on the to-be-processed image voucher with the manner of performing character recognition through selection by the user 110. In an initial recognition process, the OCR technology can be directly performed on the to-be-processed image voucher to perform character recognition, to obtain the text in the to-be-processed image voucher. When the initial recognition fails (for example, character information in the to-be-processed image voucher cannot be recognized, or recognized character information is error, or enough character

information is not recognized), the manner of performing automatic selection by the user 110 is triggered. In the to-be-processed image voucher, the user can select the character recognition area in the to-be-processed image voucher based on prompt information. In this combined manner, the user 110 can be prevented from being disturbed, thereby improving user experience while improving recognition accuracy.

[0059] In some possible embodiments, step S131 may be as follows: The user 110 selects the character recognition area, while character recognition is performed on the character recognition area selected by the user 110; or after the user 110 selects all character recognition areas, character recognition is performed on all the character recognition areas selected by the user 110. A specific recognition time of character recognition can be flexibly adjusted based on a user requirement, and is not limited to that provided in the above-mentioned embodiment.

[0060] S133: Determine, based on a character recognition result, a document type corresponding to the to-be-processed image voucher.

[0061] Step S133 can be performed in a plurality of different manners. For example, the step is performed in a manner of setting a classification rule for the character recognition result. The manner of setting the classification rule for the character recognition result can be determining a result related to a scenario from the character recognition result, and determining the document type based on a predetermined determining mechanism. For another example, the step is performed in a machine learning manner. The machine learning manner can be, for example, classifying the character recognition result based on a pre-trained classification model to obtain the document type corresponding to the to-be-processed image voucher. During training of the classification model, a character recognition result for a sample voucher can be used as a training sample, and a document type corresponding to the sample voucher can be used as a training label, to train the classification model based on the training sample and the training label.

[0062] To ensure accuracy of recognizing to-be-processed image vouchers in a plurality of scenarios, the system 130 can periodically update the classification model. For example, for a new scenario and/or a new voucher, the system 130 can obtain, by marking a keyword, a voucher in the new scenario and/or classification-related key information in the new voucher, and use the key information as training data to update and train the classification model. This setting manner can resolve a problem of iteration and cold start of a new scenario (for example, a keyword such as a small sample, a minority language, or a complex semantic meaning).

[0063] For example, a classification manner can be recognizing and classifying to-be-processed image vouchers one by one. To be specific, if a quantity of to-be-processed image vouchers is N, a quantity of times of classifying the to-be-processed image vouchers is also N. Each to-be-processed image voucher has a document type corresponding to the to-be-processed

image voucher, and document types of the to-be-processed image vouchers may be the same or different.

[0064] In this embodiment of this specification, the voucher review method provided in this specification can be applied to a plurality of domestic/overseas scenarios, and therefore, languages of the to-be-processed image voucher in different scenarios may be different. In a character recognition process, for example, the to-be-processed image voucher in different languages can be first translated into a target language. For example, the target language can be Chinese, or certainly can be English, Japanese, or another target language. Subsequently, character recognition is performed based on a translated to-be-processed image voucher.

[0065] In some possible embodiments, for example, document types corresponding to different vouchers can include an identity type. In some possible embodiments, document types corresponding to different vouchers can further include a fund type. In some other possible embodiments, for example, the document type can further include a contract type. The identity type can be a voucher that can indicate an identity of the user 110 or a relationship with a third party, such as an identity card, a passport, a marriage certificate, an identity verification document, or a pass. The fund type can be a fund-related voucher that can indicate a salary, a tax, or payment of the user 110, such as a payroll or a payment bill. The contract type can be a voucher that can indicate a trade relationship between the user 110 and a third party, such as a procurement contract. It should be understood that the above-mentioned embodiment is merely an example for description. A quantity of types included in a specific document type and specific type content can be flexibly adjusted based on a scenario requirement, and are not limited to those provided in the above-mentioned embodiment. When a scenario is updated, for example, a corresponding classification model can be trained for the updated scenario. A quantity of types included in a specific document type and specific type content can be flexibly adjusted based on a user requirement, and are not limited to those provided in the above-mentioned embodiment.

[0066] S135: Perform information extraction on the character recognition result based on the document type corresponding to the to-be-processed image voucher, to obtain to-be-reviewed information in the to-be-processed image voucher.

[0067] Step S135 can be performed in a plurality of different manners. For example, the step is performed in a manner of performing information extraction on the character recognition result by using a preset information extraction policy. For another example, the step is performed in a machine learning manner. The machine learning manner can be, for example, performing information extraction on the character recognition result based on an information extraction model corresponding to the document type, to obtain the to-be-reviewed information in the to-be-processed image voucher.

