



(51) International Patent Classification:
H04M 1/00 (2006.01)

(21) International Application Number:
PCT/CN2014/087313

(22) International Filing Date:
24 September 2014 (24.09.2014)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **INTEL CORPORATION** [US/US]; 2200 Mission College Blvd., Santa Clara, California 95054 (US).

(72) Inventors: **LI, Wenlong**; 4-502 Building 18, Tiantongyuan West 2 District, Changping District, Beijing, 11 102218 (CN). **TONG, Xiaofeng**; Room 1103, Building 1, Shuangyushu Xili, Haidian District, Beijing, 11 100086 (CN). **DU, Yangzhou**; Room 502, Building 2-21-7, XinGangZhuangYuan, LiJiaQiao, ShunYi District, Beijing, 11 101304 (CN). **SACHSON, Thomas**; 3876 Alameda de las Pulgas, Menlo Park, California 94025 (US).

(74) Agent: **CHINA PATENT AGENT (H.K.) LTD.**; 22/F., Great Eagle Center, 23 Harbour Road, Wanchai, Hong Kong (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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))

(54) Title: AVATAR AUDIO COMMUNICATION SYSTEMS AND TECHNIQUES

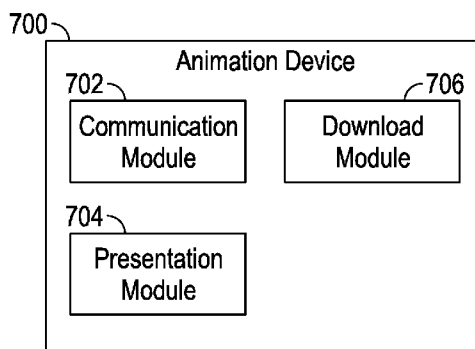


FIG. 7

(57) Abstract: Examples of systems and methods for transmitting avatar sequencing data in an audio file are generally described herein. A method can include receiving, at a second device from a first device, an audio file comprising: facial motion data, the facial motion data derived from a series of facial images captured at the first device, an avatar sequencing data structure from the first device, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream. The method can include presenting an animation of an avatar, at the second device, using the facial motion data and the audio stream.

AVATAR AUDIO COMMUNICATION SYSTEMS AND TECHNIQUES

5

BACKGROUND

[0001] Messaging services including instant messaging services and email, among others, provide users with many different types of emoticons, or emotion icons, for expressing themselves more demonstratively. Emoticons can include
10 animations where a series of images are used together to create a video or animation. These emoticons are selectable by users, and even often customizable by users. However, these approaches limit the creativity of users and limit the customizability of the animations to already created emoticons. Animations constrained by predefined emoticons are therefore not meeting user demands.

15

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] In the drawings, which are not necessarily drawn to scale, like numerals may describe similar components in different views. Like numerals having different letter suffixes may represent different instances of similar
20 components. The drawings illustrate generally, by way of example, but not by way of limitation, various embodiments discussed in the present document.

[0003] FIGs. 1A-1F are schematic diagrams illustrating a system for displaying an avatar animation, according to an embodiment.

[0004] FIG. 2 is a flow chart illustrating a system for playing an animation
25 with a selected avatar, according to an embodiment.

[0005] FIG. 3 is a flow chart illustrating a system for playing an animation using an audio file, according to an embodiment.

[0006] FIG. 4 is a flow chart illustrating a system for adjusting a frame rate of an image capture device, according to an embodiment.

30 [0007] FIG. 5 is a flow chart illustrating a system for sending an avatar in an audio file, according to an embodiment.

[0008] FIG. 6 is a block diagram of a device upon which one or more embodiments may be implemented.

[0009] FIG. 7 is a block diagram illustrating an animation device for displaying an avatar animation, according to an embodiment.

DETAILED DESCRIPTION

5 [0010] As mentioned above, existing approaches generally constrain users to predefined emoticons and animations. Various systems and techniques are proposed here to present users with an option for creating a facial gesture driven animation.

[0011] FIGs. 1A-1F are schematic diagrams illustrating a system for
10 displaying an avatar animation, according to an embodiment. In an example, FIG. 1A is a schematic diagram illustrating a system for displaying an avatar animation, including a communication module receiving, at a second device from a first device, an audio file. The audio file may include facial motion data, an avatar sequencing data structure, or an audio stream. Facial motion data may
15 be a set of facial coordinates, information about movements of a face, or the like. For example, facial motion data may be derived from a series of facial images captured at a first device. In an example, an avatar sequencing data structure may include an avatar identifier and a duration. For example, the avatar sequencing data structure may include an avatar identifier for an avatar selected
20 by a user and a duration computed using an audio recording. The avatar sequencing data structure may include multiple avatar identifiers and multiple durations or an audio file may include multiple avatar sequencing data structures, each with a single avatar identifier and a single duration. An avatar identifier in an avatar sequencing data structure may correspond to a duration in the avatar
25 sequencing data structure, or the avatar identifier and the duration may be unrelated. In an example, an audio stream can include music, speech, or any other type of audio recording.

[0012] FIG. 1B is a schematic diagram illustrating a system for displaying an avatar animation, including a presentation module that may present an animation
30 of an avatar. The presentation module may use facial motion data and an audio stream to present the animation. In an example, when an audio file is received at a device, the device may determine if the audio file includes an avatar sequencing data structure, and if so, the device may identify an avatar identifier

from the avatar sequencing data structure. The device may determine if the device has a copy of an avatar corresponding with the avatar identifier stored locally. If a copy of the avatar is stored locally, the device can display the animation using the avatar from local memory. If a copy of the avatar is not

5 stored locally, the device may present an option to a user, such as the option presented in FIG. 1B. In an example, presenting an option may include presenting an option for a user to select a download option or a play option. In an example, if the user selects the download option, the device may download the avatar from a server or directly from another device. After downloading the

10 avatar, the device may display an animation using the avatar. In an example, if the user selects the play option, the device may display an animation using an avatar from local memory of the device. In an example, the download option may include charging a fee to download an avatar, such as requiring payment, use of pre-purchased coins, or the like.

15 **[0013]** FIG. 1C is a schematic diagram illustrating a system for displaying an avatar animation, including displaying an animation using an avatar corresponding to an avatar identifier. In an example, displaying the animation may include using an avatar identifier received in an audio file, such as an avatar identifier in an avatar sequencing data structure in the audio file. For example,

20 the avatar identifier may identify an avatar that the display device has stored locally. The display device may use the locally-stored avatar, identified by the avatar identifier, to display the animation. In another example, if the display device does not have the avatar identified by the avatar identifier stored locally, the display device may download the avatar from a server or other device and

25 display the animation using the downloaded avatar.

[0014] FIG. 1D is a schematic diagram illustrating a system for displaying an avatar animation, including downloading an avatar. In an example, the avatar to download may be selected locally at the same device that downloads the avatar. The avatar may require payment, such as a specified amount of the 32 gold coins

30 shown in FIG. 1D, which represent a current balance of an account accessed by the device. In an example, the avatar may be offered or sold from an online store and downloaded from a server. In another example, the avatar may be

downloaded directly from another device and may require payment from either device or may be free.

[0015] FIG. 1E is a schematic diagram illustrating a system for displaying an avatar animation, including a communication module to send to an animation device, an audio file. The audio file may include facial motion data, an avatar sequencing data structure, or an audio stream, which are described above. The communication module may be in communication with an image capture device, which may capture a series of images of a face and send the images to a facial recognition module. The facial recognition module may compute facial motion data for each of the images in the series of images. After the audio file is sent to the animation device, the animation device may use the facial motion data and the audio stream to animate an avatar. The image capture device may include a camera. In an example, the facial recognition module may detect if a face is present in the series of images captured by the image capture device. When the facial recognition module detects that a face is present, an indication may be sent to the image capture device to operate at a normal frame rate. A frame rate may include a camera speed, a number of images to be stored, a number of images captured, a resolution, or the like. When the facial recognition module detects that a face is absent (i.e. not present), an indication may be sent to the image capture device to operate at a reduced frame rate. A reduced frame rate may be any frame rate less than a normal frame rate, such as a slower camera speed, a smaller number of images to be stored, a smaller number of images captured, a lower resolution, or the like. In an example, the image capture device may receive the indication from the facial recognition module and the image capture device may increase or decrease a frame rate using the indication.

[0016] FIG. 1F is a schematic diagram illustrating a system for displaying an avatar animation, including displaying an animation using a downloaded avatar.

[0017] FIG. 2 is a flow chart illustrating a method 200 for playing an animation with a selected avatar, according to an embodiment. At block 202, the method 200 includes a receiving an animation of an avatar. At block 204, the method 200 includes extracting an avatar identifier, such as from an avatar sequencing data structure, and determining if the avatar exists locally. When it is determined that the avatar exists locally, the local avatar may be selected without

downloading a new avatar (block 206). When it is determined that the avatar does not exist locally, the avatar may be downloaded (block 208). In an example, the downloaded avatar may be selected. In another example, a different avatar that is stored locally may be selected. At block 210, after an avatar is selected, the method 200 may include playing the animation with the selected avatar.

[0018] FIG. 3 is a flow chart illustrating a method 300 for playing an animation using an audio file, according to an embodiment. At block 302, the method 300 includes receiving an animation of an avatar. At block 304, the method 300 extracting an avatar identifier, such as from an avatar sequencing data structure, and determining if the avatar exists locally. When it is determined that the avatar exists locally, the method 300 includes selecting the local avatar (block 306). When it is determined that the avatar does not exist locally, the method 300 may include determining the avatar from an avatar identifier, such as from an avatar sequencing data structure, in an audio file (block 308). At block 310, after an avatar is selected, the method 300 includes playing the animation with the selected avatar.

