

A VIDEO PROCESSOR TRAINING METHOD, A VIDEO PROCESSOR TRAINING DEVICE, A VIDEO PROCESSING DEVICE AND A VIDEO PROCESSING METHOD

TECHNICAL FIELD

[0001] The present invention relates to the field of data processing, in particular to a video processor training method, a video processor training device, a video processing device, a video processing method, a computer device and a computer-readable storage medium.

BACKGROUND

[0002] Although data transmission is widely used in modern industrial production and modern life, data transmission is usually restrained by bandwidth and is impossible to send or receive unlimited data in a very short time. For the purpose of maximizing the quality and quantity of transmitted data, data will be compressed or encoded before transmission and be decompressed or decoded when received in some cases.

[0003] In terms of data transmission, video data transmission is of great importance, before which video data to be transmitted has to be compressed. However, even a very short video clip may occupy a large amount of data since video data contains chroma-related data and brightness-related data and usually contains stereo audio related data. Therefore, video compression technology is increasingly important in order to reduce the space occupied by video data on hard disk and reduce the network latency of video transmission.

[0004] Given that video compression will lead to loss of original video data to varying degrees, any loss of particulars in the images contained in the video data may cause inconvenience to some application scenarios that require high image resolution or fidelity. There are many methods in the prior art to solve the data loss caused by video data compression, such as the methods relying on the researched prior information of images to remedy the data loss on the basis of the video sequence, including but not

limited to the frame prediction method. The frame prediction method available in the prior art is mainly intended to predict and compensate the existing partial images by means of motion estimation and motion compensation and based on the previous partial images. When there is an intense motion area or a shot cut in video scenes and there is any new video data suddenly appearing in images, the existing coding block is unable to identify the best matching block in the sense of rate-distortion in the search area specified by the reference frame or large motion estimation residuals and coding residuals are obtained after inter-frame prediction, leading to large distortions in video restoration. In addition to the large distortions in video restoration caused, the algorithms available in the prior art are highly complex and therefore are unlikely to apply to the scenarios with high requirements on real-time rendering.

SUMMARY

[0005] In order to solve the problems existing in the prior art as well as the problems mentioned in the BACKGROUND section above, at least one embodiment of the present invention provides a video processor training method, a video processor training device, a video processing device, a video processing device and a computer device.

[0006] Firstly, the embodiments of the present invention provide a video processor training method which involves: inputting a decoded video into a video processor as a training sample, wherein the decoded video is a decoded video generated by encoding and then decoding the original video; processing the decoded video and outputting an expected difference by the video processor; determining a loss function according to the difference between the expected difference and an ideal difference, wherein the ideal difference is the difference between the original video and the decoded video; adjusting the parameters of the video processor according to the value of the loss function, and; stopping training when the value of the loss function tends to converge to obtaining a trained video processor.

[0007] In some embodiments, the ideal difference is determined according to the difference between the image frames corresponding to the original video and the

decoded video respectively.

[0008] In some embodiments, the aforesaid determination of the ideal difference according to the difference between the image frames corresponding to the original video and the decoded video respectively further involves: decomposing the original video in a preset manner to obtain each image frame of the original video; decomposing the decoded video in the same preset manner to obtain each image frame of the decoded video decomposed; obtaining a difference between each decoded video frame of the decoded video decomposed and each decoded video frame of the original video decomposed; and processing the difference between the image frames to obtain the ideal difference.

[0009] In some embodiments, the difference between the image frames is the difference between the pixel values of the image frames.

[0010] Secondly, the embodiments of the present invention also provide a video processor training device which is comprised of: an input module designed to input a decoded video into a video processor as a training sample, which decoded video is a decoded video generated by encoding and then decoding the original video; a video processor designed to process the decoded video to obtain an expected difference; a loss function determination module designed to determine a loss function according to the difference between the expected difference and an ideal difference, which ideal difference is the difference between the original video and the decoded video; and, an adjustment module designed to adjust the parameters of the video processor according to the value of the loss function and to stop training when the value of the loss function tends to converge to obtain a trained video processor.

