



(51) International Patent Classification:
G06K 9/00 (2006.01)

(21) International Application Number:
PCT/CN2019/117421

(22) International Filing Date:
12 November 2019 (12.11.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201910696640.3 30 July 2019 (30.07.2019) CN

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

(72) Inventors: **DENG, Xiang**; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN). **RU-AN, Xuewu**; No. 1187 Binan Road, Binjiang District,

Hangzhou, Zhejiang 310053 (CN). **ZHOU, Daoli**; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN). **WANG, Dian**; No. 1187 Binan Road, Binjiang District, Hangzhou, Zhejiang 310053 (CN).

(74) Agent: **METIS IP (CHENGDU) LLC**; Room 808, 8th Floor, Building B7, Block D, Tianfu Jingrong Center, No. 99 West Section of Hupan Road, Tianfu New District, 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, 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,

(54) Title: SYSTEMS AND METHODS FOR IMAGE CLUSTERING

500

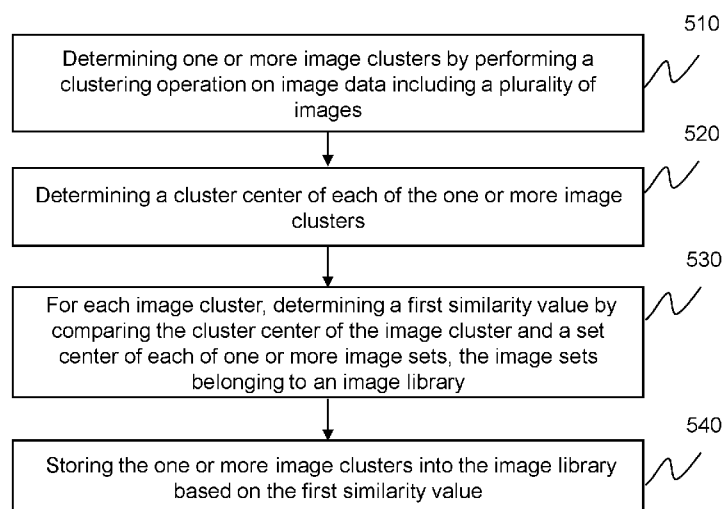


FIG. 5

(57) Abstract: A system and method for image clustering in an image library stored in a storage medium are disclosed. The method may comprise determining one or more image clusters by performing a clustering operation on image data including a plurality of images. The method may also comprise determining a cluster center of each of the one or more image clusters. For each image cluster, the method may further comprise determining a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library, and storing the image cluster into the image library based on the first similarity value.

SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))*

SYSTEMS AND METHODS FOR IMAGE CLUSTERING

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority of Chinese Patent Application No. 201910696640.3, filed on July 30, 2019, the contents of which are incorporated herein by reference.

TECHNICAL FIELD

[0002] This disclosure generally relates to image clustering, and more particularly, relates to systems, methods, electronic devices, and storage media for clustering of mass images.

BACKGROUND

[0003] In conventional image clustering technologies, images captured by multiple cameras are often grouped into various groups. An image to be analyzed can be compared to these groups and the comparison results are used to classify and store the image. However, such an approach often results in low efficiency and a waste of computing power. When there are a large number of images, e.g. captured by many cameras, the conventional approach is especially cumbersome and wasteful. Thus, there is a need for systems and methods for image clustering and analysis in order to achieve higher efficiency and lower cost.

SUMMARY

[0004] According to an aspect of the present disclosure, a method for image clustering in an image library stored in a storage medium is provided. The method may comprise determining one or more image clusters by performing a clustering operation on image data including a plurality of images, determining a cluster center of each of the one or more image clusters, and for each image

cluster, determining a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library, and storing the image cluster into the image library based on the first similarity value.

[0005] In some embodiments, determining the one or more image clusters by performing the clustering operation on the image data including the plurality of images includes for each of the plurality of images, determining a second similarity value between the image and each of the other images of the plurality of images; and determining a similar image set regarding the image by grouping images with second similarity values larger than a first threshold into the image cluster; determining whether each of the plurality of images is grouped into an image cluster; in response to a determination that an image of the plurality of images is not grouped into an image cluster, grouping the image into a new image cluster; and in response to a determination that an image of the plurality of images is grouped into an image cluster, grouping the similar image set for the image into the image cluster.

[0006] In some embodiments, the method further comprises obtaining the image data including the plurality of images, wherein the plurality of images are captured by one or more imaging devices within a predetermined time period.

[0007] In some embodiments, determining the cluster center of each of the one or more image clusters includes, for each image cluster, determining a quality value of each of the images in the image cluster; and designating an image with a larger quality value than other images in the image cluster as the cluster center of the image cluster.

[0008] In some embodiments, the quality value is associated with at least one of an occlusion coefficient, a blur coefficient, an illumination coefficient, or a three-dimensional angle.

[0009] In some embodiments, storing the image cluster into the image library based on the first similarity value includes determining whether the first similarity value between the cluster center of the image cluster and a set center of an image set is larger than a second threshold; and in response to a determination that the first similarity value between the cluster center of the image cluster and a set center of an image set is larger than the second threshold, storing the image cluster into the image set, or in response to a determination that the first similarity value between the cluster center of the image cluster and the set center of each of the one or more image sets is not larger than the second threshold, designating the image cluster as a new image set; and storing the new image set into the image library.

[0010] In some embodiments, storing the image cluster into the image set includes storing the image cluster as a subset of the image set.

[0011] In some embodiments, the method further includes updating the set center of the image set based on the quality values of images in the image set.

[0012] In some embodiments, each of the plurality of images includes at least a face of a person, and the clustering operation is based on comparing the faces.

[0013] In some embodiments, the method further includes obtaining original image data including multiple images; identifying images including at least a face of a person from the multiple images; and designating the identified images including at least a face of a person as the plurality of images.

[0014] According to another aspect of the present disclosure, a system for image clustering in an image library stored in a storage medium is provided. The system may be configured to perform operations including determining one or more image clusters by performing a clustering operation on image data including a plurality of images; determining a cluster center of each of the one or more image clusters; for each image cluster, determining a first similarity value

by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library, and storing the image cluster into the image library based on the first similarity value.

[0015] According to a further aspect of the present disclosure, a non-transitory computer readable medium comprising at least one set of instructions for image clustering in an image library stored in a storage medium is provided. When the at least one set of instructions executed by a device, the at least one set of instructions causes the device to perform one or more operations. The one or more operations comprises determining one or more image clusters by performing a clustering operation on image data including a plurality of images; determining a cluster center of each of the one or more image clusters; for each image cluster, determining a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library, and storing the image cluster into the image library based on the first similarity value.

[0016] 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

[0017] The present disclosure is further described in terms of ~~exemplary~~ embodiments. These example embodiments are described in detail with reference to the drawings. The drawings are not to scale. These embodiments are non-limiting example embodiments, in which like reference

numerals represent similar structures throughout the several views of the drawings, and wherein:

[0018] FIG. 1 is a schematic diagram illustrating an exemplary image clustering system according to some embodiments of the present disclosure;

[0019] FIG. 2 is a schematic diagram illustrating exemplary hardware and optionally software components of an example computing device according to some embodiments of the present disclosure;

[0020] FIG. 3 is a schematic diagram illustrating exemplary components of an example mobile device according to some embodiments of the present disclosure;

[0021] FIG. 4 is a block diagram illustrating an exemplary media server according to some embodiments of the present disclosure;

[0022] FIG. 5 illustrates an exemplary process for image clustering according to some embodiments of the present disclosure;

[0023] FIG. 6 illustrates an exemplary process for image clustering according to some embodiments of the present disclosure;

[0024] FIG. 7 is a flowchart illustrating an exemplary process for performing a cluster operation on image data including a plurality of images according to some embodiments of the present disclosure;

[0025] FIG. 8 is a schematic block diagram of an exemplary image clustering device according to some embodiments of the present disclosure; and

[0026] FIG. 9 is a schematic diagram illustrating an exemplary computer readable storage medium according to some embodiments of the present disclosure.

DETAILED DESCRIPTION

[0027] In order to illustrate the technical solutions related to the embodiments of the present disclosure, brief introduction of the drawings referred to in the

description of the embodiments is provided below. Obviously, drawings described below are only some examples or embodiments of the present disclosure. Those having ordinary skills in the art, without further creative efforts, may apply the present disclosure to other similar scenarios according to these drawings. Unless stated otherwise or obvious from the context, the same reference numeral in the drawings refers to the same structure and operation.

[0028] As used in the disclosure and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the content clearly dictates otherwise. It will be further understood that the terms “comprises,” “comprising,” “includes,” and/or “including” when used in the disclosure, specify the presence of stated steps and elements, but do not preclude the presence or addition of one or more other steps and elements.

[0029] Some modules of the system may be referred to in various ways according to some embodiments of the present disclosure, however, any number of different modules may be used and operated in a client terminal and/or a server. These modules are intended to be illustrative, not intended to limit the scope of the present disclosure. Different modules may be used in different aspects of the system and method.

[0030] According to some embodiments of the present disclosure, flow charts are used to illustrate the operations performed by the system. It is to be expressly understood, the operations above or below may or may not be implemented in order. Conversely, the operations may be performed in inverted order, or simultaneously. Besides, one or more other operations may be added to the flowcharts, or one or more operations may be omitted from the flowchart.

[0031] Technical solutions of the embodiments of the present disclosure be described with reference to the drawings as described below. It is obvious that the described embodiments are not exhaustive and are not limiting. Other

embodiments obtained, based on the embodiments set forth in the present disclosure, by those with ordinary skill in the art without any creative works are within the scope of the present disclosure.

[0032] In one aspect, the present disclosure is directed to a system and method for image clustering. The system provided in the present disclosure may be adapted to camera systems having a large number of cameras. The system may determine one or more image clusters by performing a clustering operation on image data including a plurality of images, and determine a cluster center of each of the one or more image clusters. The cluster center of each image cluster may be compared with a set center of each of one or more image sets in an image library. The one or more image clusters into the image library based on the comparison. In this way, the classification and storage of images may be optimized greatly.

