(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 840 393 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
23.06.2021 Bulletin 2021/25

(21) Application number: **21152338.6**

(22) Date of filing: **12.03.2014**

(51) Int Cl.:
H04N 21/43 *(2011.01)* ***H04N 21/433*** *(2011.01)*
G06F 3/147 *(2006.01)* ***G06F 1/16*** *(2006.01)*
G06F 16/50 *(2019.01)* ***G06F 16/51*** *(2019.01)*
G06F 16/583 *(2019.01)* ***G06F 16/955*** *(2019.01)*
G06T 19/00 *(2011.01)* ***H04N 1/00*** *(2006.01)*

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **15.03.2013 US 201313839818**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
14769648.8 / 2 974 338

(71) Applicant: **RPX Corporation**
San Francisco, CA 94111 (US)

(72) Inventor: **MULLINS, Brian**
Sierra Madre, CA California 91024 (US)

(74) Representative: **Schwegman Lundberg Woessner Limited**
Hillington Park Innovation Centre
1 Ainslie Road
Glasgow G52 4RU (GB)

Remarks:
This application was filed on 19.01.2021 as a divisional application to the application mentioned under INID code 62.

(54) CONTEXTUAL LOCAL IMAGE RECOGNITION DATASET

(57) A contextual local image recognition module of a device retrieves a primary content dataset from a server and then generates and updates a contextual content dataset based on an image captured with the device. The device stores the primary content dataset and the contextual content dataset. The primary content dataset comprises a first set of images and corresponding virtual object models. The contextual content dataset comprises a second set of images and corresponding virtual object models retrieved from the server.

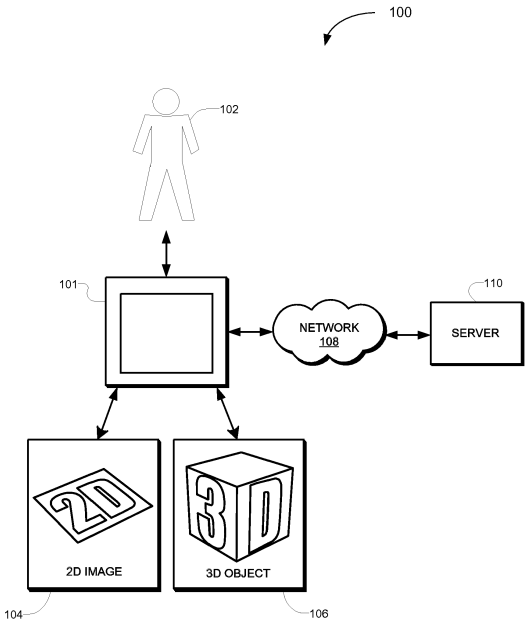

FIG. 1

EP 3 840 393 A1

Description

PRIORITY APPLICATION

[0001] This application claims the benefit of priority to U.S. Application Serial No. 13/839,818, filed on 15 March 2013, which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] The subject matter disclosed herein generally relates to the processing of data. Specifically, the present disclosure addresses systems and methods that involve a contextual local image recognition dataset.

BACKGROUND

[0003] A device can be used to generate data based on an image captured with the device. For example, augmented reality (AR) is a live, direct or indirect, view of a physical, real-world environment whose elements are augmented by computer-generated sensory input such as sound, video, graphics or GPS data. With the help of advanced AR technology (e.g. adding computer vision and object recognition), the information about the surrounding real world of the user becomes interactive. Device-generated (e.g., artificial) information about the environment and its objects can be overlaid on the real world.

[0004] The device may constantly scan and capture images and send them to an AR server for recognition. Once an image matches with a stock image in the database of the AR server, the device then downloads the three-dimensional model from the AR server. The constant uploading of scanned images from the device to the server and the constant downloading of the three-dimensional models may add data traffic to limited and valuable network bandwidth between the device and AR server.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] Some embodiments are illustrated by way of example and not limitation in the figures of the accompanying drawings.

FIG. 1 is a block diagram illustrating an example of a network suitable for operating the device, according to some example embodiments.

FIG. 2 is a block diagram illustrating modules (e.g., components) of the device, according to some example embodiments.

FIG. 3 is a block diagram illustrating modules (e.g., components) of a contextual local image recognition dataset module, according to some example embodiments.

FIG. 4 is a ladder diagram illustrating an operation of the contextual local image recognition dataset module of the device, according to some example embodiments.

FIG. 5 is a flowchart illustrating an example operation of the contextual local image recognition dataset module of the device, according to some example embodiments.

FIG. 6 is a flowchart illustrating another example operation of the contextual local image recognition dataset module of the device, according to some example embodiments.

FIG. 7 is a block diagram illustrating components of a machine, according to some example embodiments, able to read instructions from a machine-readable medium and perform any one or more of the methodologies discussed herein.

DETAILED DESCRIPTION

[0006] Example methods and systems are directed to a contextual local image recognition dataset. Examples merely typify possible variations. Unless explicitly stated otherwise, components and functions are optional and may be combined or subdivided, and operations may vary in sequence or be combined or subdivided. In the following description, for purposes of explanation, numerous specific details are set forth to provide a thorough understanding of example embodiments. It will be evident to one skilled in the art, however, that the present subject matter may be practiced without these specific details.

[0007] Augmented reality applications allow a user to experience information, such as in the form of a virtual object overlaid on a picture of a physical object captured by a camera of a device. The physical object may include a visual reference that the augmented reality application can identify. A visualization of the additional information, such as the virtual object engaged (e.g., overlaid on top of) with an image of the physical object is generated in a display of the device. The virtual object may be generated based on the recognized visual reference. A rendering of the visualization of the virtual object may be based on a position of the display relative to the visual reference.

[0008] A contextual local image recognition module in the device retrieves a primary content dataset from a server. The primary content dataset comprises a first set of images and corresponding virtual object models. For example, the first set of images may include most common images that a user of the device is likely to capture with the device. The contextual content dataset comprises a second set of images and corresponding virtual ob-

ject models retrieved from the server. The contextual local image recognition module generates and updates the contextual content dataset based an image captured with the device. A storage device of the device stores the primary content dataset and the contextual content dataset.

[0009] **FIG. 1** is a network diagram illustrating a network environment 100 suitable for operating an augmented reality application of a device, according to some example embodiments. The network environment 100 includes a device 101 and a server 110, communicatively coupled to each other via a network 108. The device 101 and the server 110 may each be implemented in a computer system, in whole or in part, as described below with respect to FIG. 7.

[0010] The server 110 may be part of a network-based system. For example, the network-based system may be or include a cloud-based server system that provides additional information such, as three-dimensional models, to the device 101.

[0011] FIG. 1 illustrates a user 102 using the device 101. The user may be a human user (e.g., a human being), a machine user (e.g., a computer configured by a software program to interact with the device 101), or any suitable combination thereof (e.g., a human assisted by a machine or a machine supervised by a human). The user 102 is not part of the network environment 100, but is associated with the device 101 and may be a user of the device 101. For example, the device 101 may be a desktop computer, a vehicle computer, a tablet computer, a navigational device, a portable media device, or a smart phone belonging to the user 102.

[0012] The user 102 may be a user of an application in the device 101. The application may include an augmented reality application configured to provide the user 102 with an experience triggered by a physical object such as, a two-dimensional physical object 104 (e.g., a picture) or a three-dimensional physical object 106 (e.g., a statue). For example, the user 102 may point a camera of the device 101 to capture an image of the two-dimensional physical object 104. The image is recognized locally in the device 101 using a local context recognition dataset module of the augmented reality application of the device 101. The augmented reality application then generates additional information corresponding to the image (e.g., a three-dimensional model) and presents this additional information in a display of the device 101 in response to recognizing the image. If the captured image is not recognized locally at the device 101, the device 101 downloads additional information (e.g., the three-dimensional model) corresponding to the captured image, from a database of the server 110 over the network 108.

[0013] Any of the machines, databases, or devices shown in FIG. 1 may be implemented in a general-purpose computer modified (e.g., configured or programmed) by software to be a special-purpose computer to perform one or more of the functions described herein for that machine, database, or device. For example, a computer system able to implement any one or more of the methodologies described herein is discussed below with respect to FIG. 7. As used herein, a "database" is a data storage resource and may store data structured as a text file, a table, a spreadsheet, a relational database (e.g., an object-relational database), a triple store, a hierarchical data store, or any suitable combination thereof. Moreover, any two or more of the machines, databases, or devices illustrated in FIG. 1 may be combined into a single machine, and the functions described herein for any single machine, database, or device may be subdivided among multiple machines, databases, or devices.