[0011] In some embodiments, the loss function determination module is further comprised of: a first decomposition unit configured to decompose the original video in a preset manner to obtain each image frame of the original video; a second decomposition unit configured to decompose the decoded video in the same preset manner to obtain each image frame of the decoded video decomposed; an image frame difference value obtaining unit configured to obtain the difference between the image frames corresponding to the decoded video decomposed and the original video

decomposed respectively; and an ideal difference determination unit configured to process the difference between the image frames to obtain the ideal difference.

[0012] In some embodiments, the video processor is comprised of a convolutional neural network which has at least one of the following characteristics: the convolutional neural network is a shallow neural network, or the parameters of the convolutional neural network are less than a preset value.

[0013] In some embodiments, the convolutional neural network is mobilenet or shufflenet.

[0014] In some embodiments, the convolutional neural network uses an unsaturated activation function.

[0015] Thirdly, the embodiments of the present invention also provide a video processing method which involves: acquiring a to-be-processed video which is a decoded video generated after encoding and decoding; processing the to-be-processed video and obtaining an expected difference by a video processor which is trained according to the training method described in the first aspect above or is trained using the training device described in the second aspect above; and, determining a new video according to the expected difference and the to-be-processed video.

[0016] In some embodiments, determining a new video according to the expected difference and the to-be-processed video as mentioned above further involves superimposing the expected difference value with the to-be-processed video to form a new video.

[0017] In some embodiments, superimposing the expected difference value with the to-be-processed video to form a new video as mentioned above further involves: decomposing the to-be-processed video in a preset manner to obtain image frames of the to-be-processed video decomposed; decomposing the expected difference in the same preset manner to obtain image frames of the expected difference decomposed; and, superimposing the image frames corresponding to the image frames of the to-be-processed video decomposed and the image frames of the expected difference decomposed to form a new video.

[0018] Fourthly, the embodiments of the present invention also provide a video

processing device which is comprised of: an acquisition module designed to acquire a to-be-processed video which is a decoded video generated after encoding and decoding; a video processor which is designed to process the to-be-processed video to obtain an expected difference and is trained according to the training method described in the first aspect above or trained using the training device described in the second aspect above; and a new video determination module designed to determine a new video based on the expected difference and the to-be-processed video.

[0019] In some embodiments, the new video determination module is used to superimpose the expected difference and the to-be-processed video to form a new video, and is further comprised of: a third decomposition unit configured to decompose the to-be-processed video in a preset manner to obtain image frames of the to-be-processed video decomposed; a fourth decomposition unit configured to decompose the expected difference in the same preset manner to obtain image frames of the expected difference decomposed; and, a new video forming unit configured to superimpose the image frames corresponding to image frames of the expected difference decomposed and image frames of the expected difference decomposed in order to form a new video.

[0020] Fifthly, the embodiments of the present invention also provide a computer device which is comprised of: at least one processor, a memory, a communication interface and a communication bus, of which the processor, the memory and the communication interface communicate with each other through the communication bus; and a memory coupled to the at least one processor, which stores an executable instruction that, when executed by the at least one processor, makes it possible to implement any method described in the first aspect or the third aspect as mentioned above.

[0021] Sixthly, the embodiments of the present invention also provide a chip used to implement the method described in the first aspect as mentioned above. Specifically, the chip is comprised of a processor designed to call and run a computer program from the memory, making it possible to implement any method described in the first aspect or the third aspect as mentioned above using any device installed with the chip.

[0022] Seventhly, the embodiments of the present invention also provide a computer-

readable storage medium which stores a computer program that, when executed by a processor, makes it possible to implement any method described in the first aspect or the third aspect as mentioned above.

[0023] Finally, the embodiments of the present invention also provide a computer program product which includes a computer program instruction that makes it possible for a computer to implement any method described in the first aspect or the third aspect as mentioned above.