[0033] FIG. 1 is a schematic diagram illustrating an exemplary image clustering system according to some embodiments of the present disclosure. The image clustering system 100 may be configured to group a plurality of images into one or more image clusters. As shown, the image clustering system 100 may include an image source 101, an encoder 104, a buffer manager 105, a buffer 106, a transmitter 107, a terminal 108 (or a plurality of terminals 108), network 112, and a network storage device 113 (or a plurality of network storages 113).

[0034] The image source 101 may provide an image through the network 112 to a user of the terminal 108. The image source 101 may generate an image itself or via an image transfer site. For example, the image source 101 may include a camera 102 and/or a media server 103.

[0035] The camera 102 may be able to capture one or more images or a video composed of a plurality of images. In some embodiments, the camera 102 may be a digital camera, a video camera, a security camera, a web camera, a

smartphone, a tablet, a laptop, a video gaming console equipped with a web camera, a camera with multiple lenses, etc.

[0036] The camera 102 may include a lens, a shutter, a sensor, a processing device, and a storage device. The lens may be an optical device that focuses a light beam by means of refraction to form an image. In some embodiments, the lens may include one or more lenses. The lens may be configured to intake a scene that a user is interested in. An aperture of the lens may refer to the size of the hole through which light passes to reach the sensor. The aperture may be adjustable to adjust the amount of light that passes through the lens. The focal length of the lens may be adjustable to adjust the coverage of the camera 102.

[0037] The shutter may be opened to allow light through the lens when an image is captured. The shutter may be controlled manually or automatically by the processing device.

[0038] The sensor may be configured to receive light passing through the lens and transform the light signals of the received light into electrical signals. The sensor may include charge coupled device (CCD) and complementary metal-oxide semiconductor (CMOS). The sensor may be configured to detect the scene from the lens, and transform the scene into electronic signals.

[0039] The processing device may be configured to process data and/or information relating to the camera 102 in the present disclosure and/or control one or more components (e.g., the lens, the shutter) in the camera 102. For example, the processing device may automatically determine target values of exposure parameters of the camera 102 such as an exposure time, an exposure gain, and an aperture. The processing device may also adjust quality images taken the camera 102, such as sharpness of the images.

[0040] In some embodiments, the processing device may be local or remote.

For example, the processing device may communicate with the camera 102 via a network. As another example, the processing device may be integrated into the camera 102.

[0041] The storage device may store data, instructions, and/or any other information. In some embodiments, the storage device may store data obtained from the processing device. For example, the storage device may store captured images. In some embodiments, the storage device may store data and/or instructions that the processing device may execute or use to perform exemplary methods described in the present disclosure. In some embodiments, the storage device may include a mass storage, 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.

[0042] In some embodiments, the camera 102 may be a pan-tilt-zoom (PTZ) camera. The camera 102 may include or be mounted on one or more movable platforms. The movable platforms may control the camera 102 to move in a certain way. For example, rotate, pivot, translational move, or the like. In some embodiments, the movable platforms may be a PTZ head. The PTZ head may

control the lateral rotation, longitudinal tilt, and/or optical zooming of the camera 102 or the lens thereof. For example, the PTZ head may be rotated in a certain range so as to switch the view of the PTZ camera or the lens thereof. For example, the PTZ head may be rotated in 360 degrees in a horizontal plane. As another example, the PTZ head may be rotated in 180 degrees in a vertical plane. As still another example, the PTZ head may be rotated both in the horizontal plane and the vertical plane. Further, the PTZ head may also be a platform configured to perform translational motion to the PTZ camera.

[0043] The media sever 103 may be a server (e.g., a computer or a group of computers) for storing, broadcasting, selling, renting, or providing images or videos. The media server 103 may also include an image and/or video processing electronic device (not shown) configured to process the images and/or video streams from the image source 101.

[0044] A video may include a plurality of frames. A frame may be one of a plurality of still images. By sequentially displaying frames (e.g., images) in a rate (frame rate), a video player installed on the terminal 108 may present the video to a user.

[0045] Before transmitting images or a video including video frames through the network 112, the image source 101 may send the images or the video to the encoder 104 for encoding, or send the images or the video to the buffer 106 through the buffer manager 105. For example, the images or the video provided by the image source 101 may be relatively large in size (e.g., encoded with low compression rate), thus before the transmission, the image source 101 may send the images or the video to the encoder 104 for compression. As another example, the images or the video provided by the image source 101 may be proper in size, and the image source 101 may directly send the images or the video to the buffer 106 through the buffer manager 105 for transmission.

[0046] An encoder 104 may be a remote device from the image source 101 or a local device interpreted in the image source 101. It may encode the images or the video provided by the image source 101 before the images or the video are transmitted through the network 112. Through encoding, the images or the video to be transmitted may be compressed and/or encrypted. For example, the encoder 104 may encode a video using an algorithm for video compression so that the cost (e.g., time cost, resource cost, financial cost) for transmitting the video may be significantly reduced. As another example, the encoder 104 may encode a video using an algorithm for video encryption so that the video may be transmitted safely and a user without permission may not watch the video. The encoder 104 may encode the images or the video, and generate encoded images or video. The encoder 104 may send the encoded images or video to the buffer 106 through the buffer manager 105. Alternatively or additionally, the buffer manager 105 may obtain the encoded images or video from the encoder 104.

[0047] The video frames and/or images to be transmitted may be stored in the buffer 106, which may be managed by the buffer manager 105. The buffer 106 may use a queue based data structure for buffering the images or the video to be transmitted.

[0048] The buffer 106 may be a storage device for buffering the images or the video to be transmitted through the network 112. It may be a remote device from the image source 101 or a local device interpreted in the image source 101, such as the storage medium of the camera 102. The buffer 106 may include a mass storage device, a removable storage device, 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), such as 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). 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.

[0049] The transmitter 107 may transmit the images or the video buffered in the buffer 106 to the network 112. The transmitter 107 may transmit the images or the video in response to instructions sent from the image source 101, the buffer manager 105, the terminal 108, or the like, or a combination thereof.

Alternatively or additionally, the transmitter 107 may spontaneously transmit the images or the video stored in the buffer 106. The transmitter 107 may transmit the images or the video through the network 112 to the terminal 108 through one or more network connections (wired and/or wireless).

[0050] In some embodiments, the transmitter 107 may be capable of determining the transmission performance of the network 112. For example, the transmitter 107 may monitor its data transmission rate for determining the transmission performance.

[0051] The terminal 108 may receive the transmitted images or video through the network 112. The terminal 108 may decode (e.g., through a video player installed on the terminal 108) the transmitted images or video using a decoding algorithm and display the images or the video to a user. The decoding algorithm may correspond to the encoding algorithm used by the encoder 104.

[0052] The terminal 108 may be various in forms. For example, the terminal

108 may include a mobile device 109, a tablet computer 110, a laptop computer 111, or the like, or any combination thereof. In some embodiments, the mobile device 109 may include, a wearable device, a mobile device, a virtual reality device, an augmented reality device, or the like, or any combination thereof. In some embodiments, the wearable device may include a bracelet, footgear, eyeglasses, a helmet, a watch, clothing, a backpack, a smart accessory, or the like, or any combination thereof. In some embodiments, the mobile device may include a mobile phone, a personal digital assistance (PDA), a laptop, a tablet computer, a desktop, or the like, or any combination thereof. In some embodiments, the virtual reality device and/or the augmented reality device may include a virtual reality helmet, virtual reality glasses, a virtual reality patch, an augmented reality helmet, augmented reality glasses, an augmented reality patch, or the like, or any combination thereof. For example, the virtual reality device and/or the augmented reality device may include a Google Glass™, an Oculus Rift™, a Hololens™, a Gear VR™, etc. In some embodiments, the terminal(s) 108 may be part of a processing engine.

[0053] The network 112 may include any suitable network that can facilitate a transmission of images or video provided by the image source 101 to the terminal(s) 108. The network 112 may be and/or include a public network (e.g., the Internet), a private network (e.g., a local area network (LAN), a wide area network (WAN)), a wired network (e.g., an Ethernet network), a wireless network (e.g., an 802.11 network, a Wi-Fi network), a cellular network (e.g., a Long Term Evolution (LTE) network), a frame relay network, a virtual private network ("VPN"), a satellite network, a telephone network, routers, hubs, switches, server computers, and/or any combination thereof. Merely by way of example, the network 112 may include a cable network, a wireline network, a fiber-optic network, a telecommunications network, an intranet, 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. In some embodiments, the network 112 may include one or more network access points. For example, the network 112 may include wired and/or wireless network access points such as base stations and/or internet exchange points through which a video provided by the image source 101 may be transmitted to the terminal 108.

[0054] In some embodiments, the network 112 may include one or more network storage devices 113. The network storage device 113 may be a device for buffering or caching data transmitted in the network 112. The images or the video transmitted by the transmitter 107 may be buffered or cached in one or more network storage devices 113 before being received by the terminal 108. The network storage device 113 may be a server, a hub, a gateway, or the like, or a combination thereof.

[0055] It may be noted that, one or more of the encoder 104, buffer manager 105, buffer 106 and transmitter may be a stand-alone device, or a module integrated into the image source 101 or another stand-alone device. For example, one or more of the encoder 104, buffer manager 105, buffer 106 and transmitter 107 may be integrated into the camera 102 or the media server 103. As another example, the encoder 104, buffer manager 105, buffer 106 and transmitter 107 may be included in a video processing engine which may communicate with the image source 101 through direct wired connection, the network 112, or another network not shown in FIG. 3A. As a further example, the encoder 104 may be a stand-alone device (e.g., a computer or a server), while the buffer manager 105, buffer 106 and transmitter 107 may be included in another stand-alone device.

[0056] 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. For example, the computing device 200 may be an implementation of the image clustering system 100 specialized in image clustering. As illustrated in FIG. 2, the computing device 200 may include a processor 222, a storage 227, an input/output (I/O) 226, and a communication port 225.

[0057] The processor 222 (e.g., logic circuits) may execute computer instructions (e.g., program code) and perform functions in accordance with techniques described herein. For example, the processor 222 may include interface circuits and processing circuits therein. The interface circuits may be configured to receive electronic signals from a bus (not shown in FIG. 2), wherein the electronic signals encode structured data and/or instructions for the processing circuits to process. The processing circuits may conduct logical operations 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.

