(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 October 2019 (24.10.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/204464 A1

(51) International Patent Classification:
G06T 13/40 (2011.01) *G06T 19/00* (2011.01)
G06T 7/20 (2017.01) *G06K 9/00* (2006.01)

(21) International Application Number:
PCT/US2019/027900

(22) International Filing Date:
17 April 2019 (17.04.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/659,337 18 April 2018 (18.04.2018) US

(71) Applicant: **SNAP INC.** [US/US]; 2772 Donald Douglas Loop North, Santa Monica, California 90405 (US).

(72) Inventors: **CAO, Chen**; 2772 Donald Douglas Loop North, Santa Monica, California 90405 (US). **GAO, Yang**; 2772 Donald Douglas Loop North, Santa Monica, California 90405 (US). **XUE, Zehao**; 2772 Donald Douglas Loop North, Santa Monica, California 90405 (US).

(74) Agent: **PERDOK, Monique M.** et al.; Schwegman Lundberg & Woessner P.A., P.O. Box 2938, Minneapolis, Minnesota 55402 (US).

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

(54) Title: AUGMENTED EXPRESSION SYSTEM

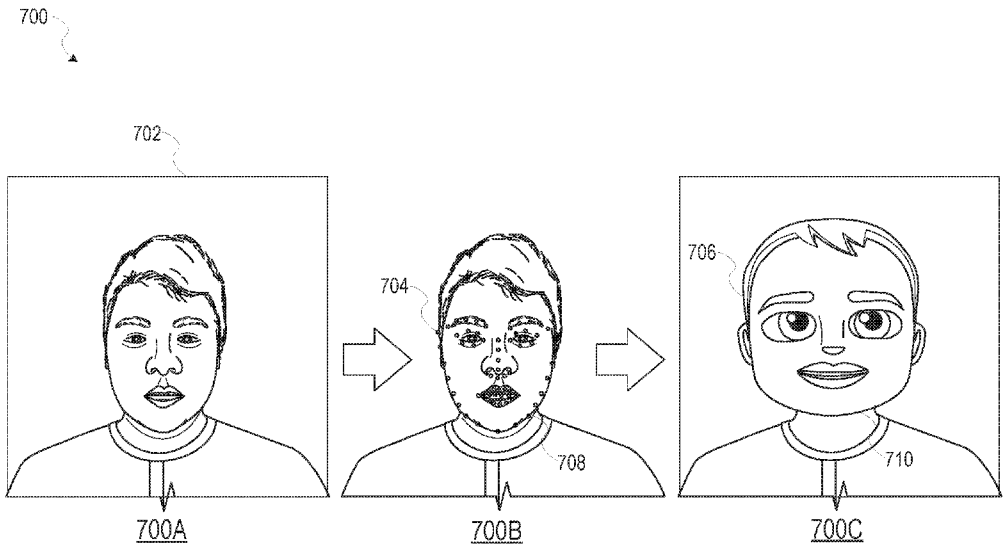


FIG. 7

(57) Abstract: The present invention relates to improvements to systems and methods for determining a current location of a client device, and for identifying and selecting appropriate geo-fences based on the current location of the client device. An improved geo-fence selection system performs operations that include associating media content with a geo-fence that encompasses a portion of a geographic region, sampling location data from a client device, defining a boundary based on the sampled location data from the client device, detecting an overlap between the boundary and the geo-fence, retrieving the media content associated with the geo-fence, and loading the media content at a memory location of the client device, in response to detecting the overlap.

WO 2019/204464 A1

[Continued on next page]

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))*

AUGMENTED EXPRESSION SYSTEM

PRIORITY

[0001] This application claims the benefit of priority to U.S. Provisional Patent Application Serial No. 62/659,337, filed on April 18, 2018, the benefit of priority which is claimed hereby, and which is incorporated by reference herein in its entirety.

TECHNICAL FIELD

[0002] Embodiments of the present disclosure relate generally to mobile computing technology and, more particularly, but not by way of limitation, to systems for generating and displaying augmented reality interfaces.

BACKGROUND

[0003] Augmented reality (AR), is a live direct or indirect view of a physical, real-world environment whose elements are augmented by computer-generated sensory inputs.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0004] To easily identify the discussion of any particular element or act, the most significant digit or digits in a reference number refer to the figure number in which that element is first introduced.

[0005] FIG. 1 is a block diagram showing an example messaging system for exchanging data (e.g., messages and associated content) over a network in accordance with some embodiments, wherein the messaging system includes an augmented expression system.

[0006] FIG. 2 is block diagram illustrating further details regarding a messaging system, according to example embodiments.

[0007] FIG. 3 is a block diagram illustrating various modules of an augmented expression system, according to certain example embodiments.

[0008] FIG. 4 is a flowchart illustrating a method for presenting an augmented reality display, according to certain example embodiments.

[0009] FIG. 5 is a flowchart illustrating a method for presenting an augmented reality display, according to certain example embodiments.

[0010] FIG. 6 is a flowchart illustrating a method for presenting an augmented reality display, according to certain example embodiments.

[0011] FIG. 7 is a flow diagram illustrating a method for generating a 3D model based on facial tracking inputs, according to certain example embodiments.

[0012] FIG. 8 is an illustration of an interface to display an augmented reality image, according to certain example embodiments.

[0013] FIG. 9 is an illustration of an interface to display an augmented reality image, according to certain example embodiments.

[0014] FIG. 10 is an illustration of an interface to display an augmented reality image, according to certain example embodiments.

[0015] FIG. 11 is a block diagram illustrating a representative software architecture, which may be used in conjunction with various hardware architectures herein described and used to implement various embodiments.

[0016] FIG. 12 is a block diagram illustrating components of a machine, according to some example embodiments, able to read instructions from a machine-readable medium (e.g., a machine-readable storage medium) and perform any one or more of the methodologies discussed herein.

DETAILED DESCRIPTION

[0017] Embodiments described herein relate to an augmented expression system to generate and cause display of a specially configured interface to present an augmented reality perspective. The augmented expression system receives image and video data of a user and tracks facial landmarks of the user based on the image and video data, in real-time to generate and present a 3-dimensional (3D) bitmoji of the user. The system presents an augmented reality display that includes a depiction of the user that includes an augmented reality 3D bitmoji head and elements, which are overlaid at locations in the presentation of the image data. For example, a user may

display, stream, or record a video that depicts themselves and their surroundings. The system detects the presence of facial landmarks, and in response, "replaces" the user's head with a 3D bitmoji head which mimics and tracks facial expressions movements of the user.

[0018] The augmented expression system receives image and video data from a camera associated with a client device (e.g., a front facing camera), wherein the image and video data includes a set of facial landmarks of a user. In response to detecting the presence of the set of facial landmarks, the system generates a 3D model (e.g., a 3D Bitmoji) based on the set of facial landmarks, overlays the 3D model at a position over the set of facial landmarks in a presentation of the image data, and dynamically animates the 3D model based on detected movements of the set of facial landmarks, such as by facial tracking.