[0019] FIG. 4 is a flow chart illustrating a method 400 for adjusting a frame rate of an image capture device, according to an embodiment. The method 400 includes using an image capture device or image capture module to capture an image (block 402). In an example, the image capture device may be a camera. The method 400 may also include computing facial motion data for each of the images in the series of images. At block 404, the method 400 includes detecting a face, such as by determining if certain features are present. The method 400 may also track a face, such as by capturing a series of images and determining if certain features are in different places in consecutive images. In an example, if a face is present (e.g., detected), the method 400 may include operating at a normal frame rate (block 406). In an example, capturing a series of images at a normal frame rate may include capturing a series of images at thirty frames per second. In an example, the method may include detecting a face using an image quickly, such as in five to ten milliseconds or the like, if the image resolution is a resolution of a mobile phone, such as substantially 192 by 144 pixels. The method 400 may use significant system resources to detect the face and may not use significant system resources to capture images with no face detected.

[0020] In another example, if a face is absent, or not present, the image capture device or image capture module may operate at a reduced frame rate (block 408). In another example, a frame rate for the image capture device may be dynamically changed, such as by changing a camera sampling rate. A face
5 may be absent from an image or a series of images, and if a threshold number of frames or duration is reached, the frame rate may be changed to five frames per second. The method 400 may include determining that the face is absent from an image or a series of images for a duration, such as thirty seconds, and the facial recognition module may send an indication to the image capture device to alter
10 the frame rate. In another example, if the image capture device is operating at a reduced frame rate, and a face is detected as present in an image or a series of images, the method 400 may include sending an indication to the image capture device to change the frame rate to a normal frame rate.

[0021] At block 410, an indication may be received indicating whether the
15 image capture device or image capture module has completed capturing images. In an example, a user may indicate that the image capture device has completed capturing images. For example, a user may hold down a button indicating that the image capture device should start capturing images and then the user may release the button indicating that the image capture device should stop capturing
20 images. In another example, the image capture device may capture a series of images until a specified number of images or a specified duration is reached (e.g., capture images until 100 images are captured, one minute has elapsed, a memory is filled, or the like). If the image capture device has completed capturing images, the method 400 may end. If the image capture device has not completed
25 capturing images, the method 400 may capture another image (block 402) and repeat.

[0022] FIG. 5 is a flow chart illustrating a method 500 for sending an avatar in an audio file, according to an embodiment. The method 500 includes an image capture device or image capture module to capture a series of images of a face
30 (block 502). In an example, the image capture device may be a camera. The method 500 also includes computing facial motion data, such as a set of facial coordinates, for each of the images in the series of images (block 504). At block 506, the method 500 includes compiling an audio file comprising the facial

motion data, such as a set of facial coordinates, and an audio stream. At block 508, the method 500 includes identifying an avatar and a duration for the avatar and adding an avatar identifier corresponding to the avatar to the audio file.

[0023] In an example, an audio file may have a proprietary format. An avatar sequencing data structure may be added to an audio file with a proprietary format, such as by adding the avatar sequencing data structure as metadata in the audio file. For example, the avatar sequencing data structure may be added as a Universal Resource Locator (URL) for a proprietary avatar communication data structure. The URL may include information for extracting the avatar sequencing data structure or other attributes about an avatar, such as a duration. The audio file may include the avatar sequencing data structure or the avatar in metadata. If the audio file includes a URL including information for extracting the avatar, the audio file may be smaller (e.g., take up less memory) than if the avatar is included directly in the metadata. In an example, the proprietary format audio file may include an audio stream in a format such as MPEG4 and metadata.

[0024] In an example, an audio file may have a commercial format such as an Apple Core Audio Format. For example, data may be stored in reserved free chunk space in a Core Audio Format file without affecting audio playback. A Core Audio Format audio file may include an avatar, an avatar sequencing data structure, or other attributes about an avatar, such as a duration. The Core Audio Format includes free chunk header fields where an avatar and information about an avatar may be stored, such as mChunkType and mChunkSize.

[0025] In an example, the method 500 may also include adding the duration to the audio file. In another example, the method 500 may include adding an avatar sequencing data structure to the audio file. The avatar sequencing data structure may include an avatar identifier and a duration. More than one avatar identifier and duration or more than one avatar sequencing data structure may be added to the audio file (block 510 repeating block 508).

[0026] In an example, a user may record a message, such as a series of images and an audio recording. The user may choose an avatar or avatar identifier to send to a remote device, and the avatar may be used to animate the series of images and played with the audio recording. In an example, more than one avatar may be chosen for a specified message. A duration for each avatar

chosen may also be chosen. The duration may be specified using timestamps, a length of time, a number of frames, or the like. The remote device may use one or more of the chosen avatars to animate the message or may use none of the avatars. In an example, the remote device may use an avatar selected at the
5 remote device and an avatar identified in the audio file. For example, the remote device may animate the message using Avatars selected at the remote device, avatars identified in the audio file, or any combination of either of these types of avatars.

[0027] After the audio file is compiled with a specified number of avatar
10 identifiers and durations (or avatar sequencing data structures), the method 500 may include sending the audio file (block 512).

[0028] FIG. 6 is a block diagram of a machine 600 upon which one or more embodiments may be implemented. In alternative embodiments, the machine 600 can operate as a standalone device or can be connected (e.g., networked) to other
15 machines. In a networked deployment, the machine 600 can operate in the capacity of a server machine, a client machine, or both in server-client network environments. In an example, the machine 600 can act as a peer machine in peer-to-peer (P2P) (or other distributed) network environment. The machine 600 can be a personal computer (PC), a tablet PC, a set-top box (STB), a personal digital
20 assistant (PDA), a mobile telephone, a web appliance, a network router, switch or bridge, or any machine capable of executing instructions (sequential 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 any collection of machines that individually or jointly execute a set (or multiple
25 sets) of instructions to perform any one or more of the methodologies discussed herein, such as cloud computing, software as a service (SaaS), other computer cluster configurations.

[0029] Examples, as described herein, can include, or can operate on, logic or a number of components, modules, or mechanisms. Modules are tangible entities
30 (e.g., hardware) capable of performing specified operations when operating. A module includes hardware. In an example, the hardware can be specifically configured to carry out a specific operation (e.g., hardwired). In an example, the hardware can include configurable execution units (e.g., transistors, circuits, etc.)

and a computer readable medium containing instructions, where the instructions configure the execution units to carry out a specific operation when in operation. The configuring can occur under the direction of the executions units or a loading mechanism. Accordingly, the execution units are communicatively coupled to the computer readable medium when the device is operating. In this example, the execution units can be a member of more than one module. For example, under operation, the execution units can be configured by a first set of instructions to implement a first module at one point in time and reconfigured by a second set of instructions to implement a second module.

10 **[0030]** Machine (e.g., a computer system) 600 can include a hardware processor 602 (e.g., a central processing unit (CPU), a graphics processing unit (GPU), a hardware processor core, or any combination thereof), a main memory 604 and a static memory 606, some or all of which can communicate with each other via an interlink (e.g., bus) 608. The machine 600 can further include a display unit 610, an alphanumeric input device 612 (e.g., a keyboard), and a user interface (UI) navigation device 614 (e.g., a mouse). In an example, the display unit 610, alphanumeric input device 612 and UI navigation device 614 can be a touch screen display. The machine 600 can additionally include a storage device (e.g., drive unit) 616, a signal generation device 618 (e.g., a speaker), a network interface device 620, and one or more sensors 621, such as a global positioning system (GPS) sensor, compass, accelerometer, or other sensor. The machine 600 can include an output controller 628, such as a serial (e.g., universal serial bus (USB), parallel, or other wired or wireless (e.g., infrared (IR), near field communication (NFC), etc.) connection to communicate or control one or more peripheral devices (e.g., a printer, card reader, etc.).

25 **[0031]** The storage device 616 can include a machine readable medium 622 that is non-transitory on which is stored one or more sets of data structures or instructions 624 (e.g., software) embodying or utilized by any one or more of the techniques or functions described herein. The instructions 624 can also reside, completely or at least partially, within the main memory 604, within static memory 606, or within the hardware processor 602 during execution thereof by the machine 600. In an example, one or any combination of the hardware

processor 602, the main memory 604, the static memory 606, or the storage device 616 can constitute machine readable media.

[0032] While the machine readable medium 622 is illustrated as a single medium, the term "machine readable medium" can include a single medium or multiple media (e.g., a centralized or distributed database, or associated caches and servers) configured to store the one or more instructions 624.

[0033] The term "machine readable medium" can include any medium that is capable of storing, encoding, or carrying instructions for execution by the machine 600 and that cause the machine 600 to perform any one or more of the techniques of the present disclosure, or that is capable of storing, encoding or carrying data structures used by or associated with such instructions. Non-limiting machine readable medium examples can include solid-state memories, and optical and magnetic media. In an example, a massed machine readable medium comprises a machine readable medium with a plurality of particles having invariant (e.g., rest) mass. Accordingly, massed machine-readable media are not transitory propagating signals. Specific examples of massed machine readable media can include: non-volatile memory, such as semiconductor memory devices (e.g., Electrically Programmable Read-Only Memory (EPROM), Electrically Erasable Programmable Read-Only Memory (EEPROM)) and flash memory devices; magnetic disks, such as internal hard disks and removable disks; magneto-optical disks; and CD-ROM and DVD-ROM disks.

[0034] The instructions 624 can further be transmitted or received over a communications network 626 using a transmission medium via the network interface device 620 utilizing any one of a number of transfer protocols (e.g., frame relay, internet protocol (IP), transmission control protocol (TCP), user datagram protocol (UDP), hypertext transfer protocol (HTTP), etc.). Example communication networks can include a local area network (LAN), a wide area network (WAN), a packet data network (e.g., the Internet), mobile telephone networks (e.g., cellular networks), Plain Old Telephone (POTS) networks, and wireless data networks (e.g., Institute of Electrical and Electronics Engineers (IEEE) 802.11 family of standards known as Wi-Fi®, IEEE 802.16 family of standards known as WiMax®, IEEE 802.15.4 family of standards, peer-to-peer (P2P) networks, among others. In an example, the network interface device 620

can include one or more physical jacks (e.g., Ethernet, coaxial, or phone jacks) or one or more antennas to connect to the communications network 626. In an example, the network interface device 620 can include a plurality of antennas to wirelessly communicate using at least one of single-input multiple-output (SIMO), multiple-input multiple-output (MIMO), or multiple-input single-output (MISO) techniques. The term “transmission medium” shall be taken to include any intangible medium that is capable of storing, encoding or carrying instructions for execution by the machine 600, and includes digital or analog communication signals or other intangible medium to facilitate communication of such software.