[0068] Different document types correspond to different information extraction models. Each information extraction model is obtained by training a large quantity of samples of a same document type. For example, an identity-type document corresponds to an identity-type information extraction model. The identity-type information extraction model focuses only on identity information (such as name information, certificate code, and certificate validity period information) and relationship information (such as a relationship between two parties) in an identity-type voucher, and does not recognize and extract other content unrelated to the identity information, such as date content or fund content. For example, after the identity information in the identity-type voucher is extracted, the following identity information and relationship information can be obtained: "name: Zhang San", "physiological identifier: male", "date of birth: month xx day xx, year xxxx", "marriage date: month xx day xx, year xxxx", "address: No. 123, Street C, District B, City A", "certificate number: 12345678", "validity period: from month xx day xx, year xxxx to month xx day xx, year xxxx", "spouse's name: Li Si", "physiological identifier: female", "date of birth: month xx day xx, year xxxx", "certificate number 87654321", and "validity period: from month xx day xx, year xxxx to month xx day xx, year xxxx".

[0069] A fund-type document corresponds to a fund-type information extraction model. The fund-type information extraction model focuses only on fund information (such as a salary, a tax, a transfer-in/transfer-out amount, transfer-in/transfer-out time, or transfer-in/transfer-out account information) in a fund-type voucher, and does not recognize and extract other content unrelated to the fund information, such as identity content or personal information content of a user. For example, after the fund information in the fund-type voucher is extracted, the following fund information can be obtained: "transfer-in time: month xx day xx, year xxxx", "transfer-in amount: 123456", and "transfer-in account: 88888888".

[0070] A contract-type document corresponds to a contract-type information extraction model. The contract-type information extraction model focuses only on contract information (a contract number, a contract participant, a contract date, contract content, a contract title, etc.) in a contract-type voucher, and does not recognize and extract other content unrelated to the contract information. For example, after the contract information in the contract-type voucher is extracted, the following contract information can be obtained: "contract number: 666666", "contract party A: Zhang San", "contract party B: Li Si", and "contract date: month xx day xx, year xxxx".

[0071] Using a remittance as an example, a remittance voucher review can include three different scenario classifications: a study-abroad scenario, a transfer-in scenario, and a transfer-out scenario. A plurality of sub-type scenarios exist in each large scenario classification. Risk conditions, payment/remittance purposes, and types of to-be-reviewed vouchers may be different in different classifications. For example, the transfer-in scenario may include a family support sub-

scenario, a salary sub-scenario, a goods trade sub-scenario, a clothing trade sub-scenario, a house renting scenario, etc.; the study-abroad scenario may include a sub-scenario of paying fees for studying abroad, a scholarship transfer-in/transfer-out sub-scenario, etc.; and the transfer-out scenario may include a family support sub-scenario, a salary sub-scenario, a goods trade sub-scenario, a clothing trade sub-scenario, a house renting scenario, etc. Therefore, in this embodiment of this specification, a common design manner is used to analyze and review a remittance voucher. Information content of the to-be-reviewed information in different target scenarios may be different. In this extraction manner, to-be-processed image vouchers of different document types all have corresponding information extraction models, which improves accuracy of recognizing the to-be-reviewed information in the to-be-processed image voucher, and enhances a complex capability of determining the to-be-reviewed information corresponding to each to-be-processed image voucher in a multi-voucher recognition scenario.

[0072] S150: Determine a key review point in the target scenario.

[0073] The voucher review system pre-stores review points in a plurality of scenarios, where the plurality of scenarios include the target scenario. A key review point corresponding to each scenario is pre-configured. A configuration basis may be pre-configuring the key review point based on expert experience, or may be integrating historical review points to obtain the key review point, or may be obtaining the key review point through arrangement based on a newly added review rule in each scenario. A specific manner of determining the key review point can be flexibly adjusted based on a user requirement, and is not limited to that provided in the above-mentioned embodiment.

[0074] In some possible embodiments of this specification, for example, when the voucher review system is applied to a plurality of scenarios, different review points can be first pre-configured for different scenarios. In some embodiments, the system 130 pre-stores determining conditions (for example, scenario conditions, voucher type conditions, or voucher quantity conditions) corresponding to different review points. The system 130 can determine a condition satisfied by the target scenario and the at least one to-be-processed image voucher (for ease of description, the condition satisfied by the target scenario and the at least one to-be-processed image voucher is defined as a target condition), and use a key review point corresponding to the target condition as the key review point of the target scenario and the at least one to-be-processed image voucher. In this way, when voucher recognition is subsequently performed on different scenarios, the key review point corresponding to the target scenario can be first determined based on the target scenario, and then a large language model recognizes the to-be-processed image voucher based on the key review point corresponding to the target scenario, thereby helping the large language model to spread in a plurality of scenarios more quickly.

