



(51) International Patent Classification:
H04N 7/18 (2006.01)

(21) International Application Number:
PCT/CN2021/070975

(22) International Filing Date:
08 January 2021 (08.01.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
202010682724.4 15 July 2020 (15.07.2020) CN

(71) Applicant: **ZHEJIANG DAHUA TECHNOLOGY CO., LTD.** [CN/CN]; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN).

(72) Inventors: **ZHANG, Jiayu**; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN). **LU, Zhenshan**; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN). **LI, Zhewei**; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN).

(74) Agent: **METIS IP (CHENGDU) LLC**; Room 401, 4th Floor, Unit 1 of Building 1, No. 269 North Section of Hupan Road, Tianfu New District, Chengdu City, China Free Trade Zone (Sichuan), Chengdu, Sichuan 610213 (CN).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: SYSTEMS AND METHODS FOR VIDEO ANALYSIS

(57) Abstract: The present disclosure relates to systems and methods for video analysis. The systems may obtain a plurality of video data streams to be analyzed. A count of the plurality of video data streams may exceed a computing power of at least one processor of the system. The computing power may indicate a count of video data streams that the at least one processor can handle simultaneously. For each of the plurality of video data streams, the system may extract at least one frame from the video data stream based on a predetermined time interval. The systems may identify, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. The systems may analyze, using at least part of remainder portions of the computing power, the one or more target video data streams.

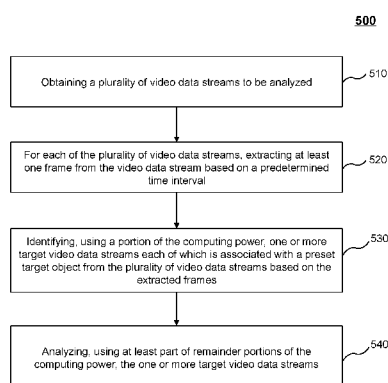


FIG. 5

SYSTEMS AND METHODS FOR VIDEO ANALYSIS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to Chinese Patent Application No. 202010682724.4 filed on July 15, 2020, the contents of which are incorporated herein by reference in their entirety.

TECHNICAL FIELD

[0002] The present disclosure generally relates to monitoring technology, and in particular, to systems and methods for video analysis.

BACKGROUND

[0003] With the development of communication and multimedia technologies, monitoring devices are widely used to monitor preset target objects (e.g., pedestrians, vehicles) in specific scenes. Commonly, a plurality of monitoring devices each of which is set in a scene may be used to capture videos of scenes in real-time, and then each of the captured videos may be analyzed to determine whether an abnormal event (e.g., a conflict) associated with the preset target object occurs in a scene corresponding to the video. However, in some situations, a portion of the captured videos may not include the preset target object or the preset target object occurs in a portion of the captured videos at a low frequency. In such cases, all captured videos are still analyzed, which may cause a waste of analysis resources. Therefore, it is desirable to provide improved systems and methods for video analysis, which can selectively analyze videos or a portion thereof including the preset target object, thereby saving analysis resources.

SUMMARY

[0004] An aspect of the present disclosure relates to a system for video analysis. The system may include at least one storage device including a set of instructions and at least one processor in communication with the at least one storage device.

When executing the set of instructions, the at least one processor may be directed to cause the system to implement operations. The operations may include obtaining a plurality of video data streams to be analyzed. A count of the plurality of video data streams may exceed a computing power of at least one processor of the system.

The computing power may indicate a count of video data streams that the at least one processor can handle simultaneously. The operations may include, for each of the plurality of video data streams, extracting at least one frame from the video data stream based on a predetermined time interval. The operations may include identifying, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. The operations may include analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams.

[0005] In some embodiments, a difference between the predetermined time interval and a time interval between two adjacent frames in each of the plurality of video data streams may be larger than a difference threshold.

[0006] In some embodiments, for each of the plurality of video data streams, the extracting at least one frame from the video data stream based on a predetermined time interval may include determining whether the predetermined time interval is less than a time interval between two adjacent frames with a preset type in the video data stream; in response to determining that the predetermined time interval is less than the time interval between two adjacent frames with the preset type in the video data stream, decoding the video data stream and extracting the at least one frame from the decoded video data stream based on the predetermined time interval.

[0007] In some embodiments, for each of the plurality of video data streams, the extracting at least one frame from the video data stream based on a predetermined time interval further may include, in response to determining that the predetermined time interval is larger than or equal to the time interval between two adjacent frames with the preset type in the video data stream, extracting at least one frame with the

preset type from the video data stream based on the predetermined time interval and determining the at least one frame by decoding the at least one frame with the preset type.

[0008] In some embodiments, the portion of the computing power may be less than or equal to a difference between the computing power of the at least one processor and a predetermined maximum count of video data streams that may include the preset target object.

[0009] In some embodiments, the portion of the computing power may be determined based on the count of the plurality of video data streams, the predetermined time interval, and a processing capacity of the at least one processor.

[0010] In some embodiments, the identifying, using a portion of the computing power, one or more target video data streams each of which includes a preset target object from the plurality of video data streams based on the extracted frames may include, for each of the plurality of video data streams, determining whether the at least one frame extracted from the video data stream includes the preset target object or a part of the preset target object; in response to determining that the at least one frame includes the preset target object or a part of the preset target object, designating the video data stream as the target video data stream.

[0011] In some embodiments, the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams may include, for each of the one or more target video data streams, extracting at least one target frame from the target video data stream based on a processing interval; determining whether an abnormal event associated with the preset target object occurs based on the at least one target frame; and in response to determining that the abnormal event associated with the preset target object occurs, providing a report associated with the abnormal event.

[0012] In some embodiments, the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams may include, for each of the one or more target video data streams, extracting at least

one target frame from the target video data stream based on a processing interval; determining whether the preset target object disappears in the target video data stream based on the at least one target frame; and in response to determining that the preset target object disappears in the target video data stream, stopping the analysis of the target video data stream.

[0013] In some embodiments, the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams may include, for each of the one or more target video data streams, determining whether the target video data stream is being analyzed; and in response to determining that the target video data stream is not being analyzed, activating an idle part of the remainder portions of the computing power to analyze the target video data stream.

[0014] In some embodiments, the operations may further include caching the plurality of video data streams while or before extracting the at least one frame from each of the plurality of video data streams.

[0015] In some embodiments, for each of the plurality of video data streams, a size of a corresponding cached video data stream may be related to the predetermined time interval and a frame rate of the video data stream.

[0016] In some embodiments, before analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams, the at least one processor is directed to perform operations. The operations may further include determining whether a count of the one or more target video data streams exceeds the remainder portions of the computing power of the at least one processor; and in response to determining that the count of the one or more target video data streams exceeds the remainder portions of the computing power of the at least one processor, providing an abnormality alert.

[0017] A further aspect of the present disclosure relates to a method for video analysis. The method may be implemented on a computing device including at least one processor, at least one storage medium, and a communication platform connected to a network. The method may include obtaining a plurality of video data

streams to be analyzed. A count of the plurality of video data streams may exceed a computing power of at least one processor of the system. The computing power may indicate a count of video data streams that the at least one processor can handle simultaneously. The method may include, for each of the plurality of video data streams, extracting at least one frame from the video data stream based on a predetermined time interval. The method may include identifying, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. The method may include analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams.

[0018] A still further aspect of the present disclosure relates to a system for video analysis. The system may include an obtaining module, an extraction module, an identification module, and an analysis module. The obtaining module may be configured to obtain a plurality of video data streams to be analyzed. A count of the plurality of video data streams may exceed a computing power of at least one processor of the system. The computing power may indicate a count of video data streams that the at least one processor can handle simultaneously. The extraction module may be configured to, for each of the plurality of video data streams, extract at least one frame from the video data stream based on a predetermined time interval. The identification module may be configured to identify, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. The analysis module may be configured to analyze, using at least part of remainder portions of the computing power, the one or more target video data streams.

[0019] A still further aspect of the present disclosure relates to a non-transitory computer readable medium including executable instructions. When the executable instructions are executed by at least one processor, the executable instructions may

direct the at least one processor to perform a method. The method may include obtaining a plurality of video data streams to be analyzed. A count of the plurality of video data streams may exceed a computing power of at least one processor of the system. The computing power may indicate a count of video data streams that the at least one processor can handle simultaneously. The method may include, for each of the plurality of video data streams, extracting at least one frame from the video data stream based on a predetermined time interval. The method may include identifying, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. The method may include analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams.

[0020] Additional features will be set forth in part in the description which follows, and in part will become apparent to those skilled in the art upon examination of the following and the accompanying drawings or may be learned by production or operation of the examples. The features of the present disclosure may be realized and attained by practice or use of various aspects of the methodologies, instrumentalities, and combinations set forth in the detailed examples discussed below.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] The present disclosure is further described in terms of exemplary embodiments. These exemplary embodiments are described in detail with reference to the drawings. These embodiments are non-limiting exemplary embodiments, in which like reference numerals represent similar structures throughout the several views of the drawings, and wherein:

[0022] FIG. 1 is a schematic diagram illustrating an exemplary video analysis system according to some embodiments of the present disclosure;

[0023] FIG. 2 is a schematic diagram illustrating exemplary hardware and/or software components of an exemplary computing device according to some embodiments of the present disclosure;

[0024] FIG. 3 is a schematic diagram illustrating exemplary hardware and/or software components of an exemplary mobile device according to some embodiments of the present disclosure;

[0025] FIG. 4 is a block diagram illustrating an exemplary processing device according to some embodiments of the present disclosure;

[0026] FIG. 5 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0027] FIG. 6 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0028] FIG. 7 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0029] FIG. 8 is a flowchart illustrating an exemplary process for forming a first image stream according to some embodiments of the present disclosure;

[0030] FIG. 9 is a flowchart illustrating an exemplary process for forming a first image stream according to some embodiments of the present disclosure;

[0031] FIG. 10 is a flowchart illustrating an exemplary process for analyzing one or more target video data streams according to some embodiments of the present disclosure;

[0032] FIG. 11 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0033] FIG. 12 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0034] FIG. 13 is a flowchart illustrating an exemplary process for analyzing one or more cached target video data streams according to some embodiments of the present disclosure;

[0035] FIG. 14 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0036] FIG. 15 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0037] FIG. 16 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0038] FIG. 17 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure;

[0039] FIG. 18 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure; and

[0040] FIG. 19 is a block diagram illustrating an exemplary video analysis device according to some embodiments of the present disclosure.

DETAILED DESCRIPTION

[0041] In the following detailed description, numerous specific details are set forth by way of examples in order to provide a thorough understanding of the relevant disclosure. However, it should be apparent to those skilled in the art that the present disclosure may be practiced without such details. In other instances, well-known methods, procedures, systems, components, and/or circuitry have been described at a relatively high-level, without detail, in order to avoid unnecessarily obscuring aspects of the present disclosure. Various modifications to the disclosed embodiments will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the present disclosure. Thus, the present disclosure is not limited to the embodiments shown, but to be accorded the widest scope consistent with the claims.

[0042] It will be understood that the terms “system,” “engine,” “unit,” “module,” and/or “block” used herein are one method to distinguish different components, elements, parts, sections, or assemblies of different levels in ascending order.

However, the terms may be displaced by other expressions if they may achieve the same purpose.

[0043] Generally, the words “module,” “unit,” or “block” used herein, refer to logic embodied in hardware or firmware, or to a collection of software instructions. A module, a unit, or a block described herein may be implemented as software and/or hardware and may be stored in any type of non-transitory computer-readable medium or other storage devices. In some embodiments, a software module/unit/block may be compiled and linked into an executable program. It will be appreciated that software modules can be callable from other modules/units/blocks or from themselves, and/or may be invoked in response to detected events or interrupts. Software modules/units/blocks configured for execution on computing devices (e.g., processor 220 illustrated in FIG. 2) may be provided on a computer-readable medium, such as a compact disc, a digital video disc, a flash drive, a magnetic disc, or any other tangible medium, or as a digital download (and can be originally stored in a compressed or installable format that needs installation, decompression, or decryption prior to execution). Such software code may be stored, partially or fully, on a storage device of the executing computing device, for execution by the computing device. Software instructions may be embedded in firmware, such as an EPROM. It will be further appreciated that hardware modules (or units or blocks) may be included in connected logic components, such as gates and flip-flops, and/or can be included in programmable units, such as programmable gate arrays or processors. The modules (or units or blocks) or computing device functionality described herein may be implemented as software modules (or units or blocks), but may be represented in hardware or firmware. In general, the modules (or units or blocks) described herein refer to logical modules (or units or blocks) that may be combined with other modules (or units or blocks) or divided into sub-modules (or sub-units or sub-blocks) despite their physical organization or storage.

[0044] It will be understood that when a unit, an engine, a module, or a block is referred to as being “on,” “connected to,” or “coupled to” another unit, engine, module, or block, it may be directly on, connected or coupled to, or communicate with the other unit, engine, module, or block, or an intervening unit, engine, module, or block may be present, unless the context clearly indicates otherwise. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items.

[0045] The terminology used herein is for the purposes of describing particular examples and embodiments only and is not intended to be limiting. As used herein, the singular forms “a,” “an,” and “the” may be intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “include” and/or “comprise,” when used in this disclosure, specify the presence of integers, devices, behaviors, stated features, steps, elements, operations, and/or components, but do not exclude the presence or addition of one or more other integers, devices, behaviors, features, steps, elements, operations, components, and/or groups thereof.

[0046] In addition, it should be understood that in the description of the present disclosure, the terms “first”, “second”, or the like, are only used for the purpose of differentiation, and cannot be interpreted as indicating or implying relative importance, nor can be understood as indicating or implying the order.

[0047] The flowcharts used in the present disclosure illustrate operations that systems implement according to some embodiments of the present disclosure. It is to be expressly understood, the operations of the flowcharts may be implemented not in order. Conversely, the operations may be implemented in an inverted order, or simultaneously. Moreover, one or more other operations may be added to the flowcharts. One or more operations may be removed from the flowcharts.

[0048] An aspect of the present disclosure relates to systems and methods for video analysis. The systems may obtain a plurality of video data streams to be analyzed. A count of the plurality of video data streams may exceed a computing

power of at least one processor (e.g., a processing device 112) of the video analysis systems. The computing power may indicate a count of video data streams that the at least one processor can handle simultaneously. For each of the plurality of video data streams, the systems may extract at least one frame from the video data stream based on a predetermined time interval. The systems may identify, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. Further, the systems may analyze the one or more target video data streams using at least part of remainder portions of the computing power.

[0049] According to the systems and methods of the present disclosure, one or more target video data streams associated with a preset target object are identified from the plurality of video data streams and may be analyzed using a portion of a computing power of the system. Accordingly, when a count of the plurality of video data streams exceeds the computing power of the system (i.e., the system can't handle all the plurality of video data streams simultaneously in real-time), the system can also realize a simultaneous monitoring of the target object, thereby improving monitoring efficiency and saving analysis resources.

[0050] FIG. 1 is a schematic diagram illustrating an exemplary video analysis system according to some embodiments of the present disclosure. In some embodiments, the video analysis system 100 may be applied in various application scenarios, such as traffic monitoring, security monitoring, etc. As shown, the video analysis system 100 may include a server 110, a network 120, a plurality of monitoring devices 130, a user device 140, and a storage device 150.

[0051] The server 110 may be a single server or a server group. The server group may be centralized or distributed (e.g., the server 110 may be a distributed system). In some embodiments, the server 110 may be local or remote. For example, the server 110 may access information and/or data stored in the monitoring devices 130, the user device 140, and/or the storage device 150 via the network 120. As another example, the server 110 may be directly connected to the monitoring devices 130,

the user device 140, and/or the storage device 150 to access stored information and/or data. In some embodiments, the server 110 may be implemented on a cloud platform. Merely by way of example, the cloud platform may include a private cloud, a public cloud, a hybrid cloud, a community cloud, a distributed cloud, an inter-cloud, a multi-cloud, or the like, or any combination thereof. In some embodiments, the server 110 may be implemented on a computing device 200 including one or more components illustrated in FIG. 2 of the present disclosure.

[0052] In some embodiments, the server 110 may include a processing device 112. The processing device 112 may process information and/or data relating to video analysis to perform one or more functions described in the present disclosure. For example, the processing device 112 may obtain a plurality of video data streams to be analyzed. For each of the plurality of video data streams, the processing device 112 may extract at least one frame from the video data stream based on a predetermined time interval. According to the extracted frames, the processing device 112 may identify, from the plurality of video data streams, one or more target video data streams each of which is associated with a preset target object. Further, the processing device 112 may analyze the one or more target video data streams. In some embodiments, the processing device 112 may include one or more processing devices (e.g., single-core processing device(s) or multi-core processor(s)). Merely by way of example, the processing device 112 may include a central processing unit (CPU), an application-specific integrated circuit (ASIC), an application-specific instruction-set processor (ASIP), a graphics processing unit (GPU), a physics processing unit (PPU), a digital signal processor (DSP), a field-programmable gate array (FPGA), a programmable logic device (PLD), a controller, a microcontroller unit, a reduced instruction set computer (RISC), a microprocessor, or the like, or any combination thereof.

[0053] In some embodiment, the server 110 may be unnecessary and all or part of the functions of the server 110 may be implemented by other components (e.g., the monitoring devices 130, the user device 140) of the video analysis system 100. For

example, the processing device 112 may be integrated into the monitoring devices 130 or the user device 140 and the functions (e.g., video analysis) of the processing device 112 may be implemented by the monitoring devices 130 or the user device 140.

[0054] The network 120 may facilitate exchange of information and/or data for the video analysis system 100. In some embodiments, one or more components (e.g., the server 110, the monitoring devices 130, the user device 140, the storage device 150) of the video analysis system 100 may transmit information and/or data to other component(s) of the video analysis system 100 via the network 120. For example, the server 110 may obtain the plurality of video data streams from the monitoring devices 130 via the network 120. As another example, the server 110 may transmit the one or more target video data streams to the user device 140 via the network 120. In some embodiments, the network 120 may be any type of wired or wireless network, or combination thereof. Merely by way of example, the network 120 may include a cable network (e.g., a coaxial cable network), a wireline network, an optical fiber network, a telecommunications network, an intranet, an Internet, a local area network (LAN), a wide area network (WAN), a wireless local area network (WLAN), a metropolitan area network (MAN), a public telephone switched network (PSTN), a Bluetooth network, a ZigBee network, a near field communication (NFC) network, or the like, or any combination thereof.

[0055] The monitoring devices 130 may be configured to capture image data or video data of a monitoring region. For example, the monitoring devices 130 may continuously capture video data to obtain a video data stream. In some embodiments, the monitoring devices 130 may include a camera 130-1, a video recorder 130-2, an image sensor 130-3, etc. The camera 130-1 may include a gun camera, a dome camera, an integrated camera, a monocular camera, a binocular camera, a multi-view camera, or the like, or any combination thereof. The video recorder 130-2 may include a PC Digital Video Recorder (DVR), an embedded DVR, or the like, or any combination thereof. The image sensor 130-3 may include a

Charge Coupled Device (CCD) image sensor, a Complementary Metal Oxide Semiconductor (CMOS) image sensor, or the like, or any combination thereof. In some embodiments, the monitoring devices 130 may include a plurality of components each of which can capture image data or video data. For example, the monitoring devices 130 may include a plurality of sub-cameras that can capture image data or video data simultaneously. In some embodiments, the monitoring devices 130 may transmit the captured image data or video data to one or more components (e.g., the server 110, the user device 140, the storage device 150) of the video analysis system 100 via the network 120.

[0056] The user device 140 may be configured to receive information and/or data from the server 110, the monitoring devices 130, and/or the storage device 150, via the network 120. For example, the user device 140 may receive the one or more target video data streams from the server 110. In some embodiments, the user device 140 may process information and/or data received from the server 110, the monitoring devices 130, and/or the storage device 150, via the network 120. In some embodiments, the user device 140 may provide a user interface via which a user may view information and/or input data and/or instructions to the video analysis system 100. For example, the user may view the one or more target video data streams via the user interface. As another example, the user may input an instruction associated with the video analysis via the user interface. In some embodiments, the user device 140 may include a mobile phone 140-1, a computer 140-2, a wearable device 140-3, or the like, or any combination thereof. In some embodiments, the user device 140 may include a display that can display information in a human-readable form, such as text, image, audio, video, graph, animation, or the like, or any combination thereof. The display of the user device 140 may include a cathode ray tube (CRT) display, a liquid crystal display (LCD), a light-emitting diode (LED) display, a plasma display panel (PDP), a three-dimensional (3D) display, or the like, or a combination thereof.