[0058] The computer instructions may include, for example, routines, programs, objects, components, data structures, procedures, modules, and functions, which perform particular functions described herein. In some embodiments, the processor 222 may include one or more hardware processors, such as a microcontroller, a microprocessor, a reduced instruction set computer (RISC), an application specific integrated circuits (ASICs), an application-specific instruction-set processor (ASIP), a central processing unit (CPU), a graphics processing unit (GPU), a physics processing unit (PPU), a microcontroller unit, a digital signal processor (DSP), a field programmable gate array (FPGA), an advanced RISC machine (ARM), a programmable logic device (PLD), any circuit or processor

capable of executing one or more functions, or the like, or any combinations thereof.

[0059] Merely for illustration, only one processor is described in the computing device 200. However, it should be noted that the computing device 200 in the present disclosure may also include multiple processors, thus operations and/or method steps that are performed by one processor 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 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 or more different processors 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).

[0060] The storage 227 may store data/information. In some embodiments, the storage 227 may include a mass storage, removable storage, a volatile read-and-write memory, a read-only memory (ROM), or the like, or any combination thereof. For example, the mass storage may include a magnetic disk, an optical disk, a solid-state drive, etc. The removable storage may include a flash drive, a floppy disk, an optical disk, a memory card, a zip disk, a magnetic tape, etc. The volatile read-and-write memory may include a random-access memory (RAM), which 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. The 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 227 may store one or more programs

and/or instructions to perform exemplary methods described in the present disclosure. For example, the storage 227 may store a program for the processing engine (e.g., the management device 110) for image clustering.

[0061] The I/O 226 may input and/or output signals, data, information, etc. In some embodiments, the I/O 226 may include an input device and an output device. Examples of the input device may include a keyboard, a mouse, a touch screen, a microphone, or the like, or a combination thereof. Examples of the output device may include a display device, a loudspeaker, a printer, a projector, or the like, or a combination thereof. Examples of the display device may include a liquid crystal display (LCD), a light-emitting diode (LED)-based display, a flat panel display, a curved screen, a television device, a cathode ray tube (CRT), a touch screen, or the like, or a combination thereof.

[0062] The communication port 225 may be connected to a network (e.g., the network 130) to facilitate data communications. The communication port 225 may establish connections between any component of the image clustering system 100. The connection may be a wired connection, a wireless connection, any other communication connection that can enable data transmission and/or reception, and/or any combination of these connections. The wired connection may include, for example, an electrical cable, an optical cable, a telephone wire, or the like, or any combination thereof. The wireless connection may include, for example, a Bluetooth™ link, a Wi-Fi™ link, a WiMax™ link, a WLAN link, a ZigBee link, a mobile network link (e.g., 3G, 4G, 5G), or the like, or a combination thereof. In some embodiments, the communication port 2600 may be and/or include a standardized communication port, such as RS232, RS485, etc. In some embodiments, the communication port 225 may be a specially designed communication port.

[0063] FIG. 3 is a schematic diagram illustrating exemplary components of an

exemplary user device according to some embodiments of the present disclosure. As illustrated in FIG. 3, the mobile device 300 may include a communication platform 320, a display 310, a graphic processing unit (GPU) 330, a central processing unit (CPU) 330, an I/O port 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. In some embodiments, a mobile operating system 370 (e.g., iOS™, Android™, Windows Phone™) and one or more applications 380 may be loaded into the memory 360 from the storage 390 in order to be executed by the processor 340. The applications 380 may include a storage manager for managing data stored in the mobile device 300.

[0064] To implement various modules, units, and their functionalities described in the present disclosure, computer hardware platforms may be used as the hardware platform(s) for one or more of the elements described herein. A computer with user interface elements may be used to implement a personal computer (PC) or any other type of work station or terminal device. A computer may also act as a server if appropriately programmed.

[0065] FIG. 4 is a block diagram illustrating an exemplary media server according to some embodiments of the present disclosure. The media server 103 may include an obtaining module 410, a clustering module 420, a matching module 430, and a storing module 440. One or more of the modules of the media server 103 may be interconnected. The connection(s) may be wireless or wired. At least a portion of the media server 103 may be implemented on a computing apparatus as illustrated in FIG. 2 or a mobile device as illustrated in FIG. 3.

[0066] The obtaining module 410 may obtain data and/or information. In some embodiments, the obtaining module 410 may obtain image data including a

plurality of images. In some embodiments, the obtaining module 410 may obtain image data including multiple images. In some embodiments, the multiple images may be captured by an image device within a predetermined time period. Exemplary imaging device may include an optical camera, an infrared camera, etc. The optical camera may be a digital camera, a video camera, a security camera, a web camera, a smartphone, a tablet, or a laptop equipped with a camera, etc. In some embodiments, the predetermined time period may be set by a user, according to default settings of the image clustering system 100, etc. For example, the predetermined time period may be 30 seconds, 1 minute, 2 minutes, 5 minutes, etc.

[0067] In some embodiments, the obtaining module 410 may obtain image data in real-time. The plurality of images may be obtained in real-time by one or more cameras. In some embodiments, the obtaining module 410 may obtain image data including a preset number of images from the database periodically. The obtained image data may be stored in a predetermined database (e.g., in the network storage device 113, the storage 227, or the storage 390).

[0068] The clustering module 420 may determine one or more image clusters. In some embodiments, the clustering module 420 may determine the one or more image clusters by performing a clustering operation on image data including a plurality of images. As used herein, an image cluster may be a cluster or set of images. Images with higher similarities (e.g., larger than a certain similarity value) may be grouped into an image cluster through the clustering operation. In some embodiments, images of a same person may be grouped into a same image cluster through the clustering operation. According to the clustering operation, the clustering module 420 may determine, for each of the plurality of images, a second similarity value between the image and each of the other images of the plurality of images, and a similar image set regarding the

image by grouping images with second similarity values larger than a first threshold into the image cluster. Then a determination may be made as to whether each of the plurality of images is grouped into an image cluster. If an image of the plurality of images is not grouped into an image cluster, the clustering module 420 may group the image into a new image cluster. If an image of the plurality of images is grouped into an image cluster, the clustering module 420 may group the similar image set for the image into the image cluster.

[0069] In some embodiments, the similarity value may be determined based on a similarity determination model. Features of the images of the one or more image clusters may be extracted and input into the similarity determination model. The output of the similarity determination model may be determined as the similarity values between the images. Merely by way of example, the similarity determination model may include a deep belief network (DBN), a Stacked Auto-Encoders (SAE), a logistic regression (LR) model, a support vector machine (SVM) model, a decision tree model, a Naive Bayesian Model, a random forest model, a Restricted Boltzmann Machine (RBM), a Gradient Boosting Decision Tree (GBDT) model, a LambdaMART model, an adaptive boosting model, a recurrent neural network (RNN) model, a convolutional network model, a hidden Markov model, a perceptron neural network model, a Hopfield network model, or the like, or any combination thereof.

[0070] In some embodiments, the clustering module 420 may determine a cluster center of each of the one or more clusters. In some embodiment, a cluster center of an image cluster may be a first image grouped into the image cluster. For example, as for an image cluster (A, (B)), the image A may be the first image grouped into the image cluster. The image A may be determined as the cluster center of the image cluster (A, (B)). In some embodiments, if an image cluster includes two or more images, the cluster center may be

determined and/or updated based on the two or more images. For example, if an image cluster includes two or more images, the cluster center may be updated based on images in the image cluster once a new image is grouped into the image cluster. As another example, if an image cluster includes two or more images, the cluster center may be updated based on images in the image cluster after the image cluster is finalized. The updated cluster center of an image cluster may be an image selected from images of the image cluster having a larger quality value than other images in the image cluster.

[0071] The matching module 430 may match the one or more image clusters with image sets in the image library. For each image cluster, the matching module 430 may determine a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to an image library.

[0072] An image set may refer to a collection of similar images. In some embodiments, an image set may include images of a same person. The image library may be a database storing the one or more image sets. Each image set may have a set center. In some embodiments, the set center of an image set may be a first image grouped into the image set. In some embodiments, if an image set includes two or more images, the set center may be updated based on the two or more images. For example, if an image set includes two or more images, the set center may be updated once a new image or an image cluster is grouped into the image cluster based on images in the image set. As another example, if an image set includes two or more images, the set center may be updated based on images in the image set after the image set is finalized. The updated set center of an image set may be an image selected from images of the image set having a larger quality value than other images in the image set.

[0073] The matching module 430 may determine the first similarity value by

comparing the cluster center of the image cluster and a set center of each of one or more image sets. A determination as to whether the first similarity value is greater than a second threshold. If the first similarity value between the cluster center of the image cluster and a set center of an image set is greater than the second threshold, it may be determined that the cluster center of the image cluster matches the set center of the image set in the image library.

[0074] The storing module 440 may store the one or more image clusters into the image library. The storing module 440 may store the one or more image clusters into the image library based on the first similarity value.

[0075] If all the similarity values between a cluster center of an image cluster and set centers of the one or more image sets are less than the second threshold, it may indicate that there may be no set center in the image library that matches the cluster center of the image cluster. In this case, the image cluster may be designated as a new image set, and stored into the image library. The image library may be a collection of a plurality of image sets.

[0076] If a similarity value between a cluster center of an image cluster and a set center of an image set is less than the second threshold, it may indicate that the set center of the image set in the image library matches the cluster center of the image cluster. In this case, the image cluster may be stored into the image set.

[0077] In some embodiments, when the image cluster is stored into the image set, the cluster center of the image cluster may be compared with the set center. For example, the quality value of the cluster center may be compared with the quality value of the set center. If the quality value of the cluster center is greater than the quality value of the set center, the cluster center may be designated as the set center of the image set after the image cluster is stored into the image set. In another word, the set center of the image set may be updated by

determining the cluster center as the set center if the quality value of the cluster center is greater than the quality value of the set center. In this way, the set center of the image set may still be an image with a largest quality value among all the images in the image set.

[0078] It should be noted that the above descriptions of the media server 103 are 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, various modifications and changes in the forms and details of the application of the above method and system may occur without departing from the principles of the present disclosure. In some embodiments, the media server 103 may include one or more other modules. In some embodiments, two or more units in the media server 103 may form one module. However, those variations and modifications also fall within the scope of the present disclosure.