[0075] For example, in a family support sub-scenario in a transfer-in scenario, after obtaining domestic and overseas marriage certificates, domestic and overseas identity verification documents, domestic and overseas notarial certificates, or domestic and overseas birth certificates that are in the family support sub-scenario, the system 130 can determine that review points corresponding to the family support scenario and the to-be-processed image voucher are: determining whether names of both parties on the above-mentioned certificates are consistent with names of a receiver/remitter, determining whether a relationship between a remitter and a receiver is direct relatives or close relatives, and determining validity of a certificate uploaded by a user that sends/receives a remittance. For example, in a salary sub-scenario in a transfer-in scenario, after obtaining a payroll, an employment contract, an employment certificate, a local personal flow bill, and an overseas bank flow, the system 130 can determine that review points corresponding to the salary scenario and the to-be-processed image voucher are: determining consistency between a name of a holder of an account entry certificate and a name of a remitter, determining whether a validity period or an issue date of the account entry certificate is within a validity period, determining whether an amount of the account entry certificate is greater than a remittance amount, determining whether bank flow time is within valid time, determining whether an amount of the account entry certificate is greater than a remittance amount, and determining whether a salary-related flow or a company name-related flow exists in a bank flow. It should be understood that the above-mentioned embodiment is merely an example for description. Specific to-be-reviewed content included in a key review point of each target scenario and a manner of determining the to-be-reviewed content can be flexibly adjusted based on a user requirement, and are not limited to those provided in the above-mentioned embodiment.

[0076] S170: Input the to-be-reviewed information and the key review point in the target scenario to a large language model to obtain a review result of the large language model for the at least one to-be-processed image voucher.

[0077] In some embodiments, the large language model includes a generative large language model. The generative large language model can automatically generate text associated with an input statement based on the input statement. The generative language model can include an autoencoding model or an autoregressive model. In some embodiments, the autoencoding model can be a series of models based on a Transformer (a language model using an attention mechanism) technology, a series of models based on a bidirectional encoder representations from transformers (BERT) technology or another similar model, or any combination thereof. In some embodiments, the autoencoding model can be a BERT, a robustly optimized BERT pre-training approach (RoBERTa), a SpanBERT (an approach for improving pre-training by representing and predicting spans) or another similar model, or any combination thereof. In some embodiments, the

autoregressive model can be a series of models based on a generative pre-trained transformer (GPT) technology, a series of models based on a ChatGLM (a universal language model based on a bidirectional attention mechanism) technology, a series of models based on an LLaMa (a universal language model using more tokens and smaller parameter sizes) technology or another similar model, or any combination thereof. In some embodiments, a large model based on the GPT technology can be ChatGPT3.5, ChatGPT4.0, etc.

[0078] In some embodiments, when reviews of a plurality of review points in the target scenario all succeed, it is determined that the review result of the large language model for the at least one to-be-processed image voucher is "review succeeds"; or if a review of at least one to-be-processed image voucher in the plurality of review points fails, it is determined that the review result of the large language model for the at least one to-be-processed image voucher is "review fails". In other words, when all review points indicate that the review succeeds, the large language model determines that reviews of all to-be-processed image vouchers submitted by the user succeed. Otherwise, it is determined that the reviews of to-be-processed image vouchers submitted by the user fail. In some embodiments, when determining that a review of at least one to-be-processed image voucher fails, the large language model further outputs the to-be-processed image voucher fails in the review and a reason for the review failure.

[0079] In some embodiments, the system 130 can use the review result of the large language model for the at least one to-be-processed image voucher as a final review result for the at least one to-be-processed image voucher. In this case, the system 130 can feed back, to a user-side client (that is, the terminal 120), the review result output by the large model. For example, when determining that the review result output by the large language model is "review succeeds", the system 130 feeds back a result indicating that the review succeeds to the user-side client. When determining that the review result output by the large language model is "review fails", the system 130 feeds back, to the user-side client, a result indicating that the review fails, a to-be-processed image voucher whose review is problematic, and a reason for the review failure. The system 130 can further prompt the user 110 to re-upload a voucher and then review the voucher again. The user 110 can re-upload the to-be-processed image voucher whose review is problematic to upload a correct to-be-processed image voucher, until a review result indicates that the review succeeds. It should be understood that the above-mentioned embodiment is merely an example for description. A review process of a review result in a specific target scenario, and how to feed back a result to the user-side client based on the review result can be flexibly adjusted based on a user requirement, and are not limited to those provided in the above-mentioned embodiment.

[0080] In some embodiments, when the system 130 determines that the review result output by the large language model is "review fails", the system 130 can output the review result of the

large language model for the at least one to-be-processed image voucher to a review-side client. In addition, the system 130 can also send the at least one to-be-processed image voucher to the review-side client. The review-side client can be a client terminal device corresponding to background review personnel. The review-side client can be communicatively connected to the system 130. The review personnel can review the at least one to-be-processed image voucher again through the review-side client to determine whether the output result of the large language model is accurate.

[0081] The system 130 pre-sets review points corresponding to different scenarios. The system 130 can invoke a general capability and a powerful context understanding capability of the large language model, and review the to-be-processed image voucher based on the key review point by using the key review point as prompt engineering data. In this way, the system 130 only needs to set review points for different scenarios, and does not need to perform code development for each voucher review rule, thereby greatly reducing a code development workload. Therefore, the system 130 can cover more scenarios.