[0035] FIG. 7 is a block diagram illustrating an animation device for displaying an avatar animation, according to an embodiment. In an example, an animation device 700 may include a communication module 702 to receive an animation of an avatar and a presentation module 704 to extract an avatar identifier, such as from an avatar sequencing data structure, and determine if the avatar exists locally. When the presentation module 704 determines that the avatar exists locally, the presentation module 704 selects the local avatar without downloading a new avatar. When the presentation module 704 determines that the avatar does not exist locally, a download module 706 may download the avatar. In an example, the presentation module 704 may select the downloaded avatar. In another example, the presentation module 704 may select a different avatar that is stored locally. After an avatar is selected, the presentation module 704 plays the animation with the selected avatar.

[0036] In an example, a communication module 702 may receive an animation of an avatar. A presentation module 704 may extract an avatar identifier, such as from an avatar sequencing data structure, and determine if the avatar exists locally. When the presentation module 704 determines that the avatar exists locally, the presentation module 704 selects the local avatar. When the presentation module 704 determines that the avatar does not exist locally, the presentation module 704 may determine the avatar from an avatar identifier, such as from an avatar sequencing data structure, in an audio file. After an avatar is selected, the presentation module 704 plays the animation with the selected avatar.

[0037] An image capture device may use an image capture module (not shown) to capture an image or series of images. In an example, the image capture device may be a camera or the image capture module may include a camera. A facial recognition module (not shown) may compute facial motion data, such as a set of facial coordinates, for each of the images in the series of images. The facial recognition module may detect a face, such as by determining if certain features are present. The facial recognition module may also track a face, such as by capturing a series of images and determining if certain features are in different places in consecutive images. In an example, if a face is present (e.g., detected), the image capture device may operate at a normal frame rate. In an example, a facial recognition module may detect a face using an image in five milliseconds, if the image resolution is 192 by 144 pixels. The facial recognition module may use significant system resources to detect the face.

[0038] In another example, if the facial recognition module determines or fails to detect that a face is absent, or not present, the image capture device may operate at a reduced frame rate. The facial recognition module may determine that the face is absent from an image or a series of images for a duration, and the facial recognition module may send an indication to the image capture device to alter the frame rate. In another example, if the image capture device is operating at a reduced frame rate, and the facial recognition module detects that a face is present in an image or a series of images, the facial recognition module may send an indication to the image capture device to change the frame rate to a normal frame rate.

Additional Notes& Examples

[0039] Example 1 includes the subject matter embodied by an animation device comprising: a communication module to receive, from an image capture device, an audio file comprising: facial motion data, the facial motion data derived from a series of facial images captured at the image capture device, an avatar sequencing data structure from the image capture device, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream, and a presentation module to present an animation of an avatar using the facial motion data and the audio stream.

[0040] In Example 2, the subject matter of Example 1 may optionally include wherein to present the animation of the avatar, the presentation module is to animate the avatar using the avatar identifier.

5 [0041] In Example 3, the subject matter of one or any combination of Examples 1-2 may optionally include wherein to present the animation of the avatar, the presentation module is to animate the avatar for a period of time lasting for the duration.

10 [0042] In Example 4, the subject matter of one or any combination of Examples 1-3 may optionally includewherein to present the animation of the avatar, the presentation module is to animate the avatar, not using the avatar identifier, for a period of time lasting for the duration.

15 [0043] In Example 5, the subject matter of one or any combination of Examples 1-4 may optionally includewherein to present the animation of the avatar, the presentation module is to animate a locally-selected avatar that is selected locally at the second device.

[0044] In Example 6, the subject matter of one or any combination of Examples 1-5 may optionally includewherein to present the animation of the avatar, the presentation module is to select the avatar from local memory of the animation device.

20 [0045] In Example 7, the subject matter of one or any combination of Examples 1-6 may optionally includefurther comprising a download module to download the avatar from a server at the second device.

[0046] In Example 8, the subject matter of one or any combination of Examples 1-7 may optionally includewherein the communication module
25 receives either a proprietary avatar communication data structure or a Core Audio Format (CAF) communication data structure, and wherein when the communication module receives the proprietary avatar communication data structure, the communication module extracts the avatar sequencing data structure using a Universal Resource Locator (URL) stored in metadata of the
30 proprietary avatar communication data structure, andwherein when the communication module receives the CAF communication data structure, the communication moduleextracts the avatar sequencing data structure stored in reserved free chunk space in the CAF communication data structure.

[0047] In Example 9, the subject matter of one or any combination of Examples 1-8 may optionally includewherein the audio file comprises a second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.

5 [0048] In Example 10, the subject matter of one or any combination of Examples 1-9 may optionally includewherein to present the animation, the presentation module is to present a first animation of the avatar corresponding to the avatar identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a
10 period of time lasting the second duration.

[0049] In Example 11, the subject matter of one or any combination of Examples 1-10 may optionally includewherein to present the animation of the avatar, the presentation module is to use two avatars.

15 [0050] In Example 12, the subject matter of one or any combination of Examples 1-11 may optionally includewherein the two avatars are selected locally at the image capture device.

[0051] In Example 13, the subject matter of one or any combination of Examples 1-12 may optionally includewherein the two avatars are selected locally at the animation device.

20 [0052] In Example 14, the subject matter of one or any combination of Examples 1-13 may optionally includewherein one of the two avatars is identified by the avatar identifier.

[0053] In Example 15, the subject matter of one or any combination of Examples 1-14 may optionally includewherein the audio file further comprises a
25 specified number of avatar sequencing data structures.

[0054] In Example 16, the subject matter of one or any combination of Examples 1-15 may optionally includewherein to present the animation of the avatar, the presentation module is to use a different number of avatars than the specified number of avatar sequencing data structures.

30 [0055] In Example 17, the subject matter of one or any combination of Examples 1-16 may optionally includewherein to present the animation of the avatar, the presentation module is to animate the avatar, using a locally-selected

avatar that is selected locally at the animation device, for a period of time lasting for a sum of all durations in the avatar sequencing data structures.

5 **[0056]** Example 18 includes the subject matter embodied by an image capture device comprising: an image capture module to capture a series of images of a face, a facial recognition module to compute facial motion data for each of the images in the series of images, and a communication module to send to an animation device, an audio file comprising: the facial motion data, an avatar sequencing data structure, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream, and wherein the animation
10 device is configured to use the facial motion data and the audio stream to animate an avatar on the animation device.

[0057] In Example 19, the subject matter of Example 18 may optionally include wherein to animate the avatar, the animation device is to animate the avatar using the avatar identifier.

15 **[0058]** In Example 20, the subject matter of one or any combination of Examples 18-19 may optionally include wherein to animate the avatar, the animation device is to animate the avatar for a period of time lasting for the duration.

[0059] In Example 21, the subject matter of one or any combination of
20 Examples 18-20 may optionally include wherein to animate the avatar, the animation device is to animate the avatar, not using the avatar identifier, for a period of time lasting for the duration.

[0060] In Example 22, the subject matter of one or any combination of
25 Examples 18-21 may optionally include wherein to animate the avatar, the animation device is to animate a locally-selected avatar that is selected locally at the animation device.

[0061] In Example 23, the subject matter of one or any combination of
30 Examples 18-22 may optionally include wherein to animate the avatar, the animation device is to select the avatar from local memory of the animation device.

[0062] In Example 24, the subject matter of one or any combination of Examples 18-23 may optionally include wherein the animation device is further configured to download the avatar from a server.

[0063] In Example 25, the subject matter of one or any combination of Examples 18-24 may optionally includewherein the communication module sends either a proprietary avatar communication data structure or a Core Audio Format (CAF) communication data structure, andwherein when the communication module sends the proprietary avatar communication data structure, the communication module stores the avatar sequencing data structure,using a Universal Resource Locator (URL), in metadata of the proprietary avatar communication data structure, andwherein when the communication module sends the CAF communication data structure, the communication module stores the avatar sequencing data structure in reserved free chunk space in the CAF communication data structure.

[0064] In Example 26, the subject matter of one or any combination of Examples 18-25 may optionally includewherein the audio file comprises a second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.

[0065] In Example 27, the subject matter of one or any combination of Examples 18-26 may optionally includewherein to animate the avatar, the animation device is to animate a first animation of the avatar corresponding to the avatar identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a period of time lasting the second duration.

[0066] In Example 28, the subject matter of one or any combination of Examples 18-27 may optionally includewherein the animation device is further configured to use two avatars.

[0067] In Example 29, the subject matter of one or any combination of Examples 18-28 may optionally includewherein the animation device is further configured to select the two avatars locally at the image capture device.

[0068] In Example 30, the subject matter of one or any combination of Examples 18-29 may optionally includewherein the animation device is further configured to select the two avatars locally at the animation device.

[0069] In Example 31, the subject matter of one or any combination of Examples 18-30 may optionally includewherein one of the avatars is identified by the first avatar identifier.

[0070] In Example 32, the subject matter of one or any combination of Examples 18-31 may optionally includewherein the audio file further comprises a specified number of avatar sequencing data structures.

5 [0071] In Example 33, the subject matter of one or any combination of Examples 18-32 may optionally includewherein to animate the avatar, the animation device is to use a different number of avatars than the specified number of avatar sequencing data structures.