[0019] In some embodiments, the system identifies a user or users based on the set of facial landmarks depicted by the image data, and generates the 3D model based on a user profile associated with the user. For example, a user may define display characteristics of a 3D bitmoji in a user profile by selecting a particular 3D bitmoji from among a set of 3D bitmoji, or may configure a 3D bitmoji based on selections of individual components of the 3D bitmoji, such as facial features, colors, hair styles, eye color, as well as accessories (e.g., glasses, hats, sunglasses, monocles). In further embodiments, the system generates the 3D bitmoji based on attributes of the set of facial landmarks of the user. For example, the attributes may include distances and slopes between various facial landmarks, size and shape of facial landmarks, as well as a complexion of the user. The system selects individual components of the 3D bitmoji based on the attributes of the facial landmarks.

[0020] The augmented expression system identifies a facial expression of the user based on facial landmark tracking techniques applied to the set of facial landmarks depicted by the image data. The augmented expression system receives inputs as movements and positions of the facial landmarks relative to one another via a facial tracking module. For example, the facial

tracking module receives inputs that include a landmark identifier (e.g., nose, mouth, left eye, right eye), an input value that indicates how much the facial landmark has moved or changed relative to a static position, and a movement direction that indicates a direction of the movement. Based on the inputs, and combinations of the inputs, the facial tracking module of the augmented expression system determines an expression of the user. For example, the system may track the user's eye brows simultaneously raise, while the user's mouth opens, and in response determine that the movements and positions of the facial landmarks correspond to a "surprised" expression.

[0021] The augmented expression system includes a data repository that includes expression definitions, wherein the expression definitions comprise a mapping of various facial tracking inputs to expressions. The expressions comprise display instructions for the 3D bitmoji, beyond simply causing the 3D bitmoji to mimic the movements and positions of the set of facial landmarks of the user. For example, display instructions may include a 3D graphical element (e.g., a balloon, confetti, hearts) to be presented within the presentation of the image data in response to identifying a particular expression, or by exaggerating one or more facial landmarks to emphasize an expression (e.g., drop jaw, eyes pop out of head, eyes change to stylized "X"). The user may additionally be able to interact with the 3D graphical elements by providing further inputs as movements of facial landmarks (e.g., winking and eye, puckering or blowing, etc.), or as tactile inputs into a touch enabled device.

[0022] The augmented expression system facilitates the sharing and distribution of content that includes the 3D bitmoji. For example, a user may stream or record a video that comprises a composite presentation that includes the 3D bitmoji and video. In some embodiments, a user may select a previously recorded image or video and cause the augmented expression system to augment the previously recorded image or video with a display of the 3D bitmoji that mimics the facial movements and expressions of the user. In further embodiments, the user may record or stream a composite presentation that includes a display of the 3D bitmoji in

real-time, such that the 3D bitmoji is rendered onto an image or video as the image data is received at the client device.

[0023] A user may additionally share or distribute the composite presentation that includes the 3D bitmoji by assigning the 3D bitmoji and the associated image data to a message to be distributed to one or more recipients. In response to assigning the 3D bitmoji to the message, the system generates a flattened presentation of the composite presentation, based on the image data and the 3D bitmoji. The user may assign or transmit the flattened presentation to one or more recipients (e.g., as a message or by adding the flattened presentation to a story associated with a user profile of the user).

[0024] In further embodiments, the system may transmit the image data that depicts the user and the 3D bitmoji separately to one or more recipients, and cause the recipients to generate a presentation of the image data that comprises a display of the 3D bitmoji at a position within the image data based on the set of facial landmarks. For example, the system may segment the 3D bitmoji into a set of regions, wherein each region corresponds to a distinct facial landmark from among the set of facial landmarks. Upon receiving the 3D bitmoji and the image data that comprises the set of facial landmarks at the recipient devices, the system causes the recipient devices to generate a presentation of the image data by presenting the 3D bitmoji at a location in the image data based on the locations of each of the facial landmarks.

[0025] In some embodiments, the user may transmit 3D bitmoji instructions that enable recipients of the 3D bitmoji instructions to display the 3D bitmoji on content created at their personal devices. For example, a user may configure a 3D bitmoji by selecting one or more bitmoji components and attributes (e.g., head shape, color, accessories, etc.). The user assigns an identifier to the 3D bitmoji (e.g., "Angry Putin"), and the augmented expression system indexes and stores the 3D bitmoji instructions and the identifier at a memory location associated with the user profile of the user. Recipients of the 3D bitmoji instructions may generate

augmented reality presentations that include a display of the 3D bitmoji (e.g., "Angry Putin").

[0026] Figure 1 is a block diagram showing an example messaging system 100 for exchanging data (e.g., messages and associated content) over a network. The messaging system 100 includes multiple client devices 102, each of which hosts a number of applications including a messaging client application 104. Each messaging client application 104 is communicatively coupled to other instances of the messaging client application 104 and a messaging server system 108 via a network 106 (e.g., the Internet).

[0027] Accordingly, each messaging client application 104 is able to communicate and exchange data with another messaging client application 104 and with the messaging server system 108 via the network 106. The data exchanged between messaging client applications 104, and between a messaging client application 104 and the messaging server system 108, includes functions (e.g., commands to invoke functions) as well as payload data (e.g., text, audio, video or other multimedia data).

[0028] The messaging server system 108 provides server-side functionality via the network 106 to a particular messaging client application 104. While certain functions of the messaging system 100 are described herein as being performed by either a messaging client application 104 or by the messaging server system 108, it will be appreciated that the location of certain functionality either within the messaging client application 104 or the messaging server system 108 is a design choice. For example, it may be technically preferable to initially deploy certain technology and functionality within the messaging server system 108, but to later migrate this technology and functionality to the messaging client application 104 where a client device 102 has a sufficient processing capacity.

[0029] The messaging server system 108 supports various services and operations that are provided to the messaging client application 104. Such operations include transmitting data to, receiving data from, and processing data generated by the messaging client application 104. In some embodiments, this data includes, message content, client device

information, geolocation information, media annotation and overlays, message content persistence conditions, social network information, and live event information, as examples. In other embodiments, other data is used. Data exchanges within the messaging system 100 are invoked and controlled through functions available via GUIs of the messaging client application 104.

[0030] Turning now specifically to the messaging server system 108, an Application Program Interface (API) server 110 is coupled to, and provides a programmatic interface to, an application server 112. The application server 112 is communicatively coupled to a database server 118, which facilitates access to a database 120 in which is stored data associated with messages processed by the application server 112.