[0014] The network 108 may be any network that enables communication between or among machines (e.g., server 110), databases, and devices (e.g., device 101). Accordingly, the network 108 may be a wired network, a wireless network (e.g., a mobile or cellular network), or any suitable combination thereof. The network 108 may include one or more portions that constitute a private network, a public network (e.g., the Internet), or any suitable combination thereof.

[0015] **FIG. 2** is a block diagram illustrating modules (e.g., components) of the device, according to some example embodiments. The device 101 may include sensors 202, a display 204, a processor 206, and a storage device 207. For example, the device 201 may be a desktop computer, a vehicle computer, a tablet computer, a navigational device, a portable media device, or a smart phone of a user. The user may be a human user (e.g., a human being), a machine user (e.g., a computer configured by a software program to interact with the device 101), or any suitable combination thereof (e.g., a human assisted by a machine or a machine supervised by a human).

[0016] The sensors 202 may include, for example, a vibration sensor, an optical sensor (e.g., a charge-coupled device (CCD)), an orientation sensor (e.g., gyroscope), an audio sensor (e.g., a microphone), or any suitable combination thereof. For example, the sensors 202 may include a rear facing camera and a front facing camera in the device 101. It is noted that the sensors described herein are for illustration purposes and the sensors 102 are thus not limited to the ones described.

[0017] The display 204 may include, for example, a touchscreen or touch sensitive display configured to receive a user input via a contact on the touchscreen display. In another example, the display 204 may include a screen or monitor configured to display images generated by the processor 206.

[0018] The processor 206 may include an experience consuming application, such as an augmented reality application 209, configured to generate a visualization of a virtual object, such as a three-dimensional virtual object, overlaid (e.g., superimposed upon, or otherwise displayed in tandem with) on an image of a physical object captured by a camera of the device 101 in the display 204 of the device 101. A visualization of the three-dimensional virtual object may be manipulated by adjusting a

position of the physical object relative to the camera of the device 101. Similarly, the visualization of the three-dimensional virtual object may be manipulated by adjusting a position of the device 101 relative to the physical object.

[0019] In one embodiment, the augmented reality application 209 communicates with a contextual local image recognition module 208 in the device 101 to retrieve three-dimensional models of virtual objects associated with a captured image (e.g., virtual object that correspond to the captured image). For example, the captured image may include a visual reference (also referred to as a marker) that consists of an identifiable image, symbol, letter, number, machine-readable code. For example, the visual reference may include a bar code, a quick response (QR) code, or an image that has been previously associated with a three-dimensional virtual object (e.g., an image that has been previously determined to correspond to the three-dimensional virtual object).

[0020] The contextual local image recognition module 208 may be configured to determine whether the captured image matches an image that is stored in a local database of images and corresponding additional information (e.g., three-dimensional model and one or more interactive features) in the storage device 207 of the device 101. In one embodiment, the contextual local image recognition module 208 retrieves a primary content dataset from the server 110 and generates and updates a contextual content dataset based an image captured by the device 101.

[0021] The storage device 207 may be configured to store a database of visual references (e.g., images) and corresponding experiences (e.g., three-dimensional virtual objects, interactive features of the three-dimensional virtual objects). For example, the visual reference may include a machine-readable code or a previously identified image (e.g., a picture of shoe). The previously identified image of the shoe may correspond to a three-dimensional virtual model of the shoe that can be viewed from different angles by manipulating the position of the device 101 relative to the picture of the shoe. Features of the three-dimensional virtual shoe may include selectable icons on the three-dimensional virtual model of the shoe. An icon may be selected or activated by tapping or moving on the device 101.

[0022] In one embodiment, the storage device 207 includes a primary content dataset 210 and a contextual content dataset 212. The primary content dataset 210 includes, for example, a first set of images and corresponding experiences (e.g., interactive three-dimensional virtual object models, downloadable content, audio, video, images, text, webpages, games, advertisements, 3D models, etc.).). The primary content dataset 210 may include a core set of images or the most popular images determined by the server 110. The core set of images may include a limited number of images identified by the server 110. For example, the core set of images may include the cover image of the ten most popular magazines and their corresponding experiences (e.g., virtual objects that represent the ten most popular magazines). In another example, the server 110 may generate the first set of images based on the most popular or often scanned images received at the server 110.

[0023] The contextual content dataset 212 includes, for example, a second set of images and corresponding experiences (e.g., three-dimensional virtual object models) retrieved from the server 110 based on usage of the device 101 and the augmented reality application 209. For example, images captured with the device 101 that are not recognized in the primary content dataset 210 are submitted to the server 110 for recognition. If the captured image is recognized by the server 110, a corresponding experience may be downloaded at the device 101 and stored in the contextual content dataset 212.

[0024] In one embodiment, the device 101 may communicate over the network 108 with the server 110 to retrieve a portion of a database that stores visual references, corresponding three-dimensional virtual objects, and corresponding interactive features of the three-dimensional virtual objects. The network 108 may be any network that enables communication between or among machines, databases, and devices (e.g., the device 101). Accordingly, the network 108 may be a wired network, a wireless network (e.g., a mobile or cellular network), or any suitable combination thereof. The network may include one or more portions that constitute a private network, a public network (e.g., the Internet), or any suitable combination thereof.

[0025] Any one or more of the modules described herein may be implemented using hardware (e.g., a processor of a machine) or a combination of hardware and software. For example, any module described herein may configure a processor to perform the operations described herein for that module. Moreover, any two or more of these modules may be combined into a single module, and the functions described herein for a single module may be subdivided among multiple modules. Furthermore, according to various example embodiments, modules described herein as being implemented within a single machine, database, or device may be distributed across multiple machines, databases, or devices.

[0026] **FIG. 3** is a block diagram illustrating modules (e.g., components) of contextual local image recognition dataset module 208, according to some example embodiments. The contextual local image recognition dataset module 208 may include an image capture module 302, a local image recognition module 304, a content request module 306, and a context content dataset update module 308.

[0027] The image capture module 302 may capture an image with a camera of the device 101. For example, the image capture module 302 may capture the image of a physical object pointed at by the device 101. In one embodiment, the image capture module 302 may capture one image or a series of snapshots. In another embodiment, the image capture module 302 may capture an

image when one of the sensors 202 (e.g., vibration, gyroscope, compass, etc.) detects that the device 101 is no longer moving or is being held steady.

[0028] The local image recognition module 304 determines that the captured image corresponds to an image stored in the primary content dataset 210 and locally renders the three-dimensional virtual object model corresponding to the image captured with the device 101 when the image captured with the device 101 corresponds to one of the set of images of the primary content dataset 210 stored in the device 101.

[0029] In another example embodiment, the local image recognition module 304 determines that the captured image corresponds to an image stored in the contextual content dataset 212 and locally renders the three-dimensional virtual object model corresponding to the image captured with the device 101 when the image captured with the device 101 corresponds to one of the set of images of the contextual content dataset 212 stored in the device 101.

[0030] The content request module 306 may request the server 110 for the three-dimensional virtual object model corresponding to the image captured with the device 101 when the image captured with the device 101 does not correspond to one of the set of images in the primary content dataset 210 and the set of images in the contextual content dataset 212 in the storage device 207.

[0031] The context content dataset update module 308 may receive the three-dimensional virtual object model corresponding to the image captured with the device 101 from the server 110 in response to the request generated by the content request module 306. In one embodiment, the context content dataset update module 308 may update the contextual content dataset 212 with the three-dimensional virtual object model corresponding to the image captured with the device 101 from the server 110 when the image captured with the device 101 does not correspond to any images stored locally in the storage device 207 of the device 101.

[0032] In another example embodiment, the content request module 306 may determine usage conditions of the device 101 and generate a request to the server 110 for a third set of images and corresponding three-dimensional virtual object models based on the usage conditions. The usage conditions may be related to when, how often, where, and how the user is using the device 101. The context content dataset update module 308 may update the contextual content dataset with the third set of images and corresponding three-dimensional virtual object models.

[0033] For example, the content request module 306 determines that the user 102 scans pages of a newspaper in the morning time. The content request module 306 then generates a request to the server 110 for a set of images and corresponding experiences that are relevant to usage of the user 102 in the morning. For example, the content request module 306 may retrieve images of sports articles that the user 102 is most likely to scan in the morning and a corresponding updated virtual score board of a sports team mentioned in one of the sports articles. The experience may include, for example, a fantasy league score board update that is personalized to the user 102.