10 [0072] In Example 34, the subject matter of one or any combination of Examples 18-33 may optionally includewherein to animate the avatar, the animation device is to animate the avatar, using a locally-selected avatar that is selected locally at the animation device, for a period of time lasting for a sum of all durations in the avatar sequencing data structures.

15 [0073] In Example 35, the subject matter of one or any combination of Examples 18-34 may optionally includewherein the facial recognition module is further configured to detect if the face is present in the series of images captured by the image capture device.

20 [0074] In Example 36, the subject matter of one or any combination of Examples 18-35 may optionally includewherein the image capture module operates at a normal frame rate if the facial recognition module detects that the face is present in the series of images.

[0075] In Example 37, the subject matter of one or any combination of Examples 18-36 may optionally includewherein the image capture module operates at a reduced frame rate if the facial recognition module detects that the face is absent in the series of images.

25 [0076] Example 38 includes the subject matter embodied by an avatar presentation method comprising: receiving, at a second device from a first device, an audio file comprising: facial motiondata, the facial motion data derived from a series of facial images captured at the first device, an avatar sequencing data structure from the first device, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream, and
30 presenting an animation of an avatar, at the second device, using the facial motion data and the audio stream.

[0077] In Example 39, the subject matter of Example 18 may optionally include further comprising, selecting the avatar using the avatar identifier.

[0078] In Example 40, the subject matter of one or any combination of Examples 38-39 may optionally include further comprising, presenting the animation of the avatar for a period of time lasting for the duration.

[0079] In Example 41, the subject matter of one or any combination of Examples 38-40 may optionally include further comprising, selecting the avatar by not using the avatar identifier.

[0080] In Example 42, the subject matter of one or any combination of Examples 38-41 may optionally include further comprising, selecting the avatar locally at the second device.

[0081] In Example 43, the subject matter of one or any combination of Examples 38-42 may optionally include wherein selecting the avatar includes selecting the avatar from local memory of the second device.

[0082] In Example 44, the subject matter of one or any combination of Examples 38-43 may optionally include further comprising, downloading the avatar from a server at the second device.

In Example 45, the subject matter of one or any combination of Examples 38-44 may optionally include wherein receiving includes receiving either a proprietary avatar communication data structure or a Core Audio Format (CAF) communication data structure, and wherein after receiving the proprietary avatar communication data structure, extracting the avatar sequencing data structure using a Universal Resource Locator (URL) stored in metadata of the proprietary avatar communication data structure, and wherein after receiving the CAF communication data structure, extracting the avatar sequencing data structure stored in reserved free chunk space in the CAF communication data structure.

[0083] In Example 46, the subject matter of one or any combination of Examples 38-45 may optionally include wherein the audio file comprises a second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.

[0084] In Example 47, the subject matter of one or any combination of Examples 38-46 may optionally include wherein presenting the animation includes presenting a first animation of the avatar corresponding to the avatar

identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a period of time lasting the second duration.

5 **[0085]** In Example 48, the subject matter of one or any combination of Examples 38-47 may optionally includewherein the audio file comprises a second avatar identifier.

[0086] In Example 49, the subject matter of one or any combination of Examples 38-48 may optionally includewherein presenting the animation of the avatar includes using two avatars.

10 **[0087]** In Example 50, the subject matter of one or any combination of Examples 38-49 may optionally includefurther comprising, selecting the two avatars locally at the first device.

[0088] In Example 51, the subject matter of one or any combination of Examples 38-50 may optionally includefurther comprising, selecting the two
15 avatars locally at the second device.

[0089] In Example 52, the subject matter of one or any combination of Examples 38-51 may optionally includewherein one of the two avatars is identified by the first avatar identifier.

[0090] In Example 53, the subject matter of one or any combination of
20 Examples 38-52 may optionally includewherein the audio file further comprises a specified number of avatar sequencing data structures.

[0091] In Example 54, the subject matter of one or any combination of Examples 38-53 may optionally includewherein presenting the animation of the avatar includes using a different number of avatars than the specified number of
25 avatar sequencing data structures.

[0092] In Example 55, the subject matter of one or any combination of Examples 38-54 may optionally includewherein presenting the animation of the avatar includes presenting the animation, using a locally-selected avatar that is selected locally at the second device, for a period of time lasting for a sum of all
30 durations in the avatar sequencing data structures.

[0093] In Example 56, the subject matter of one or any combination of Examples 38-55 may optionally includea machine-readable medium including

instructions for receiving information, which when executed by a machine, cause the machine to perform any of the methods of Examples 38-55.

[0094] Example 57 includes the subject matter embodied by a machine-readable medium including instructions for receiving information, which when executed by a machine, cause the machine to: receive, at a second device from a first device, an audio file comprising: facial motion data, the facial motion data derived from a series of facial images captured at the first device, an avatar sequencing data structure from the first device, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream; and present an animation, at the second device, of an avatar using the facial motion data and the audio stream.

[0095] In Example 58, the subject matter of Example 18 may optionally include further comprising, selecting the avatar using the avatar identifier.

[0096] In Example 59, the subject matter of one or any combination of Examples 57-58 may optionally include further comprising instructions for receiving information, which when executed by a machine, cause the machine to: present the animation of the avatar for a period of time lasting for the duration.

[0097] In Example 60, the subject matter of one or any combination of Examples 57-59 may optionally include further comprising instructions for receiving information, which when executed by a machine, cause the machine to: select the avatar by not using the avatar identifier.

[0098] In Example 61, the subject matter of one or any combination of Examples 57-60 may optionally include further comprising instructions for receiving information, which when executed by a machine, cause the machine to: select the avatar locally at the second device.

[0099] In Example 62, the subject matter of one or any combination of Examples 57-61 may optionally include wherein to select the avatar includes to select the avatar from local memory of the second device.

[0100] In Example 63, the subject matter of one or any combination of Examples 57-62 may optionally include further comprising instructions for receiving information, which when executed by a machine, cause the machine to: download the avatar from a server at the second device.

[00101] In Example 64, the subject matter of one or any combination of Examples 57-63 may optionally includewherein to receive includes to receive either a proprietary avatar communication data structure or a Core Audio Format (CAF) communication data structure, andwherein after receiving the proprietary
5 avatar communication data structure, to extract the avatar sequencing data structure using a Universal Resource Locator (URL) stored in metadata of the proprietary avatar communication data structure, andwherein afterreceiving the CAF communication data structure, to extract the avatar sequencing data structure stored in reserved free chunk space in the CAF communication data
10 structure.

[00102] In Example 65, the subject matter of one or any combination of Examples 57-64 may optionally includewherein the audio file comprises a second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.

15 [00103] In Example 66, the subject matter of one or any combination of Examples 57-65 may optionally include wherein to present the animation includes to present a first animation of the avatar corresponding to the avatar identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a period of time
20 lasting the second duration.

[00104] In Example 67, the subject matter of one or any combination of Examples 57-66 may optionally includewherein the audio file comprises a second avatar identifier.

[00105] In Example 68, the subject matter of one or any combination of
25 Examples 57-67 may optionally includewherein to present the animation of the avatar includes to present the animation of the avatar using two avatars.

[00106] In Example 69, the subject matter of one or any combination of Examples 57-68 may optionally includefurther comprisinginstructions for receiving information, which when executed by a machine, cause the machine
30 to:select the two avatars locally at the first device.

[00107] In Example 70, the subject matter of one or any combination of Examples 57-69 may optionally includefurther comprisinginstructions for

receiving information, which when executed by a machine, cause the machine to select the two avatars locally at the second device.

5 [00108] In Example 71, the subject matter of one or any combination of Examples 57-70 may optionally includewherein one of the two avatars is identified by the first avatar identifier.

[00109] In Example 72, the subject matter of one or any combination of Examples 57-71 may optionally includewherein the audio file further comprises a specified number of avatar sequencing data structures.

10 [00110] In Example 73, the subject matter of one or any combination of Examples 57-72 may optionally includewherein to present the animation of the avatar includes to present the animation of the avatar using a different number of avatars than the specified number of avatar sequencing data structures.

15 [00111] In Example 74, the subject matter of one or any combination of Examples 57-73 may optionally includewherein to present the animation of the avatar includes to present the animation of the avatar, using a locally-selected avatar that is selected locally at the second device, for a period of time lasting for a sum of all durations in the avatar sequencing data structures.

20 [00112] In Example 75, the subject matter of one or any combination of Examples 38-55 may optionally includean apparatus comprising means for performing any of the methods of Examples 38-55.

[00113] Example 76 includes the subject matter embodied by an apparatus comprising: means for receiving, at a second device from a first device, an audio file comprising: facial motion data, the facial motion data derived from a series of facial images captured at the first device, an avatar sequencing data structure from the first device, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream, and means for presenting an animation, at the second device, of an avatar using the facial motion data and the audio stream.

30 [00114] Example 77 includes the subject matter embodied byan audio file delivery method comprising: capturing a series of images of a face using an image capture device, computing facial motion data for each of the images in the series of images, and sending, to an animation device, an audio file comprising: the facial motion data, an avatar sequencing data structure, the avatar sequencing

data structure comprising an avatar identifier and a duration, and an audio stream, detecting that the animation device is configured to use the facial motion data and the audio stream to animate an avatar on the animation device.

5 [00115] In Example 78, the subject matter of Example 77 may optionally include wherein the avatar is selected using the avatar identifier.

[00116] In Example 79, the subject matter of one or any combination of Examples 77-78 may optionally include further comprising, detecting that the animation device is configured to animate the avatar for a period of time lasting for the duration.

10 [00117] In Example 80, the subject matter of one or any combination of Examples 77-79 may optionally include wherein the avatar is not selected by the avatar identifier.

[00118] In Example 81, the subject matter of one or any combination of Examples 77-80 may optionally include further comprising, detecting that the animation device is configured to select the avatar.