[0024] It can be seen that as described in the embodiments of the present application, an encoded and decoded video is taken as input by training the video processor, the video processor outputs an expected difference after video processing, a loss function is determined based on the expected difference output and the ideal difference, and parameters of the video processor are adjusted, making the difference output from the video processor close to the ideal difference. In the present application, the expected difference is output from the video processor based on the training method described herein and the loss function is determined by comparing the expected difference and the ideal difference, greatly reducing the data size to be processed and improving training efficiency. The video processing method described in the present application is applied to train the video processor through the deep learning network, obtain the expected difference output from the trained video processor, superimpose the expected difference with the video input and restore the compressed and decompressed video to the original video, helping to avoid large motion estimation residuals and coding residuals caused by inter-frame prediction using motion estimation and motion compensation in the prior art and to improve the effects of video restoration. Furthermore, the video processing method described in present application is applied to perform video restoration based on the expected difference, reducing the data size to be processed, shortening computation time and improving computing capability as well as video processing efficiency.

BRIEF DESCRIPTION OF THE DRAWINGS

[0025] In order to give a clearer explanation on the technical solutions described in the

embodiments of the present invention, the drawings required for describing the embodiments of the present invention or the prior art are briefly introduced as below. Obviously, the drawings described below only involve some embodiments of the present invention. For those of ordinary skill in the art, other drawings can also be obtained based on the following drawings without paying any creative labor.

[0026] FIG. 1 is a flowchart of a specific embodiment describing the video processor training method proposed in the present invention;

[0027] FIG. 2 is a structure block diagram of a specific embodiment describing the video processor training device proposed in the present invention;

[0028] FIG. 3 is a flowchart of a specific embodiment describing the video processing method proposed in the present invention;

[0029] FIG. 4 is a structure block diagram of a specific embodiment describing the video processing device proposed in the present invention;

[0030] FIG. 5 is a basic structure block diagram demonstrating the computer device mentioned in the embodiment.

DESCRIPTION OF THE EMBODIMENTS

[0031] A detailed description is made on the present application with reference to the accompanying drawings and specific embodiments, helping to have a clearer understanding of the purpose, characteristics and advantages of the present application. It should be understood that the following description made with reference to the accompanying drawings and specific embodiments is for illustrative purposes only in all aspects, and should not be construed as limiting the scope of protection of the present application. Unless the context clearly indicates otherwise, the terms “a/an” and “said” indicating the singular should include the plural. As used in the present application, the terms “first” and “second” can be used interchangeably to distinguish one or one type of component from another or another type, and are not intended to indicate the location or importance of any individual component.

[0032] Any reference herein to “embodiment” means that specific features, structures or characteristics described with reference to any embodiment may be included in at

least one embodiment of the present application. And the term “embodiment” referred in different places in the Specification does neither necessarily refer to the same embodiment nor serve as an independent or alternative embodiment that is mutually exclusive with other embodiments. Those skilled in the art understand explicitly and implicitly that the embodiments described herein can be used in combination with other embodiments.

[0033] The term “and/or” used herein merely describes three association relationships between associated objects. For instance, the phrase “A and/or B” may indicate three cases in which A exists alone, A and B exist simultaneously, and B exists alone. In addition, the character “/” used herein generally indicates an “or” relationship between the two associated objects placed before and after the character.

[0034] In order to enable those skilled in the art to have a better understanding of the solutions set out in the present application, a clear and complete description is made below in combination of the drawings attached on the technical solutions mentioned in the embodiments of the present application.

[0035] Un-encoded source videos usually contain large amounts of data. For example, the bitrate may reach 1.5Gbps when a one-second video with a 1920×1080 definition, a 24-bit color and 30 frames per second is output. Therefore, encoding is necessary to minimize the data size contained in videos. Although video compression using different algorithms and strategies will lead to different results and different video encoding standards, all video encoding technologies are intended to compress such three kinds of redundant data in videos as redundant space, redundant time and redundant codes. In contrast, decoding takes the opposite tack.

[0036] There are many encoding standards currently available, such as H.265 which is the best existing video compression industry standard. The basic principle of H.265 is to compress the repeated information between two adjacent frames and compress the collected video to the desired size (for example, 100M to 1M) by adjusting H.265 video compression parameters, which, however, may result in poor quality of the decoded video obtained after decoding the compressed video.

[0037] Firstly, the present application provides a video processor training method. FIG.