[0079] FIG. 5 is a flowchart illustrating an exemplary process for image clustering according to some embodiments of the present disclosure. In some embodiments, the process 500 shown in FIG. 5 may be implemented in the image clustering system 100 illustrated in FIG. 1. For example, at least a part of the process 500 may be stored in a storage (e.g., the buffer 106) as a form of instructions, and invoked and/or executed by the image clustering system 100 (e.g., the processor 222 of the computing device 200 or one or more modules illustrated in FIG. 4). In some embodiments, a part of the process 500 may be implemented on a terminal device. The operations of the illustrated process 500 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 in which the operations of the process 500 as illustrated in FIG. 5 and described below is not intended to be limiting.

[0080] In 510, one or more image clusters may be determined by performing a clustering operation on image data including a plurality of images. In some embodiments, the image clusters may be determined by the clustering module 420.

[0081] In some embodiments, the obtaining module 410 may obtain original image data including multiple images. In some embodiments, the multiple images may be captured by an image device within a predetermined time period. Exemplary imaging device may include an optical camera, an infrared camera, etc. The optical camera may be a digital camera, a video camera, a security camera, a web camera, a smartphone, a tablet, or a laptop equipped with a camera, etc. In some embodiments, the predetermined time period may be set by a user, according to default settings of the image clustering system 100, etc. For example, the predetermined time period may be 30 seconds, 1 minute, 2 minutes, 5 minutes, etc.

[0082] In some embodiments, the image cluster system 100 may identify images including at least a face of a person from the original image data including the multiple images. In some embodiments, an image identification technique may be used to identify a plurality of images including at least a face of a person from the multiple images. For example, a neural network based image identification technique may be used to identify the images including at least a face of a person from the multiple images. Exemplary neural networks may include, for example, convolutional neural network, recurrent neural network, generative adversarial network, deep neural network, etc. The identified original image data may be determined as image data, and the image clustering may be based on the faces.

[0083] As used herein, an image cluster may be a cluster or set of images. Images with higher similarities (e.g., larger than a certain similarity value) may be

grouped into an image cluster through the clustering operation. In some embodiments, images of a same person may be grouped into a same image cluster through the clustering operation. According to the clustering operation, the clustering module 420 may determine, for each of the plurality of images, a second similarity value between the image and each of the other images of the plurality of images, and a similar image set regarding the image by grouping images with second similarity values larger than a first threshold into the image cluster. Then a determination may be made as to whether each of the plurality of images is grouped into an image cluster. If an image of the plurality of images is not grouped into an image cluster, the clustering module 420 may group the image into a new image cluster. If an image of the plurality of images is grouped into an image cluster, the clustering module 420 may group the similar image set for the image into the image cluster. In some embodiments, the similarity value may be determined based on a similarity determination model. Features of the images of the one or more image clusters may be extracted and input into the similarity determination model. The output of the similarity determination model may be determined as the similarity values between the images. Merely by way of example, the similarity determination model may include a deep belief network (DBN), a Stacked Auto-Encoders (SAE), a logistic regression (LR) model, a support vector machine (SVM) model, a decision tree model, a Naive Bayesian Model, a random forest model, a Restricted Boltzmann Machine (RBM), a Gradient Boosting Decision Tree (GBDT) model, a LambdaMART model, an adaptive boosting model, a recurrent neural network (RNN) model, a convolutional network model, a hidden Markov model, a perceptron neural network model, a Hopfield network model, or the like, or any combination thereof. Details regarding the clustering operation may be described elsewhere in the present disclosure, for example, FIG. 7 and the

descriptions thereof.

[0084] In 520, a cluster center of each of the one or more image clusters may be determined. In some embodiments, the cluster center of each of the one or more clusters may be determined by the clustering module 420.

[0085] In some embodiment, a cluster center of an image cluster may be a first image grouped into the image cluster. For example, as for an image cluster (A, (B)), the image A may be the first image grouped into the image cluster. The image A may be determined as the cluster center of the image cluster (A, (B)). In some embodiments, if an image cluster includes two or more images, the cluster center may be determined and/or updated based on the two or more images. For example, if an image cluster includes two or more images, the cluster center may be updated based on images in the image cluster once a new image is grouped into the image cluster. As another example, if an image cluster includes two or more images, the cluster center may be updated based on images in the image cluster after the image cluster is finalized. The updated cluster center of an image cluster may be an image selected from images of the image cluster having a larger quality value than other images in the image cluster.

[0086] In some embodiments, the quality value of an image may reflect the quality of a face of a person in the image. The quality value may relate to at least one of an occlusion coefficient, a blur coefficient, an illumination coefficient, or a three-dimensional angle. In some embodiments, the quality value may be a weighted sum of the occlusion coefficient, the blur coefficient, the illumination coefficient, and the three-dimensional angle.

[0087] In 530, for each image cluster, a first similarity value may be determined by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to an image library. In some

embodiments, the first similarity value may be determined by a matching module 430.

[0088] An image set may refer to a collection of similar images. In some embodiments, an image set may include images of a same person. The image library may be a database storing the one or more image sets. Each image set may have a set center. In some embodiments, the set center of an image set may be a first image grouped into the image set. In some embodiments, if an image set includes two or more images, the set center may be updated based on the two or more images. For example, if an image set includes two or more images, the set center may be updated once a new image or an image cluster is grouped into the image cluster based on images in the image set. As another example, if an image set includes two or more images, the set center may be updated based on images in the image set after the image set is finalized. The updated set center of an image set may be an image selected from images of the image set having a larger quality value than other images in the image set.

[0089] The matching module 430 may determine the first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets. A determination as to whether the first similarity value is greater than a second threshold. If the first similarity value between the cluster center of the image cluster and a set center of an image set is greater than the second threshold, it may be determined that the cluster center of the image cluster matches the set center of the image set in the image library.

[0090] In some embodiments, the second threshold may be the same as the first threshold. In some embodiments, the second threshold may be different from the first threshold. For example, the second threshold may be smaller than the first threshold. The first threshold and the second threshold may be set by a user, according to default settings of the image clustering system 100, etc.

[0091] In 540, the one or more image clusters may be stored into the image library based on the first similarity value. In some embodiments, the one or more image clusters may be stored into the image library by the storing module 440.

[0092] If all the similarity values between a cluster center of an image cluster and set centers of the one or more image sets are less than the second threshold, it may indicate that there may be no set center in the image library that matches the cluster center of the image cluster. In this case, the image cluster may be designated as a new image set, and stored into the image library. The image library may be a collection of a plurality of image sets.

[0093] If a similarity value between a cluster center of an image cluster and a set center of an image set is less than the second threshold, it may indicate that the set center of the image set in the image library matches the cluster center of the image cluster. In this case, the image cluster may be stored into the image set.

[0094] In some embodiments, when the image cluster is stored into the image set, the cluster center of the image cluster may be compared with the set center. For example, the quality value of the cluster center may be compared with the quality value of the set center. If the quality value of the cluster center is greater than the quality value of the set center, the cluster center may be designated as the set center of the image set after the image cluster is stored into the image set. In another word, the set center of the image set may be updated by determining the cluster center as the set center if the quality value of the cluster center is greater than the quality value of the set center. In this way, the set center of the image set may still be an image with a largest quality value among all the images in the image set.

[0095] In some embodiments, the image library may be empty at an original

state. When a first image cluster is obtained by performing operations in 510 through 530, the image cluster may be designated as a first image set in the image library, and the cluster center may be designated as the set center of the image set since there is no other image sets in the image library.

[0096] FIG. 6 is a flowchart illustrating an exemplary process for image clustering according to some embodiments of the present disclosure. In some embodiments, the process 600 shown in FIG. 6 may be implemented in the image clustering system 100 illustrated in FIG. 1. For example, at least a part of the process 600 may be stored in a storage (e.g., the buffer 106) as a form of instructions, and invoked and/or executed by the image clustering system 100 (e.g., the processor 222 of the computing device 200 or one or more modules illustrated in FIG. 4). In some embodiments, a part of the process 600 may be implemented on a terminal device. The operations of the illustrated process 600 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 in which the operations of the process 600 as illustrated in FIG. 6 and described below is not intended to be limiting.

[0097] In some embodiments, a large number of images, each of which includes a face of a human, may need to be classified, so that images including faces of a same category may be classified into an image cluster. More particularly, images including the face of a same person may be classified into an image cluster.

[0098] In 610, image data including a plurality of images may be obtained. The image data may be obtained by the obtaining module 410.

[0099] In some embodiments, the image data may be obtained in real-time. The plurality of images may be captured in real-time by one or more cameras.

In some embodiments, feature processing may be performed on the obtained plurality of images to obtain the image data. The image data may be stored in a predetermined database (e.g., in the network storage device 113, the storage 227, or the storage 390). The obtaining module 410 may obtain image data including a preset number of images from the database periodically. In some embodiments, the obtaining module 410 may obtain the image data at a preset interval. The preset interval may be, for example, 1 hour, 12 hour, 24 hour, etc., which is not limited in the present disclosure. In some embodiments, the preset number may be determined according to a calculation capability of the image cluster system 100. If the number of images in the database is greater than or equal to the preset number, the obtaining module 410 may obtain images of the preset number. If the number of images in the database is smaller than the preset number, all the images in the database may be obtained. In some embodiments, the obtained images may be labeled, so that the obtained images may not be repeatedly obtained in a next data obtaining operation.

[0100] In 620, at least one image cluster may be determined by performing a clustering operation on the image data. In some embodiments, the at least one image cluster may be determined by the image clustering module 420.

[0101] A similarity value between each two of the set of images is greater than or equal to a certain threshold. In some embodiments, the determination of the at least one image cluster (e.g., by performing the clustering operation on the image data) may be implemented in combination of the process 700 as illustrated in FIG. 7.

[0102] In 630, a representative image of the image cluster may be determined. The representative image may be an image, or a collection of multiple images, which is not limited in the present disclosure. In some embodiments, the representative image may be a cluster center of the image cluster. In some

embodiment, the cluster center may be the first image grouped into the image cluster. For example, as for the image cluster (A, (B)) obtained in the operation 740, the image A may be the cluster center of the image cluster (A, (B)). In some embodiments, the representative image may be an image of the image cluster having a better quality value than other images in the image cluster.