15 [00119] In Example 82, the subject matter of one or any combination of Examples 77-81 may optionally include further comprising, detecting that the animation device is configured to select the avatar from local memory of the animation device.

20 [00120] In Example 83, the subject matter of one or any combination of Examples 77-82 may optionally include further comprising, detecting that the animation device is configured to download the avatar from a server.

[00121] In Example 84, the subject matter of one or any combination of Examples 77-83 may optionally include wherein sending includes sending either
25 a proprietary avatar communication data structure or a Core Audio Format (CAF) communication data structure, and wherein before sending the proprietary avatar communication data structure, storing the avatar sequencing data structure, using a Universal Resource Locator (URL), in metadata of the proprietary avatar communication data structure, and wherein before sending the CAF
30 communication data structure, storing the avatar sequencing data structure in reserved free chunk space in the CAF communication data structure.

[00122] In Example 85, the subject matter of one or any combination of Examples 77-84 may optionally include wherein the audio file comprises a

second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.

5 **[00123]** In Example 86, the subject matter of one or any combination of Examples 77-85 may optionally include further comprising, detecting that to animate the avatar, the animation device is configured to animate a first animation of the avatar corresponding to the avatar identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a period of time lasting the second duration.

10 **[00124]** In Example 87, the subject matter of one or any combination of Examples 77-86 may optionally include further comprising, detecting that the animation device is configured to use two avatars.

15 **[00125]** In Example 88, the subject matter of one or any combination of Examples 77-87 may optionally include further comprising, detecting that the animation device is configured to select the two avatars locally at the image capture device.

20 **[00126]** In Example 89, the subject matter of one or any combination of Examples 77-88 may optionally include further comprising, detecting that the animation device is configured to select the two avatars locally at the animation device.

25 **[00127]** In Example 90, the subject matter of one or any combination of Examples 77-89 may optionally include wherein one of the avatars is identified by the first avatar identifier.

30 **[00128]** In Example 91, the subject matter of one or any combination of Examples 77-90 may optionally include wherein the audio file further comprises a specified number of avatar sequencing data structures.

35 **[00129]** In Example 92, the subject matter of one or any combination of Examples 77-91 may optionally include further comprising, detecting that the animation device is configured to use a different number of avatars than the specified number of avatar sequencing data structures.

40 **[00130]** In Example 93, the subject matter of one or any combination of Examples 77-92 may optionally include further comprising, detecting that the animation device is configured to animate the avatar, using a locally-selected

avatar that is selected locally at the animation device, for a period of time lasting for a sum of all durations in the avatar sequencing data structures.

[00131] In Example 94, the subject matter of one or any combination of Examples 77-93 may optionally include further comprising, detecting if the face is present in the series of images captured by the image capture device.

[00132] In Example 95, the subject matter of one or any combination of Examples 77-94 may optionally include wherein capturing the series of images includes capturing the series of images at a normal frame rate if the face is present in the series of images.

[00133] In Example 96, the subject matter of one or any combination of Examples 77-95 may optionally include wherein capturing the series of images includes capturing the series of images at a reduced frame rate if the face is absent in the series of images.

[00134] In Example 97, the subject matter of one or any combination of Examples 77-96 may optionally include a machine-readable medium including instructions for receiving information, which when executed by a machine, cause the machine to perform any of the methods of claims 77-96.

[00135] Example 98 includes the subject matter embodied by a machine-readable medium including instructions for receiving information, which when executed by a machine, cause the machine to: capture a series of images of a face using an image capture device, compute facial motion data for each of the images in the series of images, and send, to an animation device, an audio file comprising: the facial motion data, an avatar sequencing data structure, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream, wherein the animation device is configured to use the facial motion data and the audio stream to animate an avatar on the animation device.

[00136] In Example 99, the subject matter of one or any combination of Examples 77-96 may optionally include an apparatus comprising means for performing any of the methods of claims 77-96.

[00137] Example 100 includes the subject matter embodied by an apparatus comprising: means for capturing a series of images of a face using an image capture device, means for computing facial motion data for each of the images in the series of images, and means for sending, to an animation device, an audio

file comprising: the facial motion data, an avatar sequencing data structure, the avatar sequencing data structure comprising an avatar identifier and a duration, and an audio stream, wherein the animation device is configured to use the facial motion data and the audio stream to animate an avatar on the animation device.

5 **[00138]** Each of these non-limiting examples can stand on its own, or can be combined in various permutations or combinations with one or more of the other examples.

10 **[00139]** The above detailed description includes references to the accompanying drawings, which form a part of the detailed description. The drawings show, by way of illustration, specific embodiments in which the invention can be practiced. These embodiments are also referred to herein as “examples.” Such examples can include elements in addition to those shown or described. However, the present inventors also contemplate examples in which only those elements shown or described are provided. Moreover, the present
15 inventors also contemplate examples using any combination or permutation of those elements shown or described (or one or more aspects thereof), either with respect to a particular example (or one or more aspects thereof), or with respect to other examples (or one or more aspects thereof) shown or described herein.

20 **[00140]** In the event of inconsistent usages between this document and any documents so incorporated by reference, the usage in this document controls.

25 **[00141]** In this document, the terms “a” or “an” are used, as is common in patent documents, to include one or more than one, independent of any other instances or usages of “at least one” or “one or more.” In this document, the term “or” is used to refer to a nonexclusive or, such that “A or B” includes “A but not B,” “B but not A,” and “A and B,” unless otherwise indicated. In this document, the terms “including” and “in which” are used as the plain-English equivalents of the respective terms “comprising” and “wherein.” Also, in the following claims, the terms “including” and “comprising” are open-ended, that is, a system, device, article, composition, formulation, or process that includes elements in
30 addition to those listed after such a term in a claim are still deemed to fall within the scope of that claim. Moreover, in the following claims, the terms “first,” “second,” and “third,” etc. are used merely as labels, and are not intended to impose numerical requirements on their objects.

[00142] Method examples described herein can be machine or computer-implemented at least in part. Some examples can include a computer-readable medium or machine-readable medium encoded with instructions operable to configure an electronic device to perform methods as described in the above examples. An implementation of such methods can include code, such as microcode, assembly language code, a higher-level language code, or the like. Such code can include computer readable instructions for performing various methods. The code may form portions of computer program products. Further, in an example, the code can be tangibly stored on one or more volatile, non-transitory, or non-volatile tangible computer-readable media, such as during execution or at other times. Examples of these tangible computer-readable media can include, but are not limited to, hard disks, removable magnetic disks, removable optical disks (e.g., compact disks and digital video disks), magnetic cassettes, memory cards or sticks, random access memories (RAMs), read only memories (ROMs), and the like.

[00143] The above description is intended to be illustrative, and not restrictive. For example, the above-described examples (or one or more aspects thereof) may be used in combination with each other. Other embodiments can be used, such as by one of ordinary skill in the art upon reviewing the above description. The Abstract is provided to comply with 37 C.F.R. §1.72(b), to allow the reader to quickly ascertain the nature of the technical disclosure. It is submitted with the understanding that it will not be used to interpret or limit the scope or meaning of the claims. Also, in the above Detailed Description, various features may be grouped together to streamline the disclosure. This should not be interpreted as intending that an unclaimed disclosed feature is essential to any claim. Rather, inventive subject matter may lie in less than all features of a particular disclosed embodiment. Thus, the following claims are hereby incorporated into the Detailed Description as examples or embodiments, with each claim standing on its own as a separate embodiment, and it is contemplated that such embodiments can be combined with each other in various combinations or permutations. The scope of the invention should be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled.

CLAIMS

What is claimed is:

1. An animation device comprising:
a communication module to receive, from an image capture device, an
5 audio file comprising:
facial motion data, the facial motion data derived from a series of
facial images captured at the image capture device;
an avatar sequencing data structure from the image capture device,
the avatar sequencing data structure comprising an avatar identifier and a
10 duration; and
an audio stream; and
a presentation module to present an animation of an avatar using the
facial motion data and the audio stream.
- 15 2. The animation device of claim 1, wherein to present the animation of the
avatar, the presentation module is to animate the avatar using the avatar
identifier.
3. The animation device of claim 1, wherein to present the animation of the
20 avatar, the presentation module is to animate the avatar for a period of time
lasting for the duration.
4. The animation device of claim 1, wherein to present the animation of the
avatar, the presentation module is to animate the avatar, not using the avatar
25 identifier, for a period of time lasting for the duration.
5. The animation device of claim 1, wherein to present the animation of the
avatar, the presentation module is to animate a locally-selected avatar that is
selected locally at the second device.
30

6. The animation device of claim 5, wherein to present the animation of the avatar, the presentation module is to select the avatar from local memory of the animation device.
- 5 7. The animation device of claim 1, further comprising a download module to download the avatar from a server at the second device.
8. The animation device of claim 1, wherein the communication module receives either a proprietary avatar communication data structure or a Core
- 10 Audio Format (CAF) communication data structure; and
- wherein when the communication module receives the proprietary avatar communication data structure, the communication module extracts the avatar sequencing data structure using a Universal Resource Locator (URL) stored in metadata of the proprietary avatar communication data structure; and
- 15 wherein when the communication module receives the CAF communication data structure, the communication module extracts the avatar sequencing data structure stored in reserved free chunk space in the CAF communication data structure.
- 20 9. The animation device of claim 1, wherein the audio file comprises a second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.
10. The animation device of claim 9, wherein to present the animation, the
- 25 presentation module is to present a first animation of the avatar corresponding to the avatar identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a period of time lasting the second duration.
- 30 11. The animation device of claim 1, wherein to present the animation of the avatar, the presentation module is to use two avatars.

12. The animation device of claim 11, wherein the two avatars are selected locally at the image capture device.

13. The animation device of claim 11, wherein the two avatars are selected
5 locally at the animation device.

14. The animation device of claim 11, wherein one of the two avatars is identified by the avatar identifier.

10 15. The animation device of claim 1, wherein the audio file further comprises a specified number of avatar sequencing data structures.