[0034] In another example, the content request module 306 determines that the user 102 often scans the business section of a newspaper. The content request module 306 then generates a request to the server 110 for a set of images and corresponding experiences that are relevant to the user 102. For example, the content request module 306 may retrieve images of business articles of the next issue of the newspaper as soon as the next issue's business articles are available. The experience may include, for example, a video report corresponding to an image of the next issue business article.

[0035] In yet another example embodiment, the content request module 306 may determine social information of the user 102 of the device 101 and generate a request to the server 110 for another set of images and corresponding three-dimensional virtual object models based on the social information. The social information may be obtained from a social network application in the device 101. The social information may relate to who the user 102 has interacted with, and with whom the user 102 has shared experiences using the augmented reality application 209 of the device 101. The context content dataset update module 308 may update the contextual content dataset 212 with the other set of images and corresponding three-dimensional virtual object models.

[0036] For example, the user 102 may have scanned several pages of a magazine. The content request module 306 determines from a social network application that the user 102 is friends with another user who shares similar interests and reads another magazine. As such, the content request module 306 may generate a request to the server 110 for a set of images and corresponding experiences related to the other magazine.

[0037] In another example, if the content request module 306 determines that the user 102 has scanned the cover image of a magazine, the content request module 306 may generate a request for additional content related to other images in the same magazine.

[0038] **FIG. 4** is a ladder diagram illustrating an operation of the contextual local image recognition dataset module 208 of the device 101, according to some example embodiments. At operation 402, the device 101 downloads an augmented reality application from the server 110. The augmented reality application may include the primary content dataset 210. The primary content dataset 210 may include for example, the most often scanned pictures of ten popular magazines and corresponding experiences. At operation 404, the device 101 captures an image.

[0039] At operation 406, the device 101 compares the captured image with local images from the primary content dataset 210 and from the contextual content dataset 212. If the captured image is not recognized in both the

primary content dataset and the contextual content dataset, the device 101 requests the server 110 at operation 408 to retrieve content or an experience associated with the captured image.

[0040] At operation 410, the server 110 identifies the captured image and retrieves content associated with the captured image.

[0041] At operation 412, the device 101 downloads the content corresponding to the captured image, from the server 110.

[0042] At operation 414, the device 101 updates its local storage to include the content. In one embodiment, the device 101 updates its contextual content dataset 212 with the downloaded content from operation 412.

[0043] Input conditions from the device 101 may also be submitted to the server 110 at operation 416. The input conditions may include usage time information, location information, a history of scanned images, and social information. The server 110 may retrieve content associated with the input conditions at operation 418. For example, if the input conditions indicate that the user 102 operates the device 101 mostly from location A, then content relevant to location A (e.g., restaurants nearby) may be retrieved from the server 110.

[0044] At operation 420, the device 101 downloads the content retrieved in operation 418 and updates the contextual content dataset based on the retrieved content at 422.

[0045] **FIG. 5** is a flowchart illustrating an example operation of the contextual local image recognition dataset module of the device, according to some example embodiments.

[0046] At operation 502, the contextual local image recognition dataset module 208 stores the primary content dataset 210 in the device 101.

[0047] At operation 504, the augmented reality application 209 determines that an image has been captured with the device 101.

[0048] At operation 506, the contextual local image recognition dataset module 208 compares the captured image with a set of images locally stored in the primary content dataset 210 in the device 101. If the captured image corresponds to an image from the set of images locally stored in the primary content dataset 210 in the device 101, the augmented reality application 209 generates an experience based on the recognized image at operation 508.

[0049] If the captured image does not correspond to an image from the set of images locally stored in the primary content dataset 210 in the device 101, the contextual local image recognition dataset module 208 compares the captured image with a set of images locally stored in the contextual content dataset 212 in the device 101 at operation 510.

[0050] If the captured image corresponds to an image from the set of images locally stored in the contextual content dataset 212 in the device 101, the augmented reality application 209 generates an experience based on the recognized image at operation 508.

[0051] If the captured image does not correspond to an image from the set of images locally stored in the contextual content dataset 212 in the device 101, the contextual local image recognition dataset module 208 submits a request including the captured image to the server 110 at operation 512.

[0052] At operation 514, the device 101 receives content corresponding to the captured image from the server 110.

[0053] At operation 516, the contextual local image recognition dataset module 208 updates the contextual content dataset 212 based on the received content.

[0054] **FIG. 6** is a flowchart illustrating another example operation of the contextual local image recognition dataset module of the device, according to some example embodiments.

[0055] At operation 602, the contextual local image recognition dataset module 208 captures input conditions local to the device 101. As previously described, local input conditions may include usage time information, location information, history of scanned images, and social information.

[0056] At operation 604, the contextual local image recognition dataset module 208 communicates the input conditions to the server 110. At operation 606, the server 110 retrieves new content related to the input conditions of the device 101.

[0057] At operation 608, the contextual local image recognition dataset module 208 updates the contextual content dataset 212 with the new content.

[0058] **FIG. 7** is a block diagram illustrating components of a machine 700, according to some example embodiments, able to read instructions from a machine-readable medium (e.g., a machine-readable storage medium, a computer-readable storage medium, or any suitable combination thereof) and perform any one or more of the methodologies discussed herein, in whole or in part. Specifically, FIG. 7 shows a diagrammatic representation of the machine 700 in the example form of a computer system and within which instructions 724 (e.g., software, a program, an application, an applet, an app, or other executable code) for causing the machine 700 to perform any one or more of the methodologies discussed herein may be executed, in whole or in part. In alternative embodiments, the machine 700 operates as a standalone device or may be connected (e.g., networked) to other machines. In a networked deployment, the machine 700 may operate in the capacity of a server machine or a client machine in a server-client network environment, or as a peer machine in a distributed (e.g., peer-to-peer) network environment. The machine 700 may be a server computer, a client computer, a personal computer (PC), a tablet computer, a laptop computer, a netbook, a set-top box (STB), a personal digital assistant (PDA), a cellular telephone, a smartphone, a web appliance, a network router, a network switch, a network bridge, or any machine capable of executing the instruc-

tions 724, sequentially or otherwise, that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include a collection of machines that individually or jointly execute the instructions 724 to perform all or part of any one or more of the methodologies discussed herein.

[0059] The machine 700 includes a processor 702 (e.g., a central processing unit (CPU), a graphics processing unit (GPU), a digital signal processor (DSP), an application specific integrated circuit (ASIC), a radiofrequency integrated circuit (RFIC), or any suitable combination thereof), a main memory 704, and a static memory 706, which are configured to communicate with each other via a bus 708. The machine 700 may further include a graphics display 710 (e.g., a plasma display panel (PDP), a light emitting diode (LED) display, a liquid crystal display (LCD), a projector, or a cathode ray tube (CRT)). The machine 700 may also include an alphanumeric input device 712 (e.g., a keyboard), a cursor control device 714 (e.g., a mouse, a touchpad, a trackball, a joystick, a motion sensor, or other pointing instrument), a storage unit 716, a signal generation device 718 (e.g., a speaker), and a network interface device 720.

[0060] The storage unit 716 includes a machine-readable medium 722 on which is stored the instructions 724 embodying any one or more of the methodologies or functions described herein. The instructions 724 may also reside, completely or at least partially, within the main memory 704, within the processor 702 (e.g., within the processor's cache memory), or both, during execution thereof by the machine 700. Accordingly, the main memory 704 and the processor 702 may be considered as machine-readable media. The instructions 724 may be transmitted or received over a network 726 (e.g., network 108) via the network interface device 720.

[0061] As used herein, the term "memory" refers to a machine-readable medium able to store data temporarily or permanently and may be taken to include, but not be limited to, random-access memory (RAM), read-only memory (ROM), buffer memory, flash memory, and cache memory. While the machine-readable medium 722 is shown in an example embodiment to be a single medium, the term "machine-readable medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, or associated caches and servers) able to store instructions. The term "machine-readable medium" shall also be taken to include any medium, or combination of multiple media, that is capable of storing instructions for execution by a machine (e.g., machine 700), such that the instructions, when executed by one or more processors of the machine (e.g., processor 702), cause the machine to perform any one or more of the methodologies described herein. Accordingly, a "machine-readable medium" refers to a single storage apparatus or device, as well as "cloud-based" storage systems or storage networks that include multiple storage apparatus or devices. The term "machine-readable medium" shall accordingly be taken to include, but not be limited to, one or more data repositories in the form of a solid-state memory, an optical medium, a magnetic medium, or any suitable combination thereof.