[0103] In some embodiments, the quality value may relate to at least one of an occlusion coefficient, a blur coefficient, an illumination coefficient, or a three-dimensional angle. In some embodiments, the quality value may be a weighted sum of the occlusion coefficient, the blur coefficient, the illumination coefficient, and the three-dimensional angle. Merely by way of example, the quality value may be determined according to Equation (1):

$$f = \text{Occlusion} * k_1 + \text{Blur} * k_2 + \text{Illumination} * k_3 + \text{Pitch} * k_4 + \text{Roll} * k_5 + \text{Yaw} * k_6, \quad (1)$$

where f denotes the quality value of the image, Occlusion denotes the occlusion coefficient, Blur denotes the blur coefficient, Illumination denotes the illumination coefficient, Pitch denotes an angle of pitch ranging from -90 degrees (top) to 90 degrees (bottom) (i.e., $[-90^\circ, 90^\circ]$), Roll denotes an in-plane rotation angle ranging from -180 degrees (counterclockwise) to 180 degrees (clockwise) (i.e., $[-180^\circ, 180^\circ]$), Yaw denotes a left-right rotation angle ranging from -90 degrees (left) to 90 degrees (right) (i.e., $[-90^\circ, 90^\circ]$), and k_1 through k_6 may be weight coefficients. In some embodiments, the weight coefficients k_1 through k_6 may be determined by a user, according to default settings of the image clustering system 100, etc. In some embodiments, the weight coefficients k_1 through k_6 may be determined in accordance with specific situations.

[0104] In 640, a set center in an image library that matches the representative image of the image cluster may be determined. In some embodiments, one or more image sets in the image library may be obtained. The image library may

be a database storing a plurality of image sets. Each image set may include a plurality of similar images. In some embodiments, each image set may have a set center.

[0105] A determination may be made as to whether the representative image matches a set center of an image set in the image library. In some embodiments, a similarity value between the representative image and the set center of the image set in the image library may be calculated, and a determination as to whether the similarity value is greater than a second threshold. If the similarity value is greater than the second threshold, it may be determined that the representative image matches the set center of the image set in the image library.

[0106] In some embodiments, the second threshold may be the same as the first threshold. In some embodiments, the second threshold may be different from the first threshold. For example, the second threshold may be smaller than the first threshold. The first threshold and the second threshold may be set by a user, according to default settings of the image clustering system 100, etc.

[0107] In 650, the image cluster may be stored into the image library based on the similarity value.

[0108] In some embodiments, if there is no set center in the image library that matches the representative image, the image cluster may be designated as a new image set, and stored into the image library. The image library may be a collection of a plurality of image sets.

[0109] In some embodiments, the image library may be an empty library at original. When the representative image of the image cluster of a first batch obtained by performing operations in 610 through 630 is compared with the set center of the image library, there is no set center in the image library matches the representative image. The image cluster may be designated as a new image

set in the image library, and the representative image is designated as the set center of the image set.

[0110] When representative images of image clusters of subsequent batches obtained by performing operations in 610 through 630 is compared with the set center of the image library (the image library already includes one or more image sets), similarity values between the representative images of the image clusters and set centers of the one or more image sets in the image library greater than the second threshold may be identified. Image sets corresponding to the similarity values greater than the second threshold may be determined. In some embodiments, an image set corresponding to the greatest similarity value may be selected, and the image cluster may be stored into the image set, for example, as a subset of the image set.

[0111] If the similarity values are less than the second threshold, it may indicate that there may be no set center in the image library that matches the representative image, then the image cluster may be designated as a new image set in the image library, and the representative image may be designated as a set center of the new image set.

[0112] In some embodiments, if there is a set center of an image set in the image library that matches the representative image of the image cluster, the image cluster may be stored into the image set. In some embodiments, when the image cluster is stored into the image set, the representation image of the image cluster may be compared with the set center. For example, the quality value of the representation image may be compared with the quality value of the set center. The quality values of the representation image and the set center may be determined according to Equation (1).

[0113] If the quality value of the representative image is greater than the quality value of the set center, the representative image may be designated as the set

center of the image set after the image cluster is stored into the image set. In another word, the set center of the image set may be updated by determining the representative image as the set center if the quality value of the representative image is greater than the quality value of the set center. In this way, the set center of the image set may still be an image with a greatest quality value among all the images in the image set.

[0114] According to the methods provided in the processes 600, the image data may be processed periodically, and the image data may be clustered in a primary clustering (clustered into an image cluster) and a secondary clustering (clustered into an image set) sequentially. In the primary clustering, the image data may be clustered into an image cluster so that image data with a similarity value greater than the first threshold may be grouped into an image cluster, and an image of the image data with the greatest quality value in the image cluster may be determined as the representative image. In the primary clustering, the similarity calculation may be performed for each image in sequence. Taking the images A, B, C, D, and E as an example, a similarity value between image A and image B is greater than the first threshold, a similarity value between image B and image C is greater than the first threshold, even though a similarity value between image A and image C is smaller than the first threshold, the image C and image A may still be grouped into a same image cluster. In this way, similar images may be grouped into an image cluster as many as possible, thereby minimizing errors in image clustering greatly.

[0115] In the secondary clustering, the representative image may be compared with set centers of image sets in the image library to further cluster the image data, so that the image data with a similarity value reaches the second threshold may be grouped into an image set in the image library. Since the image library includes a large number of image sets, only the representative image may be

compared with the set center in the secondary clustering, thereby reducing the amount of calculation greatly, and realizing a fast classification of image data.

[0116] In comparison with the prior art, the method provided in the present disclosure may be adapted to camera systems having a large number of cameras, and interconnected with each camera via the image library, thus optimizing the classification and storage of image data.

[0117] FIG. 7 is a flowchart illustrating an exemplary process for performing a cluster operation on image data including a plurality of images according to some embodiments of the present disclosure. In some embodiments, the process 700 shown in FIG. 7 may be implemented in the image clustering system 100 illustrated in FIG. 1. For example, at least a part of the process 700 may be stored in a storage (e.g., the buffer 106) as a form of instructions, and invoked and/or executed by the image clustering system 100 (e.g., the processor 222 of the computing device 200 or one or more modules illustrated in FIG. 4). In some embodiments, a part of the process 700 may be implemented on a terminal device. The operations of the illustrated process 700 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 in which the operations of the process 700 as illustrated in FIG. 7 and described below is not intended to be limiting.

[0118] In 710, a similarity value between a selected image of the plurality of images and each of the other of the plurality of images may be determined. The other of the plurality of images refers to images of the plurality of images other than the selected image.

[0119] In some embodiments, a similarity calculation may be performed on the image data including the plurality of images. In some embodiments, for each of

the plurality of images, a similarity value between the image and the other images of the plurality of images may be determined. Taking image data including n images as an example, $n*(n-1)$ similarity values may be determined.

[0120] Merely by ways of example, if the image data includes five images including A, B, C, D, and E, twenty similarity values of A-B, A-C, A-D, B-A, B-C, ..., and E-D may be determined. As used herein, X-Y represents a similarity value between an image X and another image Y, and the image X is the selected image. For example, A-B represents a similarity value between A and B, and A is the selected image.

[0121] In 720, an image cluster regarding the image may be determined by grouping images of the other of the plurality of images with similarity values larger than a first threshold into the image cluster.

[0122] The image cluster including the grouped images with similarity values greater than the first threshold may be determined as a similar image set regarding the selected image. The similar image set regarding the selected image may refer to a collection of images with a higher similarity to the selected image. In some embodiments, a determination as to whether a similarity value between each image and the other of the plurality of images with respect to the image is greater than or equal to the first threshold may be made. For example, if a similarity value between a selected image and one or more images among the plurality of images other than the selected image is greater than or equal to the first threshold, the one or more images may be determined as elements of the similar image set regarding the selected image.

[0123] Merely for illustration purposes, for images A, B, C, D, and E, if similarity values of A-B, B-A, B-C, C-B, D-E, and E-D are greater than the first threshold, it may indicate that a similar image set regarding image A may be (B), a similar image set regarding image B may be (A, C), a similar image set regarding image

C may be (B), a similar image set regarding image D may be (E), and a similar image set regarding image E may be (D).

[0124] In 730, a determination as to whether each of the other of the plurality of images is grouped into an image cluster may be made.

[0125] The determination may be made for each image of the plurality of images other than the selected image sequentially to determine whether each image of the plurality of images other than the selected image is classified into an image cluster. For example, the plurality of images labeled sequentially in a certain order, and the determination may be made for the images according to the order. For example, the images A, B, C, D, and E may be labeled with successive numbers 1-5, then the determination may be performed on the images A, B, C, D, and E sequentially according to the order of 1-5. In some embodiments, any manner similar to or different from the manner of labeling the plurality of images may be used so as to ensure that the determination is made for each image of the plurality of images other than the selected image.

[0126] In some embodiments, the operations 710 through 730 may be performed with respect to each image of the plurality of images. For example, for the images A, B, C, D, and E, the operations 710 through 730 may be performed 5 times, and each time an image may be selected from the images A, B, C, D, and E as the selected image.

[0127] In some embodiments, the operations 710 through 730 may be performed with respect to a cluster center of the image cluster. Detailed descriptions regarding the cluster center of the image cluster may be disclosed elsewhere in the present disclosure, for example, the operation 630 in the process 600. In some embodiments, the cluster center of the image cluster may be designated by a user, according to default settings of the image clustering system 100, etc.

[0128] In 740, if an image is not grouped into any image cluster, the image and a similar image set regarding the image may be designated as a new image cluster.

[0129] Taking images A, B, C, D, and E as an example, when the determination in 730 is performed on the images A, B, C, D, and E sequentially, the image A may be selected as the selected image at first. Since there is no image cluster yet, the image A does not belong to any image cluster, and A and a similar image set regarding A may be grouped into an image cluster. In the above example, similarity values of A-B, B-A, B-C, C-B, D-E, and E-D are greater than the first threshold, and the similar image set regarding image A may be (B), then A, (B) may be grouped into an image cluster. Similarly, when the determination in 730 is performed on the image D, and the image D is not grouped into any image cluster, then D, (E) may be grouped into a new image cluster.