1 is a flowchart of a specific embodiment describing the video processor training method proposed in the present invention. As shown in FIG. 1, the video processor training method involves:

[0038] Step 110: input a decoded video into a video processor as a training sample, wherein the decoded video is a decoded video generated by encoding and then decoding the original video;

[0039] The training mentioned in the present application is intended to enable the video processor to form an expected difference after processing the decoded video and to make it possible that the expected difference can be infinitely close to an ideal difference, so that the video obtained after superimposing the difference output by the video processor and the decoded video can be infinitely close to the original video. The training sample used in the present application is a decoded video obtained by decoding the compressed video obtained by encoding the original video. The video obtained by encoding and decoding the original video is used as a sample of the training method and input into the video processor. In the present application, the video encoding and decoding standards are subject to no limitations, and the decoded video obtained by encoding and decoding the original video using any encoding and decoding method existing in the prior art or future technologies is input into the video processor.

[0040] Step 120: process the decoded video and output an expected difference by the video processor. In the present application, a deep learning network may be applied with a decoded video as a training sample being input and an expected difference being output through the deep learning network. The expected difference may be a file in the same or similar format as the input decoded video. Specifically, the video processor may, in some embodiments, include a convolutional neural network, such as a shallow neural network or a neural network with fewer parameters. In this embodiment, the video processor is the object of training.

[0041] Step 130: determine a loss function according to the difference between the expected difference and an ideal difference, wherein the ideal difference is the difference between the original video and the decoded video. It can be understood that video is composed of consecutive images of a certain sequence, and accordingly, image

frames contained in video can be extracted according to a preset time interval. In the present application, the expected difference formed may be continuous image frame information of a certain sequence which may be two-dimensional image frame information or three-dimensional image frame information or other existing or future image frame information.

[0042] The ideal difference aforementioned is determined according to the difference between the image frames corresponding to the original video and the decoded video respectively. Specifically, the ideal difference may be obtained in the following steps: decompose the original video in a preset manner to obtain each image frame of the original video; decompose the decoded video in the same preset manner to obtain each image frame of the decoded video decomposed; obtain a difference between each image frame corresponding to the decoded video decomposed and each image frame corresponding to the original video decomposed; and, process the difference between the image frames to obtain the ideal difference. It can be understood that the image frames may be specifically expressed as pixel values, and the difference between the image frames may be expressed as the pixel value difference of the image frames. Specifically, the decomposed image frames may be marked based on time sequence or other order, and then the loss function may be determined according to the difference between the ideal difference and the expected difference.

[0043] Step 140: adjust the parameters of the video processor according to the value of the loss function. It can be understood that the loss function is the difference between the ideal difference and the expected difference, and that the expected difference is also deemed as data in the video format reflecting a certain sequence of image frames while the ideal difference video data is obtained. For example, the parameters of video processor are adjusted based on the image frame data in video data and the changes in the image frame data, after which the training sample is processed and the data of the adjusted loss function is obtained.

[0044] Step 150: stop training when the value of the loss function tends to converge to obtain a trained video processor. Training may be stopped when the value of the loss function tends to converge while continuous adjustments are made to the parameters of

the video processor. Specifically, the foregoing convergence may be either that the data curve presents a convergence state or that the value of the loss function reaches or is less than a preset value. The video processor obtained therefore is a trained video processor. For example, when the value of the loss function is equal to or is infinitely close to 0, it is indicated that the difference output from the video process is almost equal to the ideal difference between the decompressed video compressed and the original video. In this case, a video that is almost close to the original video can be obtained by superimposing the difference output and the decompressed video compressed.

[0045] If the value of the loss function has not reached the preset value or does not tend to converge, training should go on.

[0046] In the embodiments of the present application, a video processor is trained by taking an encoded and decoded video as input, an expected difference is output from the video processor after processing the video, a loss function is determined based on the expected difference output and an ideal difference, and parameters of the video processor are adjusted to ensure that the different output from the video processor is close to the ideal difference. Given that the loss function is determined by comparing the expected difference output from the video processor trained applying the training method described in the present application and the ideal difference, the amount of data to be processed is greatly reduced and the training efficiency is significantly improved.