16. The animation device of claim 15, wherein to present the animation of the avatar, the presentation module is to use a different number of avatars than the
15 specified number of avatar sequencing data structures.

17. The animation device of claim 16, wherein to present the animation of the avatar, the presentation module is to animate the avatar, using a locally-selected avatar that is selected locally at the animation device, for a period of time lasting
20 for a sum of all durations in the avatar sequencing data structures.

18. An image capture device comprising:
an image capture module to capture a series of images of a face;
a facial recognition module to compute facial motion data for each of the
25 images in the series of images; and
a communication module to send to an animation device, an audio file comprising:
the facial motion data;
an avatar sequencing data structure, the avatar sequencing data
30 structure comprising an avatar identifier and a duration; and
an audio stream; and
wherein the animation device is configured to use the facial motion data and the audio stream to animate an avatar on the animation device.

19. The image capture device of claim 18, wherein the facial recognition module is further configured to detect if the face is present in the series of images captured by the image capture device and wherein the image capture module operates at a normal frame rate if the facial recognition module detects that the face is present in the series of images and the image capture module operates at a reduced frame rate if the facial recognition module detects that the face is absent in the series of images.
20. An avatar presentation method comprising:
receiving, at a second device from a first device, an audio file comprising:
facial motion data, the motion coordinate data derived from a series of facial images captured at the first device;
an avatar sequencing data structure from the first device, the avatar sequencing data structure comprising an avatar identifier and a duration; and
an audio stream; and
presenting an animation of an avatar, at the second device, using the facial motion data and the audio stream.
21. The method of claim 20, wherein the audio file comprises a second avatar sequencing data structure, the second avatar sequencing data structure comprising a second avatar identifier and a second duration.
22. The method of claim 21, wherein presenting the animation includes presenting a first animation of the avatar corresponding to the avatar identifier for a period of time lasting the duration and a second animation of a second avatar corresponding to the second avatar identifier a period of time lasting the second duration.

23. The method of claim 20, wherein receiving includes receiving either a proprietary avatar communication data structure or a Core Audio Format (CAF) communication data structure; and

5 wherein after receiving the proprietary avatar communication data structure, extracting the avatar sequencing data structure using a Universal Resource Locator (URL) stored in metadata of the proprietary avatar communication data structure; and

10 wherein after receiving the CAF communication data structure, extracting the avatar sequencing data structure stored in reserved free chunk space in the CAF communication data structure.

24. A machine-readable medium including instructions for animating an avatar, which when executed by a machine, cause the machine to perform any one of the methods of claims 20-23.

25. An apparatus to animate an avatar, the apparatus comprising means for performing any one of the methods of claims 20-23.