[0057] The storage device 150 may be configured to store data and/or instructions. The data and/or instructions may be obtained from, for example, the server 110, the monitoring devices 130, and/or any other component of the video analysis system 100. In some embodiments, the storage device 150 may store data and/or instructions that the server 110 may execute or use to perform exemplary methods described in the present disclosure. In some embodiments, the storage device 150 may include a mass storage, a removable storage, a volatile read-and-write memory, a read-only memory (ROM), or the like, or any combination thereof. Exemplary mass storage may include a magnetic disk, an optical disk, a solid-state drive, etc. Exemplary removable storage may include a flash drive, a floppy disk, an optical disk, a memory card, a zip disk, a magnetic tape, etc. Exemplary volatile read-and-write memory may include a random access memory (RAM). Exemplary RAM may include a dynamic RAM (DRAM), a double data rate synchronous dynamic RAM (DDR SDRAM), a static RAM (SRAM), a thyristor RAM (T-RAM), and a zero-capacitor RAM (Z-RAM), etc. Exemplary ROM may include a mask ROM (MROM), a programmable ROM (PROM), an erasable programmable ROM (EPROM), an electrically erasable programmable ROM (EEPROM), a compact disk ROM (CD-ROM), and a digital versatile disk ROM, etc. In some embodiments, the storage device 150 may be implemented on a cloud platform. Merely by way of example, the cloud platform may include a private cloud, a public cloud, a hybrid cloud, a community cloud, a distributed cloud, an inter-cloud, a multi-cloud, or the like, or any combination thereof.

[0058] In some embodiments, the storage device 150 may be connected to the network 120 to communicate with one or more components (e.g., the server 110, the monitoring devices 130, the user device 140) of the video analysis system 100. One or more components of the video analysis system 100 may access the data or instructions stored in the storage device 150 via the network 120. In some embodiments, the storage device 150 may be directly connected to or communicate with one or more components (e.g., the server 110, the monitoring devices 130, the

user device 140) of the video analysis system 100. In some embodiments, the storage device 150 may be part of other components of the video analysis system 100, such as the server 110, the monitoring devices 130, or the user device 140.

[0059] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations and modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0060] FIG. 2 is a schematic diagram illustrating exemplary hardware and/or software components of an exemplary computing device according to some embodiments of the present disclosure. In some embodiments, the server 110 may be implemented on the computing device 200. For example, the processing device 112 may be implemented on the computing device 200 and configured to perform functions of the processing device 112 disclosed in this disclosure.

[0061] The computing device 200 may be used to implement any component of the video analysis system 100 as described herein. For example, the processing device 112 may be implemented on the computing device 200, via its hardware, software program, firmware, or a combination thereof. Although only one such computer is shown, for convenience, the computer functions relating to video analysis as described herein may be implemented in a distributed fashion on a number of similar platforms to distribute the processing load.

[0062] The computing device 200, for example, may include COM ports 250 connected to and from a network connected thereto to facilitate data communications. The computing device 200 may also include at least one processor (e.g., a processor 220), in the form of one or more processors (e.g., logic circuits), for executing program instructions. For example, the processor 220 may include interface circuits and processing circuits therein. The interface circuits may be configured to receive electronic signals from a bus 210, wherein the electronic signals encode structured data and/or instructions for the processing circuits to

process. The processing circuits may conduct logic calculations, and then determine a conclusion, a result, and/or an instruction encoded as electronic signals. Then the interface circuits may send out the electronic signals from the processing circuits via the bus 210.

[0063] The computing device 200 may further include program storage and data storage of different forms including, for example, a disk 270, a read-only memory (ROM) 230, or a random-access memory (RAM) 240, for storing various data files to be processed and/or transmitted by the computing device 200. The computing device 200 may also include program instructions stored in the ROM 230, RAM 240, and/or another type of non-transitory storage medium to be executed by the processor 220. The methods and/or processes of the present disclosure may be implemented as the program instructions. The computing device 200 may also include an I/O component 260, supporting input/output between the computing device 200 and other components. The computing device 200 may also receive programming and data via network communications.

[0064] Merely for illustration, only one processor is illustrated in FIG. 2. Multiple processors 220 are also contemplated; thus, operations and/or method steps performed by one processor 220 as described in the present disclosure may also be jointly or separately performed by the multiple processors. For example, if in the present disclosure the processor 220 of the computing device 200 executes both step A and step B, it should be understood that step A and step B may also be performed by two different processors 220 jointly or separately in the computing device 200 (e.g., a first processor executes step A and a second processor executes step B, or the first and second processors jointly execute steps A and B).

[0065] FIG. 3 is a schematic diagram illustrating exemplary hardware and/or software components of an exemplary mobile device according to some embodiments of the present disclosure. In some embodiments, the user device 140 may be implemented on the mobile device 300 shown in FIG. 3.

[0066] As illustrated in FIG. 3, the mobile device 300 may include a communication platform 310, a display 320, a graphic processing unit (GPU) 330, a central processing unit (CPU) 340, an I/O 350, a memory 360, and a storage 390. In some embodiments, any other suitable component, including but not limited to a system bus or a controller (not shown), may also be included in the mobile device 300.

[0067] In some embodiments, an operating system 370 (e.g., iOS™, Android™, Windows Phone™) and one or more applications (Apps) 380 may be loaded into the memory 360 from the storage 390 in order to be executed by the CPU 340. The applications 380 may include a browser or any other suitable mobile apps for receiving and rendering information relating to video analysis or other information from the processing device 112. User interactions may be achieved via the I/O 350 and provided to the processing device 112 and/or other components of the video analysis system 100 via the network 120.

[0068] FIG. 4 is a block diagram illustrating an exemplary processing device according to some embodiments of the present disclosure. The processing device 112 may include an obtaining module 410, an extraction module 420, an identification module 430, and an analysis module 440.

[0069] The obtaining module 410 may be configured to obtain a plurality of video data streams (e.g., N video data streams) to be analyzed. More descriptions regarding the obtaining of the plurality of video data streams may be found elsewhere in the present disclosure, for example, operation 510 in FIG. 5 and relevant descriptions thereof.

[0070] The extraction module 420 may be configured to, for each of the plurality of video data streams, extract at least one frame from the video data stream based on a predetermined time interval. More descriptions regarding the extraction of the at least one frame may be found elsewhere in the present disclosure, for example, operation 520 in FIG. 5 and relevant descriptions thereof.

[0071] The identification module 430 may be configured to identify, using a portion (e.g., X channels of the processing device 112) of the computing power, one or more

target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames. More descriptions regarding the identification of the one or more target video data streams may be found elsewhere in the present disclosure, for example, operation 530 in FIG. 5 and relevant descriptions thereof.

[0072] The analysis module 440 may be configured to analyze the one or more target video data streams using at least part of remainder portions (e.g., M-X channels of the processing device 112) of the computing power. More descriptions regarding the analysis of the one or more target video data streams may be found elsewhere in the present disclosure, for example, operation 540 in FIG. 5 and relevant descriptions thereof.

[0073] In some embodiments, the processing device 112 may also include a cache module (not shown) configured to cache the plurality of video data streams. More descriptions regarding the caching of the plurality of video data streams may be found elsewhere in the present disclosure, for example, operation 640 in FIG. 6 and relevant descriptions thereof. The analysis module 440 may be further configured to analyze one or more cached target video data streams using the at least part of remainder portions of the computing power. More descriptions regarding the analysis of the one or more cached target video data streams may be found elsewhere in the present disclosure, for example, operation 650 in FIG. 6 and relevant descriptions thereof.

[0074] The modules in the processing device 112 may be connected to or communicate with each other via a wired connection or a wireless connection. The wired connection may include a metal cable, an optical cable, a hybrid cable, or the like, or any combination thereof. The wireless connection may include a Local Area Network (LAN), a Wide Area Network (WAN), a Bluetooth, a ZigBee, a Near Field Communication (NFC), or the like, or any combination thereof.

[0075] In some embodiments, two or more of the modules may be combined as a single module, and any one of the modules may be divided into two or more units.

For example, the obtaining module 410 and the cache module may be combined as a single module which may both obtain the plurality of video data streams and cache the plurality of video data streams.

[0076] In some embodiments, the processing device 112 may include one or more additional modules. For example, the processing device 112 may also include a transmission module configured to transmit signals (e.g., electrical signals, electromagnetic signals) to one or more components (e.g., the monitoring devices 130, the user device 140) of the video analysis system 100. As another example, the processing device 112 may include a storage module (not shown) used to store information and/or data (e.g., the plurality of video data streams, the count of video data streams that the processing device 112 can handle simultaneously, the predetermined time interval, the one or more target video data streams) associated with the video analysis.

[0077] FIG. 5 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 500 may be executed by the video analysis system 100. For example, the process 500 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 500. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 500 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 500 illustrated in FIG. 5 and described below is not intended to be limiting.

[0078] In 510, the processing device 112 (e.g., the obtaining module 410) (e.g., the interface circuits of the processor 220) may obtain a plurality of video data streams (e.g., N video data streams) to be analyzed. As used herein, a video data stream may refer to a continuous video.

[0079] In some embodiments, the processing device 112 may obtain the plurality of video data streams (e.g., N video data streams) from a plurality of monitoring devices (e.g., the monitoring devices 130) that are used to monitor one or more monitoring regions. Merely by way of example, the processing device 112 may direct a monitoring device to continuously capture a video of a scene and obtain a video data stream of the scene accordingly.

[0080] In some embodiments, the plurality of video data streams may be previously acquired by the monitoring devices 130 and stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). The processing device 112 may obtain the plurality of video data streams from the storage device via a network (e.g., the network 120).

[0081] In some embodiments, a count (e.g., N) of the plurality of video data streams may exceed a computing power of the processing device 112. The computing power may indicate a count (e.g., M) of video data streams (e.g., real-time video data streams) that the processing device 112 can handle simultaneously. For example, the processing device 112 may include a plurality of channels (e.g., M channels) each of which may be used to handle a video data stream, accordingly, a count of the plurality of channels can be understood as the computing power of the processing device 112. In some embodiments, the count (e.g., N) of the plurality of video data streams may far exceed the computing power of the processing device 112 (e.g., $N \gg M$). For example, N may be 300 and M may be 15. As another example, N may be 1000 and M may be 50. As a further example, N may be 5000 and M may be 100.

[0082] In 520, for each of the plurality of video data streams, the processing device 112 (e.g., the extraction module 420) (e.g., the processing circuits of the processor

220) may extract at least one frame from the video data stream based on a predetermined time interval.

[0083] In some embodiments, the predetermined time interval may be a default setting (e.g., an experience value) of the video analysis system 100 or may be adjustable under different situations. For example, the predetermined time interval may be larger than a time interval between two adjacent frames in the video data stream. In some embodiments, a difference between the predetermined time interval and the time interval between two adjacent frames in the video data stream may be larger than a difference threshold. The difference threshold may be a default setting (e.g., an experience value) of the video analysis system 100 or may be adjustable under different situations. In some embodiments, the difference threshold may be relatively large to make the predetermined time interval much larger than the time interval between the two adjacent frames in the video data stream, which can save computing resources, thereby achieving performance optimization. For example, the difference threshold may be larger than or equal to a predetermined (e.g., 10, 50, 100, 500, 1000) times the time interval between two adjacent frames.

[0084] In some embodiments, for different video data streams, predetermined time intervals may be the same or different. In some embodiments, the predetermined time interval may be previously set and stored in the storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). The processing device 112 may obtain the predetermined time interval from the storage device via a network (e.g., the network 120).

[0085] In some embodiments, the processing device 112 may determine whether the predetermined time interval is less than a time interval between two adjacent frames with a preset type (e.g., I frame type) in the video data stream. In response to determining that the predetermined time interval is less than the time interval between two adjacent frames with the preset type in the video data stream, the processing device 112 may decode the video data stream and extract the at least

one frame from the decoded video data stream based on the predetermined time interval. For example, the processing device 112 may extract a frame from a portion of the decoded video data stream at a time point 1, extract another frame from a portion of the decoded video data stream at a time point 2 after the predetermined time interval, etc. In response to determining that the predetermined time interval is larger than or equal to the time interval between two adjacent frames with the preset type in the video data stream, the processing device 112 may extract at least one frame (e.g., at least one I frame) with the preset type from the video data stream based on the predetermined time interval. For example, the processing device 112 may extract a frame with the preset type from a portion of the video data stream at a time point 1, extract another frame with the preset type from a portion of the decoded video data stream at a time point 2 after the predetermined time interval, etc. Further, the processing device 112 may determine the at least one frame by decoding the at least one frame with the preset type.

[0086] That is, if the predetermined time interval is relatively large (e.g., larger than the time interval between two adjacent I frames), the processing device 112 may first directly extract at least one I frame and further decode the at least one I frame, instead of decoding the entire video data stream, which can save computing resources. In some embodiments, if the predetermined time interval is larger than the time interval between two adjacent I frames (i.e., there is at least two I frames within the predetermined time interval), the processing device 112 may randomly select at least one I frame from the I frames within the predetermined time interval. More descriptions regarding the extraction of the at least one frame may be found elsewhere in the present disclosure, for example, FIGs. 8-9 and relevant descriptions thereof.

[0087] In some embodiments, as described above, since the count of the plurality of video data streams exceeds the computing power (e.g., the count (e.g., M) of video data streams that the processing device 112 can handle simultaneously) of the processing device 112, the operation for extracting frames from the plurality of video

data streams may be dynamically performed. For example, the processing device 112 may divide the plurality of video data streams into a plurality of groups (e.g., a group 1, a group 2, a group 3, ..., and a group n) each of which may include at least one video data stream. The processing device 112 may extract frames from a portion of the group 1 at a time point 1, extract frames from a portion of the group 1 **at a time point 1' after the predetermined time interval**, etc., extract frames from a portion of the group 2 at a time point 2, extract frames from a portion of the group 2 **at a time point 2' after the predetermined time interval**, etc., ..., and extract frames from a portion of the group n at a time point x, extract frames from a portion of the **group n at a time point x' after the predetermined time interval**, etc. As another example, the processing device 112 may randomly select a video data stream (or multiple video data streams) from the plurality of video data streams and extract at least one frame from the video at a time point 1, randomly select another video data stream (or multiple video data streams) and extract at least one frame from the video at a time point 2, etc.

[0088] In some embodiments, since extracting frames from a video data stream requires a relatively low computation, the processing device 112 may extract the frames from the plurality of video data streams simultaneously or substantially simultaneously. That is, a single channel of the processing device 112 can process multiple video data streams (i.e., extract frames from the video data streams), accordingly, the extraction of the frames may be not limited to the computing power of the processing device 112.

[0089] In 530, the processing device 112 (e.g., the identification module 430) (e.g., the processing circuits of the processor 220) may identify, using a portion (e.g., X channels of the processing device 112) of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames.

[0090] As used herein, the preset target object may include a biological object and/or a non-biological object. The biological object may include a person, an

animal, a plant, or the like, or any combination thereof. The non-biological object may include a natural object (e.g., a mineral), an artifact (e.g., a vehicle), or the like, or any combination thereof.

[0091] In some embodiments, the portion (e.g., X channels of the processing device 112) of the computing power may be less than or equal to a difference between the computing power (e.g., the count (e.g., M) of video data streams that the processing device 112 can handle simultaneously) of the processing device 112 and a predetermined maximum count (e.g., Y) of video data streams that may include the preset target object. In some embodiments, the predetermined maximum count (e.g., Y) of video data streams that may include the preset target object may be a default setting of the video analysis system 100 or may be adjustable under different situations. For example, different time periods or different geographical regions may correspond to different predetermined maximum counts. In some embodiments, the predetermined maximum count may be previously determined and stored in the storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). The processing device 112 may obtain the predetermined maximum count from the storage device via a network (e.g., the network 120).

[0092] In some embodiments, as described above, the portion of the computing power can be understood as a count of channels of the processing device 112 required to process the extracted frames. In some embodiments, the portion of the computing power may be determined based on the count of the plurality of video data streams, the predetermined time interval (which influences a count of the extracted frames), and a processing capacity of the processing device 112. As used herein, the processing capacity of the processing device 112 may indicate a count of frames (or a frame frequency) that a single channel of the processing device 112 can handle simultaneously. For example, the larger the count of the plurality of video data streams is, the smaller the predetermined time interval (or the larger the count of the extracted frames) is, the larger the portion of the computing power may

be. As another example, the larger the processing capacity of the processing device 112, the smaller the portion of the computing power may be.

[0093] In some embodiments, for each of the plurality of video data streams, the processing device 112 may determine whether the at least one frame extracted from the video data stream includes the preset target object or a part of the preset target object. Merely by way of example, the processing device 112 may determine whether the at least one frame includes the preset target object or a part of the preset target object using an image identification manner. Exemplary image identification manner may include image identification based on neural network, image identification based on wavelet moments, image identification based on fractal features, or the like, or any combination thereof. In response to determining that the at least one frame includes the preset target object or a part of the preset target object, the processing device 112 may designate the video data stream as the target video data stream.

[0094] In some embodiments, for each of the plurality of video data streams, the processing device 112 may determine whether the at least one frame extracted from the video data stream includes an object closely related to the preset target object or a part of the object. The object closely related to the preset target object may be a default setting of the video analysis system 100 or may be adjustable under different situations. For example, when the preset target object is a person, the object closely related to the preset target object may include clothing and accessories worn by the person, a vehicle the person is driving, items carried by the person, or the like, or any combination thereof. In response to determining that the at least one frame extracted from the video data stream includes the object closely related to the preset target object or a part of the object, the processing device 112 may designate the video data stream as the target video data stream.

[0095] In 540, the processing device 112 (e.g., the analysis module 440) (e.g., the processing circuits of the processor 220) may analyze the one or more target video

data streams using at least part of remainder portions (e.g., M-X channels of the processing device 112) of the computing power.

[0096] In some embodiments, for each of the one or more target video data streams, the processing device 112 may extracting at least one target frame from the target video data stream based on a processing interval. For example, the processing device 112 may extract a target frame from a portion of the target video data stream at a time point 1, extract another target frame from a portion of the target video data stream at a time point 2 after the processing interval, etc.

[0097] In some embodiments, the processing interval may be a default setting (e.g., an experience value) of the video analysis system 100 or may be adjustable under different situations. In some embodiments, in order to meet the requirements of tracking an abnormal event (e.g., a conflict, an accident, a destruction) associated with the preset target object, the processing interval may be determined based on a type of the abnormal event associated with the preset target object. Different types of abnormal events may correspond to different processing intervals. The quicker a variation frequency of the abnormal event is, the smaller the processing interval may be. For example, for an abnormal event “accident,” the processing interval may be relatively small, which can ensure that relatively many target frames can be extracted and the “accident” can be accurately monitored.

[0098] In some embodiments, the processing interval may be previously determined and stored in the storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). The processing device 112 may obtain the processing interval from the storage device via a network (e.g., the network 120).

[0099] Further, the processing device 112 may determine whether the abnormal event associated with the preset target object occurs based on the at least one target frame. In some embodiments, the processing device 112 may determine whether the abnormal event associated with the preset target object occurs by comparing the at least one target frame with each other. For example, the processing device 112 may compare the at least one target frame with each other to

determine whether a status of the preset target object changes significantly. If the status of the preset target object changes significantly, the processing device 112 may determine the abnormal event associated with the preset target object occurs. In response to determining that the abnormal event associated with the preset target object occurs, the processing device 112 may provide a report associated with the abnormal event. The report may include information (e.g., name, age, job number) of the preset target object, information (e.g., time and place of occurrence) of abnormal event, the target video data stream, or the like, or any combination thereof.

[0100] In some embodiments, the processing device 112 may transmit the report associated with the abnormal event to a user device (e.g., the user device 140) for display or further processing.

[0101] In some embodiments, the processing device 112 may determine whether the preset target object disappears in the target video data stream based on the at least one target frame. In response to determining that the preset target object disappears in the target video data stream, the processing device 112 may stop the analysis of the target video data stream. Further, the processing device 112 may set a status of a corresponding part (e.g., a channel) of the remainder portions of the computing power that is used to analyze the target video data stream as idle.

[0102] In some embodiments, for each of the one or more target video data streams, the processing device 112 may determine whether the target video data stream is being analyzed. In response to determining that the target video data stream is not being analyzed, the processing device 112 may activate an idle part of the remainder portions of the computing power to analyze the target video data stream. In response to determining that the target video data stream is being analyzed, the processing device 112 may maintain the analysis status of the target video data stream.

[0103] In some embodiments, as described above, the portion (e.g., X channels of the processing device 112) of the computing power that is used to identify the one or more target video data streams may be less than or equal to a difference between

the computing power (e.g., the count (e.g., M) of video data streams that the processing device 112 can handle simultaneously) of the processing device 112 and the predetermined maximum count (e.g., Y) of video data streams that may include the preset target object. Accordingly, the remainder portions of the computing power may be larger than or equal to the predetermined maximum count (e.g., Y) of video data streams that may include the preset target object. Generally, an actual count (e.g., Y') of the one or more target video data streams that actually include the preset target object is less than or equal to the predetermined maximum count (e.g., Y) of video data streams that may include the preset target object. Accordingly, the remainder portions of the computing power may be larger than or equal to the actual count of the one or more target video data streams. That is, the remainder portions of the computing power are sufficient to handle the one or more target video data streams.

[0104] In some situations, the remainder portions of the computing power may be insufficient to handle the one or more target video data streams. In some embodiments, the processing device 112 may determine whether the count (e.g., Y') of the one or more target video data streams exceeds the remainder portions of the computing power of the processing device 112. In response to determining that the count (e.g., Y') of the one or more target video data streams exceeds the remainder portions of the computing power of the processing device 112, the processing device 112 may provide an abnormality alert. Further, the processing device 112 may transmit the abnormality alert to the user device (e.g., the user device 140) to alert a user.