[0031] Dealing specifically with the Application Program Interface (API) server 110, this server receives and transmits message data (e.g., commands and message payloads) between the client device 102 and the application server 112. Specifically, the Application Program Interface (API) server 110 provides a set of interfaces (e.g., routines and protocols) that can be called or queried by the messaging client application 104 in order to invoke functionality of the application server 112. The Application Program Interface (API) server 110 exposes various functions supported by the application server 112, including account registration, login functionality, the sending of messages, via the application server 112, from a particular messaging client application 104 to another messaging client application 104, the sending of media files (e.g., images or video) from a messaging client application 104 to the messaging server application 114, and for possible access by another messaging client application 104, the setting of a collection of media data (e.g., story), the retrieval of a list of friends of a user of a client device 102, the retrieval of such collections, the retrieval of messages and content, the adding and deletion of friends to a social graph, the location of friends within a social graph, opening and application event (e.g., relating to the messaging client application 104).

[0032] The application server 112 hosts a number of applications and subsystems, including a messaging server application 114, an image processing system 116, a social network system 122, and a augmented expression system 124. The messaging server application 114 implements a number of message processing technologies and functions, particularly related to the aggregation and other processing of content (e.g., textual and multimedia content) included in messages received from multiple instances of the messaging client application 104. As will be described in further detail, the text and media content from multiple sources may be aggregated into collections of content (e.g., called stories or galleries). These collections are then made available, by the messaging server application 114, to the messaging client application 104. Other processor and memory intensive processing of data may also be performed server-side by the messaging server application 114, in view of the hardware requirements for such processing.

[0033] The application server 112 also includes an image processing system 116 that is dedicated to performing various image processing operations, typically with respect to images or video received within the payload of a message at the messaging server application 114.

[0034] The social network system 122 supports various social networking functions services, and makes these functions and services available to the messaging server application 114. To this end, the social network system 122 maintains and accesses an entity graph 304 within the database 120. Examples of functions and services supported by the social network system 122 include the identification of other users of the messaging system 100 with which a particular user has relationships or is "following," and also the identification of other entities and interests of a particular user.

[0035] The application server 112 is communicatively coupled to a database server 118, which facilitates access to a database 120 in which is stored data associated with messages processed by the messaging server application 114.

[0036] Figure 2 is block diagram illustrating further details regarding the messaging system 100, according to example embodiments. Specifically, the messaging system 100 is shown to comprise the messaging client application 104 and the application server 112, which in turn embody a number of some subsystems, namely an ephemeral timer system 202, a collection management system 204 and an annotation system 206.

[0037] The ephemeral timer system 202 is responsible for enforcing the temporary access to content, such as 3D bitmoji, permitted by the messaging client application 104 and the messaging server application 114. To this end, the ephemeral timer system 202 incorporates a number of timers that, based on duration and display parameters associated with a message, collection of messages, or graphical element, selectively display and enable access to messages and associated content via the messaging client application 104. Further details regarding the operation of the ephemeral timer system 202 are provided below.

[0038] The collection management system 204 is responsible for managing collections of media (e.g., collections of text, image video and audio data, 3D bitmoji, flattened presentations of 3D bitmoji). In some examples, a collection of content (e.g., messages, including images, video, text and audio) may be organized into an "event gallery" or an "event story." Such a collection may be made available for a specified time period, such as the duration of an event to which the content relates. For example, content relating to a music concert may be made available as a "story" for the duration of that music concert. The collection management system 204 may also be responsible for publishing an icon that provides notification of the existence of a particular collection to the user interface of the messaging client application 104.

[0039] The collection management system 204 furthermore includes a curation interface 208 that allows a collection manager to manage and curate a particular collection of content. For example, the curation interface 208 enables an event organizer to curate a collection of content relating to a specific event (e.g., delete inappropriate content or redundant messages).

Additionally, the collection management system 204 employs machine vision (or image recognition technology) and content rules to automatically curate a content collection. In certain embodiments, compensation may be paid to a user for inclusion of user generated content into a collection. In such cases, the curation interface 208 operates to automatically make payments to such users for the use of their content.

[0040] The annotation system 206 provides various functions that enable a user to annotate or otherwise modify or edit media content associated with a message. For example, the annotation system 206 provides functions related to the generation and publishing of media overlays for messages processed by the messaging system 100. The annotation system 206 operatively supplies a media overlay to the messaging client application 104 based on a geolocation of the client device 102. In another example, the annotation system 206 operatively supplies a media overlay to the messaging client application 104 based on other information, such as, social network information of the user of the client device 102. A media overlay may include audio and visual content and visual effects. Examples of audio and visual content include pictures, texts, logos, animations, and sound effects, as well as animated facial models, such as those generated by the augmented expression system 124. An example of a visual effect includes color overlaying. The audio and visual content or the visual effects can be applied to a media content item (e.g., a photo) at the client device 102. For example, the media overlay including text that can be overlaid on top of a photograph generated taken by the client device 102. In another example, the media overlay includes an identification of a location overlay (e.g., Venice beach), a name of a live event, or a name of a merchant overlay (e.g., Beach Coffee House). In another example, the annotation system 206 uses the geolocation of the client device 102 to identify a media overlay that includes the name of a merchant at the geolocation of the client device 102. The media overlay may include other indicia associated with the merchant. The media overlays may be stored in the database 120 and accessed through the database server 118.

[0041] In one example embodiment, the annotation system 206 provides a user-based publication platform that enables users to select a geolocation on a map, and upload content associated with the selected geolocation. The user may also specify circumstances under which a particular media overlay should be offered to other users. The annotation system 206 generates a media overlay that includes the uploaded content and associates the uploaded content with the selected geolocation.

[0042] In another example embodiment, the annotation system 206 provides a merchant-based publication platform that enables merchants to select a particular media overlay associated with a geolocation via a bidding process. For example, the annotation system 206 associates the media overlay of a highest bidding merchant with a corresponding geolocation for a predefined amount of time

[0043] Figure 3 is a block diagram illustrating components of the augmented expression system 124 that configure the augmented expression system 124 to generate and display an augmented reality presentation based on image data and a 3D bitmoji, according to some example embodiments. The augmented expression system 124 is shown as including a feature tracking module 302, a modeling module 304, a presentation module 306, and a communication module 308, all configured to communicate with each other (e.g., via a bus, shared memory, or a switch). Any one or more of these modules may be implemented using one or more processors 310 (e.g., by configuring such one or more processors to perform functions described for that module) and hence may include one or more of the processors 310.

[0044] Any one or more of the modules described may be implemented using hardware alone (e.g., one or more of the processors 310 of a machine) or a combination of hardware and software. For example, any module described of the augmented expression system 124 may physically include an arrangement of one or more of the processors 310 (e.g., a subset of or among the one or more processors of the machine) configured to perform the operations described herein for that module. As another

example, any module of the augmented expression system 124 may include software, hardware, or both, that configure an arrangement of one or more processors 310 (e.g., among the one or more processors of the machine) to perform the operations described herein for that module. Accordingly, different modules of the augmented expression system 124 may include and configure different arrangements of such processors 310 or a single arrangement of such processors 310 at different points in time. Moreover, any two or more modules of the augmented expression system 124 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.