[0130] In 750, if an image is grouped into an image cluster, a similar image set regarding the image may be grouped into the image cluster.

[0131] If the image has been grouped into an image cluster, the similar image set of the image may also be grouped into the image cluster together with the image.

[0132] In the above example of the images A, B, C, D, and E, the image cluster (A, (B)) already exists. When the determination in 730 is performed on the image B, since B has been grouped into the image cluster (A, (B)), the similarity image set of B may further be grouped into image cluster (A, (B)). In this case, the similar image set regarding the image B (i.e., (A, C)) may be incorporated into the image cluster (A, (B)). Thus, an image cluster (A, (B), (A, C)) may be determined.

[0133] In some embodiments, a deduplication operation may be performed on each image cluster so as to remove same images from the image cluster. For

example, after a deduplication operation is performed on the image cluster (A, (B), (A, C)), the image cluster may become (A, B, C).

[0134] FIG. 8 is a schematic block diagram of an exemplary image clustering device according to some embodiments of the present disclosure. The image clustering device may include a processor 801 and a storage 802. The storage 802 may store a computer program for causing the processor 801 to execute the exemplary methods, processes, or operations described in the present disclosure. Referring to FIG. 9, an exemplary computer readable storage medium may be provided according to some embodiments of the present disclosure. In the embodiment, the computer readable storage medium 900 may store a computer program 910. When the computer program is executed by the processor 801, the methods, processes (e.g., the processes 500 through 700), or operations described above may be implemented.

[0135] The computer readable storage medium 900 may be a U disk, a mobile hard disk, a read-only memory (ROM), a random access memory (RAM), a magnetic disk, an optical disk, or a server that stores the computer program. The computer readable storage medium 900 may be a combination of a plurality of entities, such as a plurality of servers, a server including a storage, or a storage including a mobile hard disk, etc.

[0136] 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 example embodiments of this disclosure.

[0137] Moreover, certain terminology has been used to describe embodiments of the present disclosure. For example, the terms “one embodiment,” “an embodiment,” and “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.

[0138] 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 composition 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 “module,” “unit,” “component,” “device,” 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.

[0139] 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 electro-magnetic, 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.

[0140] Computer program code for carrying out operations for aspects of the present disclosure may be written in any 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).

[0141] Furthermore, the recited order of processing elements or sequences, or the use of numbers, letters, or other designations therefore, is 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.

[0142] 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, claim subject matter lie in smaller than all features of a single foregoing disclosed embodiment.

WE CLAIM:

1. A method for image clustering in an image library stored in a storage medium, the method comprising:

determining one or more image clusters by performing a clustering operation on image data including a plurality of images;

determining a cluster center of each of the one or more image clusters; and
for each image cluster,

determining a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library; and

storing the image cluster into the image library based on the first similarity value.

2. The method of claim 1, wherein determining the one or more image clusters by performing the clustering operation on the image data including the plurality of images includes:

for each of the plurality of images,

determining a second similarity value between the image and each of the other images of the plurality of images; and

determining a similar image set regarding the image by grouping images with second similarity values larger than a first threshold into the image cluster;

determining whether each of the plurality of images is grouped into an image cluster;

in response to a determination that an image of the plurality of images is not grouped into an image cluster,

grouping the image into a new image cluster; and

in response to a determination that an image of the plurality of images is grouped into an image cluster,

grouping the similar image set for the image into the image cluster.

3. The method of claim 1 or 2, further comprising: obtaining the image data including the plurality of images, wherein the plurality of images are captured by one or more imaging devices within a predetermined time period.

4. The method of any one of claims 1-3, wherein determining the cluster center of each of the one or more image clusters including:

for each image cluster,

determining a quality value of each of the images in the image cluster;

and

designating an image with a larger quality value than other images in the image cluster as the cluster center of the image cluster.

5. The method of claim 4, wherein the quality value is associated with at least one of an occlusion coefficient, a blur coefficient, an illumination coefficient, or a three-dimensional angle.

6. The method of any one of claims 1-5, wherein storing the image cluster into the image library based on the first similarity value includes:

determining whether the first similarity value between the cluster center of the image cluster and a set center of an image set is larger than a second threshold; and

in response to a determination that the first similarity value between the cluster center of the image cluster and a set center of an image set is larger than

the second threshold,

storing the image cluster into the image set, or

in response to a determination that the first similarity value between the cluster center of the image cluster and the set center of each of the one or more image sets is not larger than the second threshold,

designating the image cluster as a new image set; and

storing the new image set into the image library.

7. The method of claim 6, wherein storing the image cluster into the image set includes:

storing the image cluster as a subset of the image set.

8. The method of claim 7, further including:

updating the set center of the image set based on the quality values of images in the image set.

9. The method of any one of claims 1-8, wherein each of the plurality of images includes at least a face of a person, and the clustering operation is based on comparing the faces.

10. The method of claim 9, further including:

obtaining original image data including multiple images;

identifying images including at least a face of a person from the multiple images; and

designating the identified images including at least a face of a person as the plurality of images.

11. A system for image clustering in an image library stored in a storage medium, the system being configured to perform operations including:

- determining one or more image clusters by performing a clustering operation on image data including a plurality of images;

- determining a cluster center of each of the one or more image clusters; and
- for each image cluster,

- determining a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library; and

- storing the image cluster into the image library based on the first similarity value.

12. The system of claim 11, wherein determining the one or more image clusters by performing the clustering operation on the image data including the plurality of images includes:

- for each of the plurality of images,

- determining a second similarity value between the image and each of the other images of the plurality of images; and

- determining a similar image set regarding the image by grouping images with second similarity values larger than a first threshold into the image cluster;

- determining whether each of the plurality of images is grouped into an image cluster;

- in response to a determination that an image of the plurality of images is not grouped into an image cluster,

- grouping the image into a new image cluster; and

- in response to a determination that an image of the plurality of images is

grouped into an image cluster,

grouping the similar image set for the image into the image cluster.

13. The system of claim 11 or 12, further comprising: obtaining the image data including the plurality of images, wherein the plurality of images are captured by one or more imaging devices within a predetermined time period.

14. The system of any one of claims 11-13, wherein determining the cluster center of each of the one or more image clusters including:

for each image cluster,

determining a quality value of each of the images in the image cluster;

and

designating an image with a larger quality value than other images in the image cluster as the cluster center of the image cluster.

15. The system of claim 14, wherein the quality value is associated with at least one of an occlusion coefficient, a blur coefficient, an illumination coefficient, or a three-dimensional angle.

16. The system of any one of claims 11-15, wherein storing the image cluster into the image library based on the first similarity value includes:

determining whether the first similarity value between the cluster center of the image cluster and a set center of an image set is larger than a second threshold;

and

in response to a determination that the first similarity value between the cluster center of the image cluster and a set center of an image set is larger than the second threshold,

storing the image cluster into the image set, or
in response to a determination that the first similarity value between the cluster center of the image cluster and the set center of each of the one or more image sets is not larger than the second threshold,
designating the image cluster as a new image set; and
storing the new image set into the image library.

17. The system of claim 16, wherein storing the image cluster into the image set includes:

storing the image cluster as a subset of the image set.

18. The system of claim 17, the operations further including:

updating the set center of the image set based on the quality values of images in the image set.

19. The system of any one of claims 11-18, wherein each of the plurality of images includes at least a face of a person, and the clustering operation is based on comparing the faces.

20. A non-transitory computer readable medium, comprising at least one set of instructions for image clustering in an image library stored in a storage medium, wherein when executed by a device, the at least one set of instructions causes the device to perform one or more operations, the one or more operations comprising:

determining one or more image clusters by performing a clustering operation on image data including a plurality of images;

determining a cluster center of each of the one or more image clusters; and

for each image cluster,

 determining a first similarity value by comparing the cluster center of the image cluster and a set center of each of one or more image sets, the image sets belonging to the image library; and

 storing the image cluster into the image library based on the first similarity value.