[0062] Throughout this specification, plural instances may implement components, operations, or structures described as a single instance. Although individual operations of one or more methods are illustrated and described as separate operations, one or more of the individual operations may be performed concurrently, and nothing requires that the operations be performed in the order illustrated. Structures and functionality presented as separate components in example configurations may be implemented as a combined structure or component. Similarly, structures and functionality presented as a single component may be implemented as separate components. These and other variations, modifications, additions, and improvements fall within the scope of the subject matter herein.

[0063] Certain embodiments are described herein as including logic or a number of components, modules, or mechanisms. Modules may constitute either software modules (e.g., code embodied on a machine-readable medium or in a transmission signal) or hardware modules. A "hardware module" is a tangible unit capable of performing certain operations and may be configured or arranged in a certain physical manner. In various example embodiments, one or more computer systems (e.g., a standalone computer system, a client computer system, or a server computer system) or one or more hardware modules of a computer system (e.g., a processor or a group of processors) may be configured by software (e.g., an application or application portion) as a hardware module that operates to perform certain operations as described herein.

[0064] In some embodiments, a hardware module may be implemented mechanically, electronically, or any suitable combination thereof. For example, a hardware module may include dedicated circuitry or logic that is permanently configured to perform certain operations. For example, a hardware module may be a special-purpose processor, such as a field programmable gate array (FPGA) or an ASIC. A hardware module may also include programmable logic or circuitry that is temporarily configured by software to perform certain operations. For example, a hardware module may include software encompassed within a general-purpose processor or other programmable processor. It will be appreciated that the decision to implement a hardware module mechanically, in dedicated and permanently configured circuitry, or in temporarily configured circuitry (e.g., configured by software) may be driven by cost and time considerations.

[0065] Accordingly, the phrase "hardware module" should be understood to encompass a tangible entity, be that an entity that is physically constructed, permanently configured (e.g., hardwired), or temporarily configured (e.g., programmed) to operate in a certain manner or to

perform certain operations described herein. As used herein, "hardware-implemented module" refers to a hardware module. Considering embodiments in which hardware modules are temporarily configured (e.g., programmed), each of the hardware modules need not be configured or instantiated at any one instance in time. For example, where a hardware module comprises a general-purpose processor configured by software to become a special-purpose processor, the general-purpose processor may be configured as respectively different special-purpose processors (e.g., comprising different hardware modules) at different times. Software may accordingly configure a processor, for example, to constitute a particular hardware module at one instance of time and to constitute a different hardware module at a different instance of time.

[0066] Hardware modules can provide information to, and receive information from, other hardware modules. Accordingly, the described hardware modules may be regarded as being communicatively coupled. Where multiple hardware modules exist contemporaneously, communications may be achieved through signal transmission (e.g., over appropriate circuits and buses) between or among two or more of the hardware modules. In embodiments in which multiple hardware modules are configured or instantiated at different times, communications between such hardware modules may be achieved, for example, through the storage and retrieval of information in memory structures to which the multiple hardware modules have access. For example, one hardware module may perform an operation and store the output of that operation in a memory device to which it is communicatively coupled. A further hardware module may then, at a later time, access the memory device to retrieve and process the stored output. Hardware modules may also initiate communications with input or output devices, and can operate on a resource (e.g., a collection of information).

[0067] The various operations of example methods described herein may be performed, at least partially, by one or more processors that are temporarily configured (e.g., by software) or permanently configured to perform the relevant operations. Whether temporarily or permanently configured, such processors may constitute processor-implemented modules that operate to perform one or more operations or functions described herein. As used herein, "processor-implemented module" refers to a hardware module implemented using one or more processors.

[0068] Similarly, the methods described herein may be at least partially processor-implemented, a processor being an example of hardware. For example, at least some of the operations of a method may be performed by one or more processors or processor-implemented modules. Moreover, the one or more processors may also operate to support performance of the relevant operations in a "cloud computing" environment or as a "software as a service" (SaaS). For example, at least some of the operations may be performed by a group of computers (as examples of machines including processors), with these operations being accessible via a network (e.g., the Internet) and via one or more appropriate interfaces (e.g., an application program interface (API)).

[0069] The performance of certain of the operations may be distributed among the one or more processors, not only residing within a single machine, but deployed across a number of machines. In some example embodiments, the one or more processors or processor-implemented modules may be located in a single geographic location (e.g., within a home environment, an office environment, or a server farm). In other example embodiments, the one or more processors or processor-implemented modules may be distributed across a number of geographic locations.

[0070] Some portions of the subject matter discussed herein may be presented in terms of algorithms or symbolic representations of operations on data stored as bits or binary digital signals within a machine memory (e.g., a computer memory). Such algorithms or symbolic representations are examples of techniques used by those of ordinary skill in the data processing arts to convey the substance of their work to others skilled in the art. As used herein, an "algorithm" is a self-consistent sequence of operations or similar processing leading to a desired result. In this context, algorithms and operations involve physical manipulation of physical quantities. Typically, but not necessarily, such quantities may take the form of electrical, magnetic, or optical signals capable of being stored, accessed, transferred, combined, compared, or otherwise manipulated by a machine. It is convenient at times, principally for reasons of common usage, to refer to such signals using words such as "data," "content," "bits," "values," "elements," "symbols," "characters," "terms," "numbers," "numerals," or the like. These words, however, are merely convenient labels and are to be associated with appropriate physical quantities.

[0071] Unless specifically stated otherwise, discussions herein using words such as "processing," "computing," "calculating," "determining," "presenting," "displaying," or the like may refer to actions or processes of a machine (e.g., a computer) that manipulates or transforms data represented as physical (e.g., electronic, magnetic, or optical) quantities within one or more memories (e.g., volatile memory, non-volatile memory, or any suitable combination thereof), registers, or other machine components that receive, store, transmit, or display information. Furthermore, unless specifically stated otherwise, the terms "a" or "an" are herein used, as is common in patent documents, to include one or more than one instance. Finally, as used herein, the conjunction "or" refers to a non-exclusive "or," unless specifically stated otherwise.

CLAUSES:

[0072] It should also be noted that the present disclo-

sure can also take configurations in accordance with the following numbered clauses:

Clause 1. A device comprising:

a processor configured by a contextual local image recognition module to retrieve a primary content dataset from a server, to generate and to update a contextual content dataset based on an image captured with the device; and
a storage device coupled to the processor, the storage device comprising the primary content dataset and the contextual content dataset, the primary content dataset comprising a first set of images and corresponding virtual object models, the contextual content dataset comprising a second set of images and corresponding virtual object models retrieved from the server.

Clause 2. The device of Clause 1, wherein the contextual local image recognition module is configured to locally render the virtual object model corresponding to the image captured with the device when the image captured with the device corresponds to one of the first set of images in the storage device.

Clause 3. The device of Clause 1, wherein the contextual local image recognition module is configured to request the server for the virtual object model corresponding to the image captured with the device when the image captured with the device does not correspond to one of the first set of images in the storage device.

Clause 4. The device of Clause 3, wherein the contextual local image recognition module is configured to receive the virtual object model corresponding to the image captured with the device from the server, and to update the contextual content dataset when the image captured with the device does not correspond to one of the first set of images in the storage device.

Clause 5. The device of Clause 1, wherein the contextual local image recognition module is configured to request the server for the virtual object model corresponding to the image captured with the device when the image captured with the device does not correspond to one of the second set of images in the storage device.

Clause 6. The device of Clause 5, wherein the contextual local image recognition module is configured to receive the virtual object model corresponding to the image captured with the device from the server, and to update the contextual content dataset when the image captured with the device does not correspond to one of the second set of images in the storage device.

Clause 7. The device of Clause 1, wherein the contextual local image recognition module is configured to determine usage conditions of the device, and to generate a request to the server for a third set of images and corresponding virtual object models based on the usage conditions.

Clause 8. The device of Clause 7, wherein the contextual local image recognition module is configured to update the contextual content dataset with the third set of images and corresponding virtual object models.

Clause 9. The device of Clause 7, wherein the usage conditions of the device comprises social information of a user of the device and location and time information of the device using a three-dimensional object generation application based on the contextual local image recognition module.