[0047] Secondly, the present invention also provides a video processor training device. FIG. 2 is a structure block diagram of a specific embodiment describing the video processor training device proposed in the present invention. As shown in FIG. 2, the video processor training device is comprised of an input module 201, a video processor 202, a loss function determination module 203 and an adjustment module 204, of which:

[0048] The input module 201 is designed to input a decoded video into a video processor as a training sample, which decoded video is a decoded video generated by encoding and then decoding the original video; the video processor 202 is designed to process the decoded video to obtain an expected difference; the loss function determination module 203 is designed to determine a loss function according to the

difference between the expected difference and an ideal difference, which ideal difference is the difference between the original video and the decoded video; and the adjustment module 204 is designed to adjust the parameters of the video processor according to the value of the loss function and to stop training when the value of the loss function tends to converge to obtain a trained video processor.

[0049] For the operations of the modules included in the video processor training device as described in this embodiment, reference may be made to the methods demonstrated in FIG. 1 to achieve the same technical effect.

[0050] In some embodiments, the loss function determination module is further comprised of: a first decomposition unit configured to decompose the original video in a preset manner to obtain each image frame of the original video; a second decomposition unit configured to decompose the decoded video in the same preset manner to obtain each image frame of the decoded video decomposed; an image frame difference value obtaining unit configured to obtain the difference between the image frames corresponding to the decoded video decomposed and the original video decomposed respectively; and an ideal difference determination unit configured to process the difference between the image frames to obtain the ideal difference.

[0051] In some embodiments, the video processor is comprised of a convolutional neural network which has at least one of the following characteristics: the convolutional neural network is a shallow neural network, or the parameters of the convolutional neural network are less than a preset value. For example, the convolutional neural network may be mobilenet or shufflenet, and the convolutional neural network applied uses an unsaturated activation function. Of course, the choice of network depends on the actual situation. For example, if more importance is attached to the real-time performance of video processing, it is advisable to consider using CNN (Convolutional Neural Networks) with shallow network layers and small parameters, such as shuffle net; and if more importance is attached to the restoration accuracy of video processing, it is advisable to consider using the networks with deeper network layers and more parameters. In short, the choice of network depends on the actual needs.

[0052] Thirdly, the present application also provides a video processing method. FIG.

3 is a flowchart of a specific embodiment describing the video processing method proposed in the present invention. As shown in FIG. 3, the video processing method involves the following steps:

[0053] Step 310: acquire a to-be-processed video which is a decoded video generated after encoding and decoding;

[0054] Step 320: process the to-be-processed video and obtain an expected difference by a video processor which is trained according to the training method described in the first aspect above or is trained using the training device described in the second aspect above;

[0055] Step 330: determine a new video according to the expected difference and the to-be-processed video.

[0056] The video processing method described in the present application is applied to train a video processor through a deep learning network, use the trained video processor to output an expected difference, superimpose the expected difference and the video input, and restore the compressed and decompressed video to the original video, avoiding large motion estimation residuals and coding residuals caused by adopting motion estimation and motion compensation for inter-frame prediction in the prior art and improving the effect of video restoration. Besides, the video processing method described in the present application is intended to restore video based on the expected difference, reducing the data size to be processed, shortening computation time and improving computing capability as well as video processing efficiency.

[0057] In some embodiments, determining a new video according to the expected difference and the to-be-processed video as mentioned above further involves superimposing the expected difference value with the to-be-processed video to form a new video.

[0058] In some embodiments, superimposing the expected difference value with the to-be-processed video to form a new video as mentioned above further involves: decomposing the to-be-processed video in a preset manner to obtain image frames of the to-be-processed video decomposed; decomposing the expected difference in the same preset manner to obtain image frames of the expected difference decomposed;

and, superimposing the image frames corresponding to the image frames of the to-be-processed video decomposed and the image frames of the expected difference decomposed to form a new video.

[0059] Fourthly, the present application also provides a video processing device. FIG. 4 is a structure block diagram of a specific embodiment describing the video processing device proposed in the present invention. As shown in FIG. 4, the video processing device is comprised of an acquisition module 401, a video processor 402 and a new video determination module 403, of which:

[0060] The acquisition module 401 is designed to acquire a to-be-processed video which is a decoded video generated after encoding and decoding;

[0061] The video processor 402 is designed to process the to-be-processed video to obtain an expected difference and is trained according to the training method described in the first aspect above or trained using the training device described in the second aspect above;

[0062] The new video determination module 403 is designed to determine a new video based on the expected difference and the to-be-processed video.