[0105] More descriptions regarding the video analysis may be found elsewhere in the present disclosure, for example, FIGs. 6-18 and relevant descriptions thereof.

[0106] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations

and modifications do not depart from the scope of the present disclosure. For example, one or more other optional operations (e.g., a storing operation, a transmitting operation) may be added elsewhere in the process 500. In the storing operation, the processing device 112 may store information and/or data (e.g., the plurality of video data streams, the count of video data streams that the processing device 112 can handle simultaneously, the predetermined time interval, the one or more target video data streams) associated with the video analysis in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390) disclosed elsewhere in the present disclosure. In the transmitting operation, the processing device 112 may transmit the one or more target video data streams, the report associated with the abnormal event, and/or the abnormality alert to the user device 140.

[0107] FIG. 6 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 600 may be executed by the video analysis system 100. For example, the process 600 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 600. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 600 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 600 illustrated in FIG. 6 and described below is not intended to be limiting.

[0108] In 610, the processing device 112 (e.g., the obtaining module 410) (e.g., the interface circuits of the processor 220) may obtain a plurality of video data streams to be analyzed. Operation 610 may be performed in a similar manner as operation 510, and relevant descriptions are not repeated here.

[0109] In 620, for each of the plurality of video data streams, the processing device 112 (e.g., the extraction module 420) (e.g., the processing circuits of the processor 220) may extract at least one frame from the video data stream based on a predetermined time interval. Operation 620 may be performed in a similar manner as operation 520, and relevant descriptions are not repeated here.

[0110] In 630, the processing device 112 (e.g., the identification module 430) (e.g., the processing circuits of the processor 220) may identify, using a portion of the computing power, one or more target video data streams each of which is associated with the preset target object from the plurality of video data streams based on the extracted frames. Operation 630 may be performed in a similar manner as operation 530, and relevant descriptions are not repeated here.

[0111] In 640, the processing device 112 (e.g., the cache module) (e.g., the processing circuits of the processor 220) may cache the plurality of video data streams.

[0112] In some embodiments, the processing device 112 may cache the plurality of video data streams simultaneously or substantially simultaneously with the obtaining of the plurality of video data streams. In some embodiments, the processing device 112 may cache the plurality of video data streams after obtaining the plurality of video data streams. Accordingly, when the one or more target video data streams are identified, the processing device 112 can quickly and efficiently load the cached video data streams and can analyze the one or more target video data streams accordingly.

[0113] In some situations, before a target video data stream is identified, the preset target object may have already appeared in the target video data stream. In order to analyze the target object accurately, previous information associated with the

target object in the target video data stream also needs to be taken into consideration. Accordingly, the caching of the plurality of video data streams can ensure that previous information can be loaded efficiently and quickly.

[0114] In some embodiments, the processing device 112 may cache the plurality of video data streams between performing operation 610 and operation 620, while performing operation 610 or operation 620, or during performing operations 610-640. In some embodiments, the processing device 112 may cache the plurality of video data streams in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390) disclosed elsewhere in the present disclosure and/or an external storage device. In some embodiments, the video analysis system 100 may include a cache device that is used to store the plurality of cached video data streams. In some embodiments, each of the plurality of video data streams may have an identity (ID). According to an ID of each of the one or more target video data streams, the processing device 112 may obtain a cached target video data stream from the storage device, the external storage device, or the cache device.

[0115] In some embodiments, for each of the plurality of video data streams, a size (also referred to as a “cache size”) of a corresponding cached video data stream may be a default setting of the video analysis system 100 or may be adjustable under different situations. In some embodiments, the cache size may be larger than a first threshold, which can ensure that the target video data stream can be efficiently and accurately analyzed. In some embodiments, the cache size may be less than a second threshold, which can reduce storage occupancy. In some embodiments, different video data streams may correspond to different cache sizes.

[0116] In some embodiments, for each of the plurality of video data streams, a size of a corresponding cached video data stream may be related to the predetermined time interval (according to which at least one frame is extracted from the video data stream) and a frame rate of the video data stream. Merely by way of example, the size of a cached video data stream corresponding to a video data stream may be

equal to a product of the predetermined time interval and the frame rate of the video data stream. For example, when the predetermined time interval is 2 seconds, the size of the cached video data stream corresponding to the video data stream may be 2 times the frame rate of the video data stream.

[0117] In some embodiments, the caching may be a dynamic process. Take a specific video data stream as an example, it is assumed that a cache size of the video data stream is 60 seconds, since the video data stream is continuously captured, the processing device 112 may discard cached data corresponding to previous 60 seconds while caching new data at a current time point, which can reduce storage occupancy.

[0118] In 650, the processing device 112 (e.g., the analysis module 440) (e.g., the processing circuits of the processor 220) may analyze one or more cached target video data streams using at least part of remainder portions of the computing power, which may maintain a normal abnormal event detection rate under a premise of optimizing performance. Operation 650 may be performed in a similar manner as operation 540, and relevant descriptions are not repeated here.

[0119] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0120] FIG. 7 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 700 may be executed by the video analysis system 100. For example, the process 700 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device

1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 700. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 700 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 700 illustrated in FIG. 7 and described below is not intended to be limiting.

[0121] The process 700 may be applied to a situation where a video analysis device (e.g., a video analysis device 1900 illustrated in FIG. 19) with M monitoring channels monitors N scenes, wherein a predetermined maximum count Y of scenes that may include a preset target object in the N scenes is known or predefined. It should be noted that N is larger than M and M is larger than Y . In some embodiments, N may be much larger than M . For example, N may be 300, M may be 15, and Y may be 13. More descriptions regarding N , M , and Y may be found elsewhere in the present disclosure, for example, FIG. 5 and relevant descriptions thereof.

[0122] In 710, video data streams of N scenes may be obtained. In some embodiments, operation 710 may be performed by the processing device 112 (e.g., the obtaining module 410 illustrated in FIG. 4) (e.g., the interface circuits of the processor 220) and/or the video analysis device 1900 (e.g., a video obtaining module 1910 illustrated in FIG. 19).

[0123] In some embodiments, the video obtaining module 1910 may obtain the video data streams of N scenes. Specifically, the video obtaining module 1910 may include a plurality of video obtaining units (e.g., N video obtaining units) each of which may correspond to one of the N scenes. For example, for each of the N scenes, a video obtaining unit may be set in the scene and obtain a video data stream of the scene. Accordingly, N video obtaining units may obtain video data streams of N scenes. Merely by way of example, a video obtaining unit may include

a monitoring device (e.g., the monitoring devices 130), for example, a camera (e.g., the camera 130-1).

[0124] In some embodiments, the video obtaining module 1910 or the obtaining module 410 may obtain the video data streams of N scenes from N monitoring devices (e.g., the monitoring devices 130).

[0125] In 720, a predetermined maximum count Y of scenes that may include a preset target object in the N scenes may be obtained. In some embodiments, operation 720 may be performed by the processing device 112 (e.g., the obtaining module 410 illustrated in FIG. 4) (e.g., the interface circuits of the processor 220) and/or the video analysis device 1900 (e.g., a receiving module 1920 illustrated in FIG. 19). In some embodiments, Y may be a default setting of the video analysis system 100 or may be adjustable under different situations.

[0126] In 730, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. In some embodiments, operation 730 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., a first image stream forming module 1930 illustrated in FIG. 19). More descriptions regarding the forming of the first image stream may be found elsewhere in the present disclosure, for example, FIG. 8, FIG. 9, and relevant descriptions thereof.

[0127] In 740, X monitoring channels of the video analysis device 1900 (e.g., an analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include the preset target object. In some embodiments, operation 740 may be performed by the processing device 112 (e.g., the identification module 430 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940).

[0128] In some embodiments, each of the X monitoring channels may include an object recognition unit configured to recognize the preset target object in the frames in the first image stream and determine whether the frames in the first image stream include the preset target object. If one or more of the frames in the first image stream include the preset target object, operation 750 may be performed. If the frames in the first image stream do not include the preset target object, operations 710-740 may be performed again.

[0129] In some embodiments, X may be set according to the first image stream, the predetermined time interval, a hardware processing capability of each monitoring channel, and a hardware processing capability of the object recognition unit of each monitoring channel. For example, the larger a count of frames in the first image stream is, the larger X may be. As another example, the smaller the predetermined time interval is, the larger X may be. Merely by way of example, X may be less than or equal to a difference between M and Y. In some embodiments, the larger N is, the larger the first image stream is, the larger a count (i.e., X) of monitoring channels required may be. In some embodiments, X may be less than or equal to 3.

[0130] In 750, one or more target video data streams of one or more scenes corresponding to one or more frames each of which includes the preset target object may be analyzed (or monitored). In some embodiments, operation 750 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0131] In some embodiments, when one or more frames in the first image stream include the preset target object, a count Y' of the one or more frames may be obtained. Further, whether Y' is larger than a difference between M and X may be determined. If Y' is larger than the difference between M and X, an abnormality alert may be transmitted to a user device (e.g., the user device 140). If Y' is less than or equal to the difference between M and X, operation 750 may be performed. In some embodiments, Y' monitoring channels of the video analysis device 1900

(e.g., an analysis monitoring module 1940 illustrated in FIG. 19) may analyze (or monitor) Y' target video data streams of Y' scenes corresponding to Y' frames each of which includes the preset target object. Each of the Y' monitoring channels may analyze (or monitor) one of the Y' target video data streams. More descriptions regarding the analysis (or monitoring) of the one or more target video data streams may be found elsewhere in the present disclosure, for example, FIG. 10 and relevant descriptions thereof.

[0132] The present disclosure relates to a method for video analysis. The method may be applied to a situation where the video analysis device with M monitoring channels monitors N scenes, wherein the predetermined maximum count Y of scenes that may include the preset target object in the N scenes is known or predefined. Specifically, the video data streams of N scenes and the predetermined maximum count Y of scenes that may include the preset target object in the N scenes may be obtained. For the video data stream of each of N scenes, the at least one frame may be extracted from the video data stream based on the predetermined time interval, and then the plurality of frames corresponding to the N scenes may form the first image stream. Further, X monitoring channels of the video analysis device may determine whether frames in the first image stream include the preset target object. If one or more of the frames in the first image stream include the preset target object, one or more target video data streams of one or more scenes corresponding to one or more frames each of which includes the preset target object may be analyzed (or monitored). As a result, it realizes the analysis of abnormal situations in multiple scenes (e.g., N scenes) and the real-time analysis of the video data streams of N ($N > M$) scenes in the case of only M monitoring channels. Compared with that N monitoring channels are needed to perform real-time analysis on the video data streams of N scenes in the prior art, fewer monitoring channels (e.g., M) are used to implement the real-time analysis of the video data streams of N scenes in the method in the present disclosure, which

can save analysis resources and greatly improve a performance of intelligent analysis.

[0133] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0134] FIG. 8 is a flowchart illustrating an exemplary process for forming a first image stream according to some embodiments of the present disclosure. In some embodiments, process 800 may be executed by the video analysis system 100. For example, the process 800 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 800. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 800 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 800 illustrated in FIG. 8 and described below is not intended to be limiting. In some embodiments, process 800 may be used to implement operation 730 in FIG. 7.

[0135] In 810, a predetermined time interval may be obtained. In some embodiments, operation 810 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the interface circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0136] In some embodiments, for a video data stream of each of N scenes, the predetermined time interval may be checked to ensure that the predetermined time interval is much larger than a frame interval duration of the video data stream, which may achieve performance optimization. The frame interval duration may refer to a time interval between two adjacent frames in the video data stream. In some embodiments, the predetermined time interval may be an empirical value (e.g., 1- 5 seconds) that may be determined according to an actual scene. Merely by way of example, the predetermined time interval may be 2 seconds. More descriptions regarding the predetermined time interval may be found elsewhere in the present disclosure, for example, FIG. 5 and relevant descriptions thereof.

[0137] In 820, for a video data stream of each of N scenes, whether the predetermined time interval is less than a time interval between two adjacent I frames in the video data stream. In some embodiments, operation 820 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0138] In some embodiments, if the predetermined time interval is less than the time interval between two adjacent I frames, operation 830 may be performed. If the predetermined time interval is larger than or equal to the time interval between two adjacent I frames, operation 840 may be performed.

[0139] In 830, the video data stream may be decoded and at least one frame may be extracted from the decoded video data stream based on the predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. In some embodiments, operation 830 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0140] In 840, at least one I frame may be extracted from the video data stream based on the predetermined time interval and the extracted I frame(s) may be decoded and converted. Further, the plurality of decoded and converted I frames corresponding to the N scenes may form the first image stream. In some embodiments, the frame may be a YUV image. In some embodiments, operation 840 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0141] Compared with directly decoding entire video data stream, only at least one I frame in the video data stream is decoded in operation 840, which saves more analysis resources. More descriptions regarding the forming of the first image stream may be found elsewhere in the present disclosure, for example, FIG. 9 and relevant descriptions thereof.

[0142] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0143] FIG. 9 is a flowchart illustrating an exemplary process for forming a first image stream according to some embodiments of the present disclosure. In some embodiments, process 900 may be executed by the video analysis system 100. For example, the process 900 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to

perform the process 900. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 900 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 900 illustrated in FIG. 9 and described below is not intended to be limiting. In some embodiments, process 900 may be used to implement operation 840 in FIG. 8.

[0144] In 910, whether a portion of a video data stream within the predetermined time interval only includes one I frame may be determined. In some embodiments, operation 910 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0145] In some embodiments, if the portion of the video data stream within the predetermined time interval only includes one I frame, operation 920 may be performed. If the portion of the video data stream within the predetermined time interval includes a plurality of I frames, operation 930 may be performed.

[0146] In 920, the I frame may be decoded and converted. Further, a plurality of decoded and converted I frames corresponding to N scenes may form a first image stream. In some embodiments, operation 920 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0147] In 930, an I frame may be extracted from the plurality of I frames and the extracted I frame may be decoded and converted. Further, the plurality of decoded and converted I frames corresponding to the N scenes may form the first image stream. In some embodiments, operation 930 may be performed by the processing device 112 (e.g., the extraction module 420 illustrated in FIG. 4) (e.g., the processing

circuits of the processor 220) and/or the video analysis device 1900 (e.g., the first image stream forming module 1930 illustrated in FIG. 19).

[0148] In some embodiments, a new first image stream may be formed every predetermined time interval.

[0149] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0150] FIG. 10 is a flowchart illustrating an exemplary process for analyzing (or monitoring) one or more target video data streams according to some embodiments of the present disclosure. In some embodiments, process 1000 may be executed by the video analysis system 100. For example, the process 1000 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1000. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1000 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1000 illustrated in FIG. 10 and described below is not intended to be limiting. In some embodiments, process 1000 may be used to implement operation 750 in FIG. 7.

[0151] In 1010, for each of one or more target video data streams of one or more scenes corresponding to one or more frames each of which includes a preset target object, the target video data stream may be decoded and at least one target frame

may be extracted from the decoded target video data stream based on a processing interval to form a second image stream. In some embodiments, operation 1010 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0152] In some embodiments, the processing interval may be less than the predetermined time interval disclosed elsewhere (e.g., FIGs. 7-9) in the present disclosure.

[0153] In some embodiments, when a count of the one or more scenes corresponding to the one or more frames each of which includes the preset target object is Y , for each of the Y scenes, a decoding and frame extraction unit in a monitoring channel of the video analysis device may decode a target video data stream of the scene. Further, at least one target frame may be extracted from the decoded target video data stream of the scene based on the processing interval to form a second image stream corresponding to the scene. It is understandable that the second image stream may be composed of at least one target frame and the Y scenes may correspond to Y second image streams.

[0154] In some embodiments, the count (e.g., Y') of the one or more scenes corresponding to the one or more frames each of which includes the preset target object may be checked to ensure that the count (e.g., Y') of the one or more scenes that may be analyzed (or monitored) at the same time is not larger than a difference between M (a count of monitoring channels of the video analysis device) and X (a count of monitoring channels of the video analysis device that are used to identify the preset target object in frames in a first image stream). When the count (e.g., Y') of the one or more scenes is larger than the difference between M and X , which indicates that an analysis capability of the video analysis device **can't handle this** situation successfully, an abnormality alert may be provided.

[0155] In 1020, whether an abnormal event associated with the preset target object occurs may be determined by analyzing the at least one target frame in the second image stream. In some embodiments, operation 1020 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0156] In some embodiments, when the abnormal event associated with the preset target object does not occur, operation 1030 may be performed. When the abnormal event associated with the preset target object occurs, operation 1040 may be performed.

[0157] In some embodiments, the abnormal event may refer to an abnormal behavior (e.g., a conflict) of the preset target object.

[0158] In 1030, operation may be returned to perform operations 710-740 in FIG. 7 and further perform operations 1010-1020 in FIG. 10. In some embodiments, operation 1030 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0159] In 1040, a report associated with the abnormal event may be provided. In some embodiments, the report may include an alarm signal. In some embodiments, operation 1040 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0160] In some embodiments, for each of the one or more target video data streams, whether the preset target object disappears in the target video data stream may be determined by analyzing the at least one target frame in the second image stream. When the preset target object disappears in the target video data stream, a

monitoring channel of the video analysis device that is used to analyze the target video data stream may be closed.

[0161] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0162] FIG. 11 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1100 may be executed by the video analysis system 100. For example, the process 1100 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1100. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1100 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1100 illustrated in FIG. 11 and described below is not intended to be limiting.

[0163] The process 1100 may be applied to a situation where a video analysis device (e.g., a video analysis device 1900 illustrated in FIG. 19) with M monitoring channels analyzes video data streams of N scenes in the case of a predetermined maximum count Y of scenes that may include a preset target object in the N scenes is known or predefined.

[0164] In 1110, video data streams of N scenes may be obtained. Operation 1110 may be performed in a similar manner as operation 710, and relevant descriptions are not repeated here.

[0165] In 1120, the predetermined maximum count Y of scenes that may include the preset target object in the N scenes may be obtained. Operation 1120 may be performed in a similar manner as operation 720, and relevant descriptions are not repeated here.

[0166] In 1130, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. Operation 1130 may be performed in a similar manner as operation 730, and relevant descriptions are not repeated here.

[0167] In 1140, X monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include the preset target object. In some embodiments, each of the X monitoring channels may include an object recognition unit configured to recognize the preset target object in the frames in the first image stream and determine whether the frames in the first image stream include the preset target object. If one or more of the frames in the first image stream include the preset target object, the object recognition unit may notify identities (IDs) of one or more target video data streams corresponding to the one or more frames to a switching module (not shown) of the video analysis device 1900, and then operation 750 may be performed. If the frames in the first image stream do not include the preset target object, operations 1110-1140 may be performed again. Operation 1140 may be performed in a similar manner as operation 740, and relevant descriptions are not repeated here.

[0168] In 1150, for each of one or more target video data streams of one or more scenes corresponding to one or more frames each of which includes the preset target object, whether the target video data stream is being analyzed may be

determined. In some embodiments, operation 1150 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the switching module (not shown)).

[0169] In some embodiments, the switching module of the video analysis device 1900 may perform operation 1150. Specifically, after receiving the IDs of the one or more target video data streams, the switching module may perform operation 1150. If the target video data stream is being analyzed, operation 1160 may be performed. If the target video data stream is not being analyzed, operation 1170 may be performed.

[0170] In 1160, the target video data stream may be continuously analyzed. In some embodiments, operation 1160 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0171] In 1170, an idle monitoring channel of the video analysis device may be activated to analyze the target video data stream. In some embodiments, operation 1170 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0172] In some embodiments, the switching module may switch the target video data stream to any idle monitoring channel of the video analysis device 1900 to analyze (or monitor) the target video data stream. The analysis of the target video data stream may be performed in a similar manner as operation 750, and relevant descriptions are not repeated here. In these embodiments, the switching of the target video data stream may avoid the repeated analysis of the target video data stream, thereby saving analysis resources, improving a performance of intelligent

analysis, and maintaining a normal abnormal event detection rate under a premise of optimizing performance.

[0173] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0174] FIG. 12 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1200 may be executed by the video analysis system 100. For example, the process 1200 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1200. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1200 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1200 illustrated in FIG. 12 and described below is not intended to be limiting.

[0175] It should be noted that although the process for determining whether frames in a first image stream include a preset target object is fast and takes a short time, it is still time-consuming. In some embodiments, before a frame in the first image stream including the preset target object is identified, the preset target object has already appeared in a target video data stream corresponding to the frame. In order to ensure that the target video data stream before the frame may be recorded

and analyzed, a cache module (not shown) of the video analysis device 1900 may be used to cache video data streams while or before extracting a frame from each of the video data streams.

[0176] In 1210, video data streams of N scenes may be obtained. Operation 1210 may be performed in a similar manner as operation 710, and relevant descriptions are not repeated here.

[0177] In 1220, a predetermined maximum count Y of scenes that may include the preset target object in the N scenes may be obtained. Operation 1220 may be performed in a similar manner as operation 720, and relevant descriptions are not repeated here.

[0178] In 1230, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. Operation 1230 may be performed in a similar manner as operation 730, and relevant descriptions are not repeated here.

[0179] In 1240, X monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include the preset target object. In some embodiments, if one or more of the frames in the first image stream include the preset target object, operation 1250 may be performed. If the frames in the first image stream do not include the preset target object, operations 1210-1240 may be performed again. Operation 1240 may be performed in a similar manner as operation 740, and relevant descriptions are not repeated here.