[0082] In some embodiments, the system 130 can further determine the review result of the large language model for the at least one to-be-processed image voucher, to determine whether the review result of the large language model for the at least one to-be-processed image voucher is accurate. In this case, the method P100 provided in this specification can further include the following step.

[0083] P180: Determine a target review result based on a confidence level of the review result of the large language model for the at least one to-be-processed image voucher.

[0084] A process of determining the target review result can include: when the confidence level is greater than a preset threshold, outputting the review result of the large language model for the at least one to-be-processed image voucher to the user-side client as the target review result; or when the confidence level is less than the preset threshold, outputting an alert to the review-side client, where the target review result includes the alert. After receiving the review result including the alert, the review-side client can further review the target review result including the alert through, for example, manual access. A review result with a high confidence level can be directly confirmed and output by the large language model, and a review result with a low confidence level is output to the review-side client, so that the review-side client confirms the target review result including the alert through a manual review. This setting manner can further improve accuracy of an automated review. In addition, a manual review is triggered only when the confidence level of the review result output by the large model is low, which greatly reduces manpower costs and time costs required for review, and further improves user experience.

[0085] In some possible embodiments, for example, results output by the classification model

and the information extraction model both have corresponding confidence levels. When obtaining the to-be-reviewed information, the large language model can also obtain a confidence level corresponding to the to-be-reviewed information. The system 130 can determine, based on a confidence level of each piece of to-be-reviewed information, the confidence level of the review result output by the large language model.

[0086] According to the voucher review method P100 and system 130 provided in this specification, after the at least one to-be-processed image voucher in the target scenario in the plurality of scenarios is obtained, the to-be-reviewed information in the at least one to-be-processed image voucher can be extracted; then based on the key review point that is in the target scenario and that is pre-stored in the voucher review system, the to-be-reviewed information and the key review point are input to the large language model; and a general capability and a powerful context understanding capability of the large model are invoked to review the to-be-reviewed information based on the key review point, thereby implementing an automated review, improving voucher review efficiency, avoiding error-prone problems in mutual review, and improving user experience. In addition, the system 130 only needs to pre-configure review points in different scenarios, and does not need to perform code development for each voucher review rule, so that code development costs are low, and a development period is short. Therefore, the voucher review method P100 and system 130 provided in this specification can more easily cover more scenarios.

[0087] FIG. 5 is a flowchart of a voucher review method according to another embodiment of this specification. As shown in FIG. 5, before step S170 is performed, the method further includes the following step.

[0088] S190: Perform image detection on the at least one to-be-processed image voucher.

[0089] In this embodiment of this specification, image detection includes performing anti-counterfeiting detection and/or seal detection on the at least one to-be-processed image voucher.

[0090] For example, a manner of performing image detection on the at least one to-be-processed image voucher is performing image detection on each to-be-processed image voucher, and a detection manner can be, for example, sequentially detecting to-be-processed image vouchers, or can be detecting a plurality of to-be-processed image vouchers in parallel. A specific manner of detecting the to-be-processed image can be flexibly adjusted based on a user requirement, and is not limited to that provided in the above-mentioned embodiment.

[0091] Anti-counterfeiting detection is, for example, detecting whether the to-be-processed image voucher is a spliced image, or detecting whether the to-be-processed image voucher is a synthesized image. When a detection result indicates that the to-be-processed image voucher is a spliced image/synthesized image, it is determined that anti-counterfeiting detection on the to-be-processed image voucher fails; or when a detection result indicates that the to-be-processed image

voucher is not a spliced image/synthesized image, it is determined that anti-counterfeiting detection on the to-be-processed image voucher succeeds. In other words, anti-counterfeiting detection is used to determine whether the to-be-processed image voucher is an actually captured or scanned image, and when a detection result is that the to-be-processed image voucher is an actually captured/scanned image, it is determined that anti-counterfeiting detection on the to-be-processed image voucher succeeds; or when a detection result is that the to-be-processed image voucher is not an actually photographed/scanned images, it is determined that anti-counterfeiting detection on the to-be-processed image voucher fails. This detection manner ensures authenticity of the to-be-processed image voucher.

[0092] Seal detection is detecting validity and accuracy of a seal in a to-be-processed image voucher with a public seal/private seal. A detection manner includes, for example, but is not limited to performing detection based on a shape feature of a seal (detection based on feature extraction), performing detection based on a difference between a color and a background color of a seal (detection based on an edge), or performing detection by using a texture feature specific to a seal.

A specific detection manner can be flexibly adjusted based on a scenario requirement, and can be the above-mentioned detection on a single feature, or can be the above-mentioned comprehensive detection on a plurality of features. Seal detection is used to detect authenticity of the seal in the to-be-processed image voucher, and recognize whether a forged seal exists in the to-be-processed image voucher (for example, the seal is added to the to-be-processed image voucher through later image processing), or whether a fake seal exists in the to-be-processed image voucher (for example, the seal is added to the to-be-processed image voucher through later image processing). This further ensures that a specific manner of detecting the to-be-processed image and content included in the detection are not limited to those provided in the above-mentioned embodiment.