[0045] Figure 4 is a flowchart illustrating a method 400 for presenting an augmented reality display, according to certain example embodiments. Operations of the method 400 may be performed by the modules described above with respect to Figure 3. As shown in figure 4, the method 400 includes one or more operations 402, 404, 406, 408, and 410.

[0046] At operation 402, the presentation module 306 causes display of a presentation of image data at a client device (e.g., client device 102A), wherein the image data depicts a set of facial landmarks of a user. For example, the image data may be collected by a camera associated with the client device 102A, such as a front facing camera.

[0047] At operation 404, the modeling module 304 generates a 3D model (e.g., a 3D bitmoji) based on the set of facial landmarks, in response to the presentation module 306 displaying the presentation of the image data. In some embodiments, the modeling module 304 generates the 3D model in response to detecting the presence of facial landmarks within the image data. For example, the facial tracking module 302 may employ various facial motion capture methodologies including marker-based, as well as markerless technologies. Markerless facial motion capture methodologies identify and track facial movements and expressions based on facial

landmarks, such as nostrils, corners of the lips and eyes, pupils, and wrinkles. In response to detecting the presence of such facial landmarks within the image data, the facial tracking module 302 causes the modeling module to generate the 3D model based on the identifies facial landmarks.

[0048] In some embodiments, the facial tracking module 302 identifies the user based on attributes of the identified set of facial landmarks, such as distances and slopes between each facial landmark, a size and shape of each facial landmark, as well as device attributes of the client device 102A (e.g., a device identifier), and retrieves a user profile associated with the user, wherein the user profile include display instructions for a 3D bitmoji.

[0049] In further embodiments, the modeling module 304 generates a 3D bitmoji on the fly, based on the attributes of the facial landmarks, such that the 3D bitmoji resembles the user. For example, the modeling module 304 may access a bitmoji repository that contains bitmoji elements, select a set of bitmoji elements based on the attributes of the facial landmarks, and generate the 3D bitmoji based on the selected bitmoji elements.

[0050] In some embodiments, the 3D bitmoji generated by the modeling module 304 comprises a set of bitmoji regions, wherein each region corresponds to a distinct facial landmark from among the set of facial landmarks (e.g., a region corresponding to the facial landmark that represents the user's mouth, a region corresponding to the facial landmark that represents the user's left eye, etc.). In further embodiments, the 3D bitmoji generated by the modeling module 304 comprises a blendshape that represents the set of facial landmarks as a series of vertex positions.

[0051] At operation 406, the presentation module 306 overlays the 3D bitmoji generated by the modeling module 304 at a position on the image data displayed at the client device 102A based on the locations of the set of facial landmarks. For example, the facial tracking module 302 may identify a set of reference features based on the set of facial landmarks and cause the presentation module 306 to overlay a presentation of the 3D bitmoji on the image data at the client device 102A.

[0052] At operation 408, the feature tracking module 302 receive a facial tracking input (e.g., a movement of a facial landmark from among the set of facial landmarks). For example, the user may open or close their mouth, raise an eyebrow, or turn their head. In response to the feature tracking module 302 receiving the facial tracking input, at operation 410 the presentation module 302 dynamically animates a region of the 3D bitmoji that corresponds to the facial landmark at the client device 102A. The 3D bitmoji displayed at the client device 102A mimics the user's movements and facial expressions.

[0053] Figure 5 is a flowchart illustrating a method 500 for presenting an augmented reality display, according to certain example embodiments. Operations of the method 500 may be performed by the modules described above with respect to Figure 3. As shown in figure 5, the method 500 includes one or more operations 502, 504, 506, 508, and 510.

[0054] At operation 502, the facial tracking module 302 identifies an expression (e.g., a facial expression) of the user based on inputs received as movements and positions of the set of facial landmarks. For example, the facial tracking module 302 receives inputs that include a landmark identifier (e.g., nose, mouth, left eye, right eye), an input value that indicates how much the facial landmark has moved or changed relative to a static position, and a movement direction that indicates a direction of the movement. In response to receiving the inputs, the facial tracking module 302 accesses a repository that includes a set of expression definitions, wherein the expression definitions comprise a mapping of the various facial tracking inputs to predefined expressions.

[0055] In some embodiments, a user of the client device 102A explicitly provides the expression definitions via one or more user inputs. For example, a user may form their face into a particular expression and capture an image that depicts the expression at the client device 102A. The facial tracking module 302 determines relative positions of each facial landmark among the set of facial landmarks in the image, and assign the relative positions of the facial landmarks to a particular expression, which the user

may assign an expression identifier as well as display instructions to. For example, the user may indicate that in response to detecting a particular expression based on facial tracking inputs, the augmented expression system 124 displays a particular graphical element within the presentation of the image data.

[0056] At operation 504, the modeling module 304 selects a graphical element from among a set of graphical elements based on the expression identified by the facial tracking module 302. For example, responsive to the facial tracking module 302 identifying an expression, the modeling module 304 retrieves the corresponding expression definition which includes instructions to display a particular graphical element, such as a 3D object, along with the presentation of the image data at the client device 102A.

[0057] At operation 506, the presentation module 306 causes display of the graphical element within the presentation of the image data at the client device 102A. In some embodiments, a location and orientation of the graphical element is based on the expression of the user. For example, each graphical element may include display instructions that define how and where in the presentation of the image data the graphical element is to be displayed, based on the locations, movements, and orientations of the set of facial landmarks tracked by the facial tracking module 302.

[0058] At operation 508, the facial tracking module 302 receives a second facial tracking input (e.g., a movement distinct from the previous movement). At operation 510 the presentation module 306 applies a change to the graphical element based on attributes of the second facial tracking input, such as an input value.

[0059] For example, the graphical element displayed within the presentation of the image data may include a 3D balloon displayed proximate to a region of the 3D bitmoji that represents the user's lips. The user may pucker their lips to indicate that they are blowing up the balloon. In response to detecting the movement of the user's lips, the presentation module 306 causes the 3D balloon to inflate as if the user was inflating it.

[0060] Figure 6 is a flowchart illustrating a method 600 for presenting an augmented reality display, according to certain example embodiments. Operations of the method 600 may be performed by the modules described above with respect to Figure 3. As shown in figure 6, the method 600 includes one or more operations 602, 604, 606, and 608.

[0061] At operation 602, the communication module 308 generates a flattened presentation based on the image data and the 3D bitmoji. The flattened presentation comprises a merging of the image data and the 3D bitmoji. After merging the image data and the 3D bitmoji into a single layer, at operation 604 the user of the client device 102A assigns the flattened presentation to a message. For example, the message may be addressed to one or more recipients, including a second client device 102B.

[0062] At operation 606, the communication module 308 distributes the message that includes the flattened presentation to the recipients including the second client device 102B. At operation 608, the presentation module 306 causes display of the flattened presentation at the second client device 102B, for example as an ephemeral message.

[0063] Figure 7 is a flow diagram 700 illustrating a method for generating a 3D model based on facial tracking inputs, according to certain example embodiments. As seen in figure 7, the flow diagram 700 comprises steps 700A, 700B, and 700C.