Clause 10. The device of Clause 1, wherein the processor further comprises a three-dimensional object generation module configured to generate, in a display of the device, a visualization of a virtual object engaged with the image of a physical object captured with the device, the virtual object corresponding to the image, with a rendering of the visualization of the virtual object based a position of the device relative to the physical object.

Clause 11. A computer-implemented method comprising:

retrieving a primary content dataset from a server;
generating and updating, at a device, a contextual content dataset based on an image captured with the device; and
storing, at the device, the primary content dataset and the contextual content dataset, the primary content dataset comprising a first set of images and corresponding virtual object models, the contextual content dataset comprising a second set of images and corresponding virtual object models retrieved from the server.

Clause 12. The computer-implemented method of Clause 11, further comprising:
locally rendering the virtual object model corresponding to the image captured with the device when the image captured with the device corresponds to one of the first set of images stored in the device.

Clause 13. The computer-implemented method of Clause 11, further comprising:
requesting the server for the virtual object model cor-

responding to the image captured with the device when the image captured with the device does not correspond to one of the first set of images stored in device.

Clause 14. The computer-implemented method of Clause 13, further comprising:

receiving the virtual object model corresponding to the image captured with the device from the server; and
updating the contextual content dataset when the image captured with the device does not correspond to one of the first set of images in the device.

Clause 15. The computer-implemented method of Clause 11, further comprising:
requesting the server for the virtual object model corresponding to the image captured with the device when the image captured with the device does not correspond to one of the second set of images in the device.

Clause 16. The computer-implemented method of Clause 15, further comprising:

receiving the virtual object model corresponding to the image captured with the device from the server; and
updating the contextual content dataset when the image captured with the device does not correspond to one of the second set of images in the device.

Clause 17. The computer-implemented method of Clause 11, further comprising:

determining usage conditions of the device;
generating a request to the server for a third set of images and corresponding virtual object models; and
updating the contextual content dataset with the third set of images and corresponding virtual object models.

Clause 18. The computer-implemented method of Clause 17, wherein the usage conditions of the device comprises social information of a user of the device; and location and time information of the device using a three-dimensional object generation application based on the contextual local image recognition module.

Clause 19. The computer-implemented method of Clause 11, further comprising:
generating, in a display of the device, a visualization of a virtual object engaged with the image of a physical object captured with the device, the virtual object corresponding to the image, with a rendering of the visualization of the virtual object based on a position of the device relative to the physical object.

Clause 20. A non-transitory machine-readable medium comprising instructions that, when executed by one or more processors of a machine, cause the machine to perform operations comprising:

retrieving a primary content dataset from a server;
generating and updating, at a device, a contextual content dataset based on an image captured with the device; and
storing, at the device, the primary content dataset and the contextual content dataset, the primary content dataset comprising a first set of images and corresponding virtual object models, the contextual content dataset comprising a second set of images and corresponding virtual object models retrieved from the server.

Claims

1. A method comprising:

generating, at a server, a primary content dataset comprising a first set of images and corresponding virtual content, the first set of images including a core set of images;
accessing a database to retrieve an image captured by a device and received by the server;
comparing the image captured by the device with the first set of images;
based on the comparison, retrieving virtual content associated with the image in response to determining that the first set of images does not include the image captured by the device; and
communicating, to the device, the virtual content associated with the image captured by the device, the device being configured to update a contextual content dataset with the virtual content associated with the image captured by the device; and
storing, at the server, the primary content dataset, a secondary content dataset corresponding to the contextual content dataset of the plurality of devices, and a third content dataset comprising virtual content not included in the primary and secondary content datasets.

2. The method of claim 1, wherein the core set of images includes the most common images captured from the plurality of devices.

3. The method of claim 1 or claim 2, further comprising:

providing the primary content dataset to a plurality of devices;
generating a corresponding secondary content dataset for each device based on images captured from the corresponding device; and
generating a corresponding third content dataset for each device based on usage conditions of the corresponding device.

4. The method of claim 1, further comprising:

determining usage conditions of the device;
identifying virtual content based on the usage conditions and not included in the primary and secondary content datasets; and
communicating the virtual content based on the usage conditions and not included in the primary and secondary content datasets to the device.

5. The method of claim 4, further comprising:
updating the third content dataset of the device with the virtual content based on the usage conditions and not included in the primary and secondary content datasets.

6. The method of claim 4, wherein the usage conditions of the device include at least one of a time, duration, location, and orientation of the device at the time of capture of the image.

7. The method of claim 4, wherein the usage conditions of the device include images from other users in a social network of the user of the device.

8. The method of claim 1, wherein the device includes an augmented reality application configured to display the virtual content in a display of the device.

9. The method of claim 8, further comprising:

uploading the primary content dataset to the augmented reality application of the device in response to detecting an initialization of the augmented reality application in the device,
wherein the usage conditions of the device include at least one of a time and location of the device while the augmented reality application operates in the device.

10. A server comprising:

a network communication interface to communicate with a plurality of devices;
one or more hardware processors configured to:

generate a primary content dataset comprising a first set of images and corresponding virtual content, the first set of images including a core set of images,
access a database to retrieve an image captured by a device and received by the server via the network communication interface,
compare the image captured by the device with the first set of images,
based on the comparison, retrieve virtual content associated with the image in response to determining that the first set of images does not include the image captured by the device, and
communicate, via the network communication interface, to the device the virtual content associated with the image captured by the device, the device being configured to update a contextual content dataset with the virtual content associated with the image captured by the device; and

a storage device coupled to the one or more hardware processors, the storage device comprising the primary content dataset, a secondary content dataset corresponding to the contextual content dataset of the plurality of devices, and a third content dataset comprising virtual content not included in the primary and secondary content datasets.

11. A system comprising:

a plurality of devices;
a server having: a network communication interface to communicate with the plurality of devices; a storage device; and one or more hardware processors configured to perform the method of any one of claims 1 to 9;
wherein at least one of the devices comprises: a processor configured by a contextual local image recognition module to retrieve the primary content dataset from the server, to generate and to update a contextual content dataset based on an image captured with the device; and a local storage device coupled to the processor, the local storage device comprising the primary content dataset and the contextual content dataset, the contextual content dataset comprising a second set of images and corresponding virtual object models retrieved from the server, and
wherein the storage device is coupled to the one or more hardware processors, the storage device comprising the primary content dataset, a secondary content dataset corresponding to the contextual content dataset of the plurality of devices, and a third content dataset comprising virtual content not included in the primary and secondary content datasets.

12. The system of claim 11, wherein the device includes

an augmented reality application configured to display the virtual content in a display of the device.

13. The system of claim 12, wherein the one or more hardware processors are further configured to:
upload the primary content dataset to the augmented reality application of the device in response to detecting an initialization of the augmented reality application in the device.

14. A machine-readable medium comprising instructions that, when executed by one or more processors of a server, cause the server to perform the method of any one of claims 1 to 9.