[0180] In 1250, the video data streams of N scenes may be cached. In some embodiments, operation 1250 may be performed by the processing device 112 (e.g., the obtaining module 410 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the cache module (not shown)). In some embodiments, operation 1250 may be performed while performing operations 1210-1240.

[0181] In some embodiments, N buffer queues may be set to store the video data streams of N scenes, separately. For each of the video data streams of N scenes, a size (also referred to as a “cache size”) of a corresponding cached video data stream may be a product of the predetermined time interval and a frame rate of the video data stream. For example, when the predetermined time interval is 2 seconds, a size of a cached video data stream corresponding to a video data stream may be a product of 2 and a frame rate of the video data stream. More descriptions regarding the cache size may be found elsewhere in the present disclosure, for example, FIG. 5 and relevant descriptions thereof.

[0182] In 1260, one or more cached target video data streams of one or more scenes corresponding to one or more frames each of which includes the preset target object may be analyzed (or monitored). The analysis of the one or more cached target video data streams may be performed in a similar manner as operation 750, and relevant descriptions are not repeated here. More descriptions regarding the analysis of the one or more cached target video data streams may be found elsewhere in the present disclosure, for example, FIG. 13 and relevant descriptions thereof.

[0183] In these embodiments, the caching of the video data streams of N scenes can ensure that the video data streams can be recorded before corresponding frames including the preset target object are identified. Accordingly, the analysis of the one or more cached target video data streams can maintain a normal abnormal event detection rate under a premise of optimizing performance.

[0184] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0185] FIG. 13 is a flowchart illustrating an exemplary process for analyzing (or monitoring) one or more cached target video data streams according to some

embodiments of the present disclosure. In some embodiments, process 1300 may be executed by the video analysis system 100. For example, the process 1300 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1300. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1300 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1300 illustrated in FIG. 13 and described below is not intended to be limiting. In some embodiments, process 1300 may be used to implement operation 1260 in FIG. 12.

[0186] In 1310, for each of one or more cached target video data streams of one or more scenes corresponding to one or more frames each of which includes a preset target object, the cached target video data stream may be decoded and at least one target frame may be extracted from the decoded target video data stream based on a processing interval to form a second image stream. Operation 1310 may be performed in a similar manner as operation 1010, and relevant descriptions are not repeated here.

[0187] In 1320, whether an abnormal event associated with the preset target object occurs may be determined by analyzing the at least one target frame in the second image stream. In some embodiments, when the abnormal event associated with the preset target object does not occur, operation 1330 may be performed. When the abnormal event associated with the preset target object occurs, operation 1340 may be performed. Operation 1320 may be performed in a similar manner as operation 1020, and relevant descriptions are not repeated here.

[0188] In 1330, operation may be returned to perform operations 1210-1250 in FIG. 12 and further perform operations 1310-1320 in FIG. 13. Operation 1330 may be performed in a similar manner as operation 1030, and relevant descriptions are not repeated here.

[0189] In 1340, a report associated with the abnormal event may be provided. Operation 1340 may be performed in a similar manner as operation 1040, and relevant descriptions are not repeated here.

[0190] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0191] FIG. 14 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1400 may be executed by the video analysis system 100. For example, the process 1400 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1400. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1400 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1400 illustrated in FIG. 14 and described below is not intended to be limiting.

[0192] In 1410, video data streams of N scenes may be obtained. Operation 1410 may be performed in a similar manner as operation 710, and relevant descriptions are not repeated here.

[0193] In 1420, a predetermined maximum count Y of scenes that may include a preset target object in the N scenes may be obtained. Operation 1420 may be performed in a similar manner as operation 720, and relevant descriptions are not repeated here.

[0194] In 1430, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. Operation 1430 may be performed in a similar manner as operation 730, and relevant descriptions are not repeated here.

[0195] In 1440, X monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include the preset target object. If one or more of the frames in the first image stream include the preset target object, operation 1450 may be performed. If the frames in the first image stream do not include the preset target object, operations 1410-1440 may be performed again. Operation 1440 may be performed in a similar manner as operation 740 or 1140, and relevant descriptions are not repeated here.

[0196] In 1450, the video data streams of N scenes may be cached. Operation 1450 may be performed in a similar manner as operation 1250, and relevant descriptions are not repeated here.

[0197] In 1460, for each of one or more cached target video data streams of one or more scenes corresponding to one or more frames each of which includes the preset target object, whether the target video data stream is being analyzed may be determined. If the cached target video data stream is being analyzed, operation 1470 may be performed. If the cached target video data stream is not being

analyzed, operation 1480 may be performed. Operation 1460 may be performed in a similar manner as operation 1150, and relevant descriptions are not repeated here.

[0198] In 1470, the cached target video data stream may be continuously analyzed. Operation 1470 may be performed in a similar manner as operation 1160, and relevant descriptions are not repeated here.

[0199] In 1480, an idle monitoring channel of the video analysis device may be activated to analyze the cached target video data stream. In some embodiments, a switching module (not shown) of the video analysis device 1900 may switch the cached target video data stream to any idle monitoring channel of the video analysis device to analyze (or monitor) the cached target video data stream. Operation 1480 may be performed in a similar manner as operation 1170, and relevant descriptions are not repeated here.

[0200] In these embodiments, the caching of the video data streams of N scenes can ensure that the video data streams can be recorded before corresponding frames including the preset target object are identified. Accordingly, the analysis of the one or more cached target video data streams can maintain a normal abnormal event detection rate under a premise of optimizing performance. The switching of the cached target video data stream may avoid the repeated analysis of the cached target video data stream, thereby saving analysis resources, and improving a performance of intelligent analysis.

[0201] It should be noted that processes 700-1400 may be applied to a situation where a video analysis device with M monitoring channels analyzes video data streams of N scenes in the case of the predetermined maximum count Y of scenes that may include the preset target object in the N scenes is known or predefined, which may achieve an analysis capability far beyond a current hardware capability when the N scenes do not include the preset target object or include the preset target object at a low frequency, thereby greatly reducing costs. The technology of combining image stream and video stream is used, that is, a technology with low performance occupancy (e.g., the identification of the preset target object in frames

in the first image stream) is used to filter a large number of analysis requirements (e.g., the video data streams of N scenes), and most of limited resources (e.g., M monitoring channels) may be applied to a technology with high performance occupancy (e.g., the analysis of the target video data stream), which may improve intelligent analysis performance. In addition, the caching of the video data streams and the switching of the cached target video data stream based on notification may maintain a normal abnormal event detection rate under a premise of optimizing performance.

[0202] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0203] FIG. 15 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1500 may be executed by the video analysis system 100. For example, the process 1500 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1500. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1500 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1500 illustrated in FIG. 15 and described below is not intended to be limiting.

[0204] The process 1500 may be applied to a situation where a video analysis device (e.g., a video analysis device 1900 illustrated in FIG. 19) with M monitoring channels monitors N scenes, wherein a predetermined maximum count Y of scenes that may include a preset target object in the N scenes is unknown. It should be noted that N is larger than M and M is larger than Y . In some embodiments, N may be much larger than M .

[0205] In 1510, video data streams of N scenes may be obtained. Operation 1510 may be performed in a similar manner as operation 710, and relevant descriptions are not repeated here.

[0206] In 1520, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. Operation 1520 may be performed in a similar manner as operation 730, and relevant descriptions are not repeated here.

[0207] In 1530, X monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include the preset target object. If one or more of the frames in the first image stream include the preset target object, operation 1540 may be performed. If the frames in the first image stream do not include the preset target object, operations 1510-1530 may be performed again.

[0208] In these embodiments, X may be much smaller than N , which may improve intelligent analysis performance under limited resource conditions. Merely by way of example, X may be set according to the first image stream, the predetermined time interval, a hardware processing capability of each monitoring channel, and a hardware processing capability of an object recognition unit of each monitoring channel. It should be noted that X may be not restricted by Y and M . Operation 1530 may be performed in a similar manner as operation 740, and relevant descriptions are not repeated here.

[0209] In 1540, a count Y' of the one or more frames each of which includes the preset target object may be obtained. In some embodiments, operation 1540 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0210] In some embodiments, whether Y' is larger than a difference between M and X may be determined. If Y' is larger than the difference between M and X , which indicates that an analysis capability of the video analysis device is exceeded, therefore an abnormality alert may be transmitted to a user device (e.g., the user device 140). If Y' is less than or equal to the difference between M and X , operation 1550 may be performed.

[0211] In 1550, Y' monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may analyze (or monitor) Y' target video data streams of Y' scenes corresponding to Y' frames each of which includes the preset target object. Each of the Y' monitoring channels may analyze (or monitor) one of the Y' target video data streams. For example, if Y' is 10, 10 monitoring channels of the video analysis device may analyze (or monitor) 10 target video data streams of 10 scenes corresponding to 10 frames each of which includes the preset target object. Operation 1550 may be performed in a similar manner as operation 750, and relevant descriptions are not repeated here.

[0212] The present disclosure relates to a method for video analysis. The method may analyze or monitor abnormal situations in a plurality of scenes (e.g., N scenes). Specifically, a plurality of frames each of which is extracted from a video data stream of each of N scenes may form a first image stream. Further, one or more target video data streams of one or more scenes in the plurality of scenes corresponding to one or more frames each of which includes the preset target object may be analyzed (or monitored), instead of analyzing all scenes (e.g., N scenes), which may save analysis resources and greatly improve a performance of intelligent analysis. In

addition, the predetermined maximum count Y of scenes that may include the preset target object in N scenes does not need to be set in advance, which has a higher degree of automation.

[0213] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0214] FIG. 16 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1600 may be executed by the video analysis system 100. For example, the process 1600 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1600. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1600 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1600 illustrated in FIG. 16 and described below is not intended to be limiting. The process 1600 may be applied to a situation where a video analysis device with M monitoring channels analyzes video data streams of N scenes. It should be noted that N is larger than M .

[0215] In 1610, video data streams of N scenes may be obtained. Operation 1610 may be performed in a similar manner as operation 710, and relevant descriptions are not repeated here.

[0216] In 1620, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. Operation 1620 may be performed in a similar manner as operation 730, and relevant descriptions are not repeated here.

[0217] In 1630, X monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include a preset target object. If one or more of the frames in the first image stream include the preset target object, operation 1640 may be performed. If the frames in the first image stream do not include the preset target object, operations 1610-1630 may be performed again. Operation 1630 may be performed in a similar manner as operation 740, and relevant descriptions are not repeated here.

[0218] In 1640, a count Y' of the one or more frames each of which includes the preset target object may be obtained. Operation 1640 may be performed in a similar manner as operation 1540, and relevant descriptions are not repeated here.

[0219] In 1650, whether the count Y' has changed may be determined. In some embodiments, operation 1650 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0220] In some embodiments, whether the count Y' has changed may be determined by comparing the count Y' and a count Y' determined when process 1600 was executed last time. The change may include an increase or decrease. When the count Y' has changed, operation 1660 may be performed. When the count Y' has not changed, operation 1690 may be performed.

[0221] In 1660, whether the count Y' increases or decreases may be determined. When the count Y' increases, operation 1670 may be performed. When the count Y' decreases, operation 1680 may be performed. In some embodiments, operation

1660 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0222] In 1670, at least one new monitoring channel of the video analysis device may be activated and used to analyze (or monitor) at least one target video data stream corresponding to at least one increased frame in the first image stream. In some embodiments, operation 1670 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0223] In 1680, at least one monitoring channel that is used to analyze at least one target video data stream corresponding to at least one decreased frame in the first image stream may be closed. In some embodiments, operation 1680 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0224] In 1690, Y' monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may analyze (or monitor) Y' target video data streams of Y' scenes corresponding to Y' frames each of which includes the preset target object. Operation 1690 may be performed in a similar manner as operation 750, and relevant descriptions are not repeated here.

[0225] It should be noted that when it is first determined that one or more of the frames in the first image stream include the preset target object, operation 1690 may be performed after operation 1640 without performing operations 1650-1680.

[0226] In these embodiments, at least one monitoring channel may be activated or closed according to whether the count Y' **increases or decreases**, which may further save analysis resources and further improve the degree of automation.

[0227] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0228] FIG. 17 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1700 may be executed by the video analysis system 100. For example, the process 1700 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1700. The operations of the illustrated process presented below are intended to be illustrative. In some embodiments, the process 1700 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1700 illustrated in FIG. 17 and described below is not intended to be limiting. The process 1700 may include operations 1610-1690 in process 1600 illustrated in FIG. 16 and operations 1710-1720 illustrated in FIG. 17.

[0229] In 1710, for each of Y' target video data streams of Y' scenes corresponding to Y' frames each of which includes the preset target object, whether the preset target object disappears in the target video data stream. When the preset target object disappears in the target video data stream, operation 1720 may be performed. When the preset target object does not disappear in the target video data stream, operations 1610-1690 in FIG. 16, and operation 1710 may be performed again. In some embodiments, operation 1710 may be performed by the processing device 112

(e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0230] In 1720, a monitoring channel that is used to analyze the target video data stream may be closed. In some embodiments, operation 1720 may be performed by the processing device 112 (e.g., the analysis module 440 illustrated in FIG. 4) (e.g., the processing circuits of the processor 220) and/or the video analysis device 1900 (e.g., the analysis monitoring module 1940 illustrated in FIG. 19).

[0231] In 1730, operation may be returned to perform operations 1610-1690 in FIG. 16 and further perform operation 1710 in FIG. 17.

[0232] In these embodiments, the monitoring channel that is used to analyze the target video data stream may be closed after the preset target object disappears in the target video data stream, thereby further saving analysis resources.

[0233] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0234] FIG. 18 is a flowchart illustrating an exemplary process for video analysis according to some embodiments of the present disclosure. In some embodiments, process 1800 may be executed by the video analysis system 100. For example, the process 1800 may be implemented as a set of instructions (e.g., an application) stored in a storage device (e.g., the storage device 150, the ROM 230, the RAM 240, and/or the storage 390). In some embodiments, the processing device 112 (e.g., the processor 220 of the computing device 200, the CPU 340 of the mobile device 300, and/or one or more modules illustrated in FIG. 4) and/or a video analysis device 1900 (e.g., one or more modules illustrated in FIG. 19) illustrated in FIG. 19 may execute the set of instructions and may accordingly be directed to perform the process 1800. The operations of the illustrated process presented below are

intended to be illustrative. In some embodiments, the process 1800 may be accomplished with one or more additional operations not described and/or without one or more of the operations discussed. Additionally, the order of the operations of process 1800 illustrated in FIG. 18 and described below is not intended to be limiting.

[0235] In 1810, video data streams of N scenes may be obtained. Operation 1810 may be performed in a similar manner as operation 710, and relevant descriptions are not repeated here.

[0236] In 1820, for a video data stream of each of N scenes, at least one frame may be extracted from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream. Operation 1820 may be performed in a similar manner as operation 730, and relevant descriptions are not repeated here.

[0237] In 1830, X monitoring channels of the video analysis device (e.g., an analysis monitoring module 1940 illustrated in FIG. 19) may determine whether frames in the first image stream include a preset target object. If one or more of the frames in the first image stream include the preset target object, operation 1850 may be performed. If the frames in the first image stream do not include the preset target object, operations 1810-1830 may be performed again. Operation 1830 may be performed in a similar manner as operation 740, and relevant descriptions are not repeated here. It should be noted that when it is first determined that one or more of the frames in the first image stream include the preset target object, operations 1840-1860 may be performed after operation 1830 without performing operations 1870-1880. When it is determined again that one or more of the frames in the first image stream include the preset target object, operations 1840-1880 may be performed after operation 1830.

[0238] In 1840, the video data streams of N scenes may be cached. In some embodiments, operation 1840 may be performed while performing operations 1810-

1830. Operation 1840 may be performed in a similar manner as operation 1250, and relevant descriptions are not repeated here.

[0239] In 1850, a count Y' of the one or more frames each of which includes the preset target object may be obtained. Operation 1850 may be performed in a similar manner as operation 1540, and relevant descriptions are not repeated here.

[0240] In 1860, Y' monitoring channels of the video analysis device (e.g., the analysis monitoring module 1940 illustrated in FIG. 19) may analyze (or monitor) Y' cached target video data streams of Y' scenes corresponding to Y' frames each of which includes the preset target object. Operation 1860 may be performed in a similar manner as operation 750, and relevant descriptions are not repeated here.

[0241] In 1870, for each of Y' cached target video data streams, whether the cached target video data stream is being analyzed may be determined. If the cached target video data stream is being analyzed, operation 1860 may be performed. If the cached target video data stream is not being analyzed, operation 1880 may be performed. Operation 1870 may be performed in a similar manner as operation 1150, and relevant descriptions are not repeated here.

[0242] In 1880, an idle monitoring channel of the video analysis device may be activated to analyze the cached target video data stream. In some embodiments, a switching module (not shown) of the video analysis device 1900 may switch the cached target video data stream to any idle monitoring channel of the video analysis device to analyze (or monitor) the cached target video data stream. Operation 1880 may be performed in a similar manner as operation 1170, and relevant descriptions are not repeated here.

[0243] In some embodiments, after operation 1850, whether Y' is larger than a difference between M and X may be determined. If Y' is larger than the difference between M and X , which indicates that an analysis capability of the video analysis device is exceeded, therefore an abnormality alert may be transmitted to a user device (e.g., the user device 140). If Y' is less than or equal to the difference between M and X , operation 1860 may be performed. In some embodiments, after

operation 1850, process 1800 may include contents of operations 1650-1680 illustrated in FIG. 16. In some embodiments, after operation 1880, process 1800 may include contents of operations 1710-1720 illustrated in FIG. 17.

[0244] In these embodiments, the caching of the video data streams of N scenes can ensure that the video data streams can be recorded before corresponding frames including the preset target object are identified. Accordingly, the analysis of the one or more cached target video data streams can maintain a normal abnormal event detection rate under a premise of optimizing performance. The switching of the cached target video data stream may avoid the repeated analysis of the cached target video data stream, thereby saving analysis resources, and improving a performance of intelligent analysis.

[0245] It should be noted that processes 1500-1800 may be applied to a situation where Y' monitoring channels that are selected from the M monitoring channels of the video analysis device analyze the video data streams of N scenes in the case of a predetermined maximum count Y of scenes that may include the preset target object in N scenes is unknown, which may achieve an analysis capability far beyond a current hardware capability when the N scenes include the preset target object at a low frequency, thereby greatly reducing costs. The technology of combining image stream and video stream is used, that is, a technology with low performance occupancy (e.g., the identification of the preset target object in frames in the first image stream) is used to filter a large number of analysis requirements (e.g., the video data streams of N scenes), and most of limited resources (e.g., M monitoring channels) may be applied to a technology with high performance occupancy (e.g., the analysis of the target video data stream), which may improve intelligent analysis performance. The caching of the video data streams and the switching of the cached target video data stream based on notification may maintain a normal abnormal event detection rate under a premise of optimizing performance. In addition, at least one monitoring channel may be activated or closed according to

whether the count Y' increases or decreases, which may further save analysis resources and further improve the degree of automation.

[0246] It should be noted that the above description is merely provided for the purposes of illustration, and not intended to limit the scope of the present disclosure. For persons having ordinary skills in the art, multiple variations or modifications may be made under the teachings of the present disclosure. However, those variations and modifications do not depart from the scope of the present disclosure.

[0247] FIG. 19 is a block diagram illustrating an exemplary video analysis device according to some embodiments of the present disclosure. In some embodiments, the video analysis device 1900 or a portion thereof may be integrated into the processing device 112. The video analysis device 1900 may include a video obtaining module 1910, a receiving module 1920, a first image stream forming module 1930, an analysis monitoring module 1940, and a control module 1950.

[0248] The video obtaining module 1910 may be configured to obtain video data streams of N scenes. In some embodiments, the video obtaining module 1910 may be a stand-alone device rather than a module of the video analysis device 1900. More descriptions regarding the obtaining of the video data streams of the N scenes may be found elsewhere in the present disclosure, for example, operation 710 in FIG. 7 and relevant descriptions thereof.

[0249] The receiving module 1920 may be configured to obtain a predetermined maximum count Y of scenes that may include a preset target object in the N scenes. More descriptions regarding the obtaining of the predetermined maximum count Y may be found elsewhere in the present disclosure, for example, operation 720 in FIG. 7 and relevant descriptions thereof.

[0250] The first image stream forming module 1930 may be configured to form a first image stream. In some embodiments, for a video data stream of each of N scenes, the first image stream forming module 1930 may extract at least one frame from the video data stream based on a predetermined time interval. Further, a plurality of frames corresponding to the N scenes may form a first image stream.

More descriptions regarding the forming of the first image stream may be found elsewhere in the present disclosure, for example, FIG. 8, FIG. 9, and relevant descriptions thereof.