[0064] At step 700A, image data 702 is received at a client device 102A, as described at operation 402 of the method 400, in figure 4. For example, the image data 702 may be collected by a camera associated by the client device 102A, such as a front facing camera.

[0065] At step 700B, the facial tracking module 302 identifies a set of facial landmarks 704 within the image data 702. As seen in figure 7, the set of facial landmarks 704 may be represented as a distribution of points that represent facial landmarks of a user's face depicted in the image data 702. The facial landmarks 704 may for example include the tip of a nose, corners of the eyes, corners of eyebrows, corners of the mouth, and pupils. The distribution of points may be sub-divided into collections of points,

such that each collection of points corresponds to a region of the 3D bitmoji 706. For example, the collection of points 708 that line a lower portion of the face depicted in the image data 702 may correspond with the lower portion 710 of the 3D bitmoji 706.

[0066] At step 700C, the modeling module 304 generate a 3D bitmoji 706 based on the set of facial landmarks 704, as described at operation 404 of the method 400, in figure 4. As seen at step 700C of figure 7, the 3D bitmoji 706 is overlaid upon the image data 702, at a location based on the set of facial landmarks 704.

[0067] Figure 8 is an illustration 800 of an interface to display an augmented reality image, according to certain example embodiments. As seen in figure 8, the illustration 800 includes a depiction of a 3D bitmoji 802, overlaid upon image data 804, according to certain example embodiments such as those described in the method 400 of figure 4.

[0068] Figure 9 is an illustration 900 of an interface to display an augmented reality image, according to certain example embodiments. As seen in figure 9, the illustration 900 includes a depiction of a 3D bitmoji 902 overlaid upon image data 904. The illustration 900 also includes an interface item 906, to receive a user input to record content that includes composite presentations comprising the image data 804 and the 3D bitmoji 902.

[0069] As seen in figure 9, the 3D bitmoji 902 comprises an exaggerated depiction of an expression identified by the augmented expression system 124, as described in the method 500 of figure 5. In response to detecting a particular expression based on one or more facial tracking inputs received at the client device 102A, the augmented expression system 124 causes the 3D bitmoji to present of the exaggerated depiction of the expression.

[0070] Figure 10 is an illustration 1000 of an interface to display an augmented reality image, according to certain example embodiments. As seen in figure 10, the illustration10 includes a 3D bitmoji 1002, a graphical element 1004, overlaid upon image data 1006, according to certain example embodiments such as those described in the method 500 of figure 5.

[0071] For example, in response to the facial tracking module 302 receiving a facial tracking input corresponding to a particular expression (e.g., the user puckering his lips), the presentation module 306 causes display of the graphical element 1004 at a position in the image data 1006.

SOFTWARE ARCHITECTURE

[0072] Figure 11 is a block diagram illustrating an example software architecture 1106, which may be used in conjunction with various hardware architectures herein described. Figure 11 is a non-limiting example of a software architecture and it will be appreciated that many other architectures may be implemented to facilitate the functionality described herein. The software architecture 1106 may execute on hardware such as machine 1200 of Figure 12 that includes, among other things, processors 1204, memory 1214, and I/O components 1218. A representative hardware layer 1152 is illustrated and can represent, for example, the machine 1100 of Figure 11. The representative hardware layer 1152 includes a processing unit 1154 having associated executable instructions 1104. Executable instructions 1104 represent the executable instructions of the software architecture 1106, including implementation of the methods, components and so forth described herein. The hardware layer 1152 also includes memory and/or storage modules memory/storage 1156, which also have executable instructions 1104. The hardware layer 1152 may also comprise other hardware 1158.

[0073] In the example architecture of Figure 11, the software architecture 1106 may be conceptualized as a stack of layers where each layer provides particular functionality. For example, the software architecture 1106 may include layers such as an operating system 1102, libraries 1120, applications 1116 and a presentation layer 1114. Operationally, the applications 1116 and/or other components within the layers may invoke application programming interface (API) API calls 1108 through the software stack and receive a response as in response to the API calls 1108. The layers illustrated are representative in nature and not all software

architectures have all layers. For example, some mobile or special purpose operating systems may not provide a frameworks/middleware 1118, while others may provide such a layer. Other software architectures may include additional or different layers.

[0074] The operating system 1102 may manage hardware resources and provide common services. The operating system 1102 may include, for example, a kernel 1122, services 1124 and drivers 1126. The kernel 1122 may act as an abstraction layer between the hardware and the other software layers. For example, the kernel 1122 may be responsible for memory management, processor management (e.g., scheduling), component management, networking, security settings, and so on. The services 1124 may provide other common services for the other software layers. The drivers 1126 are responsible for controlling or interfacing with the underlying hardware. For instance, the drivers 1126 include display drivers, camera drivers, Bluetooth® drivers, flash memory drivers, serial communication drivers (e.g., Universal Serial Bus (USB) drivers), Wi-Fi® drivers, audio drivers, power management drivers, and so forth depending on the hardware configuration.

[0075] The libraries 1120 provide a common infrastructure that is used by the applications 1116 and/or other components and/or layers. The libraries 1120 provide functionality that allows other software components to perform tasks in an easier fashion than to interface directly with the underlying operating system 1102 functionality (e.g., kernel 1122, services 1124 and/or drivers 1126). The libraries 1120 may include system libraries 1144 (e.g., C standard library) that may provide functions such as memory allocation functions, string manipulation functions, mathematical functions, and the like. In addition, the libraries 1120 may include API libraries 1146 such as media libraries (e.g., libraries to support presentation and manipulation of various media format such as MPREG4, H.264, MP3, AAC, AMR, JPG, PNG), graphics libraries (e.g., an OpenGL framework that may be used to render 2D and 3D in a graphic content on a display), database libraries (e.g., SQLite that may provide various relational database functions), web libraries (e.g., WebKit that may provide web browsing

functionality), and the like. The libraries 1120 may also include a wide variety of other libraries 1148 to provide many other APIs to the applications 1116 and other software components/modules.

[0076] The frameworks/middleware 1118 (also sometimes referred to as middleware) provide a higher-level common infrastructure that may be used by the applications 1116 and/or other software components/modules. For example, the frameworks/middleware 1118 may provide various graphic user interface (GUI) functions, high-level resource management, high-level location services, and so forth. The frameworks/middleware 1118 may provide a broad spectrum of other APIs that may be utilized by the applications 1116 and/or other software components/modules, some of which may be specific to a particular operating system 1102 or platform.

[0077] The applications 1116 include built-in applications 1138 and/or third-party applications 1140. Examples of representative built-in applications 1138 may include, but are not limited to, a contacts application, a browser application, a book reader application, a location application, a media application, a messaging application, and/or a game application. Third-party applications 1140 may include an application developed using the ANDROID™ or IOS™ software development kit (SDK) by an entity other than the vendor of the particular platform, and may be mobile software running on a mobile operating system such as IOS™, ANDROID™, WINDOWS® Phone, or other mobile operating systems. The third-party applications 1140 may invoke the API calls 1108 provided by the mobile operating system (such as operating system 1102) to facilitate functionality described herein.