[0093] In the above-mentioned manner of performing image detection on the to-be-processed image voucher, validity and authenticity of the to-be-processed image voucher uploaded by the user are ensured. For a to-be-processed image voucher fails in image detection, a subsequent review step is directly stopped, a review procedure returns to the user-side client, and it is prompted that the to-be-processed image voucher uploaded by the user fails in verification, and the user is prompted to re-upload a to-be-processed image voucher. After the to-be-processed image voucher re-uploaded by the user is obtained, image verification is performed again on the to-be-processed image voucher uploaded by the user, until the to-be-processed image voucher uploaded by the user succeeds in verification. In this case, it is determined that the to-be-processed image voucher uploaded by the user succeeds in verification and is a valid voucher, and the to-be-processed image voucher can be further reviewed.

[0094] In some possible embodiments, to improve detection efficiency, for example, after the

to-be-processed image voucher is obtained, image detection is first performed on the to-be-processed image; the to-be-reviewed information in the to-be-processed image voucher is extracted only after a detection result of the image detection indicates that the image detection succeeds; and after the to-be-reviewed information is extracted, the to-be-reviewed information and the key review point in the target scenario are input to the large language model to obtain the review result of the large language model for the at least one to-be-processed image voucher. In some other possible embodiments, certainly, extraction of the to-be-reviewed information and image detection on the to-be-processed image voucher can be performed synchronously, and after the extracted to-be-reviewed information and an image detection result are obtained, the image detection result, the to-be-reviewed information, and the key review point in the target scenario are input to the large language model to obtain the review result of the large language model for the at least one to-be-processed image voucher. It should be understood that the above-mentioned embodiment is merely an example for description. Specific steps for reviewing the to-be-processed image voucher and a sequence between the steps are not specifically limited herein in this specification, and can be specifically adjusted based on a user requirement.

[0095] In some embodiments, step S170 includes: inputting the to-be-reviewed information, an image detection result, and the key review point in the target scenario to the large language model to obtain the review result of the large language model for the at least one to-be-processed image voucher.

[0096] In this embodiment provided in this specification, when the large language model reviews the to-be-processed image voucher, the large language model comprehensively considers the to-be-reviewed information in the to-be-processed image voucher, the image detection result, and the key review point in the target scenario, and therefore, the review result for the to-be-processed image voucher is more accurate. In addition, in this specification, after the to-be-processed image voucher is obtained, extraction of the to-be-reviewed information in the to-be-processed image voucher, determining of image detection, and determining of the key review point in the target scenario are all automatically performed. After the above-mentioned information is automatically obtained, the above-mentioned information is automatically input to the large language model to obtain the review result determined by the large language model based on the above-mentioned information. Therefore, efficiency of reviewing the to-be-processed image voucher is improved, consumption of manpower costs are avoided. In addition, because the key review point in the target scenario in this specification is pre-configured, when the to-be-reviewed information is reviewed based on the key review point in the target scenario, accuracy of the voucher review method provided in this specification is improved.

[0097] In some other possible embodiments, the method provided in this specification further

includes performing fine-tuning training on the large language model to improve accuracy and reliability of a result output by the large language model. A fine-tuning training manner can be, for example, performing fine-tuning training on the large language model by using data that is output by the large language model and whose review result is incorrect as a training sample.

5 **[0098]** In some embodiments, when the confidence level of the review result output by the large language model is less than the preset threshold, after the review-side client reviews the review result, it can be determined whether the review result output by the large language model is correct or incorrect. If a review result indicates "correct", it indicates that the review result output by the large language model is correct, but the confidence level is low. In this case, the current
10 review data can be used as sample data for fine-tuning training of the large language model. When the review-side client determines, through review, that the review result output by the large language model is incorrect, it indicates that the review result output by the large language model is incorrect, and the current review data can also be used as sample data for fine-tuning training of the large language model.

15 **[0099]** Fine-tuning training of the large language model can be performed periodically, or can be performed only when a quantity of times that the confidence level of the review result output by the large language model is less than the preset threshold are accumulated to a preset quantity.

20 **[0100]** The above-mentioned manner of performing fine-tuning training on the large language model can be understood as follows: A solution for linking the large language model and the review-side client is set, and fine-tuning training is performed on the large language model by using a manual review result, to improve accuracy of a result output by the large language model.

25 **[0101]** In some embodiments, fine-tuning training of the large language model can be performed in synchronization with update training of the above-mentioned classification model and information extraction model.