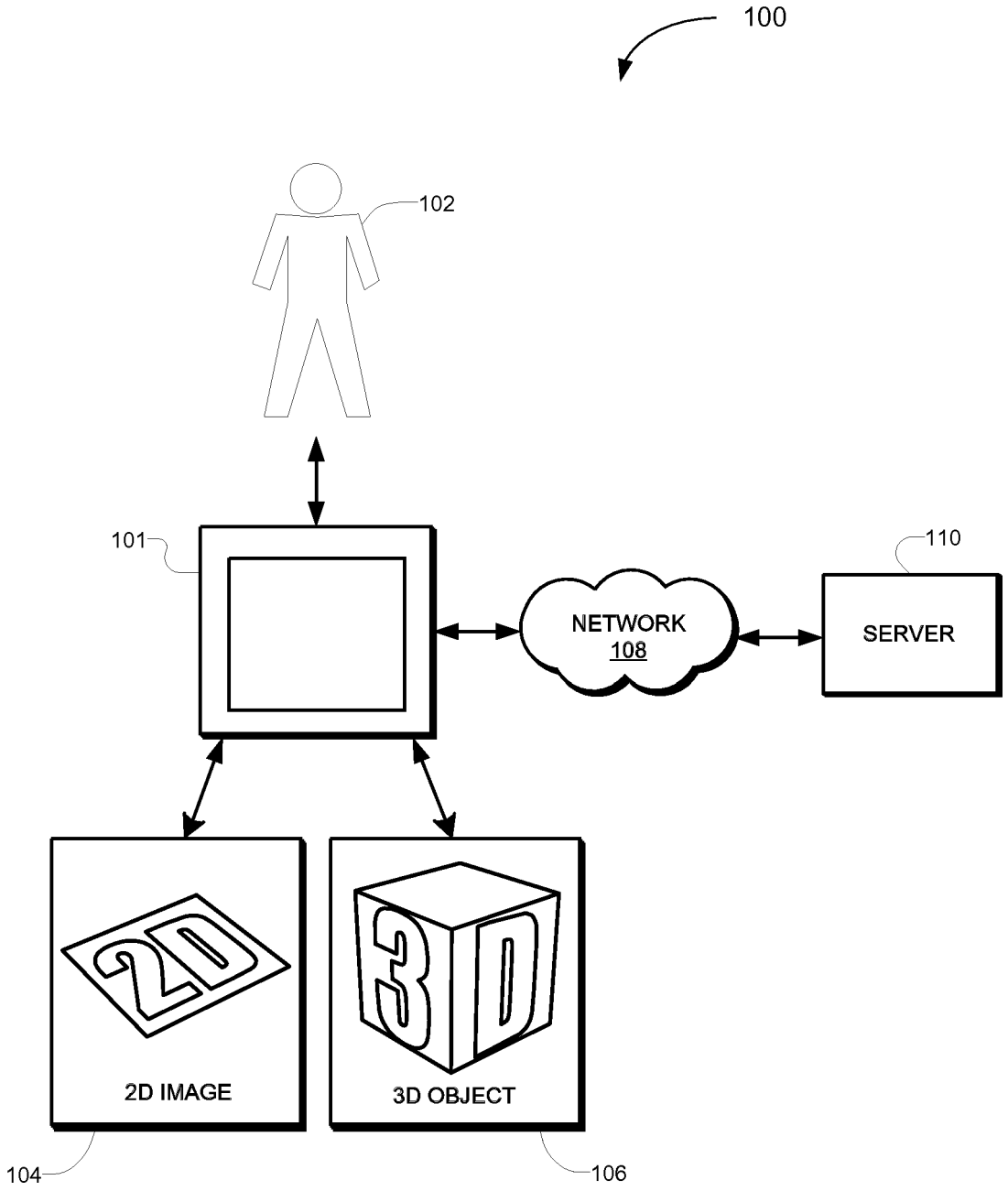
100
102
101
110
NETWORK
108
SERVER
2D
3D
2D IMAGE
3D OBJECT
104
106

FIG. 1

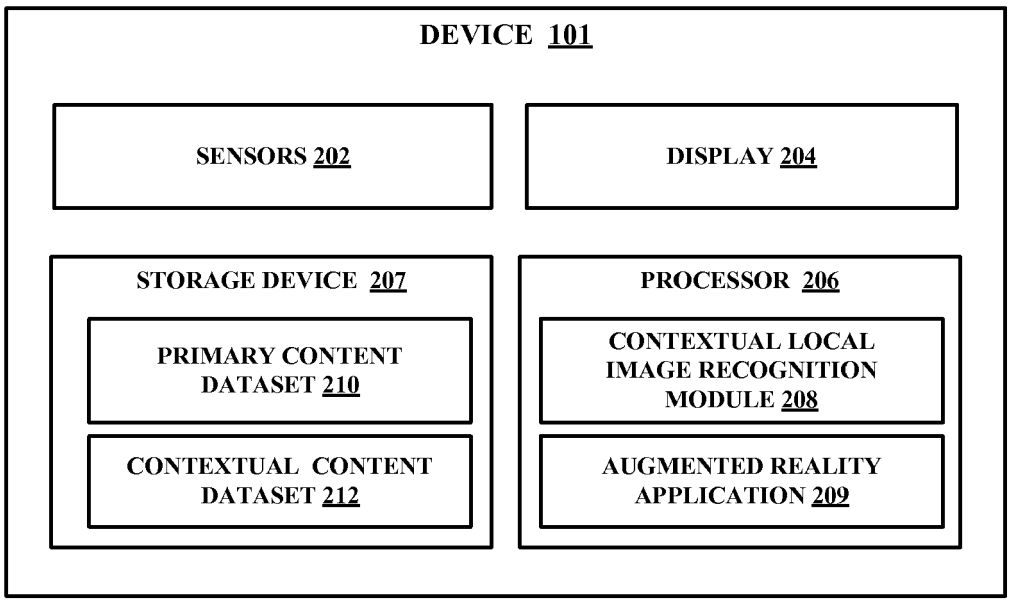
DEVICE 101
SENSORS 202
DISPLAY 204
STORAGE DEVICE 207
PRIMARY CONTENT DATASET 210
CONTEXTUAL CONTENT DATASET 212
PROCESSOR 206
CONTEXTUAL LOCAL IMAGE RECOGNITION MODULE 208
AUGMENTED REALITY APPLICATION 209

FIG. 2

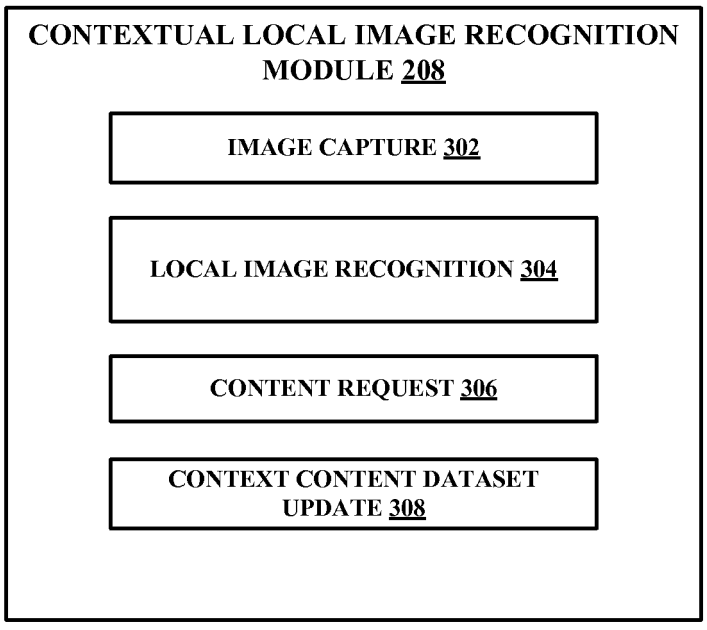
CONTEXTUAL LOCAL IMAGE RECOGNITION MODULE 208
IMAGE CAPTURE 302
LOCAL IMAGE RECOGNITION 304
CONTENT REQUEST 306
CONTEXT CONTENT DATASET UPDATE 308