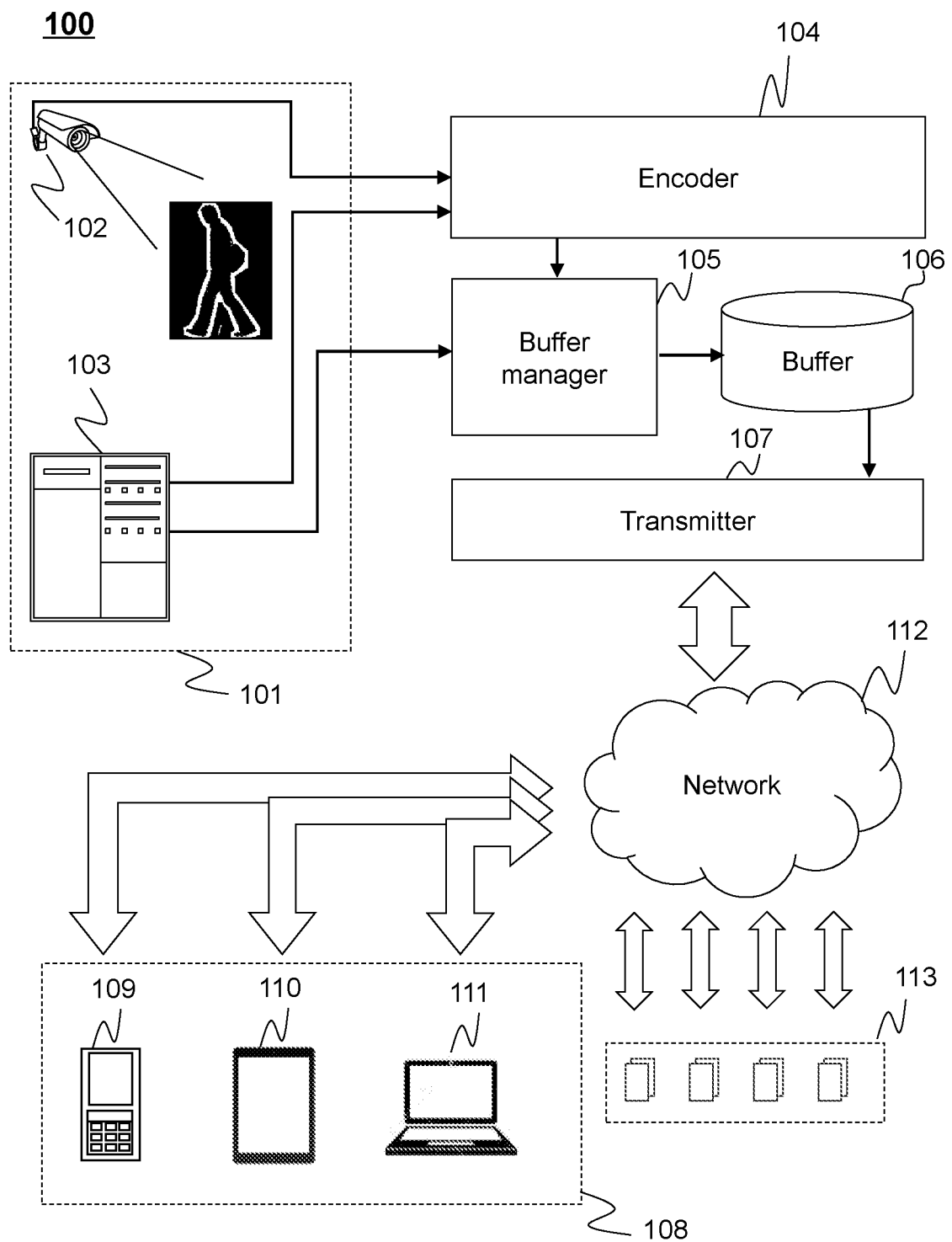
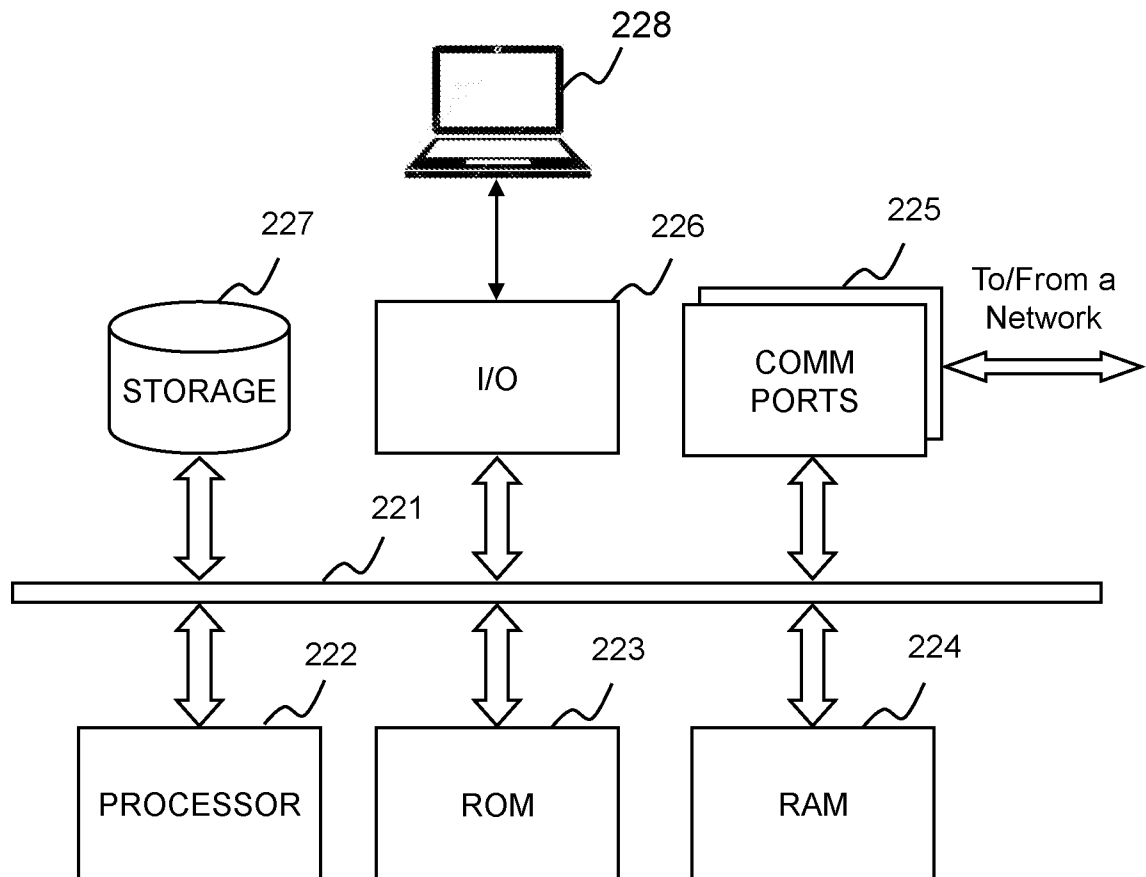
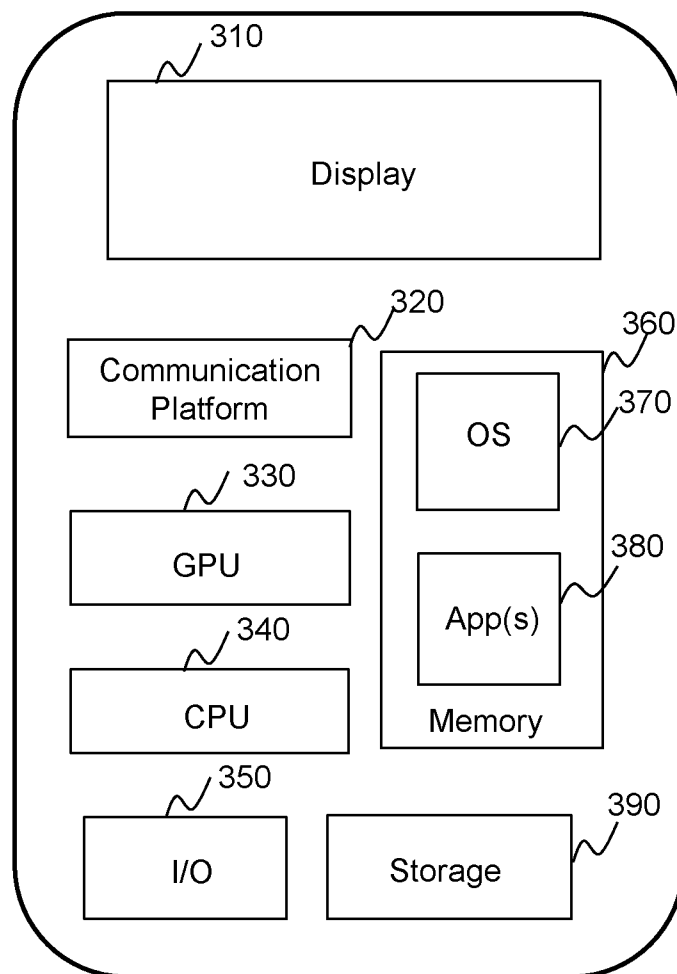