[0251] The analysis monitoring module 1940 may be configured to determine whether frames in the first image stream include the preset target object. In some embodiments, the analysis monitoring module 1940 may include an object recognition unit (not shown) configured to recognize the preset target object in the frames in the first image stream and determine whether the frames in the first image stream include the preset target object. If one or more of the frames in the first image stream include the preset target object, the analysis monitoring module 1940 may analyze one or more target video data streams of one or more scenes corresponding to one or more frames each of which includes the preset target object. In some embodiments, the analysis monitoring module 1940 may also include a switching module (not shown). If one or more of the frames in the first image stream include the preset target object, the object recognition unit may notify identities (IDs) of one or more target video data streams corresponding to the one or more frames to the switching module. After receiving the IDs of the one or more target video data streams, the switching module may determine, for each of the one or more target video data streams, whether the target video data stream is being analyzed. If the target video data stream is being analyzed, the analysis monitoring module 1940 may continuously analyze the target video data stream. If the target video data stream is not being analyzed, the switching module may switch the target video data stream to any idle monitoring channel of the analysis monitoring module 1940 to analyze (or monitor) the target video data stream.

[0252] In some embodiments, the analysis monitoring module 1940 may also include a decoding and frame extraction unit (not shown) configured to, for each of one or more target video data streams, decode the target video data stream and extract at least one target frame from the decoded target video data stream based on a processing interval to form a second image stream.

[0253] In some embodiments, the analysis monitoring module 1940 may also include an analysis monitoring unit (not shown) configured to determine whether an abnormal event associated with the preset target object occurs by analyzing the at least one target frame in the second image stream. When the abnormal event associated with the preset target object occurs, the analysis monitoring unit may provide a report associated with the abnormal event.

[0254] The control module 1950 may be configured to control the video obtaining module 1910, the receiving module 1920, the first image stream forming module 1930, and the analysis monitoring module 1940.

[0255] In some embodiments, the video analysis device 1900 may also include a cache module (not shown) configured to cache the video data streams of N scenes. More descriptions regarding the caching of the video data streams of N scenes may be found elsewhere in the present disclosure, for example, operations 1250 in FIG. 12 and relevant descriptions thereof.

[0256] The modules in the video analysis device 1900 may be connected to or communicate with each other via a wired connection or a wireless connection. The wired connection may include a metal cable, an optical cable, a hybrid cable, or the like, or any combination thereof. The wireless connection may include a Local Area Network (LAN), a Wide Area Network (WAN), a Bluetooth, a ZigBee, a Near Field Communication (NFC), or the like, or any combination thereof. In some embodiments, two or more of the modules may be combined as a single module, and any one of the modules may be divided into two or more units. In some embodiments, the processing device 120 may include one or more additional units. In some embodiments, one or more of the units may be omitted. For example, the video obtaining module 1910 and the receiving module 1920 may be combined as a single module. As another example, the video obtaining module 1910 or the receiving module 1920 may be omitted.

[0257] Having thus described the basic concepts, it may be rather apparent to those skilled in the art after reading this detailed disclosure that the foregoing detailed

disclosure is intended to be presented by way of example only and is not limiting. Various alterations, improvements, and modifications may occur and are intended to those skilled in the art, though not expressly stated herein. These alterations, improvements, and modifications are intended to be suggested by this disclosure, and are within the spirit and scope of the exemplary embodiments of this disclosure.

[0258] Moreover, certain terminology has been used to describe embodiments of the present disclosure. For example, the terms “one embodiment,” “an embodiment,” and/or “some embodiments” mean that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present disclosure. Therefore, it is emphasized and should be appreciated that two or more references to “an embodiment” or “one embodiment” or “an alternative embodiment” in various portions of this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined as suitable in one or more embodiments of the present disclosure.

[0259] Further, it will be appreciated by one skilled in the art, aspects of the present disclosure may be illustrated and described herein in any of a number of patentable classes or context including any new and useful process, machine, manufacture, or comlocation of matter, or any new and useful improvement thereof. Accordingly, aspects of the present disclosure may be implemented entirely hardware, entirely software (including firmware, resident software, micro-code, etc.) or combining software and hardware implementation that may all generally be referred to herein as a “unit,” “module,” or “system.” Furthermore, aspects of the present disclosure may take the form of a computer program product embodied in one or more computer readable media having computer-readable program code embodied thereon.

[0260] A computer readable signal medium may include a propagated data signal with computer readable program code embodied therein, for example, in baseband or as part of a carrier wave. Such a propagated signal may take any of a variety of

forms, including electromagnetic, optical, or the like, or any suitable combination thereof. A computer readable signal medium may be any computer readable medium that is not a computer readable storage medium and that may communicate, propagate, or transport a program for use by or in connection with an instruction execution system, apparatus, or device. Program code embodied on a computer readable signal medium may be transmitted using any appropriate medium, including wireless, wireline, optical fiber cable, RF, or the like, or any suitable combination of the foregoing.

[0261] Computer program code for carrying out operations for aspects of the present disclosure may be written in a combination of one or more programming languages, including an object oriented programming language such as Java, Scala, Smalltalk, Eiffel, JADE, Emerald, C++, C#, VB. NET, Python or the like, conventional procedural programming languages, such as the "C" programming language, Visual Basic, Fortran 2003, Perl, COBOL 2002, PHP, ABAP, dynamic programming languages such as Python, Ruby, and Groovy, or other programming languages. The program code may execute entirely on the user's computer, partly on the user's computer, as a stand-alone software package, partly on the user's computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user's computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider) or in a cloud computing environment or offered as a service such as a Software as a Service (SaaS).

[0262] Furthermore, the recited order of processing elements or sequences, or the use of numbers, letters, or other designations thereof, are not intended to limit the claimed processes and methods to any order except as may be specified in the claims. Although the above disclosure discusses through various examples what is currently considered to be a variety of useful embodiments of the disclosure, it is to

be understood that such detail is solely for that purpose, and that the appended claims are not limited to the disclosed embodiments, but, on the contrary, are intended to cover modifications and equivalent arrangements that are within the spirit and scope of the disclosed embodiments. For example, although the implementation of various components described above may be embodied in a hardware device, it may also be implemented as a software only solution, e.g., an installation on an existing server or mobile device.

[0263] Similarly, it should be appreciated that in the foregoing description of embodiments of the present disclosure, various features are sometimes grouped together in a single embodiment, figure, or description thereof for the purpose of streamlining the disclosure aiding in the understanding of one or more of the various embodiments. This method of disclosure, however, is not to be interpreted as reflecting an intention that the claimed subject matter requires more features than are expressly recited in each claim. Rather, claimed subject matter may lie in less than all features of a single foregoing disclosed embodiment.

WHAT IS CLAIMED IS:**1. A system for video analysis, comprising:**

at least one storage device including a set of instructions; and
at least one processor in communication with the at least one storage device,
wherein when executing the set of instructions, the at least one processor is directed
to perform operations including:

obtaining a plurality of video data streams to be analyzed, wherein a count of
the plurality of video data streams exceeds a computing power of the at least
one processor, the computing power indicating a count of video data streams
that the at least one processor can handle simultaneously;

for each of the plurality of video data streams, extracting at least one frame
from the video data stream based on a predetermined time interval;

identifying, using a portion of the computing power, one or more target video
data streams each of which is associated with a preset target object from the
plurality of video data streams based on the extracted frames; and

analyzing, using at least part of remainder portions of the computing power,
the one or more target video data streams.

**2. The system of claim 1, wherein a difference between the predetermined time
interval and a time interval between two adjacent frames in each of the plurality of
video data streams is larger than a difference threshold.**

**3. The system of claim 1 or claim 2, wherein for each of the plurality of video data
streams, the extracting at least one frame from the video data stream based on a
predetermined time interval includes:**

determining whether the predetermined time interval is less than a time interval
between two adjacent frames with a preset type in the video data stream;

in response to determining that the predetermined time interval is less than the
time interval between two adjacent frames with the preset type in the video data

stream,

decoding the video data stream; and

extracting the at least one frame from the decoded video data stream based on the predetermined time interval.

4. The system of claim 3, wherein for each of the plurality of video data streams, the extracting at least one frame from the video data stream based on a predetermined time interval further includes:

in response to determining that the predetermined time interval is larger than or equal to the time interval between two adjacent frames with the preset type in the video data stream,

extracting at least one frame with the preset type from the video data stream based on the predetermined time interval; and

determining the at least one frame by decoding the at least one frame with the preset type.

5. The system of any of claims 1-4, wherein the portion of the computing power is less than or equal to a difference between the computing power of the at least one processor and a predetermined maximum count of video data streams that may include the preset target object.

6. The system of any of claims 1-5, wherein the portion of the computing power is determined based on the count of the plurality of video data streams, the predetermined time interval, and a processing capacity of the at least one processor.

7. The system of any of claims 1-6, wherein the identifying, using a portion of the computing power, one or more target video data streams each of which includes a preset target object from the plurality of video data streams based on the extracted frames includes:

for each of the plurality of video data streams,

determining whether the at least one frame extracted from the video data stream includes the preset target object or a part of the preset target object;

in response to determining that the at least one frame includes the preset target object or a part of the preset target object, designating the video data stream as the target video data stream.

8. The system of any of claims 1-7, wherein the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams includes:

for each of the one or more target video data streams,

extracting at least one target frame from the target video data stream based on a processing interval;

determining whether an abnormal event associated with the preset target object occurs based on the at least one target frame; and

in response to determining that the abnormal event associated with the preset target object occurs, providing a report associated with the abnormal event.

9. The system of any of claims 1-7, wherein the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams includes:

for each of the one or more target video data streams,

extracting at least one target frame from the target video data stream based on a processing interval;

determining whether the preset target object disappears in the target video data stream based on the at least one target frame; and

in response to determining that the preset target object disappears in the

target video data stream, stopping the analysis of the target video data stream.

10. The system of any of claims 1-9, wherein the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams includes:

for each of the one or more target video data streams,

determining whether the target video data stream is being analyzed; and

in response to determining that the target video data stream is not being analyzed, activating an idle part of the remainder portions of the computing power to analyze the target video data stream.

11. The system of any of claims 1-10, wherein the operations further include:

caching the plurality of video data streams while or before extracting the at least one frame from each of the plurality of video data streams.

12. The system of claim 11, wherein for each of the plurality of video data streams, a size of a corresponding cached video data stream is related to the predetermined time interval and a frame rate of the video data stream.

13. The system of any of claims 1-12, wherein before analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams, the at least one processor is directed to perform operations further including:

determining whether a count of the one or more target video data streams exceeds the remainder portions of the computing power of the at least one processor; and

in response to determining that the count of the one or more target video data streams exceeds the remainder portions of the computing power of the at least one processor, providing an abnormality alert.

14. A method implemented on a computing device including at least one processor, at least one storage medium, and a communication platform connected to a network, the method comprising:

obtaining a plurality of video data streams to be analyzed, wherein a count of the plurality of video data streams exceeds a computing power of the at least one processor, the computing power indicating a count of video data streams that the at least one processor can handle simultaneously;

for each of the plurality of video data streams, extracting at least one frame from the video data stream based on a predetermined time interval;

identifying, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames; and

analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams.

15. The method of claim 14, wherein a difference between the predetermined time interval and a time interval between two adjacent frames in each of the plurality of video data streams is larger than a difference threshold.

16. The method of claim 14 or claim 15, wherein for each of the plurality of video data streams, the extracting at least one frame from the video data stream based on a predetermined time interval includes:

determining whether the predetermined time interval is less than a time interval between two adjacent frames with a preset type in the video data stream;

in response to determining that the predetermined time interval is less than the time interval between two adjacent frames with the preset type in the video data stream,

decoding the video data stream; and

extracting the at least one frame from the decoded video data stream based on the predetermined time interval.

17. The method of claim 16, wherein for each of the plurality of video data streams, the extracting at least one frame from the video data stream based on a predetermined time interval further includes:

in response to determining that the predetermined time interval is larger than or equal to the time interval between two adjacent frames with the preset type in the video data stream,

extracting at least one frame with the preset type from the video data stream based on the predetermined time interval; and

determining the at least one frame by decoding the at least one frame with the preset type.

18. The method of any of claims 14-17, wherein the portion of the computing power is less than or equal to a difference between the computing power of the at least one processor and a predetermined maximum count of video data streams that may include the preset target object.

19. The method of any of claims 14-18, wherein the portion of the computing power is determined based on the count of the plurality of video data streams, the predetermined time interval, and a processing capacity of the at least one processor.

20. The method of any of claims 14-19, wherein the identifying, using a portion of the computing power, one or more target video data streams each of which includes a preset target object from the plurality of video data streams based on the extracted frames includes:

for each of the plurality of video data streams,

determining whether the at least one frame extracted from the video data

stream includes the preset target object or a part of the preset target object;

in response to determining that the at least one frame includes the preset target object or a part of the preset target object, designating the video data stream as the target video data stream.

21. The method of any of claims 14-20, wherein the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams includes:

for each of the one or more target video data streams,

extracting at least one target frame from the target video data stream based on a processing interval;

determining whether an abnormal event associated with the preset target object occurs based on the at least one target frame; and

in response to determining that the abnormal event associated with the preset target object occurs, providing a report associated with the abnormal event.

22. The method of any of claims 14-20, wherein the analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams includes:

for each of the one or more target video data streams,

extracting at least one target frame from the target video data stream based on a processing interval;

determining whether the preset target object disappears in the target video data stream based on the at least one target frame; and

in response to determining that the preset target object disappears in the target video data stream, stopping the analysis of the target video data stream.

23. The method of any of claims 14-22, wherein the analyzing, using at least part of

remainder portions of the computing power, the one or more target video data streams includes:

for each of the one or more target video data streams,
determining whether the target video data stream is being analyzed; and
in response to determining that the target video data stream is not being analyzed, activating an idle part of the remainder portions of the computing power to analyze the target video data stream.

24. The method of any of claims 14-23, wherein the method further includes:

caching the plurality of video data streams while or before extracting the at least one frame from each of the plurality of video data streams.

25. The method of claim 24, wherein for each of the plurality of video data streams, a size of a corresponding cached video data stream is related to the predetermined time interval and a frame rate of the video data stream.

26. The method of any of claims 14-25, wherein before analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams, the method further including:

determining whether a count of the one or more target video data streams exceeds the remainder portions of the computing power of the at least one processor; and

in response to determining that the count of the one or more target video data streams exceeds the remainder portions of the computing power of the at least one processor, providing an abnormality alert.

27. A system, comprising:

an obtaining module configured to obtain a plurality of video data streams to be analyzed, wherein a count of the plurality of video data streams exceeds a

computing power of the at least one processor, the computing power indicating a count of video data streams that the at least one processor can handle simultaneously;

an extraction module configured to, for each of the plurality of video data streams, extract at least one frame from the video data stream based on a predetermined time interval;

an identification module configured to identify, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames; and

an analysis module configured to analyze, using at least part of remainder portions of the computing power, the one or more target video data streams.

38. A non-transitory computer readable medium, comprising executable instructions that, when executed by at least one processor, direct the at least one processor to perform a method, the method comprising:

obtaining a plurality of video data streams to be analyzed, wherein a count of the plurality of video data streams exceeds a computing power of the at least one processor, the computing power indicating a count of video data streams that the at least one processor can handle simultaneously;

for each of the plurality of video data streams, extracting at least one frame from the video data stream based on a predetermined time interval;

identifying, using a portion of the computing power, one or more target video data streams each of which is associated with a preset target object from the plurality of video data streams based on the extracted frames; and

analyzing, using at least part of remainder portions of the computing power, the one or more target video data streams.

100

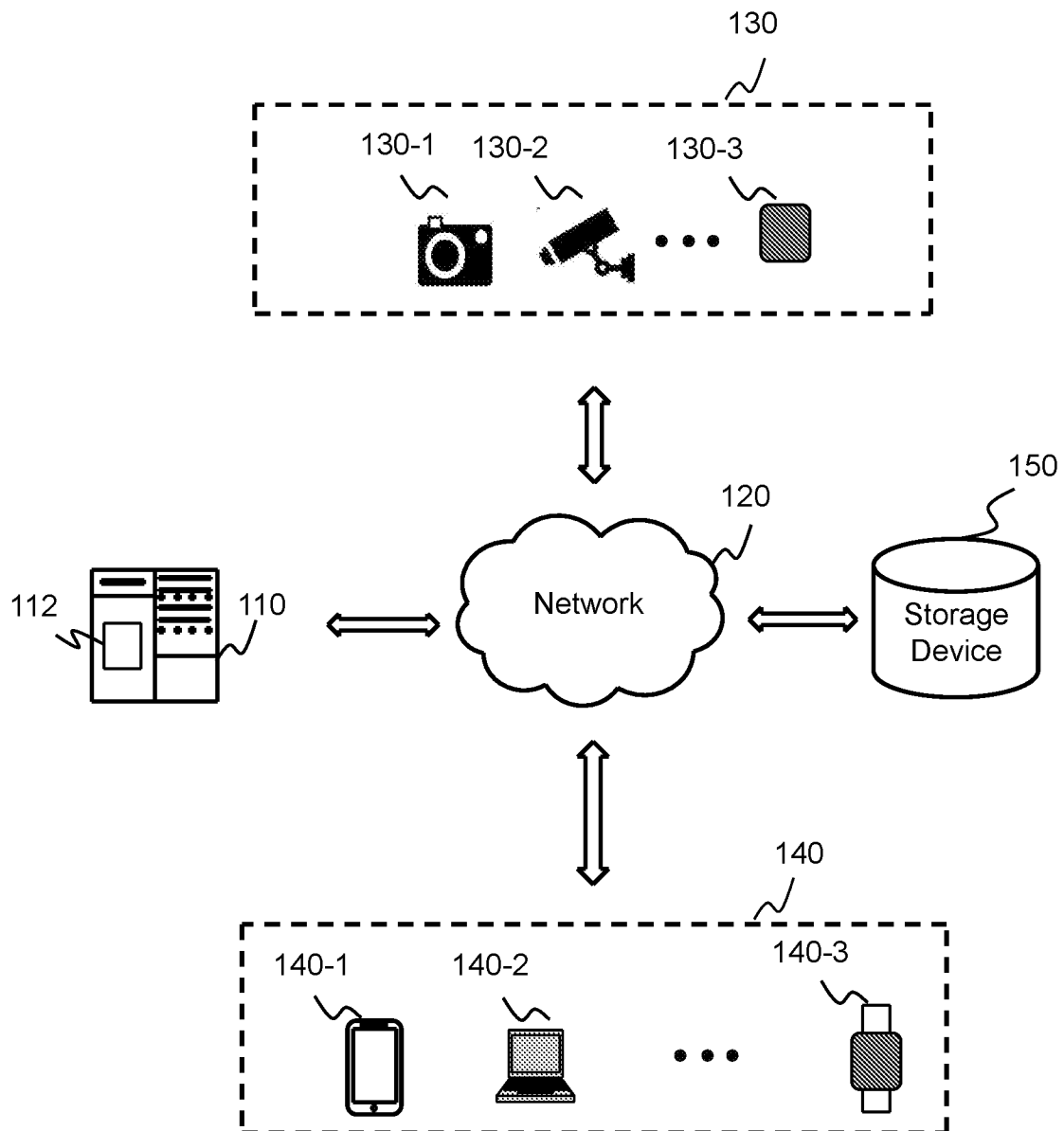
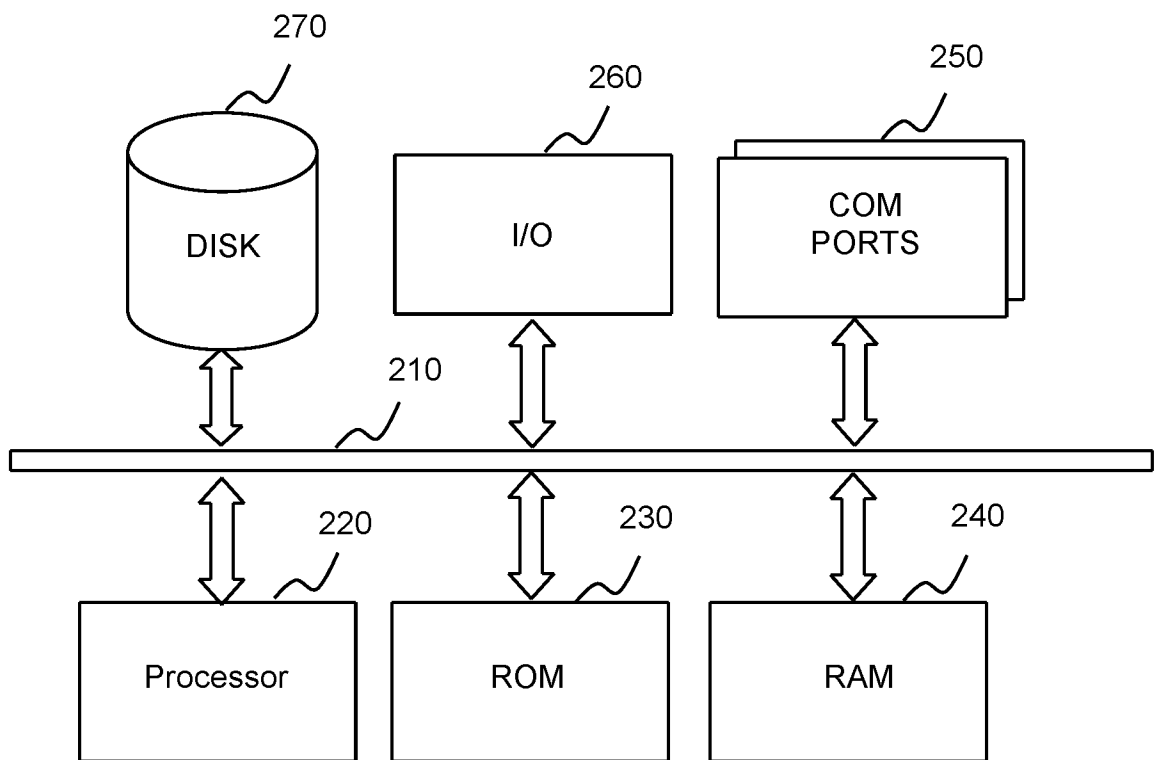
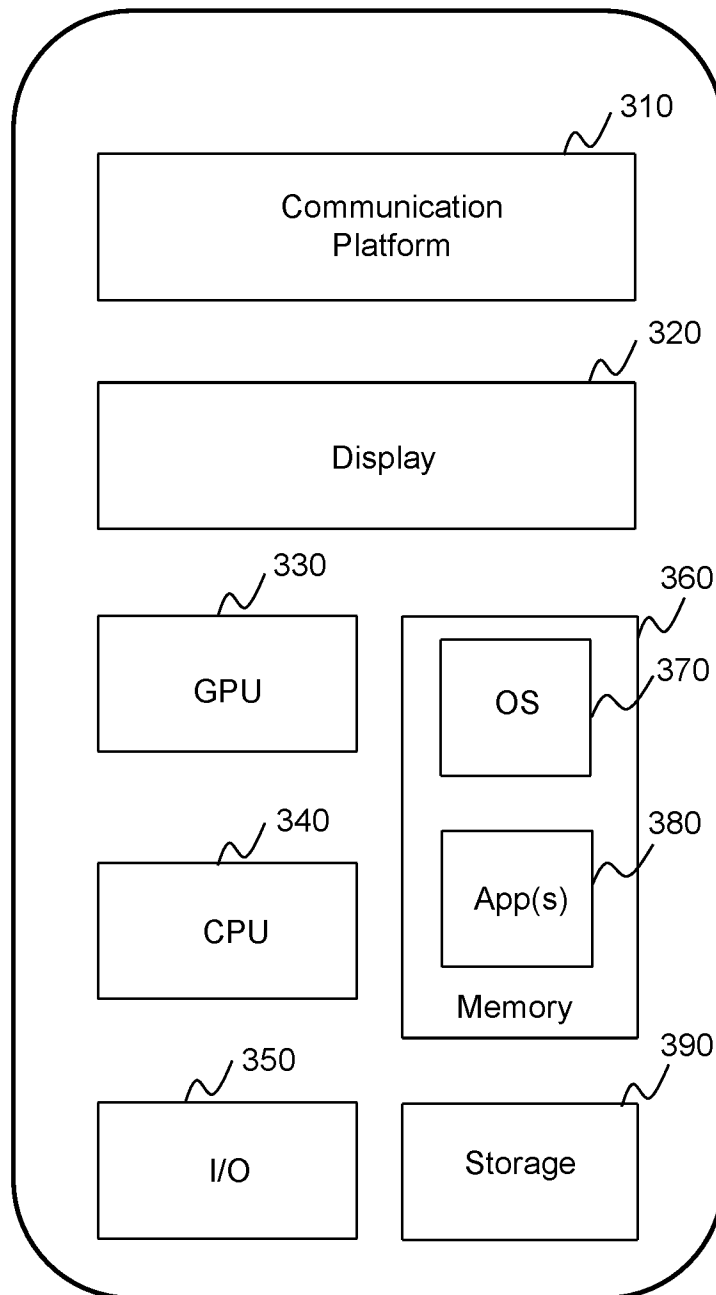
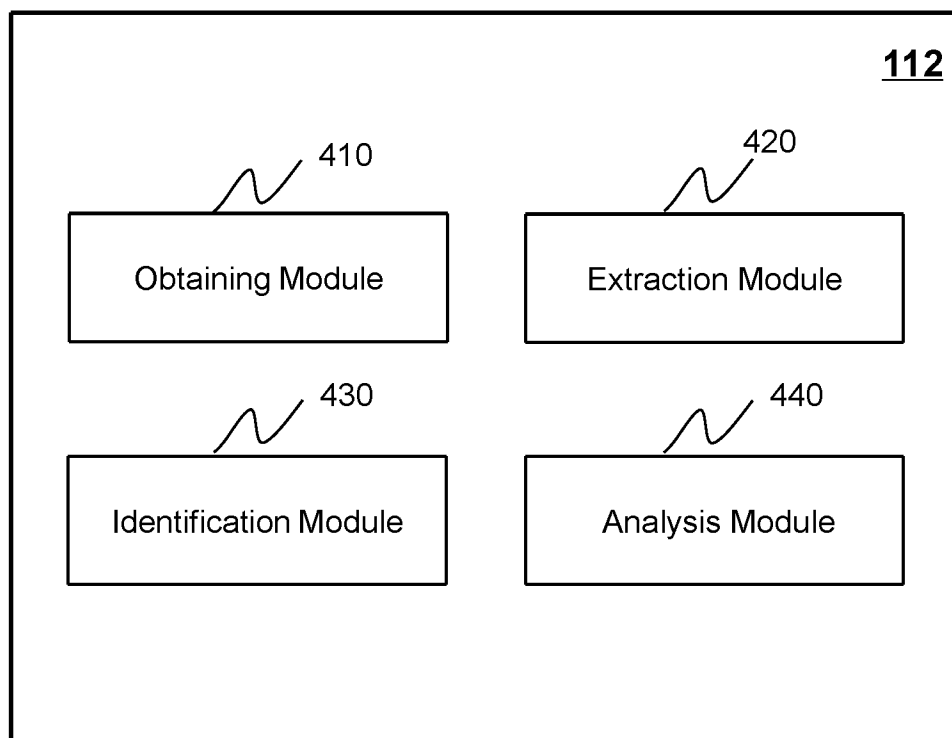