FIG. 3

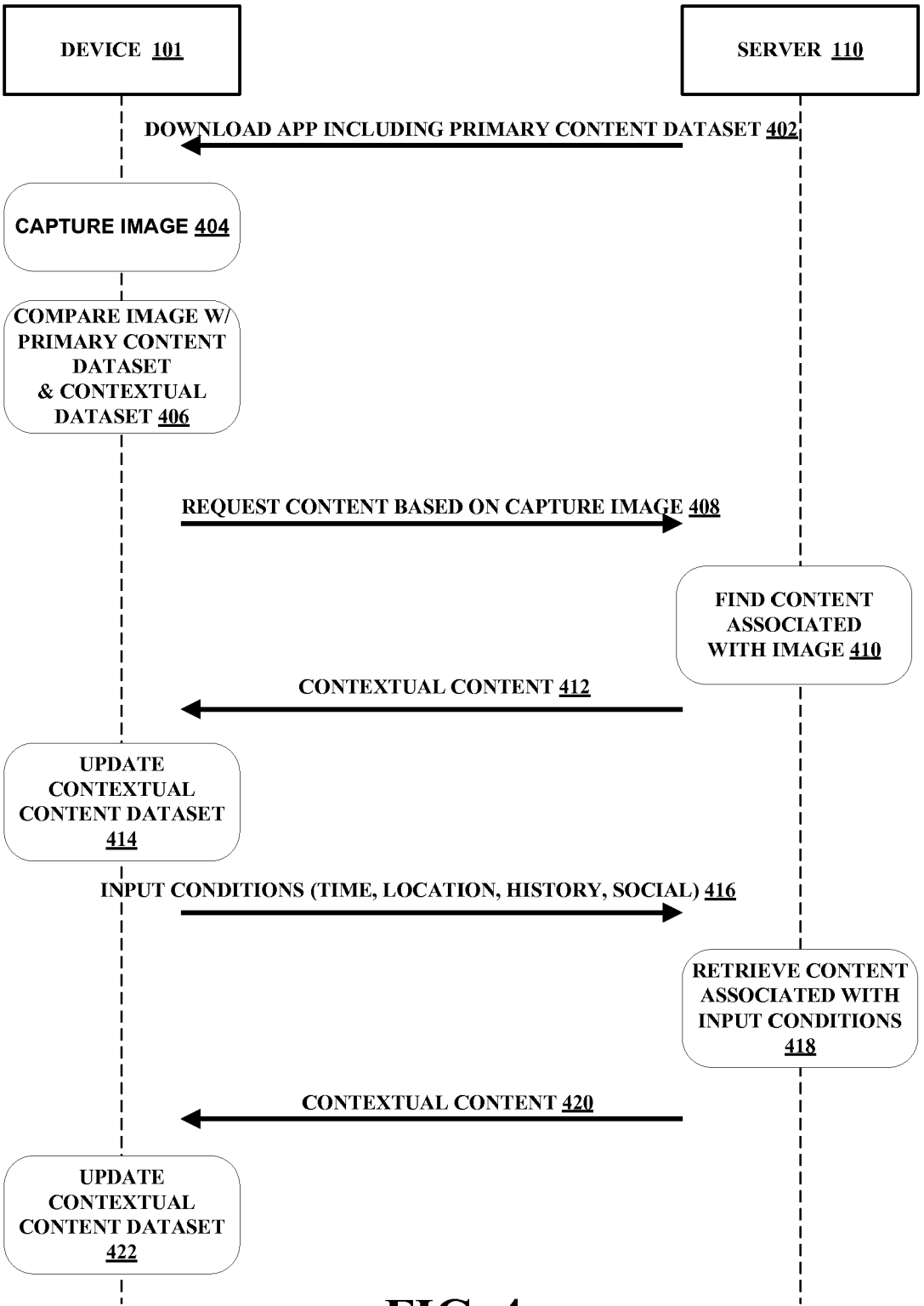
DEVICE 101
SERVER 110
DOWNLOAD APP INCLUDING PRIMARY CONTENT DATASET 402
CAPTURE IMAGE 404
COMPARE IMAGE W/ PRIMARY CONTENT DATASET & CONTEXTUAL DATASET 406
REQUEST CONTENT BASED ON CAPTURE IMAGE 408
FIND CONTENT ASSOCIATED WITH IMAGE 410
CONTEXTUAL CONTENT 412
UPDATE CONTEXTUAL CONTENT DATASET 414
INPUT CONDITIONS (TIME, LOCATION, HISTORY, SOCIAL) 416
RETRIEVE CONTENT ASSOCIATED WITH INPUT CONDITIONS 418
CONTEXTUAL CONTENT 420
UPDATE CONTEXTUAL CONTENT DATASET 422

FIG. 4

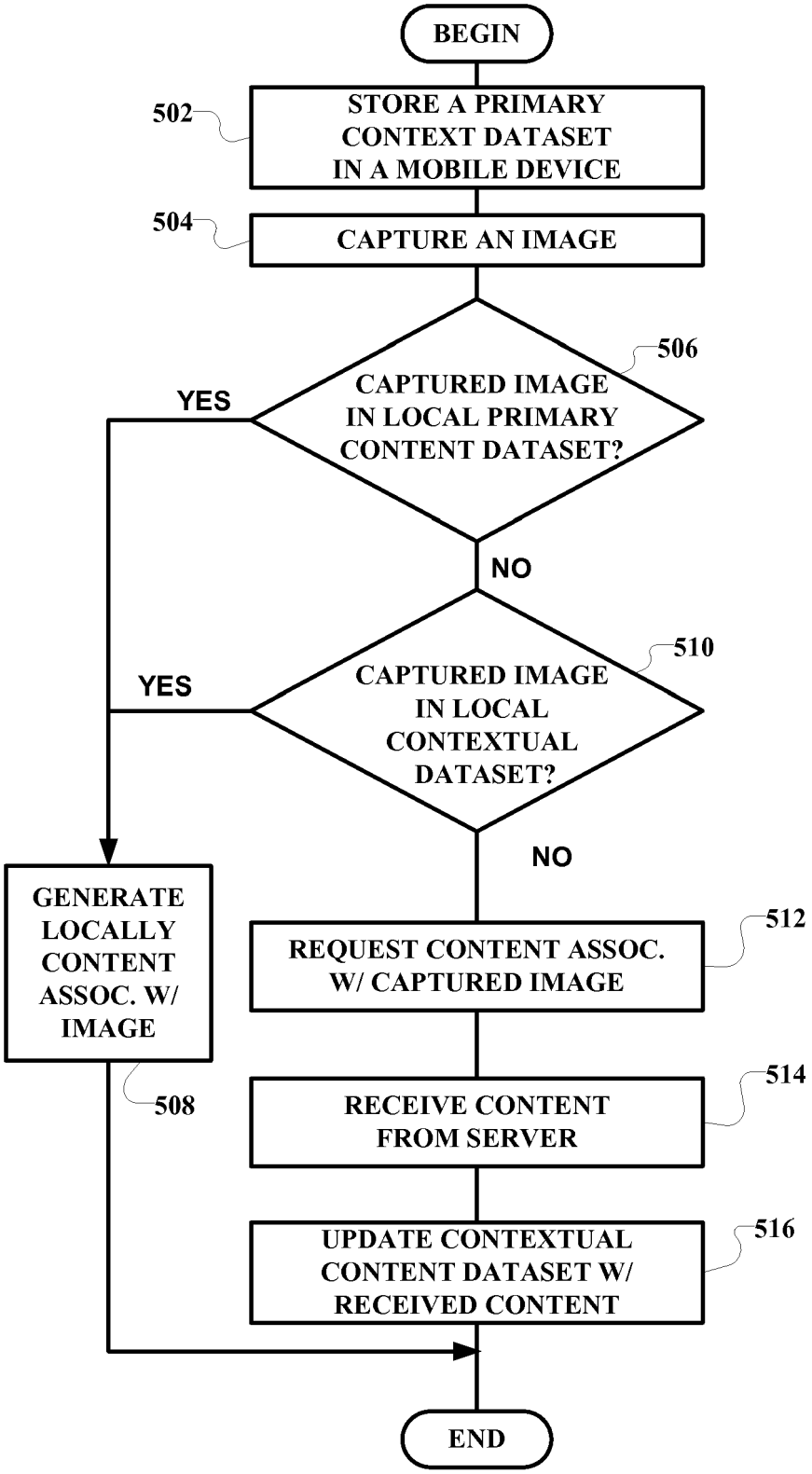
BEGIN
502
STORE A PRIMARY CONTEXT DATASET IN A MOBILE DEVICE
504
CAPTURE AN IMAGE
506
CAPTURED IMAGE IN LOCAL PRIMARY CONTENT DATASET?
YES
NO
510
CAPTURED IMAGE IN LOCAL CONTEXTUAL DATASET?
YES
NO
GENERATE LOCALLY CONTENT ASSOC. W/ IMAGE
508
REQUEST CONTENT ASSOC. W/ CAPTURED IMAGE
512
RECEIVE CONTENT FROM SERVER
514
UPDATE CONTEXTUAL CONTENT DATASET W/ RECEIVED CONTENT
516
END