30 **[0102]** FIG. 6 is an example flowchart of a voucher recognition method according to some embodiments of this specification. As shown in FIG. 6, after to-be-processed image vouchers uploaded by a user are obtained, the N to-be-processed images are input to a recognition and parsing module and an image detection module sequentially or in parallel in a manner of 1: N. The recognition and parsing module includes an OCR module, a document classification module, and
35 an information extraction module. In this embodiment of this specification, character recognition is performed on the to-be-processed image voucher by using an OCR technology, and document classification is performed on each to-be-processed image voucher based on a character recognition result and a pre-trained classification model. Then, information extraction is performed on the character recognition result for each to-be-processed image based on a classification result and an information extraction model corresponding to each document type, to

obtain to-be-reviewed information in each to-be-processed image voucher. In addition, the image detection module performs anti-counterfeiting detection and/or seal detection on each to-be-processed image voucher to determine authenticity of the to-be-processed image voucher. Subsequently, the extracted to-be-reviewed information and the detected image detection information are aggregated into integrated to-be-reviewed information and image detection information in a target scenario in a manner of N: 1. Afterwards, the to-be-reviewed information, the image detection information, and review point information corresponding to the target scenario are jointly input to a large language model, so that the large language model performs information processing on the to-be-reviewed information. An information processing manner includes but is not limited to information integration, basic logic, and model determining. The information integration is to integrate a plurality of pieces of to-be-reviewed information corresponding to the N to-be-processed image vouchers, perform basic logic arrangement based on a key review point corresponding to the target scenario, and finally perform large model determining based on the above-mentioned content, to obtain and output a review result for the at least one to-be-processed image voucher.

[0103] In conclusion, according to the voucher review method P100 and system 130 provided in this specification, at least one to-be-processed image voucher is obtained, a plurality of to-be-recognized areas are recognized from each to-be-processed image voucher, and the at least one candidate area corresponds to at least one piece of to-be-reviewed information. In addition, image detection further needs to be performed on each to-be-processed image voucher to obtain a detection result corresponding to the to-be-processed image voucher. A key review point corresponding to a target scenario is determined based on the target scenario from pre-stored review points in a plurality of scenarios, so that a large language model reviews the to-be-reviewed voucher based on the plurality of pieces of to-be-reviewed information, each image detection result, and the key review point corresponding to the target scenario, to obtain a review result for the to-be-processed image voucher. In this determining manner, a to-be-reviewed voucher is automatically reviewed. In addition, an applicable scenario of the voucher review system is diversified, and has strong scalability. Moreover, fine-tuning training is further performed on the large language model based on a confidence level of each review result and information that is fed back by a review-side client for a review result with a low confidence level, to obtain a trained large language model. In this setting manner, the method P100 provided in this embodiment of this specification can implement an automated review of a to-be-reviewed voucher and has an applicable scenario with strong scalability, and accuracy of the large language model can be improved through continuous iterative training, to maintain accuracy of a result output by the large language model.

[0104] Another aspect of this specification provides a computer-readable non-transitory storage medium storing at least one instruction set that includes executable instructions for performing voucher review. When the at least one instruction set is executed by a processor, the at least one instruction set instructs the processor to implement the steps of the voucher review method P100 in this specification. In some possible implementations, various aspects of this specification can be further implemented in a form of a program product, including program code. When the program product runs in the system 130, the program code is used to enable the system 130 to perform the steps of the method P100 described in this specification. The program product used to implement the above-mentioned method can be a portable compact disk read-only memory (CD-ROM), including program code, and can run in the system 130. However, the program product in this specification is not limited thereto. In this specification, a readable storage medium can be any tangible medium that includes or stores a program, and the program can be used by or in combination with an instruction execution system. The program product can be any combination of one or more readable media. The readable medium can be a readable signal medium or a readable storage medium. For example, the readable storage medium can be but is not limited to an electrical system, apparatus, or device, a magnetic system, apparatus, or device, an optical system, apparatus, or device, an electromagnetic system, apparatus, or device, an infrared system, apparatus, or device, or a semiconductor system, apparatus, or device, or any combination thereof. More specific examples of the readable storage medium include an electrical connection including one or more conducting wires, a portable disk, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM) or a flash memory, an optical fiber, a portable compact disc read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any proper combination thereof. The computer-readable storage medium can include a data signal propagated in a baseband or as part of a carrier, and carries readable program code. The propagated data signal can be in a plurality of forms, and includes but is not limited to an electromagnetic signal, an optical signal, or any proper combination thereof. The readable storage medium can alternatively be any readable medium other than the readable storage medium, and the readable medium can send, propagate, or transmit a program used by or in combination with an instruction execution system, apparatus, or device. The program code included in the readable storage medium can be transmitted by using any proper medium, including but not limited to a wireless medium, a wired medium, an optical cable, an RF medium, or any proper combination thereof. Program code for performing an operation of this specification can be written in any combination of one or more program design languages. The program design language includes an object-oriented program design language, for example, Java or C++, and further includes a conventional procedural program design language, for example, a

"C" language, or a similar program design language.

[0105] In conclusion, the voucher review method P100 and the system 130 provided in this specification can implement an automated review of a to-be-reviewed voucher. In addition, a scenario to which the method P100 and the system 130 provided in this specification has strong
5 extensibility, and accuracy of a large language model can be improved through continuous iterative training, to maintain accurate determining of an output result of the large language model.