[0078] The applications 1116 may use built in operating system functions (e.g., kernel 1122, services 1124 and/or drivers 1126), libraries 1120, and frameworks/middleware 1118 to create user interfaces to interact with users of the system. Alternatively, or additionally, in some systems interactions with a user may occur through a presentation layer, such as presentation layer 1114. In these systems, the application/component "logic" can be

separated from the aspects of the application/component that interact with a user.

[0079] Figure 12 is a block diagram illustrating components of a machine 1200, according to some example embodiments, able to read instructions from a machine-readable medium (e.g., a machine-readable storage medium) and perform any one or more of the methodologies discussed herein. Specifically, Figure 12 shows a diagrammatic representation of the machine 1200 in the example form of a computer system, within which instructions 1210(e.g., software, a program, an application, an applet, an app, or other executable code) for causing the machine 1200 to perform any one or more of the methodologies discussed herein may be executed. As such, the instructions 1210 may be used to implement modules or components described herein. The instructions 1210 transform the general, non-programmed machine 1200 into a particular machine 1200 programmed to carry out the described and illustrated functions in the manner described. In alternative embodiments, the machine 1200 operates as a standalone device or may be coupled (e.g., networked) to other machines. In a networked deployment, the machine 1200 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 peer-to-peer (or distributed) network environment. The machine 1200 may comprise, but not be limited to, 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), an entertainment media system, a cellular telephone, a smart phone, a mobile device, a wearable device (e.g., a smart watch), a smart home device (e.g., a smart appliance), other smart devices, a web appliance, a network router, a network switch, a network bridge, or any machine capable of executing the instructions 1210, sequentially or otherwise, that specify actions to be taken by machine 1200. Further, while only a single machine 1200 is illustrated, the term "machine" shall also be taken to include a collection of machines that individually or jointly execute the instructions 1210 to perform any one or more of the methodologies discussed herein.

[0080] The machine 1200 may include processors 1204, memory memory/storage 1206, and I/O components 1218, which may be configured to communicate with each other such as via a bus 1202. The memory/storage 1206 may include a memory 1214, such as a main memory, or other memory storage, and a storage unit 1216, both accessible to the processors 1204 such as via the bus 1202. The storage unit 1216 and memory 1214 store the instructions 1210 embodying any one or more of the methodologies or functions described herein. The instructions 1210 may also reside, completely or partially, within the memory 1214, within the storage unit 1216, within at least one of the processors 1204 (e.g., within the processor's cache memory), or any suitable combination thereof, during execution thereof by the machine 1200. Accordingly, the memory 1214, the storage unit 1216, and the memory of processors 1204 are examples of machine-readable media.

[0081] The I/O components 1218 may include a wide variety of components to receive input, provide output, produce output, transmit information, exchange information, capture measurements, and so on. The specific I/O components 1218 that are included in a particular machine 1200 will depend on the type of machine. For example, portable machines such as mobile phones will likely include a touch input device or other such input mechanisms, while a headless server machine will likely not include such a touch input device. It will be appreciated that the I/O components 1218 may include many other components that are not shown in Figure 12. The I/O components 1218 are grouped according to functionality merely for simplifying the following discussion and the grouping is in no way limiting. In various example embodiments, the I/O components 1218 may include output components 1226 and input components 1228. The output components 1226 may include visual components (e.g., a display such as a plasma display panel (PDP), a light emitting diode (LED) display, a liquid crystal display (LCD), a projector, or a cathode ray tube (CRT)), acoustic components (e.g., speakers), haptic components (e.g., a vibratory motor, resistance mechanisms), other signal generators, and so forth. The input components 1228 may include alphanumeric input components (e.g., a

keyboard, a touch screen configured to receive alphanumeric input, a photo-optical keyboard, or other alphanumeric input components), point based input components (e.g., a mouse, a touchpad, a trackball, a joystick, a motion sensor, or other pointing instrument), tactile input components (e.g., a physical button, a touch screen that provides location and/or force of touches or touch gestures, or other tactile input components), audio input components (e.g., a microphone), and the like.

[0082] In further example embodiments, the I/O components 1218 may include biometric components 1230, motion components 1234, environmental environment components 1236, or position components 1238 among a wide array of other components. For example, the biometric components 1230 may include components to detect expressions (e.g., hand expressions, facial expressions, vocal expressions, body gestures, or eye tracking), measure biosignals (e.g., blood pressure, heart rate, body temperature, perspiration, or brain waves), identify a person (e.g., voice identification, retinal identification, facial identification, fingerprint identification, or electroencephalogram based identification), and the like. The motion components 1234 may include acceleration sensor components (e.g., accelerometer), gravitation sensor components, rotation sensor components (e.g., gyroscope), and so forth. The environment components 1236 may include, for example, illumination sensor components (e.g., photometer), temperature sensor components (e.g., one or more thermometer that detect ambient temperature), humidity sensor components, pressure sensor components (e.g., barometer), acoustic sensor components (e.g., one or more microphones that detect background noise), proximity sensor components (e.g., infrared sensors that detect nearby objects), gas sensors (e.g., gas detection sensors to detection concentrations of hazardous gases for safety or to measure pollutants in the atmosphere), or other components that may provide indications, measurements, or signals corresponding to a surrounding physical environment. The position components 1238 may include location sensor components (e.g., a Global Position system (GPS) receiver component), altitude sensor components (e.g., altimeters or barometers that detect air pressure from which altitude

may be derived), orientation sensor components (e.g., magnetometers), and the like.

[0083] Communication may be implemented using a wide variety of technologies. The I/O components 1218 may include communication components 1240 operable to couple the machine 1200 to a network 1232 or devices 1220 via coupling 1222 and coupling 1224 respectively. For example, the communication components 1240 may include a network interface component or other suitable device to interface with the network 1232. In further examples, communication components 1240 may include wired communication components, wireless communication components, cellular communication components, Near Field Communication (NFC) components, Bluetooth® components (e.g., Bluetooth® Low Energy), Wi-Fi® components, and other communication components to provide communication via other modalities. The devices 1220 may be another machine or any of a wide variety of peripheral devices (e.g., a peripheral device coupled via a Universal Serial Bus (USB)).

[0084] Moreover, the communication components 1240 may detect identifiers or include components operable to detect identifiers. For example, the communication components 1240 may include Radio Frequency Identification (RFID) tag reader components, NFC smart tag detection components, optical reader components (e.g., an optical sensor to detect one-dimensional bar codes such as Universal Product Code (UPC) bar code, multi-dimensional bar codes such as Quick Response (QR) code, Aztec code, Data Matrix, Dataglyph, MaxiCode, PDF417, Ultra Code, UCC RSS-2D bar code, and other optical codes), or acoustic detection components (e.g., microphones to identify tagged audio signals). In addition, a variety of information may be derived via the communication components 1240, such as, location via Internet Protocol (IP) geo-location, location via Wi-Fi® signal triangulation, location via detecting a NFC beacon signal that may indicate a particular location, and so forth.