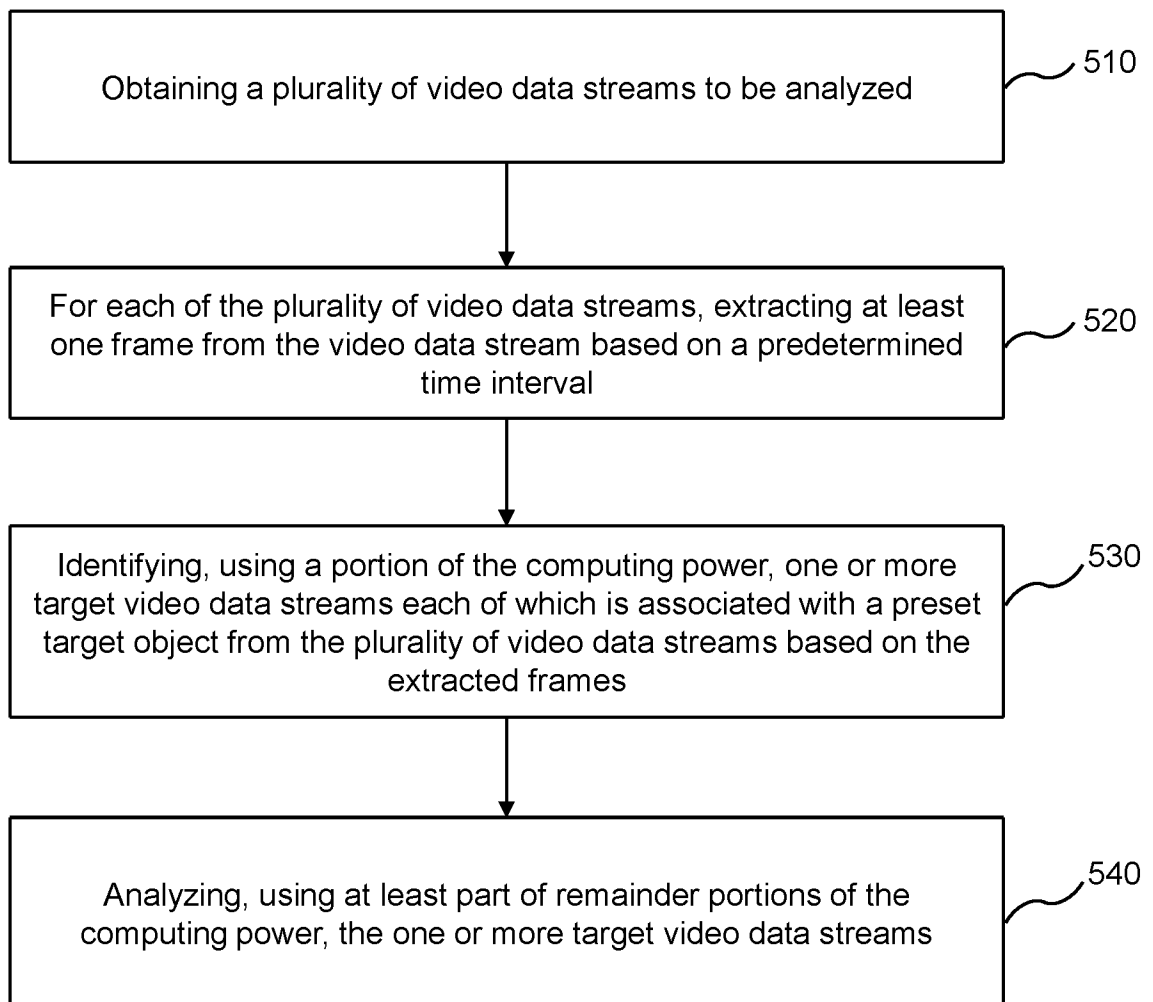
FIG. 1

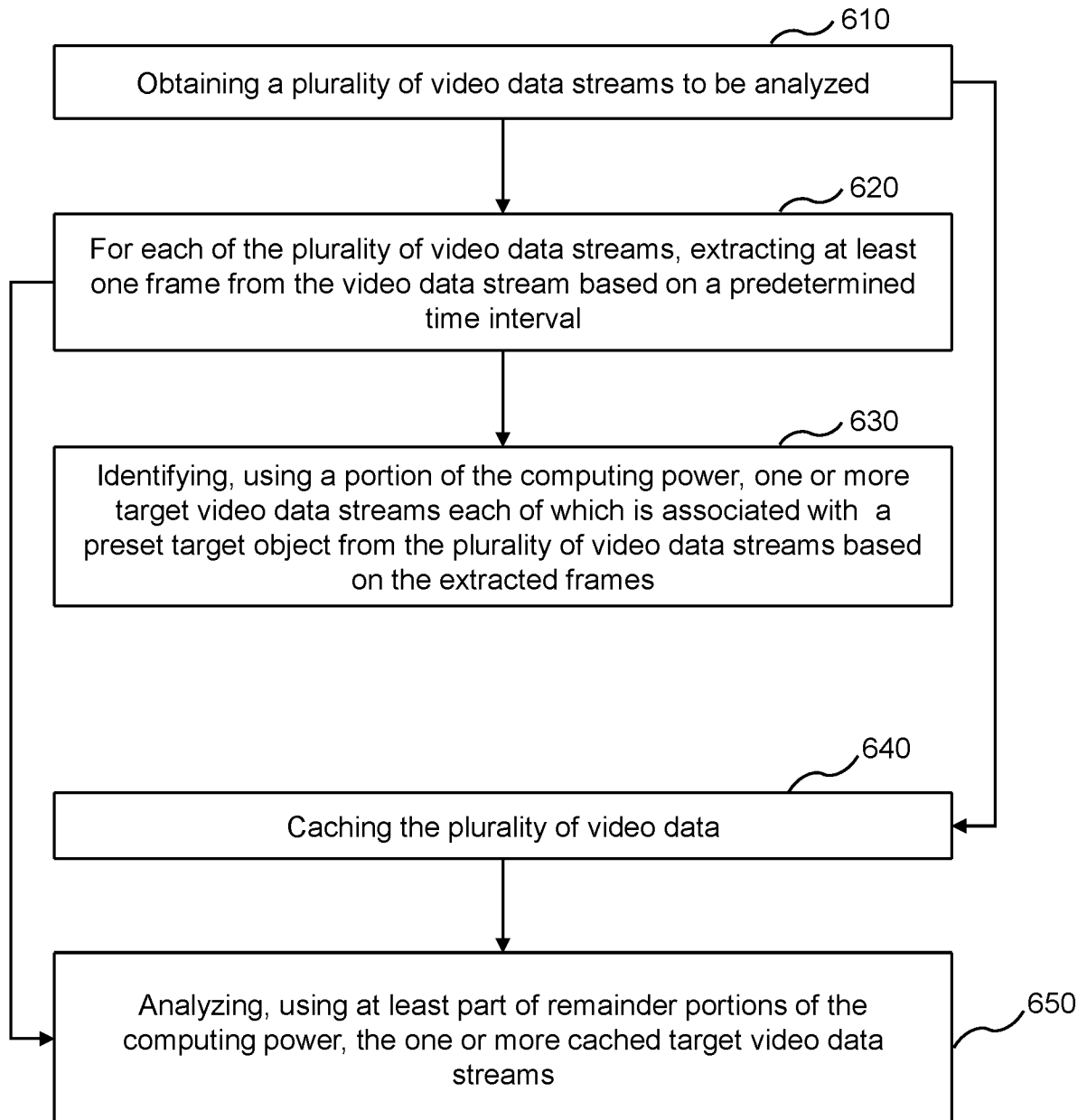
200**FIG. 2**

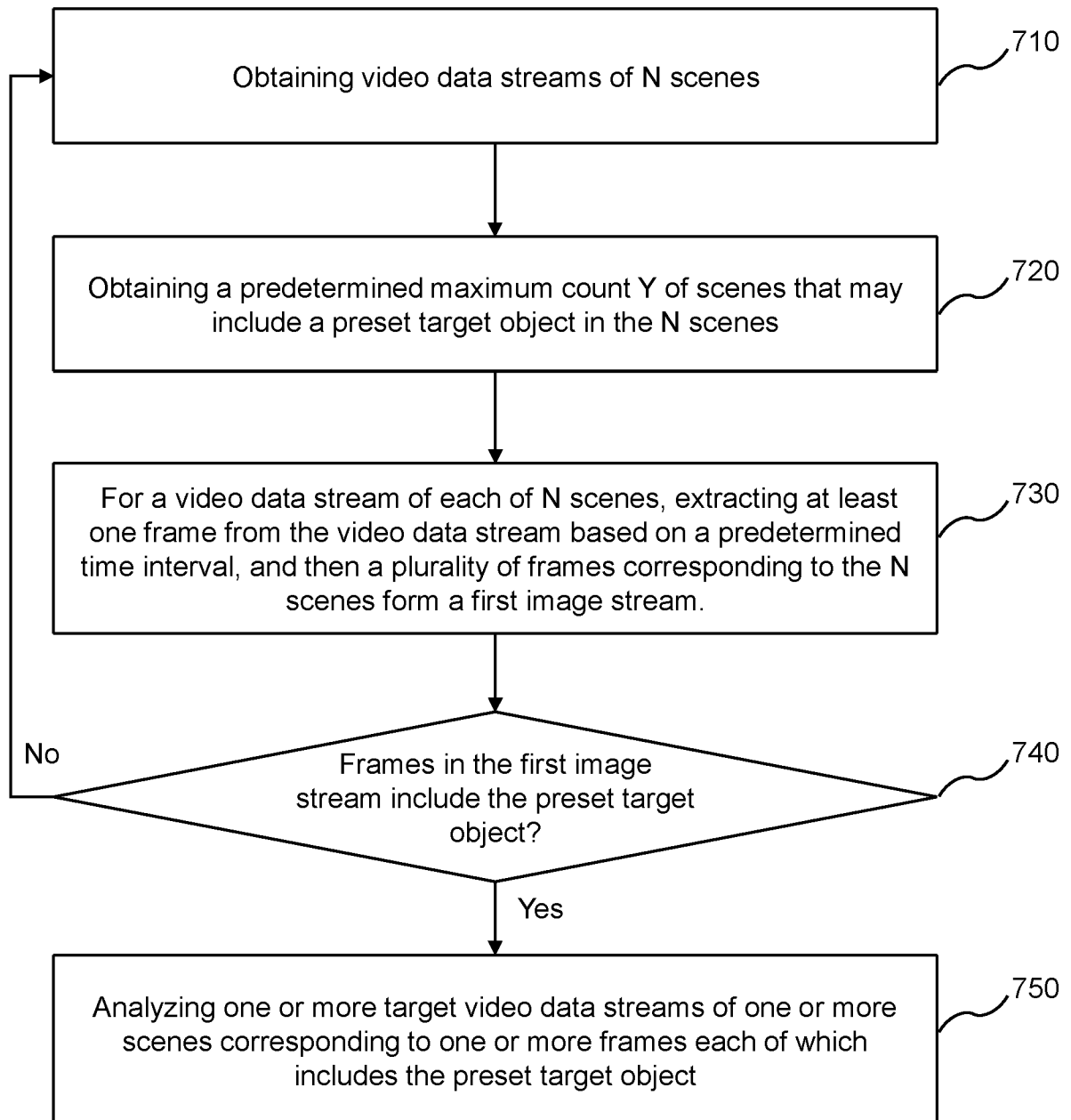
300**FIG. 3**

Sheet 4/19

**FIG. 4**

500**FIG. 5**

**FIG. 6**

700**FIG. 7**

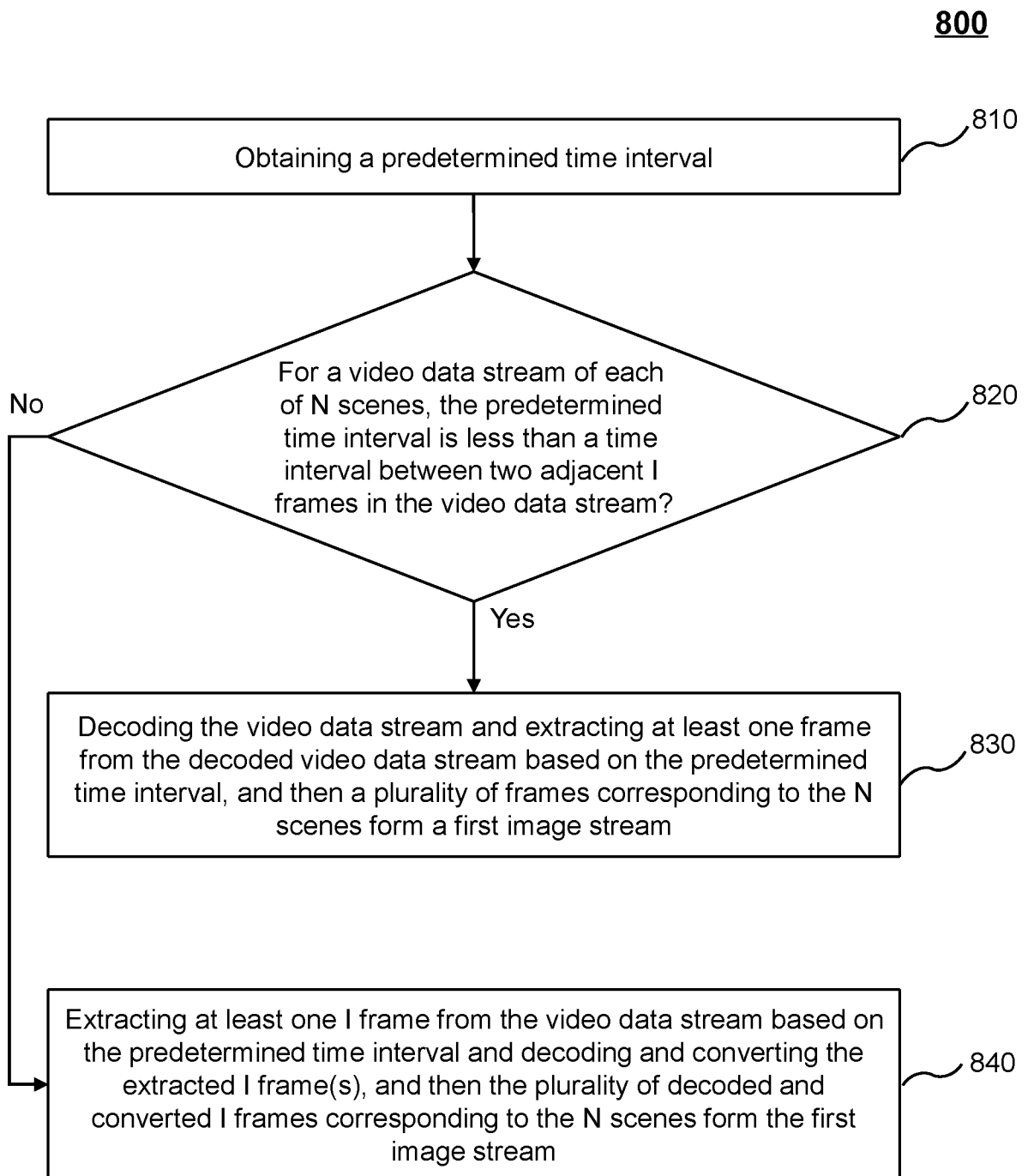
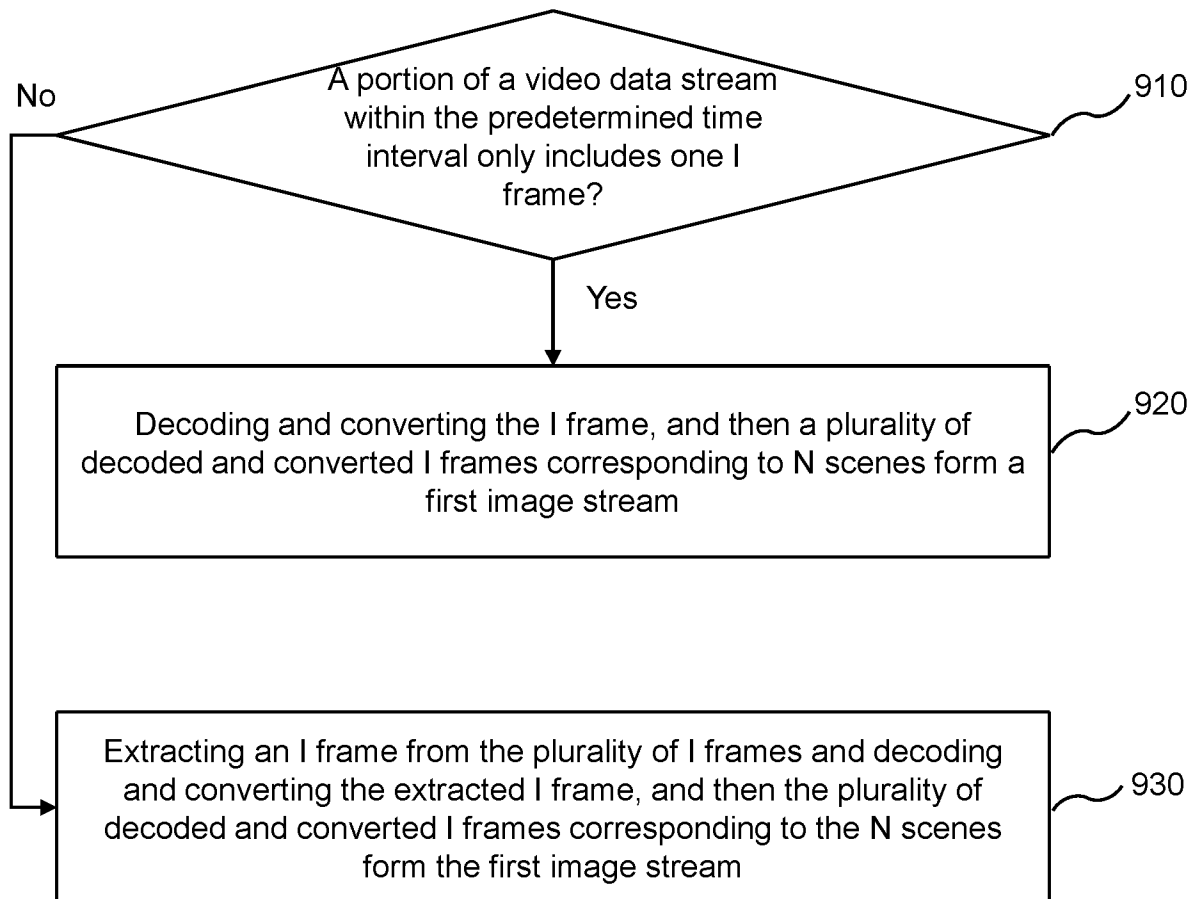
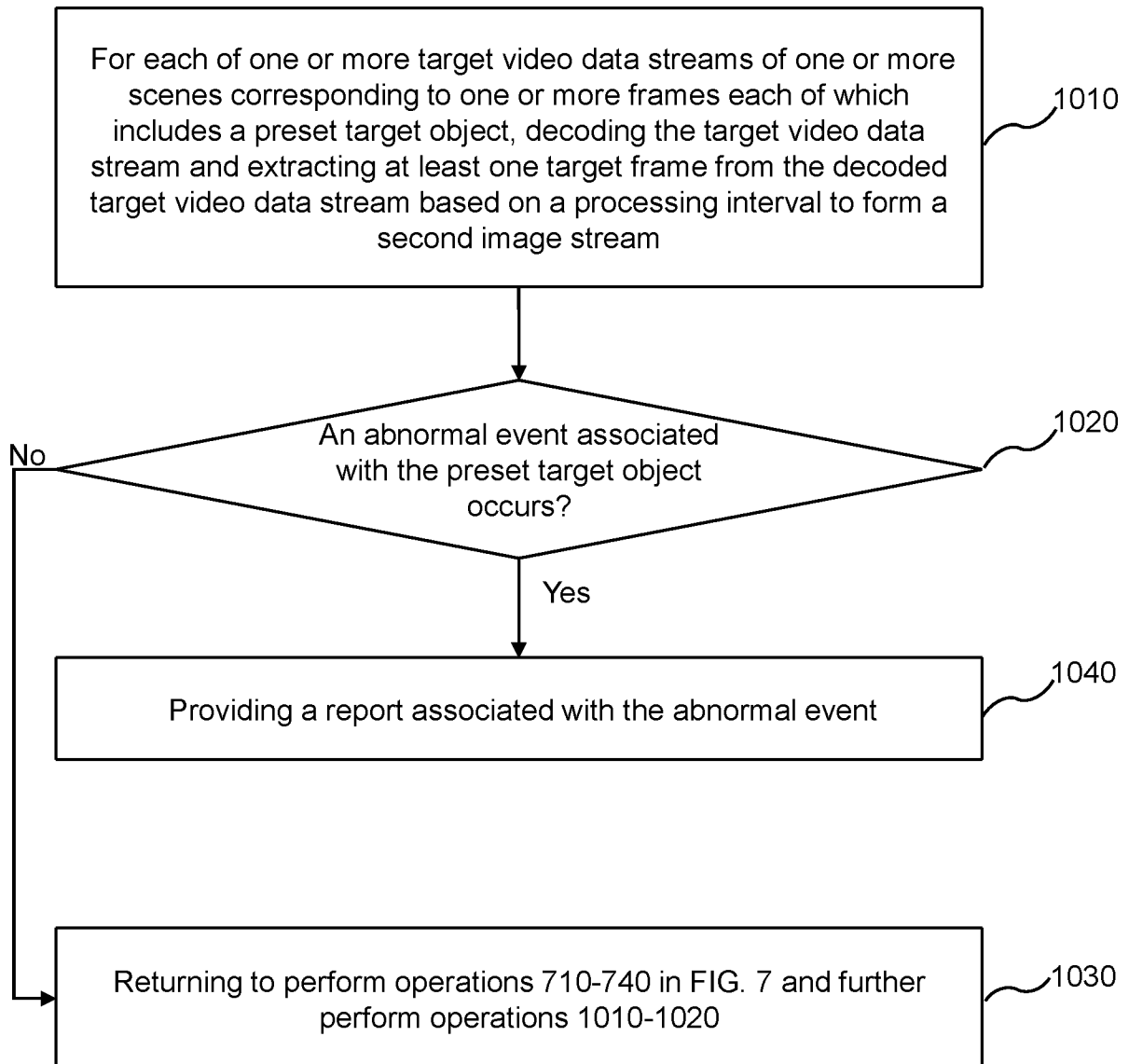
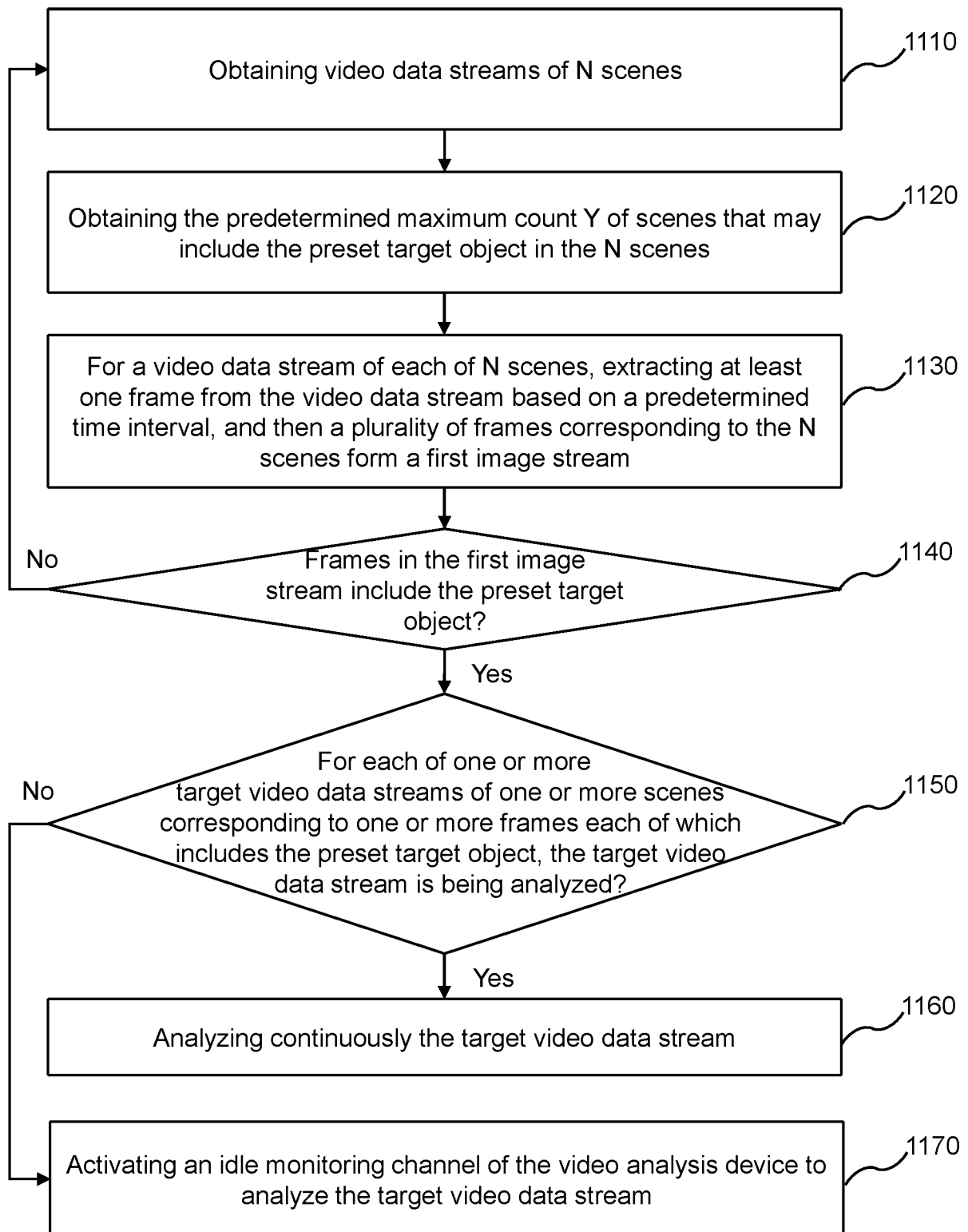