[0063] For the operations of the modules included in the video processing device as described in this embodiment, reference may be made to the methods demonstrated in FIG. 3 to achieve the same technical effect.

[0064] In some embodiments, the new video determination module is used to superimpose the expected difference and the to-be-processed video to form a new video, and is further comprised of: a third decomposition unit configured to decompose the to-be-processed video in a preset manner to obtain image frames of the to-be-processed video decomposed; a fourth decomposition unit configured to decompose the expected difference in the same preset manner to obtain image frames of the expected difference decomposed; and, a new video forming unit configured to superimpose the image frames corresponding to image frames of the expected difference decomposed and image frames of the expected difference decomposed in order to form a new video.

[0065] Fifthly, the embodiments of the present application also provide a computer device which may be a neural network server. For details, please refer to FIG. 5. FIG.

5 is a basic structure block diagram demonstrating the computer device mentioned in the embodiment.

[0066] The computer device is comprised of a memory 51, a processor 52 and a communication interface 53 that are communicatively connected through a system bus. It should be pointed out that only the computer device 5 with the components 51-53 is shown in the figure, but it should be understood that not all components shown are required to be implemented and more or fewer components may be implemented instead. Those skilled in the art can understand that the computer device 5 mentioned here is a device that can automatically perform numerical calculation and/or information processing according to a preset or stored instruction with its hardware units including but not limited to microprocessor, Application Specific Integrated Circuit (ASIC), Field-Programmable Gate Array (FPGA), Digital Signal Processor (DSP), embedded device, etc..

[0067] The computer device 5 may be a computing device such as a desktop computer, a notebook, a palmtop computer and a cloud server. And the computer device 5 can interact with users through a keyboard, a mouse, a remote control, a touchpad or a voice control device.

[0068] The memory 51 includes at least one type of readable storage medium, and can store executable instruction controllers, microcontrollers, microprocessors or other data processing chips, including at least one type of readable storage medium. The readable storage medium includes non-volatile memory or volatile memory, such as flash memory, hard disk, multimedia card, card-type memory (for example, SD or DX memory, etc.), random access memory (RAM), read-only memory (ROM), erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), programmable read-only memory (PROM), magnetic memory, magnetic disk, optical disk, etc.. The aforesaid RAM may include static RAM or dynamic RAM. In some embodiments, the memory 51 may be an internal storage unit of the computer device 5, for example, a hard disk or a memory of the computer device 5. In other embodiments, the memory 51 may also be an external storage unit of the computer device 5, for example, a plug-in hard disk, a Smart Media Card (SMC), a

Secure Digital (SD) card or a Flash Card equipped on the computer device 5. Of course, the memory 51 may also include both the internal storage unit and the external storage unit of the computer device 5. In this embodiment, the memory 51 is generally used to store the operating system and other application software installed on the computer device 5, such as the program codes for executing the first aspect or the fourth aspect described in the present application above. In addition, the memory 51 may also be used to temporarily store various types of data that have been or will be output.

[0069] In the embodiments of the present application, the processor 52 may be a Central Processing Unit (CPU) or a Graphic Processing Unit (GPU), and may also be other general-purpose processor, Digital Signal Processing (DSP), Application Specific Integrated Circuit (ASIC), Field-Programmable Gate Array (FPGA) or other programmable logic device, discrete gate or transistor logic device, discrete hardware component, etc.. And the aforementioned general-purpose processor may be a microprocessor or any conventional processor, such as a single-chip microcomputer.

[0070] The processor 52 is generally used to control the overall operation of the computer device 5. In this embodiment, the memory 51 is used to store program codes or instructions, with the program codes stored including computer operation instructions and the program codes or instructions or processing data stored by processor 52 for executing the memory 51, such as the program codes for executing the first aspect or the fourth aspect described in the present application above.

[0071] The bus described herein may be an Industry Standard Architecture (ISA) bus, a Peripheral Component Interconnect (PCI) bus or an Extended Industry Standard Architecture (EISA) bus, etc., and can be divided into address bus, data bus, control bus and so on. For ease of representation, only one thick line is used in the accompanying drawings to represent the buses involved, which does not indicate that there is only one or one type of bus.