FIG. 5

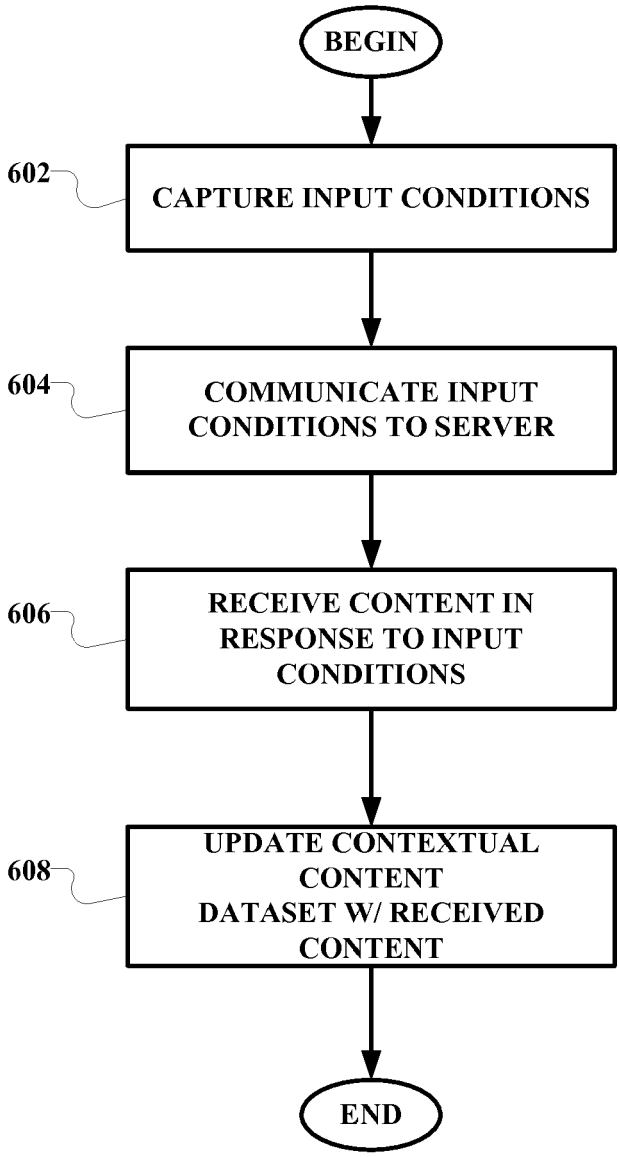
BEGIN
602
CAPTURE INPUT CONDITIONS
604
COMMUNICATE INPUT CONDITIONS TO SERVER
606
RECEIVE CONTENT IN RESPONSE TO INPUT CONDITIONS
608
UPDATE CONTEXTUAL CONTENT DATASET W/ RECEIVED CONTENT
END

FIG. 6

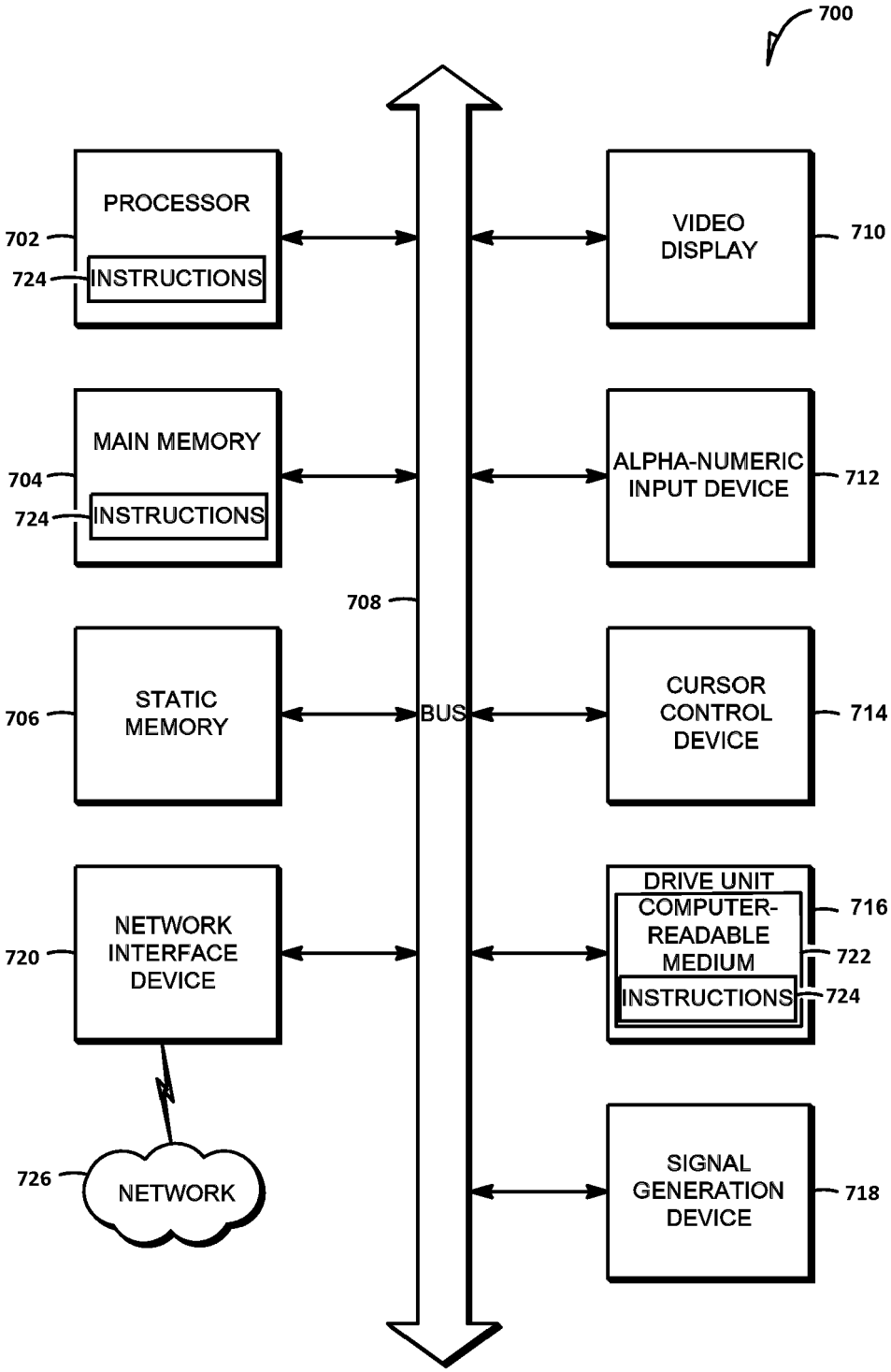
700
PROCESSOR
702
724
INSTRUCTIONS
MAIN MEMORY
704
724
INSTRUCTIONS
708
STATIC MEMORY
706
BUS
NETWORK INTERFACE DEVICE
720
726
NETWORK
VIDEO DISPLAY
710
ALPHA-NUMERIC INPUT DEVICE
712
CURSOR CONTROL DEVICE
714
DRIVE UNIT
716
COMPUTER-READABLE MEDIUM
722
INSTRUCTIONS
724
SIGNAL GENERATION DEVICE
718

FIG. 7

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 21 15 2338

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
Y	US 2010/302143 A1 (SPIVACK NOVA T [US]) 2 December 2010 (2010-12-02) * paragraph [0079] * * paragraph [0081] * * paragraph [0093] - paragraph [0100] * * paragraph [0101] - paragraph [0102] * * paragraph [0178] * * paragraph [0037] - paragraph [0039] * * paragraph [0050] * * paragraph [0195] - paragraph [0197] * * paragraphs [0045], [0065], [0078], [0117], [0134], [0123], [0162] * * paragraphs [0070], [0080], [0188] *	1-14	INV. H04N21/43 H04N21/433 G06F3/147 G06F1/16 G06F16/50 G06F16/51 G06F16/583 G06F16/955 G06T19/00 H04N1/00
Y A	US 2012/243732 A1 (SWAMINATHAN ASHWIN [US] ET AL) 27 September 2012 (2012-09-27) * paragraphs [0031], [0041], [0044], [0067] *	1,4-14 2,3	
Y A	US 8 341 245 B1 (ROSKIND JAMES [US] ET AL) 25 December 2012 (2012-12-25) * column 3, line 42 - line 53 *	2,3 1,4-14	TECHNICAL FIELDS SEARCHED (IPC) G06F G06T G09G
A	US 2013/044128 A1 (LIU JAMES C [US] ET AL) 21 February 2013 (2013-02-21) * the whole document *	1-14	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	12 May 2021	Ong, Bee

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 21 15 2338

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

12-05-2021

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010302143 A1	02-12-2010	NONE	
US 2012243732 A1	27-09-2012	CN 103119627 A	22-05-2013
		EP 2619728 A1	31-07-2013
		ES 2745739 T3	03-03-2020
		HU E047021 T2	28-04-2020
		JP 5989832 B2	07-09-2016
		JP 6000954 B2	05-10-2016
		JP 6290331 B2	07-03-2018
		JP 2013541096 A	07-11-2013
		JP 2015144474 A	06-08-2015
		JP 2017011718 A	12-01-2017
		KR 20130060339 A	07-06-2013
		US 2012243732 A1	27-09-2012
		US 2016284099 A1	29-09-2016
		WO 2012040099 A1	29-03-2012
US 8341245 B1	25-12-2012	US 8341245 B1	25-12-2012
		US 8762490 B1	24-06-2014
		US 9602620 B1	21-03-2017
US 2013044128 A1	21-02-2013	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 83981813 **[0001]**