GLOSSARY

[0085] "CARRIER SIGNAL" in this context refers to any intangible medium that is capable of storing, encoding, or carrying instructions for execution by the machine, and includes digital or analog communications signals or other intangible medium to facilitate communication of such instructions. Instructions may be transmitted or received over the network using a transmission medium via a network interface device and using any one of a number of well-known transfer protocols.

[0086] "CLIENT DEVICE" in this context refers to any machine that interfaces to a communications network to obtain resources from one or more server systems or other client devices. A client device may be, but is not limited to, a mobile phone, desktop computer, laptop, portable digital assistants (PDAs), smart phones, tablets, ultra books, netbooks, laptops, multi-processor systems, microprocessor-based or programmable consumer electronics, game consoles, set-top boxes, or any other communication device that a user may use to access a network.

[0087] "COMMUNICATIONS NETWORK" in this context refers to one or more portions of a network that may be an ad hoc network, an intranet, an extranet, a virtual private network (VPN), a local area network (LAN), a wireless LAN (WLAN), a wide area network (WAN), a wireless WAN (WWAN), a metropolitan area network (MAN), the Internet, a portion of the Internet, a portion of the Public Switched Telephone Network (PSTN), a plain old telephone service (POTS) network, a cellular telephone network, a wireless network, a Wi-Fi® network, another type of network, or a combination of two or more such networks. For example, a network or a portion of a network may include a wireless or cellular network and the coupling may be a Code Division Multiple Access (CDMA) connection, a Global System for Mobile communications (GSM) connection, or other type of cellular or wireless coupling. In this example, the coupling may implement any of a variety of types of data transfer technology, such as Single Carrier Radio Transmission Technology (1xRTT), Evolution-Data Optimized (EVDO) technology, General Packet Radio Service (GPRS) technology, Enhanced Data rates for GSM Evolution (EDGE) technology, third Generation Partnership Project (3GPP) including 3G, fourth

generation wireless (4G) networks, Universal Mobile Telecommunications System (UMTS), High Speed Packet Access (HSPA), Worldwide Interoperability for Microwave Access (WiMAX), Long Term Evolution (LTE) standard, others defined by various standard setting organizations, other long range protocols, or other data transfer technology.

[0088] "EMPHEMERAL MESSAGE" in this context refers to a message that is accessible for a time-limited duration. An ephemeral message may be a text, an image, a video and the like. The access time for the ephemeral message may be set by the message sender. Alternatively, the access time may be a default setting or a setting specified by the recipient. Regardless of the setting technique, the message is transitory.

[0089] "MACHINE-READABLE MEDIUM" in this context refers to a component, device or other tangible media able to store instructions and data temporarily or permanently and may include, but is not be limited to, random-access memory (RAM), read-only memory (ROM), buffer memory, flash memory, optical media, magnetic media, cache memory, other types of storage (e.g., Erasable Programmable Read-Only Memory (EEPROM)) and/or any suitable combination thereof. 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 (e.g., code) for execution by a machine, such that the instructions, when executed by one or more processors of the machine, 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" excludes signals per se.

[0090] "COMPONENT" in this context refers to a device, physical entity or logic having boundaries defined by function or subroutine calls, branch

points, application program interfaces (APIs), or other technologies that provide for the partitioning or modularization of particular processing or control functions. Components may be combined via their interfaces with other components to carry out a machine process. A component may be a packaged functional hardware unit designed for use with other components and a part of a program that usually performs a particular function of related functions. Components may constitute either software components (e.g., code embodied on a machine-readable medium) or hardware components. A "hardware component" 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 components 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 component that operates to perform certain operations as described herein. A hardware component may also be implemented mechanically, electronically, or any suitable combination thereof. For example, a hardware component may include dedicated circuitry or logic that is permanently configured to perform certain operations. A hardware component may be a special-purpose processor, such as a Field-Programmable Gate Array (FPGA) or an Application Specific Integrated Circuit (ASIC). A hardware component may also include programmable logic or circuitry that is temporarily configured by software to perform certain operations. For example, a hardware component may include software executed by a general-purpose processor or other programmable processor. Once configured by such software, hardware components become specific machines (or specific components of a machine) uniquely tailored to perform the configured functions and are no longer general-purpose processors. It will be appreciated that the decision to implement a hardware component 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. Accordingly,

the phrase "hardware component"(or "hardware-implemented component") 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. Considering embodiments in which hardware components are temporarily configured (e.g., programmed), each of the hardware components need not be configured or instantiated at any one instance in time. For example, where a hardware component 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 components) at different times. Software accordingly configures a particular processor or processors, for example, to constitute a particular hardware component at one instance of time and to constitute a different hardware component at a different instance of time. Hardware components can provide information to, and receive information from, other hardware components. Accordingly, the described hardware components may be regarded as being communicatively coupled. Where multiple hardware components 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 components. In embodiments in which multiple hardware components are configured or instantiated at different times, communications between such hardware components may be achieved, for example, through the storage and retrieval of information in memory structures to which the multiple hardware components have access. For example, one hardware component may perform an operation and store the output of that operation in a memory device to which it is communicatively coupled. A further hardware component may then, at a later time, access the memory device to retrieve and process the stored output. Hardware components may also initiate communications with input or output devices, and can operate on a resource (e.g., a collection of information). 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 components that operate to perform one or more operations or functions described herein. As used herein, "processor-implemented component" refers to a hardware component implemented using one or more processors. Similarly, the methods described herein may be at least partially processor-implemented, with a particular processor or processors 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 components. 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)). The performance of certain of the operations may be distributed among the processors, not only residing within a single machine, but deployed across a number of machines. In some example embodiments, the processors or processor-implemented components 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 processors or processor-implemented components may be distributed across a number of geographic locations.

[0091] "PROCESSOR" in this context refers to any circuit or virtual circuit (a physical circuit emulated by logic executing on an actual processor) that manipulates data values according to control signals (e.g., "commands", "op codes", "machine code", etc.) and which produces corresponding output signals that are applied to operate a machine. A processor may, for example, be a Central Processing Unit (CPU), a Reduced Instruction Set Computing (RISC) processor, a Complex Instruction Set Computing (CISC) processor, a Graphics Processing Unit (GPU), a Digital Signal

Processor (DSP), an Application Specific Integrated Circuit (ASIC), a Radio-Frequency Integrated Circuit (RFIC) or any combination thereof. A processor may further be a multi-core processor having two or more independent processors (sometimes referred to as "cores") that may execute instructions contemporaneously.

[0092] "TIMESTAMP" in this context refers to a sequence of characters or encoded information identifying when a certain event occurred, for example giving date and time of day, sometimes accurate to a small fraction of a second.