FIG. 1

200**FIG. 2**

300**FIG. 3**

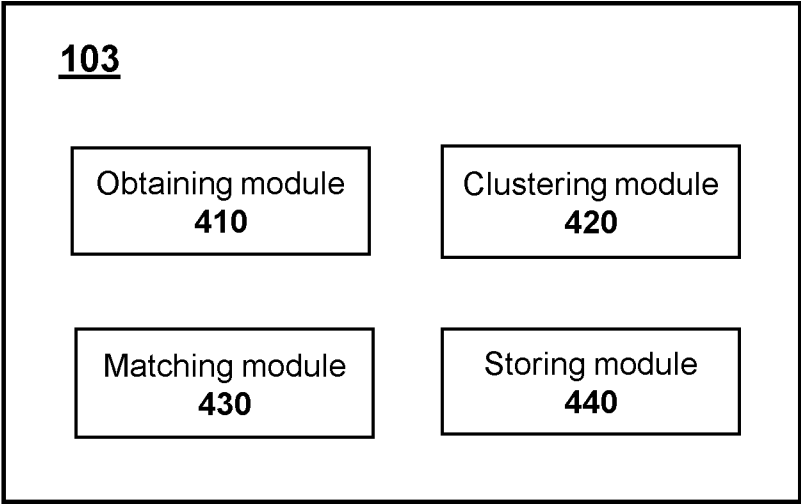
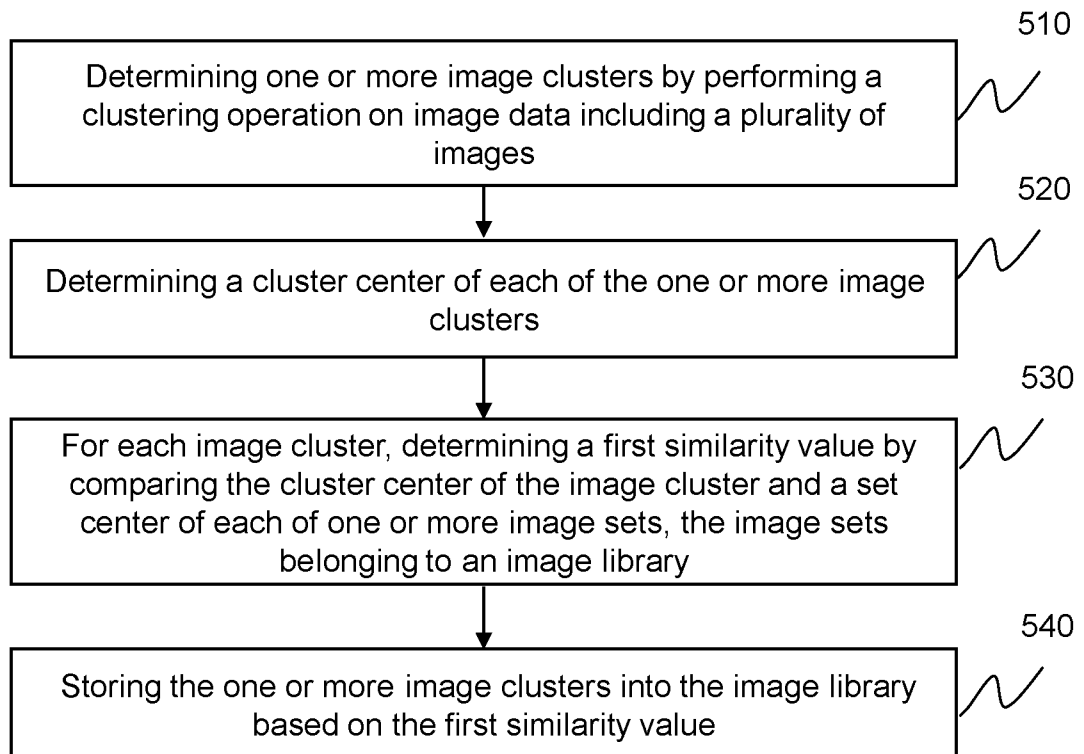
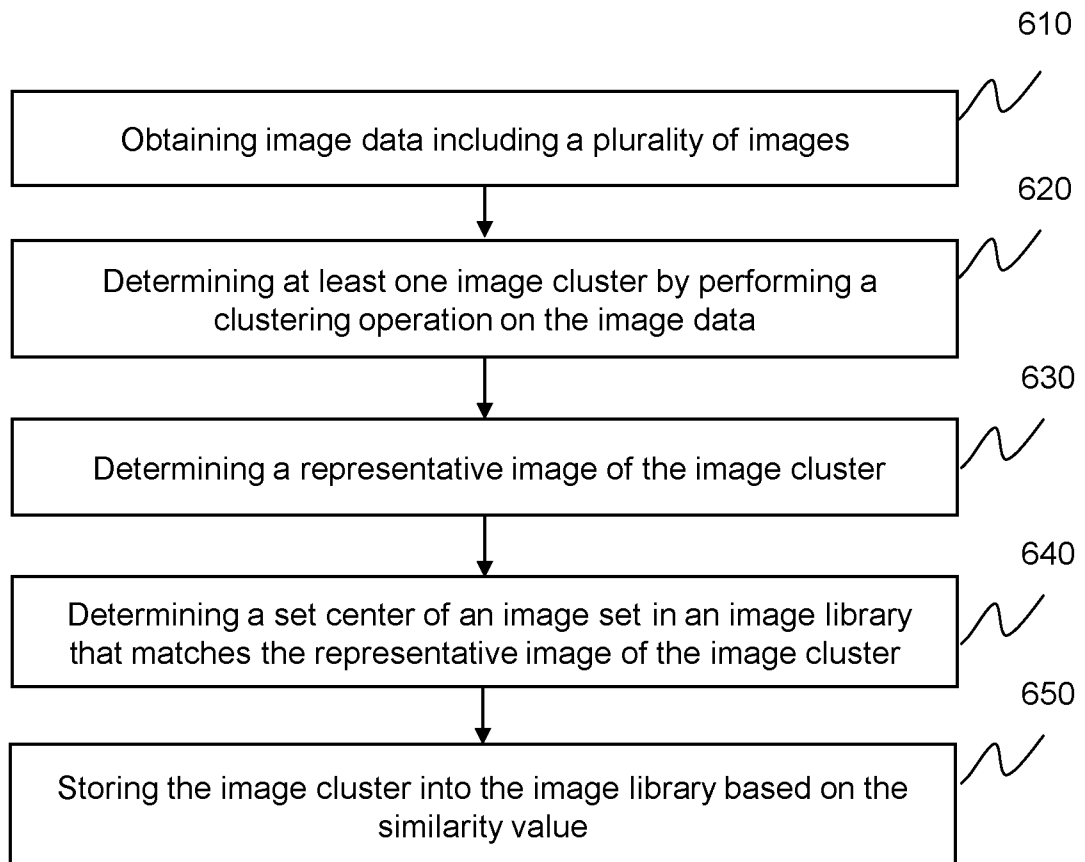
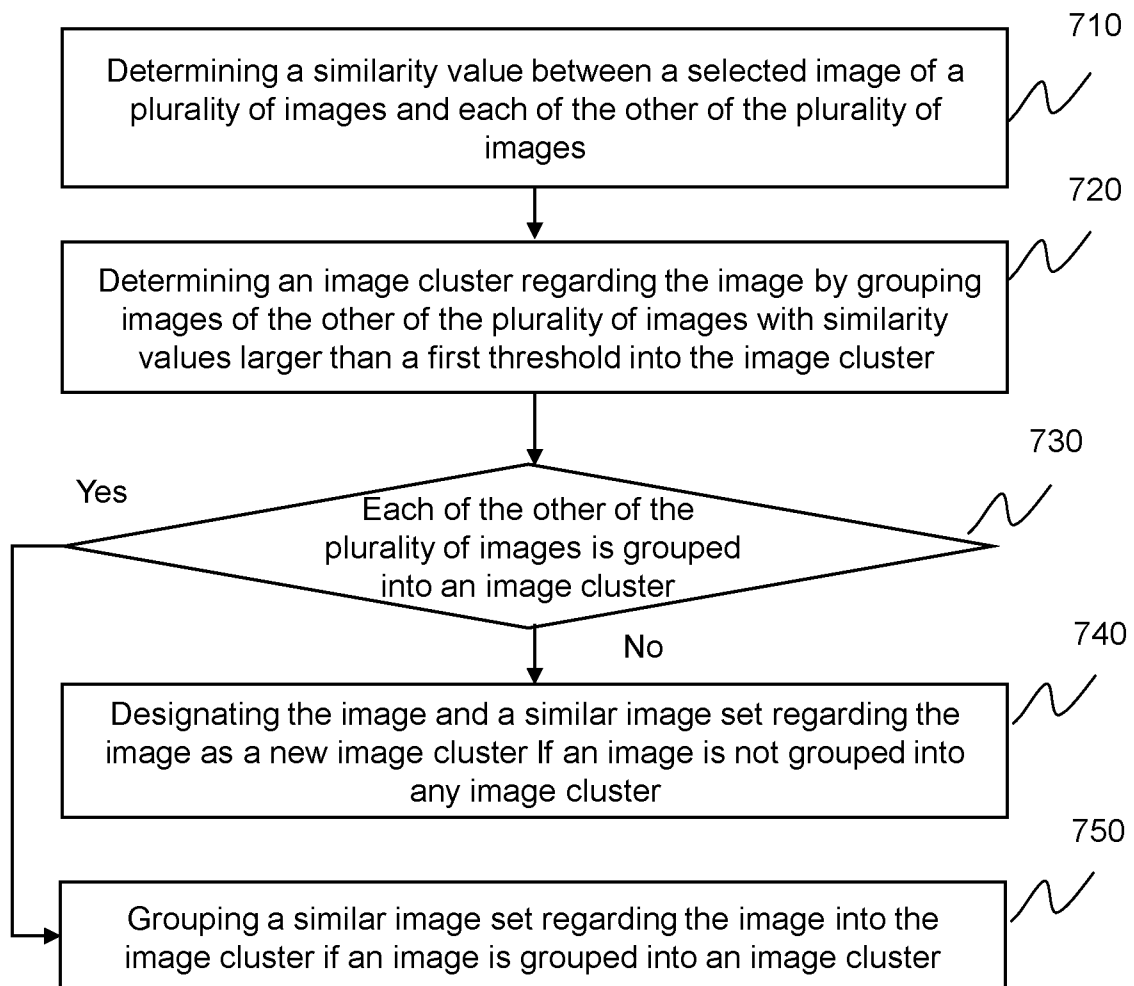


FIG. 4

500**FIG. 5**

600**FIG. 6**

700**FIG. 7**

800

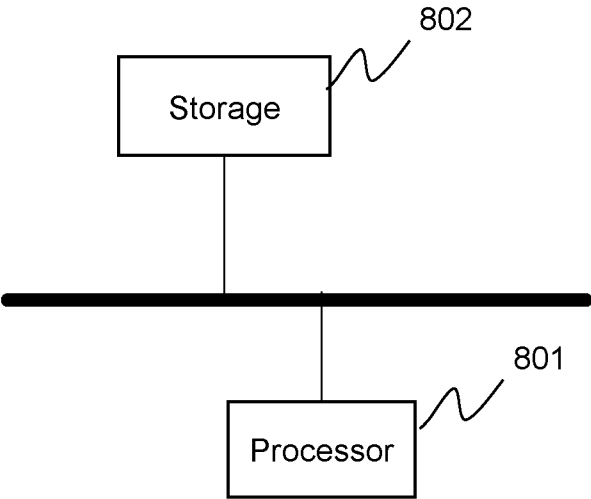


FIG. 8

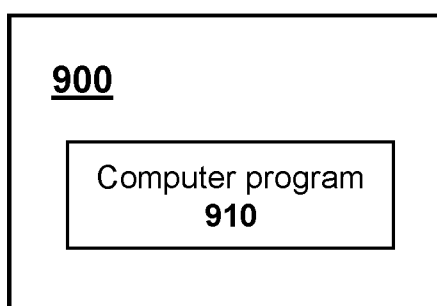


FIG. 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/117421

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(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)

G06K

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)

CNPAT, EPODOC, WPI, CNKI, IEEE: cluster, clustering, center, compar+, set, representative w image, key 1d frame, centroid, quality, summary, library, similarity, match, distance

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2011211764 A1 (MICROSOFT CORPORATION) 01 September 2011 (2011-09-01) description, paragraphs 0054-0060, 0169-0184, figure 12	1-20
Y	US 2006200475 A1 (ASTMAN KODAK COMPANY) 07 September 2006 (2006-09-07) description, paragraphs 0035-0042, figure 3	1-20
A	US 2011129126 A1 (AT&T INTELLECTUAL PROPERTY I, L.P.) 02 June 2011 (2011-06-02) the whole document	1-20
A	US 9613455 B1 (GOOGLE INC.) 04 April 2017 (2017-04-04) the whole document	1-20



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

19 April 2020

Date of mailing of the international search report

26 April 2020

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

LI,Nan

Facsimile No. (86-10)62019451

Telephone No. 86-(10)-53961438

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/117421

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2011211764	A1	01 September 2011	CN	102200988	A	28 September 2011
US	2006200475	A1	07 September 2006	EP	1854033	A1	14 November 2007
				JP	2008532180	A	14 August 2008
				WO	2006096384	A1	14 September 2006
				CN	101167076	A	23 April 2008
US	2011129126	A1	02 June 2011	US	2013070975	A1	21 March 2013
US	9613455	B1	04 April 2017	None			