1/7

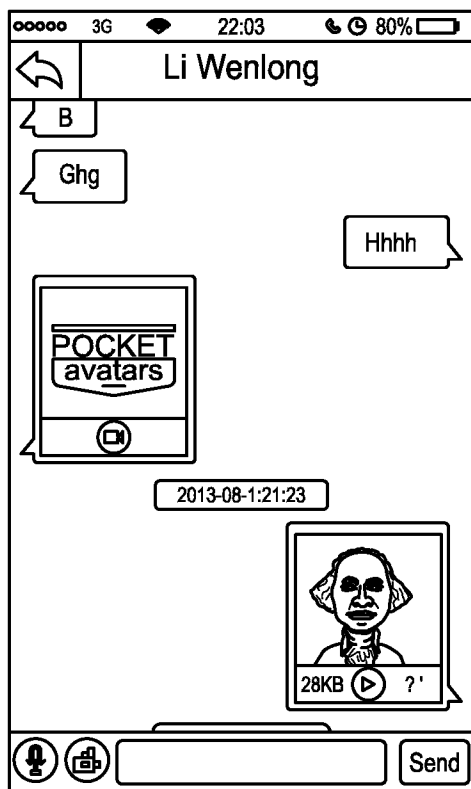


FIG. 1A

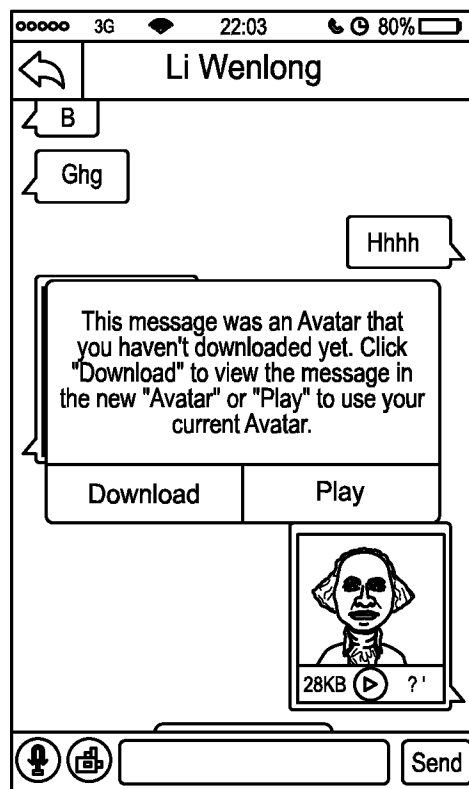


FIG. 1B

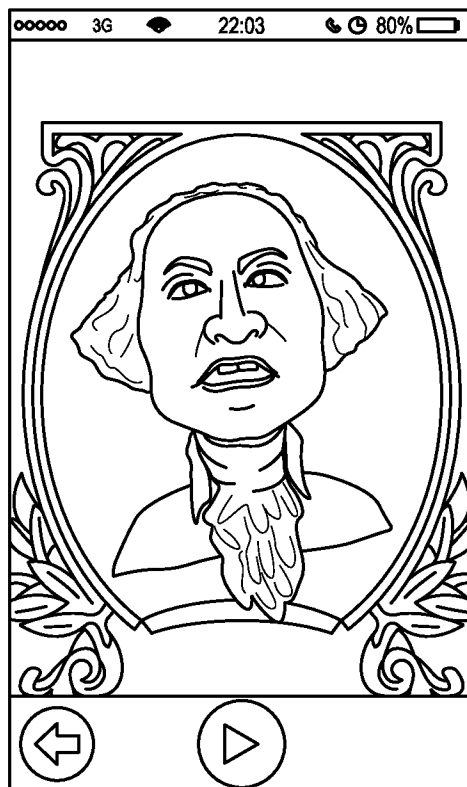


FIG. 1C

2/7

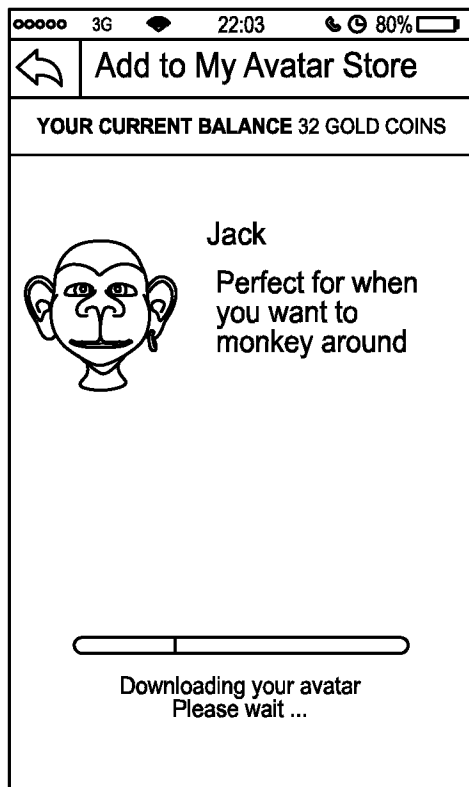


FIG. 1D

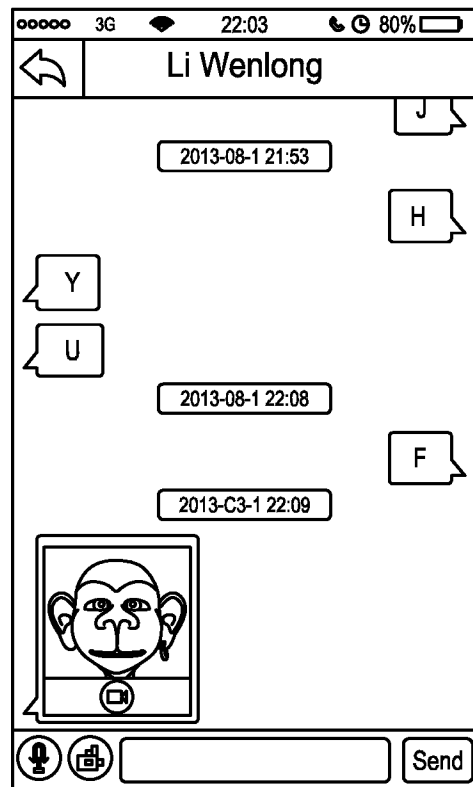


FIG. 1E

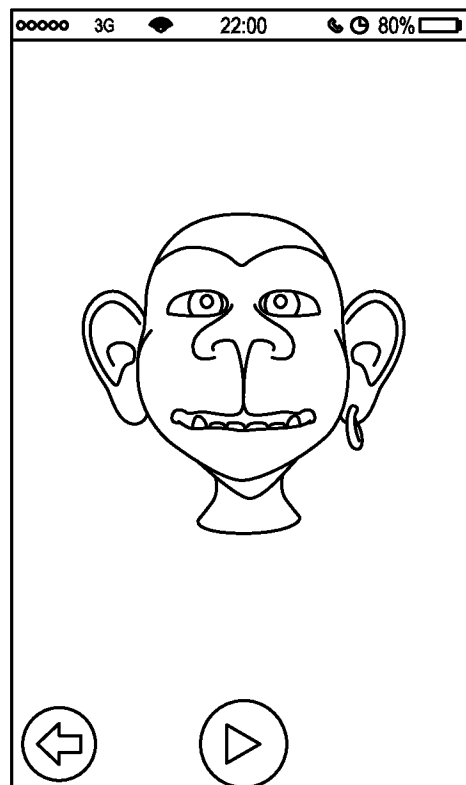


FIG. 1F

3/7

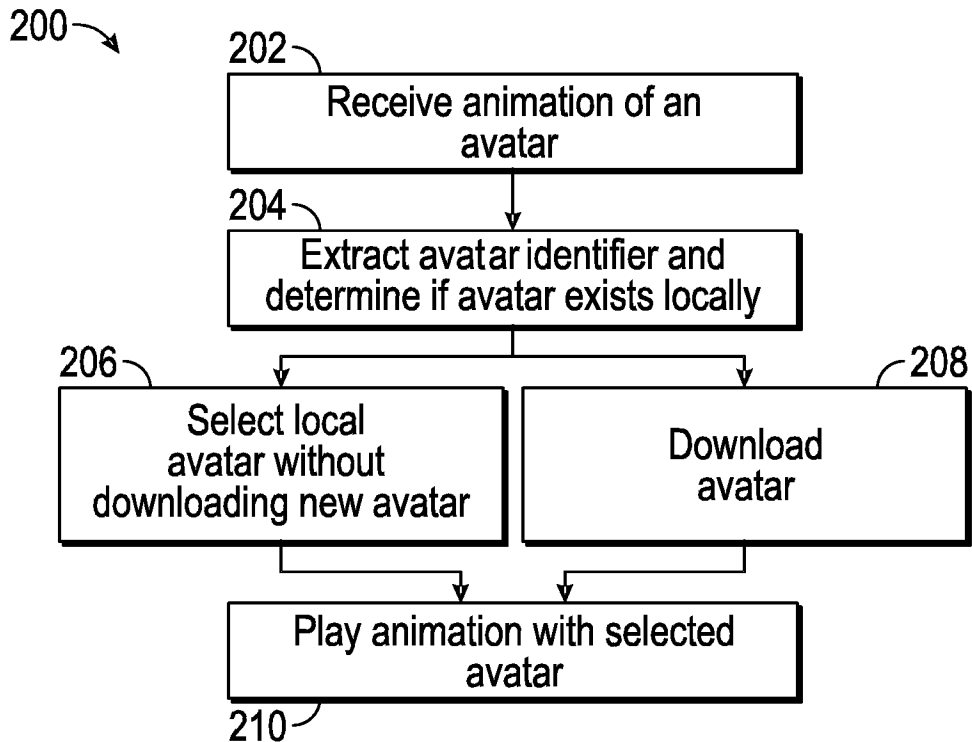


FIG. 2

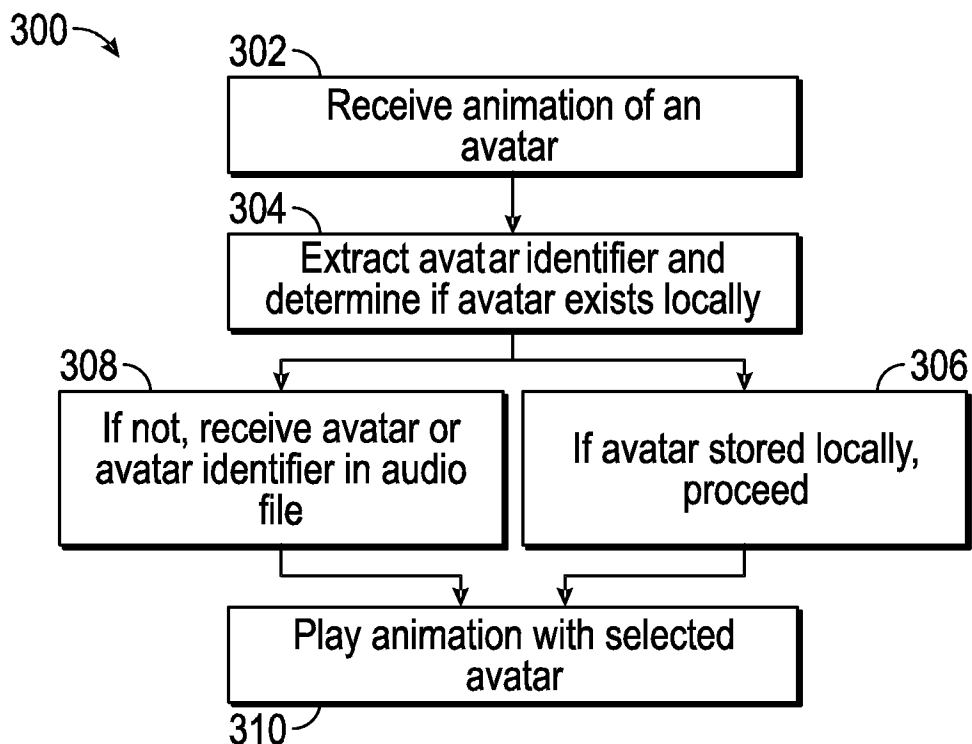


FIG. 3

4/7

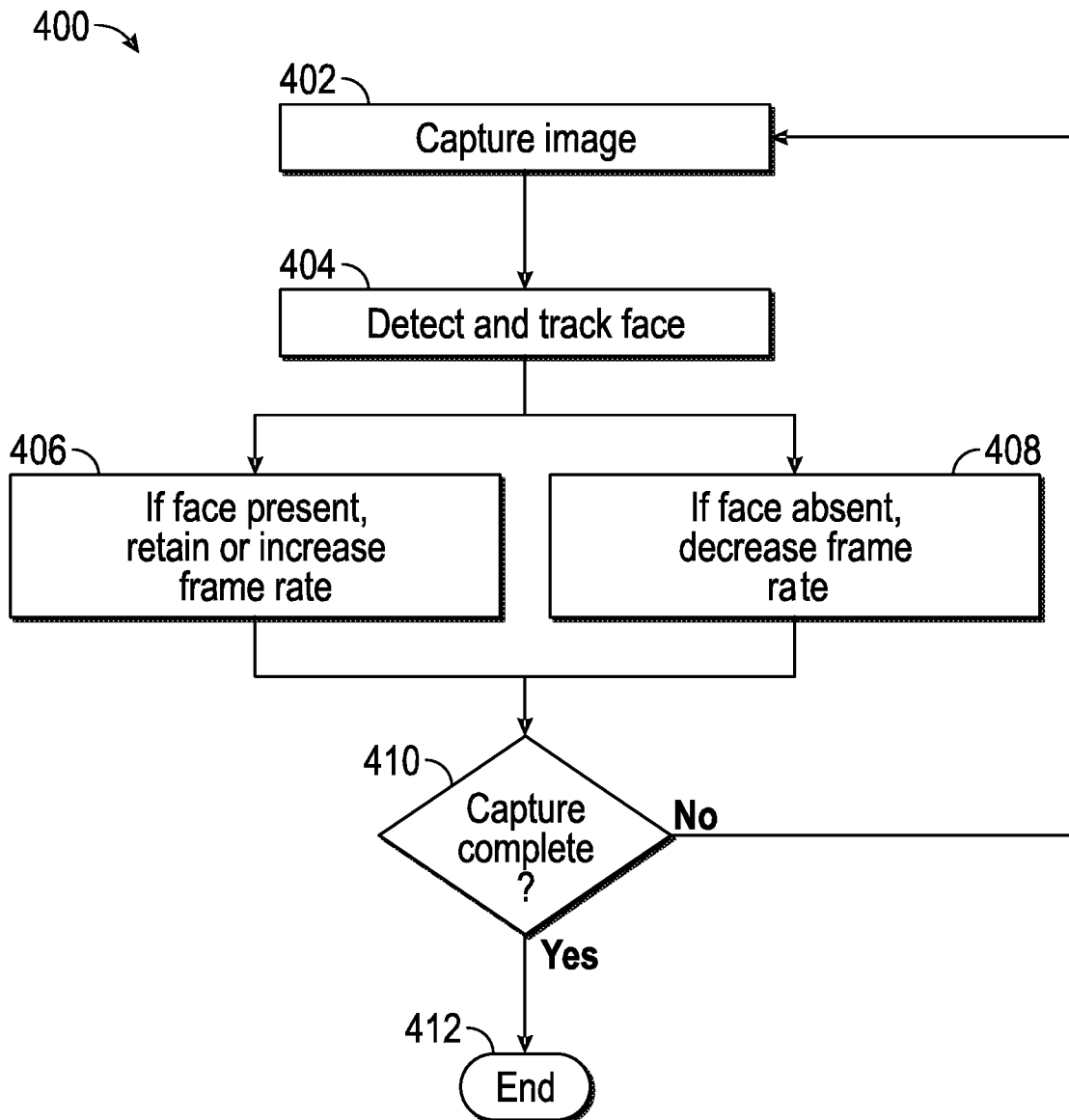


FIG. 4

5/7

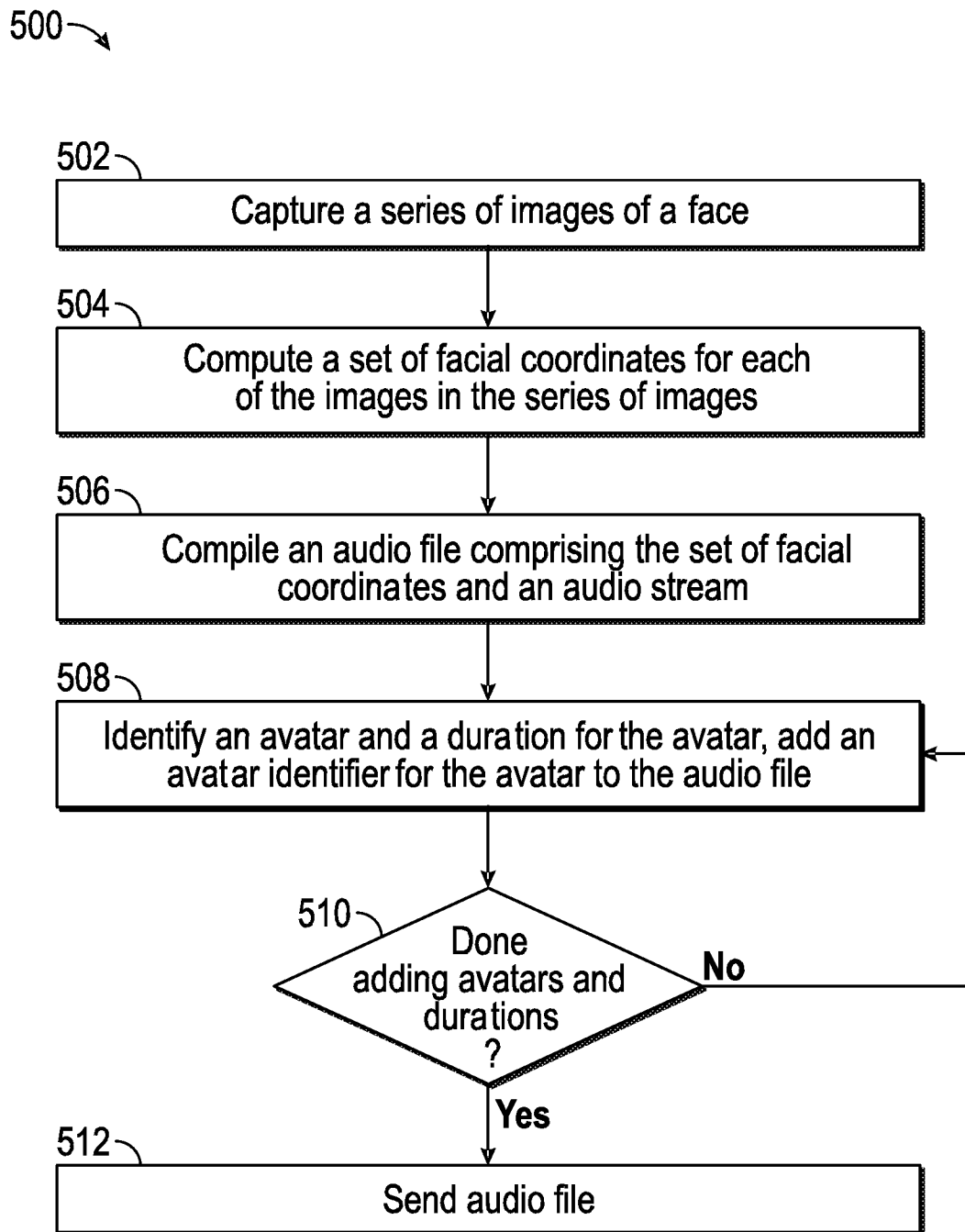


FIG. 5

6/7

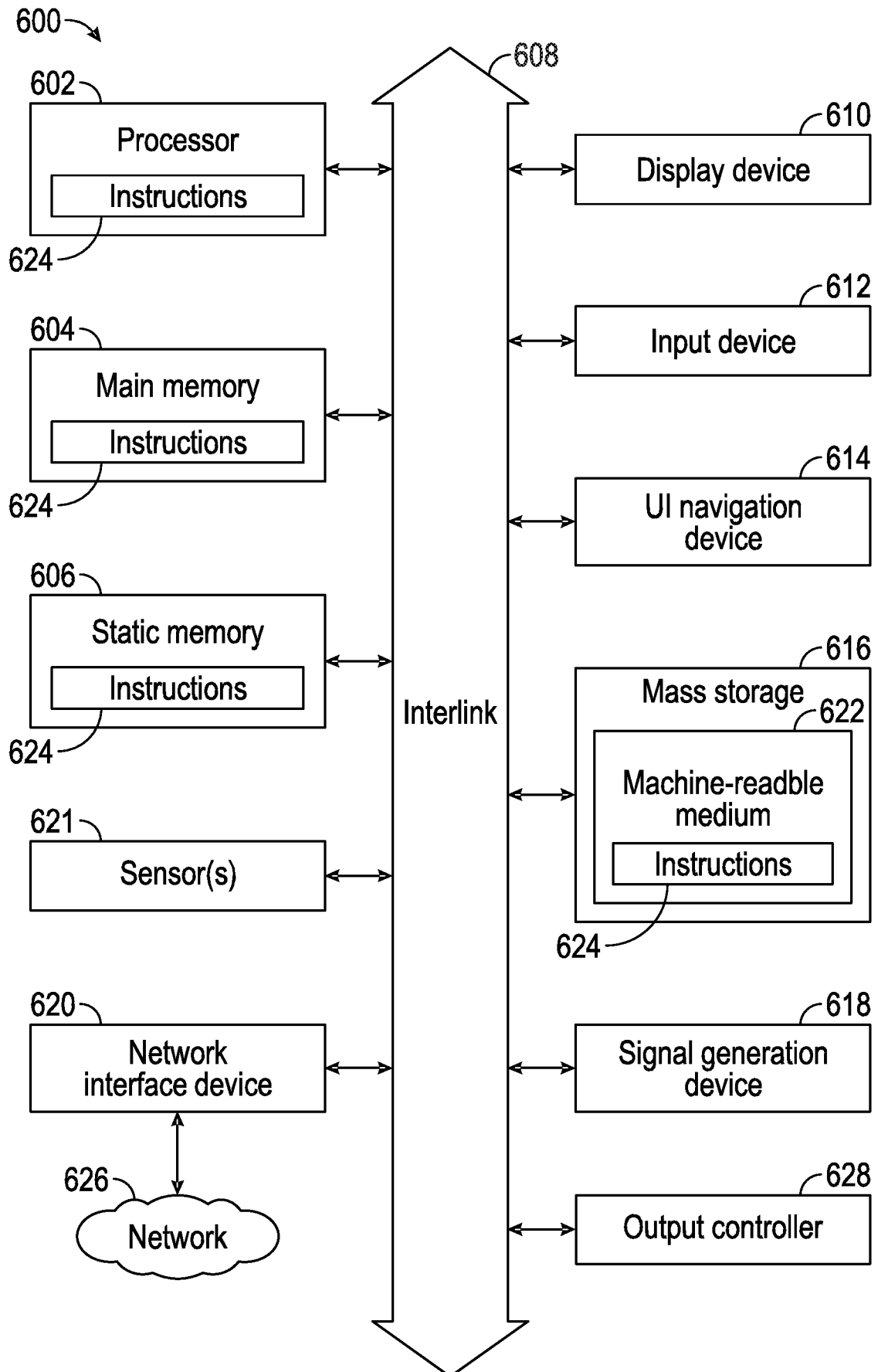


FIG. 6

7/7

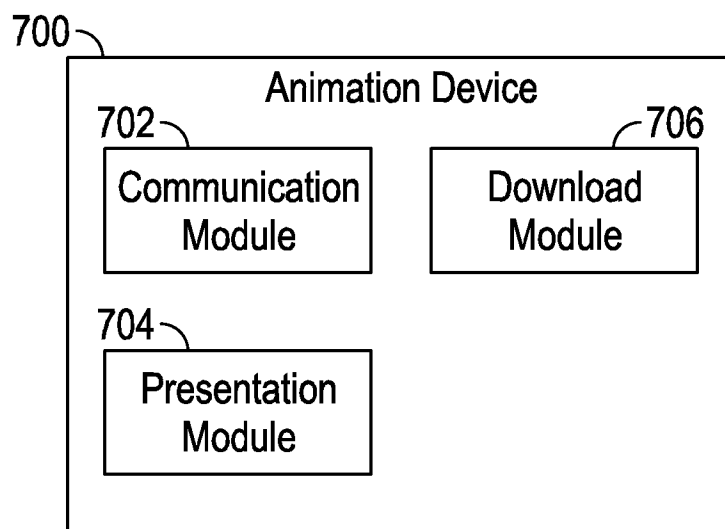


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/087313**A. CLASSIFICATION OF SUBJECT MATTER**

H04M 1/00(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)

H04M; H04N; G06T

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, CNKI, WPI, EPODOC: avatar, anima+, emoticon?, emotion, motion, expression, capture, facial, face, audio, video, sound, tone, messag+, email

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014267544 A1 (INTEL CORPORATION) 18 September 2014 (2014-09-18) claims 1-50; description, paragraphs 24-30, 40-43; figures 1-7	1-25