[0072] Sixthly, the embodiments of the present invention also provide a chip used to implement the method described in the first aspect as mentioned above. Specifically, the chip is comprised of a processor designed to call and run a computer program from the memory, making it possible to implement any method described in the first aspect

or the fourth aspect as mentioned above using any device installed with the chip.

[0073] Seventhly, the embodiments of the present invention also provide a computer-readable storage medium which may be a computer-readable signal medium or a computer-readable medium. The computer-readable program codes stored in the computer-readable medium can be read by the processor in computer, allowing the processor to perform the functional tasks set out in each step or any combination of steps as shown in FIG. 1 or FIG. 3 and generate a device that perform the functional tasks set out in each block or any combination of blocks as shown in the diagram.

[0074] The computer-readable medium includes, but is not limited to, an electronic, magnetic, optical, electromagnetic or infrared memory, or a semiconductor system, device or unit, or any appropriate combination of the foregoing. The memory is used to store program codes or instructions, which program codes stored contain computer operation instructions. And the processor is used to execute the program codes or instructions stored in the memory.

[0075] For the definitions of the memory and the processor, reference may be made to the description of the foregoing embodiment of computer device, and no further details are explained herein. The computer-readable storage medium stores a computer program which makes it possible to implement the method described in any paragraph in the first aspect or in the fourth aspect above when executed by a processor.

[0076] Sixthly, the embodiments of the present invention also provide a computer program product containing a computer program instruction which makes it possible to implement the method described in any paragraph in the first aspect or in the fourth aspect above by computer.

[0077] Those skilled in the art can understand that the exemplary units and algorithms described in conjunction with the embodiments disclosed herein can be implemented by electronic hardware or a combination of software and electronic hardware. And whether those functions are implemented by hardware or software depends on the specific application and design constraints of technical solutions. Those skilled in the art may use different ways to implement the described functions for each specific application, which implementation should not be deemed to go beyond the scope of the

present application.

[0078] In the embodiments of the present application, the system, device and method described may be implemented in other ways. For example, unit division is only about the division of logical functions and there may be other divisions in actual implementation. As another example, multiple units or components may be combined or may be integrated into another system. In addition, different units may be directly or indirectly coupled. And the functional units in the embodiments of the present application may either be integrated into one processing unit or independently exist in the form of physical presence.

[0079] It should be understood that in the embodiments of the present application, the serial numbers of each process do not indicate the sequence in which those processes are executed, and the sequence in which those processes are executed should be determined by their respective functions and internal logics and should not constitute any limitation on the implementation of the embodiments of the present application.

[0080] If the functions aforesaid are implemented in the form of a software functional unit and are sold or used as independent products, they may be stored in a machine-readable storage medium. Therefore, the technical solutions mentioned in the present application may be embodied in the form of a software product, which software product may be stored in a machine-readable storage medium and may include several instructions to enable the electronic device to perform all or part of the technical solutions described in the embodiments of the present application. The aforesaid storage medium may include various media that can store program codes, such as ROM, RAM, removable disk, hard disk, magnetic disk or optical disk.

[0081] The above gives a detailed description of the implementation of the present application, but imposes no limitation on the protection scope of the present application. Those skilled in the art may make changes or replacements within the technical scope disclosed in the present application, provided that those changes or replacements shall fall within the protection scope of the present application.

Claims

WHAT IS CLAIMED IS:

1. A video processor training method, comprising:

inputting a decoded video into a video processor as a training sample, wherein the decoded video is generated by encoding and then decoding an original video;

processing the decoded video and outputting an expected difference by the video processor;

determining a loss function according to a difference between the expected difference and an ideal difference, wherein the ideal difference is a difference between the original video and the decoded video;

adjusting parameters of the video processor according to the value of the loss function;

stopping training when the value of the loss function tends to converge, to obtain a trained video processor.

2. The method according to claim 1, wherein the ideal difference is determined according to a difference between image frames corresponding to the original video and the decoded video respectively.

3. The method according to claim 2, wherein the said determination of the ideal difference according to the difference between the image frames corresponding to the original video and the decoded video respectively comprises:

decomposing the original video in a preset manner to obtain each image frame of the original video;

decomposing the decoded video in the same preset manner to obtain each image frame of the decoded video decomposed;

obtaining a difference between each decoded video frame of the decoded video decomposed and each decoded video frame of the original video decomposed; and

processing the difference between the image frames to obtain the ideal difference.