FIG. 8

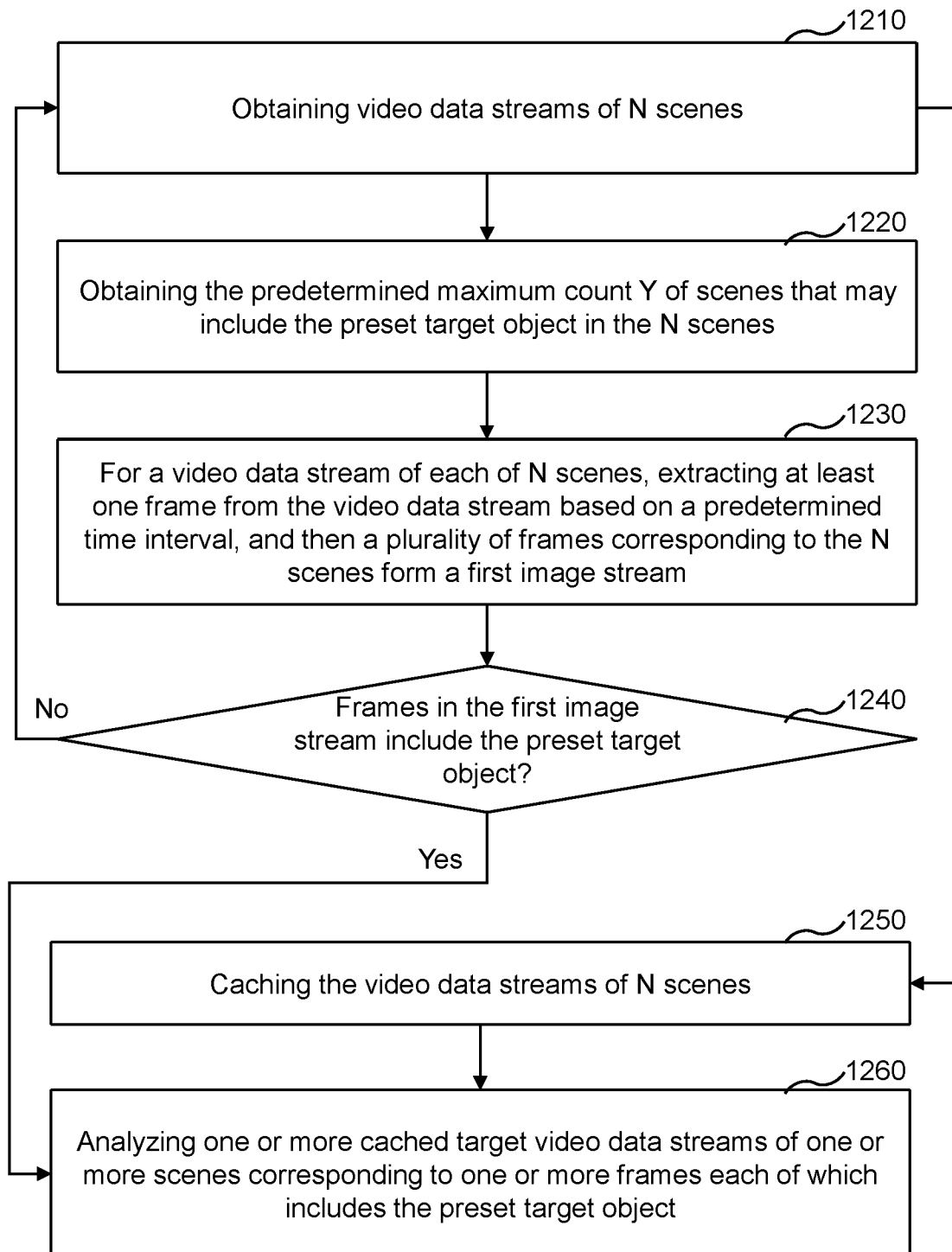
900**FIG. 9**

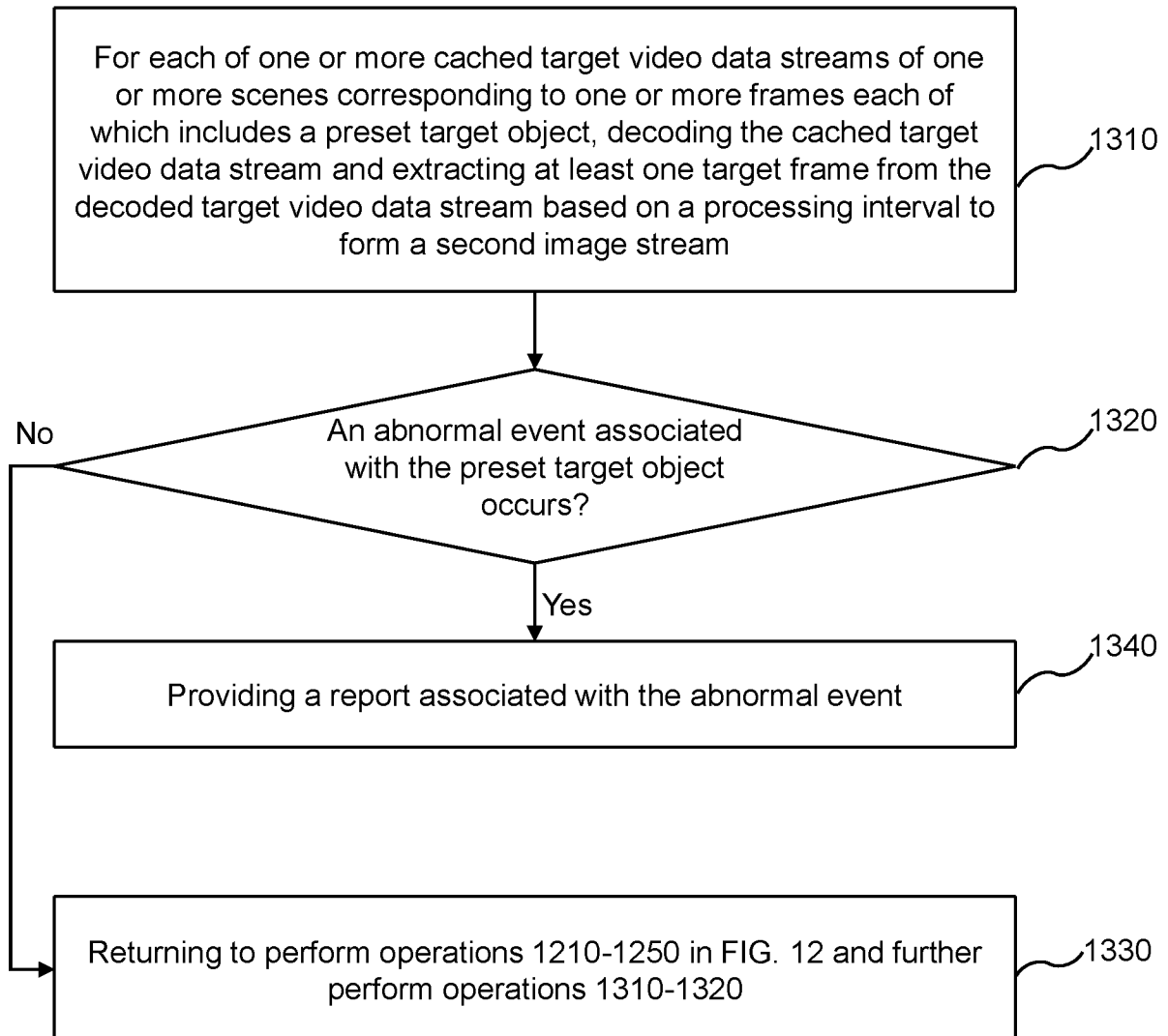
1000**FIG. 10**

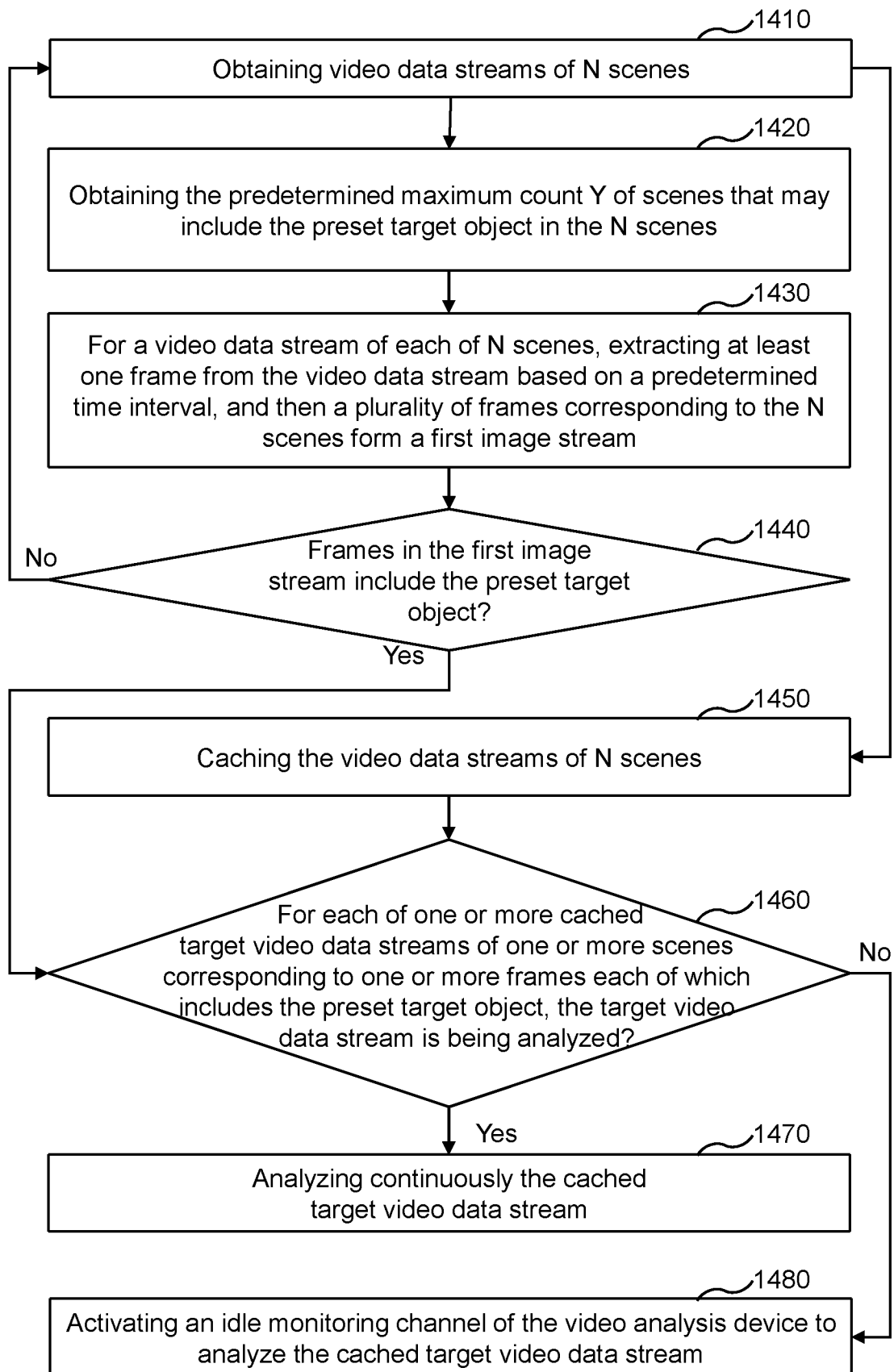
Sheet 11/19

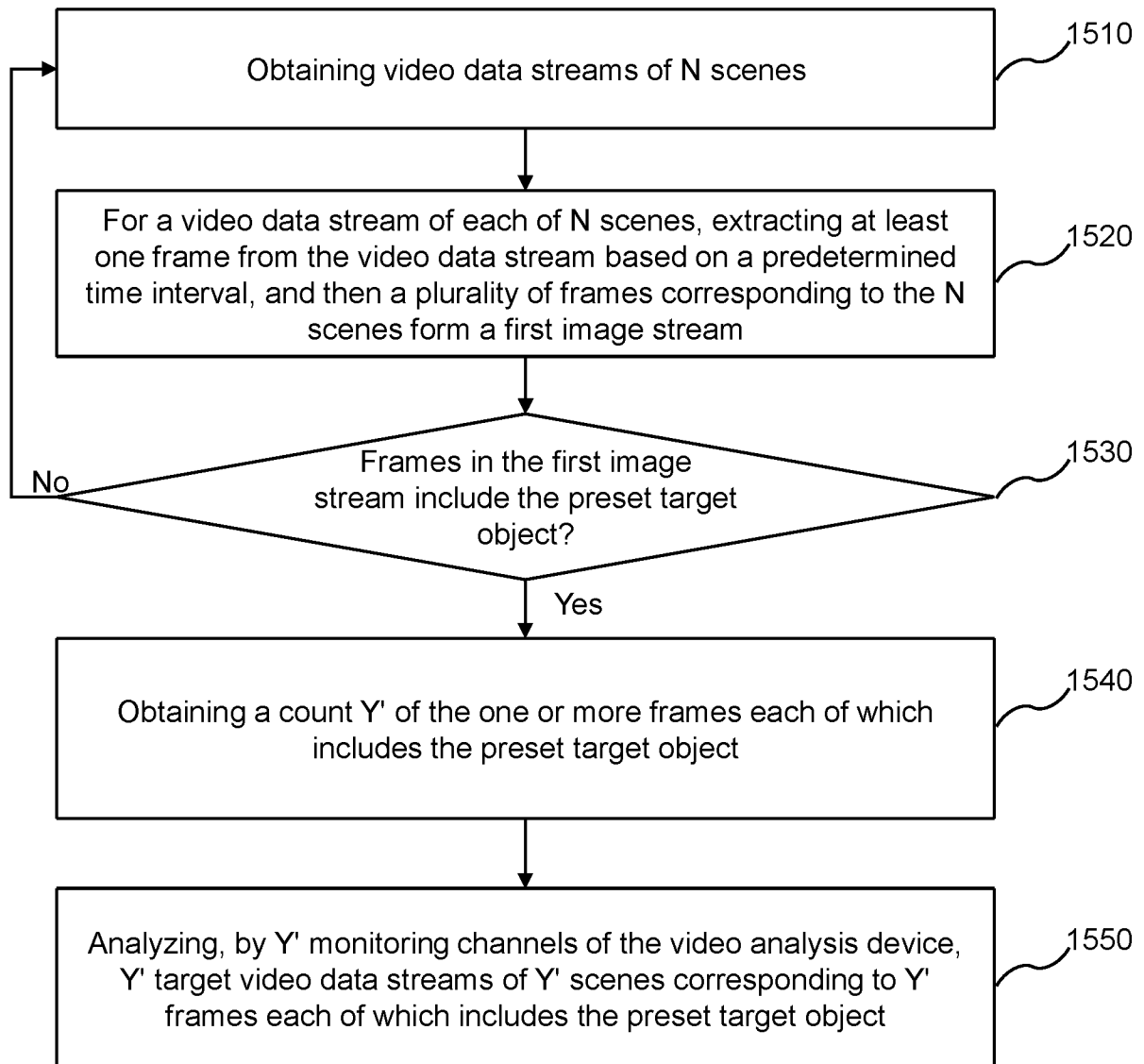
1100**FIG. 11**

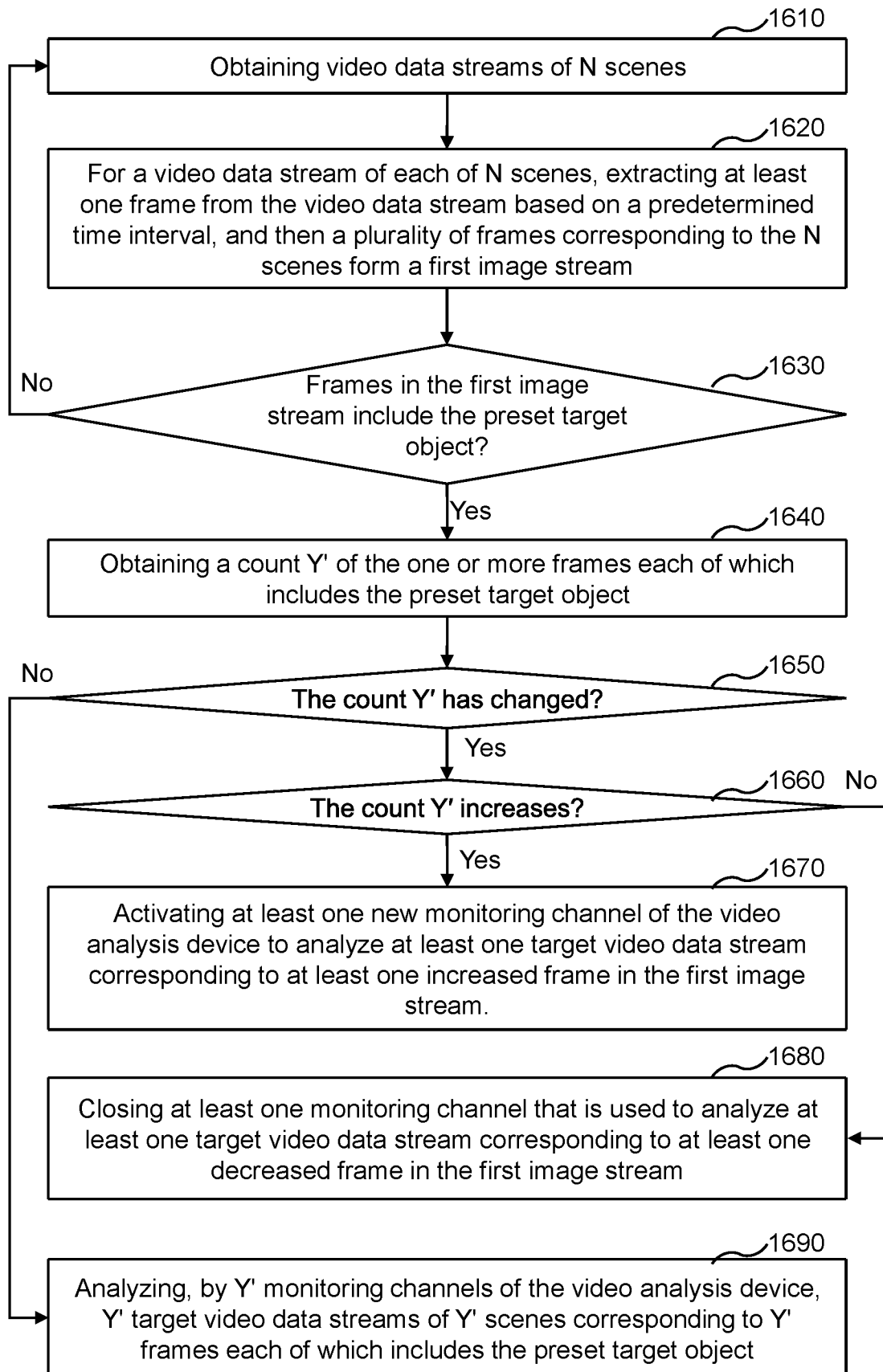
Sheet 12/19

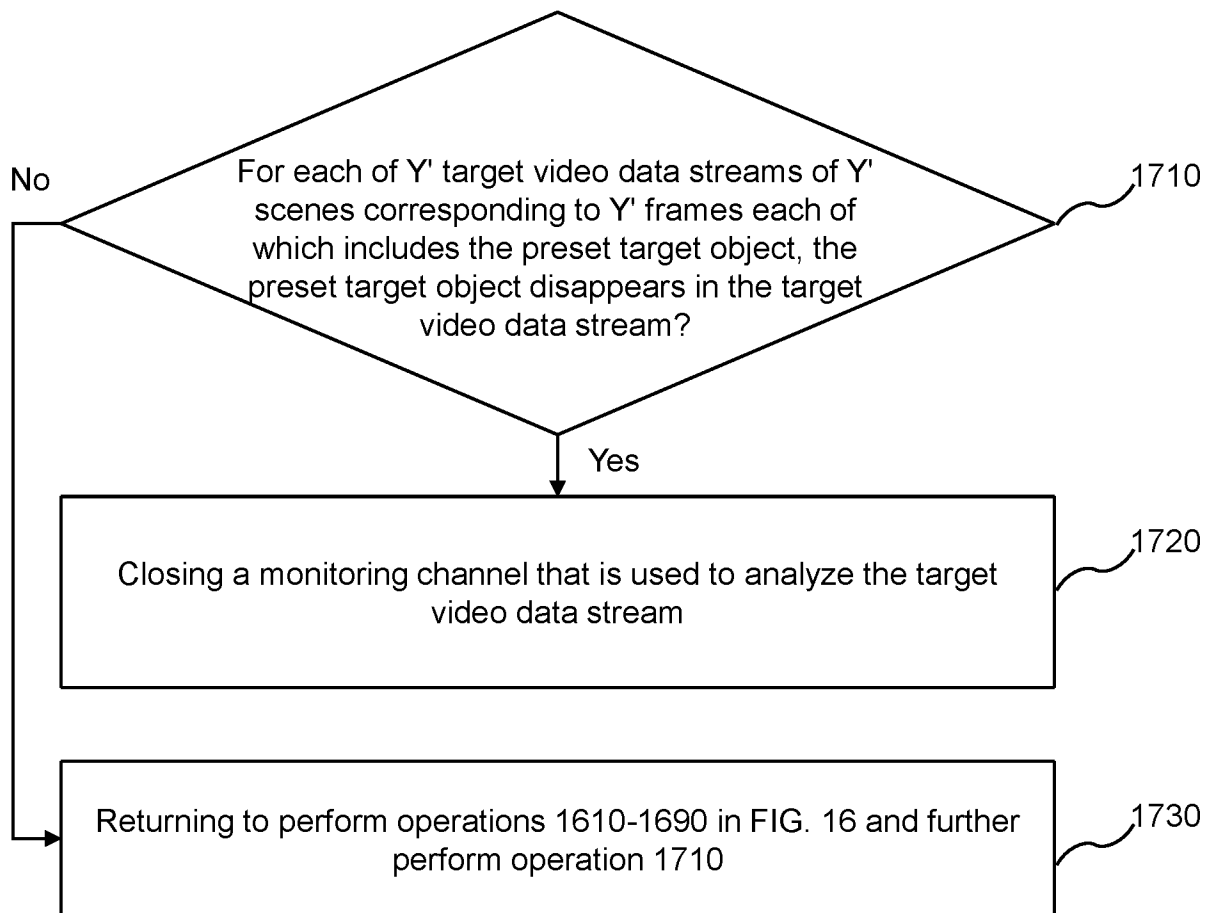
1200**FIG. 12**

1300**FIG. 13**

**FIG. 14**

1500**FIG. 15**

**FIG. 16**

1700**FIG. 17**

Sheet 18/19

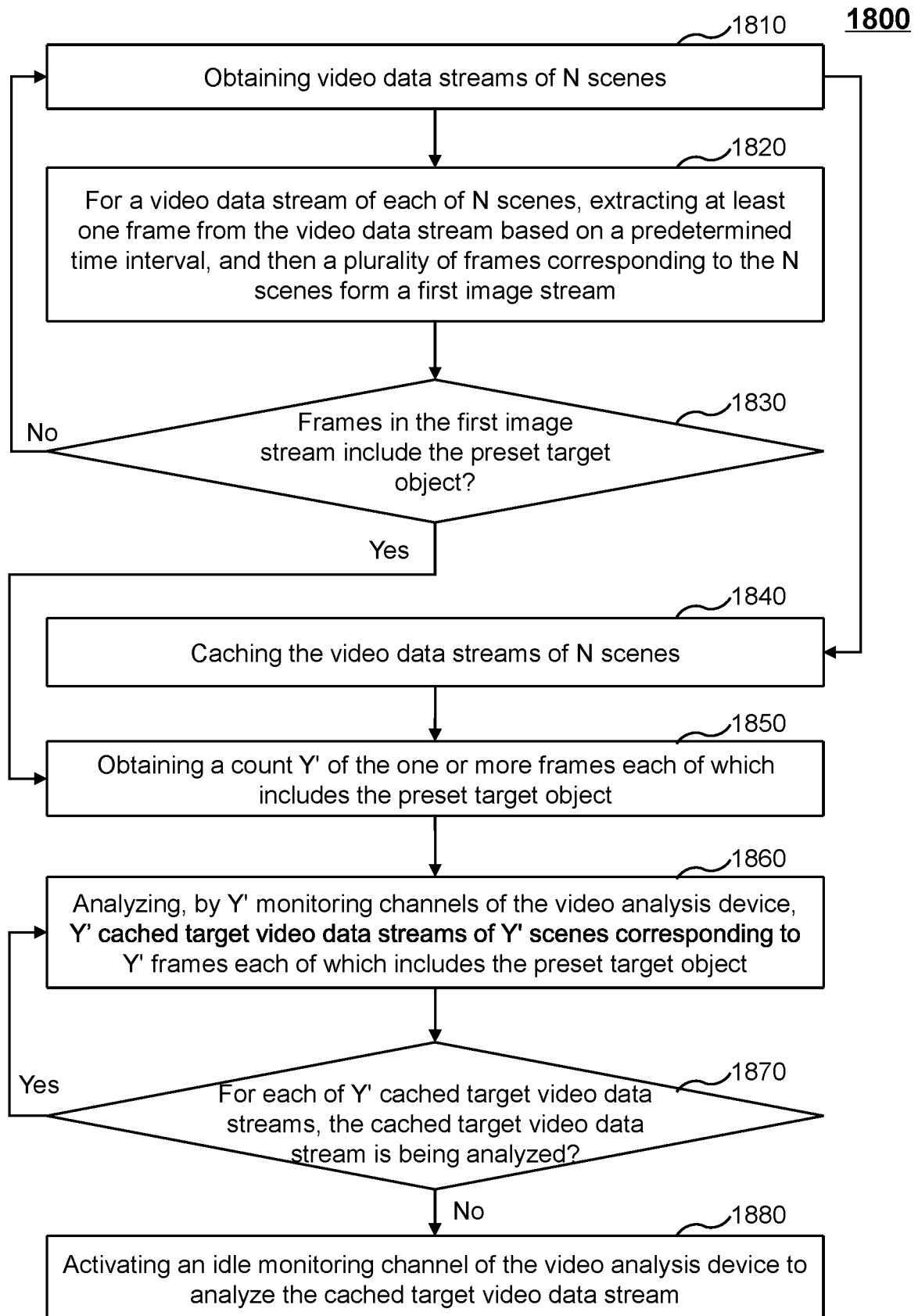
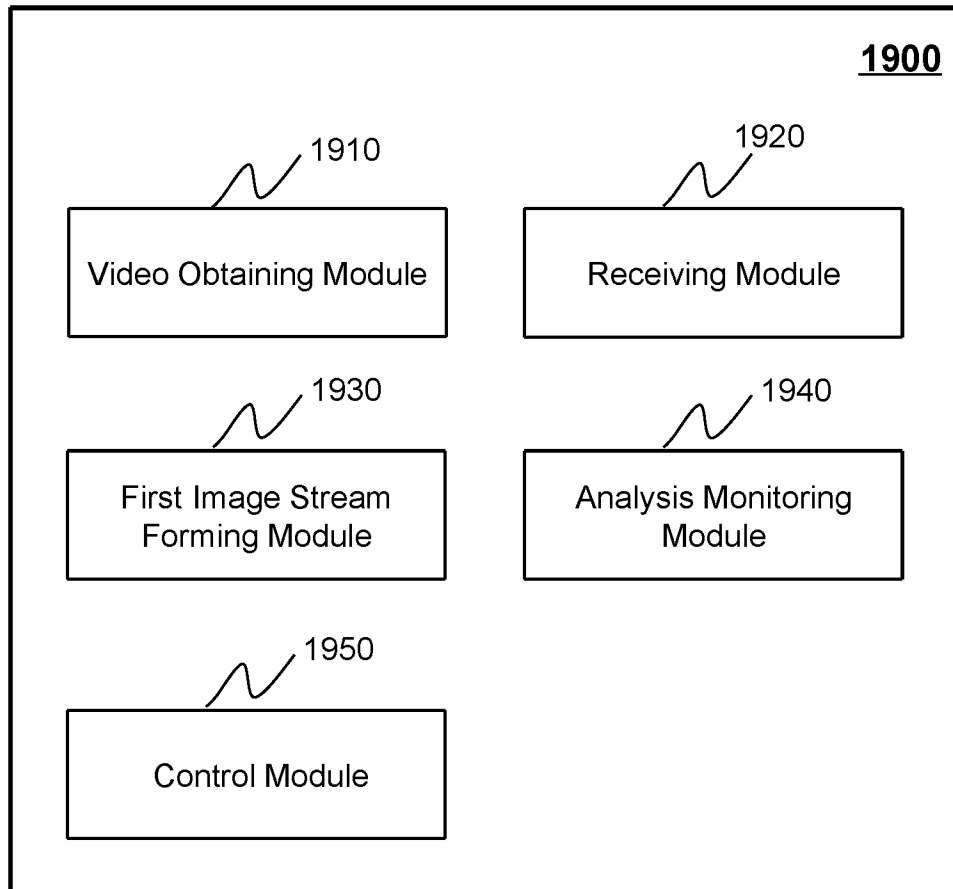


FIG. 18

Sheet 19/19

**FIG. 19**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/070975

A. CLASSIFICATION OF SUBJECT MATTER

H04N 7/18(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS,CNKI,CNTXT,VEN,WOTXT,USTXT,EPTXT: multi, scene, channel, path, monitoring, video, stream, predetermined, target, extract, frame, every, duration, period, interval

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 109167979 A (DEEPBLUE TECHNOLOGY SHANGHAI CO., LTD.) 08 January 2019 (2019-01-08) the description paragraphs [0012]-[0057]	1-28
Y	CN 110659627 A (SHANDONG INSPUR ARTIFICIAL INTELLIGENCE) 07 January 2020 (2020-01-07) the description paragraphs [0002]-[0079]	1-28
A	CN 111147768 A (BEIJING HENGFENG ZHIYUAN TECH CO., LTD.) 12 May 2020 (2020-05-12) the whole document	1-28