[0106] Specific embodiments of this specification are described above. Other embodiments fall within the scope of the appended claims. In some cases, the actions or steps described in the claims can be performed in an order different from that in the embodiments and desired results can
10 still be achieved. In addition, the process depicted in the accompanying drawings does not necessarily need a particular order or consecutive order to achieve the desired results. In some implementations, multitasking and parallel processing are possible or may be advantageous.

[0107] In conclusion, after reading the detailed disclosure, a person skilled in the art can understand that the detailed disclosure can be presented only in an example manner, and can
15 impose no limitation. Although it is not explicitly stated herein, a person skilled in the art can understand that the requirements of this specification include various proper changes, improvements, and modifications to the embodiments. These changes, improvements, and modifications are intended to be provided in this specification, and fall within the spirit and scope of the example embodiments of this specification.

[0108] In addition, some terms in this specification are used to describe the embodiments of this specification. For example, "one embodiment", "embodiments", and/or "some embodiments" mean/means that specific features, structures, or characteristics described with reference to the embodiment can be included in at least one embodiment of this specification. Therefore, it can be emphasized and understood that two or more references to "embodiments", "one embodiment", or
25 "alternative embodiments" in various parts of this specification do not necessarily refer to a same embodiment. In addition, specific features, structures, or characteristics can be properly combined in one or more embodiments of this specification.

[0109] It should be understood that in the above-mentioned descriptions of the embodiments of this specification, to help understand a feature, for the purpose of simplifying this specification, various features are combined in a single embodiment, accompanying drawing, or description thereof in this specification. However, this does not mean that the combination of these features is necessary. It is entirely possible for a person skilled in the art to mark some devices and understand the devices as separate embodiments when reading this specification. In other words, the embodiments in this specification can also be understood as an integration of a plurality of
35 secondary embodiments. Content of each secondary embodiment is also true when features are

less than all features in one of the above-mentioned disclosed embodiments.

[0110] Each patent, patent application, publication of a patent application, and other materials, for example, articles, books, instructions, publications, documents, and products, that are cited in this specification and that are other than any document that is inconsistent or conflict with this document or that imposes limiting impact on the widest scope of the claims can be incorporated herein by reference and used for all purposes now or later associated with this document. In addition, the terms in this document shall prevail if there is any inconsistency or conflict between the descriptions, definitions, and/or use of the related terms in any material and the descriptions, definitions, and/or use of the related terms in this document.

[0111] Finally, it should be understood that the principles of the implementation solutions of this specification are described in the implementation solutions of the application disclosed in this specification. Other modified embodiments also fall within the scope of this specification. Therefore, the embodiments disclosed in this specification are merely examples and not limitations. A person skilled in the art can use an alternative configuration according to the embodiments of this specification to implement the application in this specification. Therefore, the embodiments of this specification are not limited to the embodiments accurately described in the application.

CLAIMS

What is claimed is:

1. A voucher review method, applied to a voucher review system, wherein the method comprises:

- 5 obtaining at least one to-be-processed image voucher in a target scenario;
extracting to-be-reviewed information in the at least one to-be-processed image voucher;
determining a key review point in the target scenario, wherein the voucher review system pre-stores review points in a plurality of scenarios, and the plurality of scenarios comprise the target scenario; and
10 inputting the to-be-reviewed information and the key review point in the target scenario to a large language model to obtain a review result of the large language model for the at least one to-be-processed image voucher.

2. The method according to claim 1, wherein the extracting to-be-reviewed information in the at least one to-be-processed image voucher comprises performing the following operations on a
15 to-be-processed image voucher:

- performing character recognition on the to-be-processed image voucher;
determining, based on the character recognition result, a document type corresponding to the to-be-processed image voucher; and
performing information extraction on the character recognition result based on the document
20 type corresponding to the to-be-processed image voucher, to obtain to-be-reviewed information in the to-be-processed image voucher.

3. The method according to claim 2, wherein the performing character recognition on the to-be-processed image voucher data comprises:

- determining, based on a selection operation performed by a user on the to-be-reviewed
25 information in the to-be-processed image voucher, a character recognition area corresponding to the to-be-reviewed information in the to-be-processed image voucher; and
performing character recognition on the character recognition area to obtain the character recognition result.

4. The method according to claim 2, wherein the determining, based on the character
30 recognition result, a document type corresponding to the to-be-processed image voucher comprises:

- classifying the character recognition result based on a pre-trained classification model to obtain the document type corresponding to the to-be-processed image voucher.

5. The method according to claim 2, wherein the performing information extraction on the character recognition result based on the document type corresponding to the to-be-processed

image voucher, to obtain to-be-reviewed information in the to-be-processed image voucher comprises:

performing information extraction on the character recognition result based on an information extraction model corresponding to the document type, to obtain the to-be-reviewed information in the to-be-processed image voucher.