4. The method according to claim 3, wherein the difference between the image frames

is a difference between pixel values of the image frames.

5. A video processor training device, comprising:

an input module designed to input a decoded video into a video processor as a training sample, wherein the decoded video is generated by encoding and then decoding an original video;

a video processor designed to process the decoded video to obtain an expected difference;

a loss function determination module designed to determine a loss function according to a difference between the expected difference and an ideal difference, wherein the ideal difference is a difference between the original video and the decoded video;

an adjustment module designed to adjust parameters of the video processor according to the value of the loss function and to stop training when the value of the loss function tends to converge to obtain a trained video processor.

6. The device according to claim 5, wherein the loss function determination module comprises:

a first decomposition unit configured to decompose the original video in a preset manner to obtain each image frame of the original video;

a second decomposition unit configured to decompose the decoded video in the same preset manner to obtain each image frame of the decoded video decomposed;

an image frame difference value obtaining unit configured to obtain the difference between the image frames corresponding to the decoded video decomposed and the original video decomposed respectively;

an ideal difference determination unit configured to process the difference between the image frames to obtain the ideal difference.

7. The device according to claim 5, wherein the video processor is comprised of a convolutional neural network which has at least one of the following characteristics: the convolutional neural network is a shallow neural network, or the parameters of the convolutional neural network are less than a preset value.

8. The device according to claim 7, wherein the convolutional neural network is

mobilenet or shufflenet.

9. The method according to claim 7 or 8, wherein the convolutional neural network uses an unsaturated activation function.

10. A video processing method, comprising:

acquiring a to-be-processed video which is a decoded video generated after encoding and decoding;

processing the to-be-processed video and obtaining an expected difference by a video processor which is trained according to any method described in claims 1-4 or is trained using any device described in claims 5-9;

determining a new video according to the expected difference and the to-be-processed video.

11. The method according to claim 10, wherein said determining a new video according to the expected difference and the to-be-processed video comprising: superimposing the expected difference value with the to-be-processed video to form a new video.

12. The method according to claim 11, wherein said superimposing the expected difference value with the to-be-processed video to form a new video comprises:

decomposing the to-be-processed video in a preset manner to obtain image frames of the to-be-processed video decomposed;

decomposing the expected difference in the same preset manner to obtain image frames of the expected difference decomposed;

superimposing the image frames corresponding to the image frames of the to-be-processed video decomposed and the image frames of the expected difference decomposed to form a new video.

13. A video processing device, comprising:

an acquisition module designed to acquire a to-be-processed video which is a decoded video generated after encoding and decoding;

a video processor which is designed to process the to-be-processed video to obtain an expected difference and is trained according to any method described in claims 1-4 or trained using any device described in claims 5-9; and

a new video determination module designed to determine a new video based on the

expected difference and the to-be-processed video.

14. The device according to Claim 13, wherein the new video determination module is used to superimpose the expected difference and the to-be-processed video to form a new video, and comprises:

 a third decomposition unit configured to decompose the to-be-processed video in a preset manner to obtain image frames of the to-be-processed video decomposed;

 a fourth decomposition unit configured to decompose the expected difference in the same preset manner to obtain image frames of the expected difference decomposed;

 a new video forming unit configured to superimpose the image frames corresponding to image frames of the expected difference decomposed and image frames of the expected difference decomposed in order to form a new video.

15. A computer device, comprising: at least one processor, a memory, a communication interface and a communication bus, of which the processor, the memory and the communication interface communicate with each other through the communication bus; and a memory coupled to the at least one processor, which stores an executable instruction that, when executed by the at least one processor, makes it possible to implement any method described in claims 1-4 or claims 10-12.

16. A chip, comprising: a processor designed to call and run a computer program from the memory, making it possible to implement any method described in claims 1-4 or claims 10-12 using any device installed with the chip.

17. A computer program product, wherein the computer program product stores a computer program that when executed by a processor, makes it possible to implement the steps involved in any method described in claims 1-4 or claims 10-12 above.