A	CN 110971880 A (BEIJING BOYA TIANAN INFORMATION TECHNOLOGY CO., LTD.) 07 April 2020 (2020-04-07) the whole document	1-28
A	CN 105005772 A (UNIV PEKING) 28 October 2015 (2015-10-28) the whole document	1-28
A	US 2009290020 A1 (CANON KK) 26 November 2009 (2009-11-26) the whole document	1-28

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

29 March 2021

Date of mailing of the international search report

06 April 2021

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088
China

Authorized officer

QI,Jingwei

Facsimile No. (86-10)62019451

Telephone No. 86- (010) -62411482

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/070975

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111988563 A (ZHEJIANG DAHUA TECHNOLOGY CO., LTD.) 24 November 2020 (2020-11-24) description paragraphs [0027]-[0202]	1-28

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/070975

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109167979	A	08 January 2019	CN	109167979	B	11 September 2020
CN	110659627	A	07 January 2020	None			
CN	111147768	A	12 May 2020	US	2020151888	A1	14 May 2020
CN	110971880	A	07 April 2020	None			
CN	105005772	A	28 October 2015	CN	105005772	B	12 June 2018
US	2009290020	A1	26 November 2009	AU	2008200966	A1	17 September 2009
				US	8305440	B2	06 November 2012
				AU	2008200966	B2	15 March 2012
CN	111988563	A	24 November 2020	None			