A	US 2001051535 A1 (MINOLTA CO., LTD.) 13 December 2001 (2001-12-13) the whole document	1-25
A	CN 1460232 A (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 03 December 2003 (2003-12-03) the whole document	1-25
A	WO 2012139276 A1 (INTEL CORPORATION) 18 October 2012 (2012-10-18) the whole document	1-25

☐ 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

13 May 2015

Date of mailing of the international search report

29 May 2015

Name and mailing address of the ISA/CN

**STATE INTELLECTUAL PROPERTY OFFICE OF THE
P.R.CHINA**
**6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088, China**

Authorized officer

WANG, Yujing

Facsimile No. (86-10)62019451

Telephone No. (86-10)62413418

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2014/087313

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2014267544	A1	18 September 2014	WO	2014139142	A1	18 September 2014
US	2001051535	A1	13 December 2001	JP	2001357414	A	26 December 2001
				JP	2001357413	A	26 December 2001
CN	1460232	A	03 December 2003	WO	02080107	A1	10 October 2002
				US	2002194006	A1	19 December 2002
				KR	20030007726	A	23 January 2003
				JP	2004519787	A	02 July 2004
				EP	1374179	A1	02 January 2004
WO	2012139276	A1	18 October 2012	TW	201303747	A	16 January 2013
				US	2014035934	A1	06 February 2014