6. The method according to claim 2, wherein the document type comprises an identity type and a fund type.

7. The method according to claim 6, wherein the document type further comprises a contract type.

8. The method according to claim 1, wherein the method further comprises:
determining a target review result based on a confidence level of the review result of the large language model for the at least one to-be-processed image voucher, comprising:

when the confidence level is greater than a preset threshold, outputting the review result of the large language model for the at least one to-be-processed image voucher to a user-side client as the target review result; or

when the confidence level is less than the preset threshold, outputting an alert to a review-side client, wherein the target review result comprises the alert.

9. The method according to claim 1, further comprising:
performing image detection on the at least one to-be-processed image voucher, wherein the inputting the to-be-reviewed information and the key review point in the target scenario to a large language model to obtain a review result of the large language model for the at least one to-be-processed image voucher comprises:

inputting the to-be-reviewed information, an image detection result, and the key review point in the target scenario to the large language model to obtain the review result of the large language model for the at least one to-be-processed image voucher.

10. The method according to claim 9, wherein the performing image detection on the at least one to-be-processed image voucher comprises:

performing anti-counterfeiting detection and/or seal detection on the at least one to-be-processed image voucher.

11. The method according to claim 1, further comprising:
performing fine-tuning training on the large language model, comprising:
performing fine-tuning training on the large language model by using data that is output by the large language model and whose review result is incorrect as a training sample.

12. A voucher review system, comprising:
at least one storage medium, storing at least one instruction set used to perform a voucher

review method; and

at least one processor, communicatively connected to the at least one storage medium, wherein when running, the at least one processor reads the at least one instruction set, and performs the following operations based on an instruction of the at least one instruction set:

- 5 obtaining at least one to-be-processed image voucher in a target scenario;
 extracting to-be-reviewed information in the at least one to-be-processed image voucher;
 determining a key review point in the target scenario, wherein the voucher review system pre-stores review points in a plurality of scenarios, and the plurality of scenarios comprise the target scenario; and
- 10 inputting the to-be-reviewed information and the key review point in the target scenario to a large language model to obtain a review result of the large language model for the at least one to-be-processed image voucher.

13. The system according to claim 12, wherein to extract the to-be-reviewed information in the at least one to-be-processed image voucher, the at least one processor performs the following

15 operations on a to-be-processed image voucher:

- performing character recognition on the to-be-processed image voucher;
 determining, based on the character recognition result, a document type corresponding to the to-be-processed image voucher; and
 performing information extraction on the character recognition result based on the document
- 20 type corresponding to the to-be-processed image voucher, to obtain to-be-reviewed information in the to-be-processed image voucher.

14. The system according to claim 13, wherein to perform character recognition on the to-be-processed image voucher data, the at least one processor performs the following operations:

- determining, based on a selection operation performed by a user on the to-be-reviewed
- 25 information in the to-be-processed image voucher, a character recognition area corresponding to the to-be-reviewed information in the to-be-processed image voucher; and
 performing character recognition on the character recognition area to obtain the character recognition result.

15. The system according to claim 13, wherein to determine, based on the character

30 recognition result, the document type corresponding to the to-be-processed image voucher, the at least one processor performs the following operation:

 classifying the character recognition result based on a pre-trained classification model to obtain the document type corresponding to the to-be-processed image voucher.

16. The system according to claim 13, wherein to determine, based on the character

35 recognition result, the document type corresponding to the to-be-processed image voucher, the at

least one processor performs the following operation:

classifying the character recognition result based on a pre-trained classification model to obtain the document type corresponding to the to-be-processed image voucher.

17. The system according to claim 13, wherein to perform information extraction on the character recognition result based on the document type corresponding to the to-be-processed image voucher, to obtain the to-be-reviewed information in the to-be-processed image voucher, the at least one processor performs the following operation:

performing information extraction on the character recognition result based on an information extraction model corresponding to the document type, to obtain the to-be-reviewed information in the to-be-processed image voucher.

18. The system according to claim 12, wherein the at least one processor performs the following operation:

determining a target review result based on a confidence level of the review result of the large language model for the at least one to-be-processed image voucher, comprising:

when the confidence level is greater than a preset threshold, outputting the review result of the large language model for the at least one to-be-processed image voucher to a user-side client as the target review result; or

when the confidence level is less than the preset threshold, outputting an alert to a review-side client, wherein the target review result comprises the alert.

19. The system according to claim 12, wherein the at least one processor performs the following operation:

performing image detection on the at least one to-be-processed image voucher, and

to input the to-be-reviewed information and the key review point in the target scenario to the large language model to obtain the review result of the large language model for the at least one to-be-processed image voucher, the at least one processor performs the following operation:

inputting the to-be-reviewed information, an image detection result, and the key review point in the target scenario to the large language model to obtain the review result of the large language model for the at least one to-be-processed image voucher.

20. A computer-readable nonvolatile storage medium, wherein the computer-readable nonvolatile storage medium stores at least one instruction set, and the at least one instruction set is executed by at least one processor to implement the voucher review method according to any one of claims 1 to 11.