[0093] "LIFT" in this context is a measure of the performance of a targeted model at predicting or classifying cases as having an enhanced response (with respect to a population as a whole), measured against a random choice targeting model.

[0094] "PHONEME ALIGNMENT" in this context, a phoneme is a unit of speech that differentiates one word from another. One phoneme may consist of a sequence of closure, burst, and aspiration events; or, a dipthong may transition from a back vowel to a front vowel. A speech signal may therefore be described not only by what phonemes it contains, but also the locations of the phonemes. Phoneme alignment may therefore be described as a "time-alignment" of phonemes in a waveform, in order to determine an appropriate sequence and location of each phoneme in a speech signal.

[0095] "AUDIO-TO-VISUAL CONVERSION" in this context refers to the conversion of audible speech signals into visible speech, wherein the visible speech may include a mouth shape representative of the audible speech signal.

[0096] "TIME DELAYED NEURAL NETWORK (TDNN)" in this context, a TDNN is an artificial neural network architecture whose primary purpose is to work on sequential data. An example would be converting continuous audio into a stream of classified phoneme labels for speech recognition.

[0097] "BI-DIRECTIONAL LONG-SHORT TERM MEMORY (BLSTM)" in this context refers to a recurrent neural network (RNN)

architecture that remembers values over arbitrary intervals. Stored values are not modified as learning proceeds. RNNs allow forward and backward connections between neurons. BLSTM are well-suited for the classification, processing, and prediction of time series, given time lags of unknown size and duration between events.

CLAIMS

What is claimed is:

1. A method comprising:

causing display of a presentation of image data at a client device, the image data depicting a set of facial landmarks of a user;

generating a 3D model based on the set of facial landmarks in response to the causing display of the presentation of the image data;

overlaying the 3D model at a position on the presentation of the image data at the client device, the position based on the set of facial landmarks;

receiving a facial tracking input that comprises a movement of a facial landmark from among the set of facial landmarks of the user; and

dynamically animating the 3D model at the client device based on the facial tracking input.

2. The method of claim 1, wherein the method further comprises:

identifying a user profile of the user based on the set of facial landmarks and an identifier of the client device; and

generating the 3D model based on the user profile and the set of facial landmarks.

3. The method of claim 1, wherein the method further comprises:

identifying an expression based on the facial tracking input;

selecting a graphical element from among a set of graphical elements based on the expression; and

causing display of the graphical element within the presentation of the image data at the client device.

4. The method of claim 3, wherein the facial tracking input is a first facial tracking input, and the method further comprises:

receiving a second facial tracking input, the second facial tracking input comprising a second movement of a facial landmark from among the set of facial landmarks; and

applying a change to the graphical element based on the second facial tracking input.

5. The method of claim 1, wherein the method further comprises:

detecting a rotation of a head of the user based on at least the facial tracking input; and

rotating the 3D model based on the rotation of the head of the user.

6. The method of claim 1, wherein the client device is a first client device, and the method further comprises:

generating a flattened presentation based on the 3D model and the image data;

assigning the flattened presentation to a message at the first client device;

distributing the message to a second client device; and

causing display of the flattened presentation ta at the second client device.

7. The method of claim 6, wherein the message includes an ephemeral message.

8. The method of claim 1, wherein the 3D model comprises a set of regions, each region of the 3D model corresponding to a facial landmark from among the set of facial landmarks of the user, and the method further comprises:

identifying a region of the 3D model based on the facial tracking input; and

dynamically animating the region in response to the receiving the facial tracking input.

9. A system comprising:

a memory; and

at least one hardware processor coupled to the memory and comprising instructions that causes the system to perform operations comprising:

causing display of a presentation of image data at a client device, the image data depicting a set of facial landmarks of a user;

generating a 3D model based on the set of facial landmarks in response to the causing display of the presentation of the image data;

overlaying the 3D model at a position on the presentation of the image data at the client device, the position based on the set of facial landmarks;

receiving a facial tracking input that comprises a movement of a facial landmark from among the set of facial landmarks of the user; and

dynamically animating the 3D model at the client device based on the facial tracking input.

10. The system of claim 9, wherein the instructions cause the system to perform operations further comprising:

identifying a user profile of the user based on the set of facial landmarks and an identifier of the client device; and

generating the 3D model based on the user profile and the set of facial landmarks.

11. The system of claim 9, wherein the instructions cause the system to perform operations further comprising:

identifying an expression based on the facial tracking input;

selecting a graphical element from among a set of graphical elements based on the expression; and

causing display of the graphical element within the presentation of the image data at the client device.

12. The system of claim 11, wherein the facial tracking input is a first facial tracking input, and the instructions cause the system to perform operations further comprising:

receiving a second facial tracking input, the second facial tracking input comprising a second movement of a facial landmark from among the set of facial landmarks; and

applying a change to the graphical element based on the second facial tracking input.

13. The system of claim 9, wherein the instructions cause the system to perform operations further comprising:

detecting a rotation of a head of the user based on at least the facial tracking input; and

rotating the 3D model based on the rotation of the head of the user.

14. The system of claim 9, wherein the client device is a first client device, and the instructions cause the system to perform operations further comprising:

generating a flattened presentation based on the 3D model and the image data;

assigning the flattened presentation to a message at the first client device;

distributing the message to a second client device; and

causing display of the flattened presentation ta at the second client device.

15. The system of claim 14, wherein the message includes an ephemeral message.

16. The system of claim 9, wherein the 3D model comprises a set of regions, each region of the 3D model corresponding to a facial landmark from among the set of facial landmarks of the user, and the instructions cause the system to perform operations further comprising:

identifying a region of the 3D model based on the facial tracking input; and

dynamically animating the region in response to the receiving the facial tracking input.

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

causing display of a presentation of image data at a client device, the image data depicting a set of facial landmarks of a user;

generating a 3D model based on the set of facial landmarks in response to the causing display of the presentation of the image data;

overlaying the 3D model at a position on the presentation of the image data at the client device, the position based on the set of facial landmarks;

receiving a facial tracking input that comprises a movement of a facial landmark from among the set of facial landmarks of the user; and

dynamically animating the 3D model at the client device based on the facial tracking input.

18. The non-transitory machine-readable storage medium of claim 17, further comprising instructions that cause the machine to perform operations further comprising:

identifying a user profile of the user based on the set of facial landmarks and an identifier of the client device; and

generating the 3D model based on the user profile and the set of facial landmarks.

19. The non-transitory machine-readable storage medium of claim 17, further comprising instructions that cause the machine to perform operations further comprising:

identifying an expression based on the facial tracking input;

selecting a graphical element from among a set of graphical elements based on the expression; and

causing display of the graphical element within the presentation of the image data at the client device.

20. The non-transitory machine-readable storage medium of claim 19, further comprising instructions that cause the machine to perform operations further comprising:

receiving a second facial tracking input, the second facial tracking input comprising a second movement of a facial landmark from among the set of facial landmarks; and

applying a change to the graphical element based on the second facial tracking input.

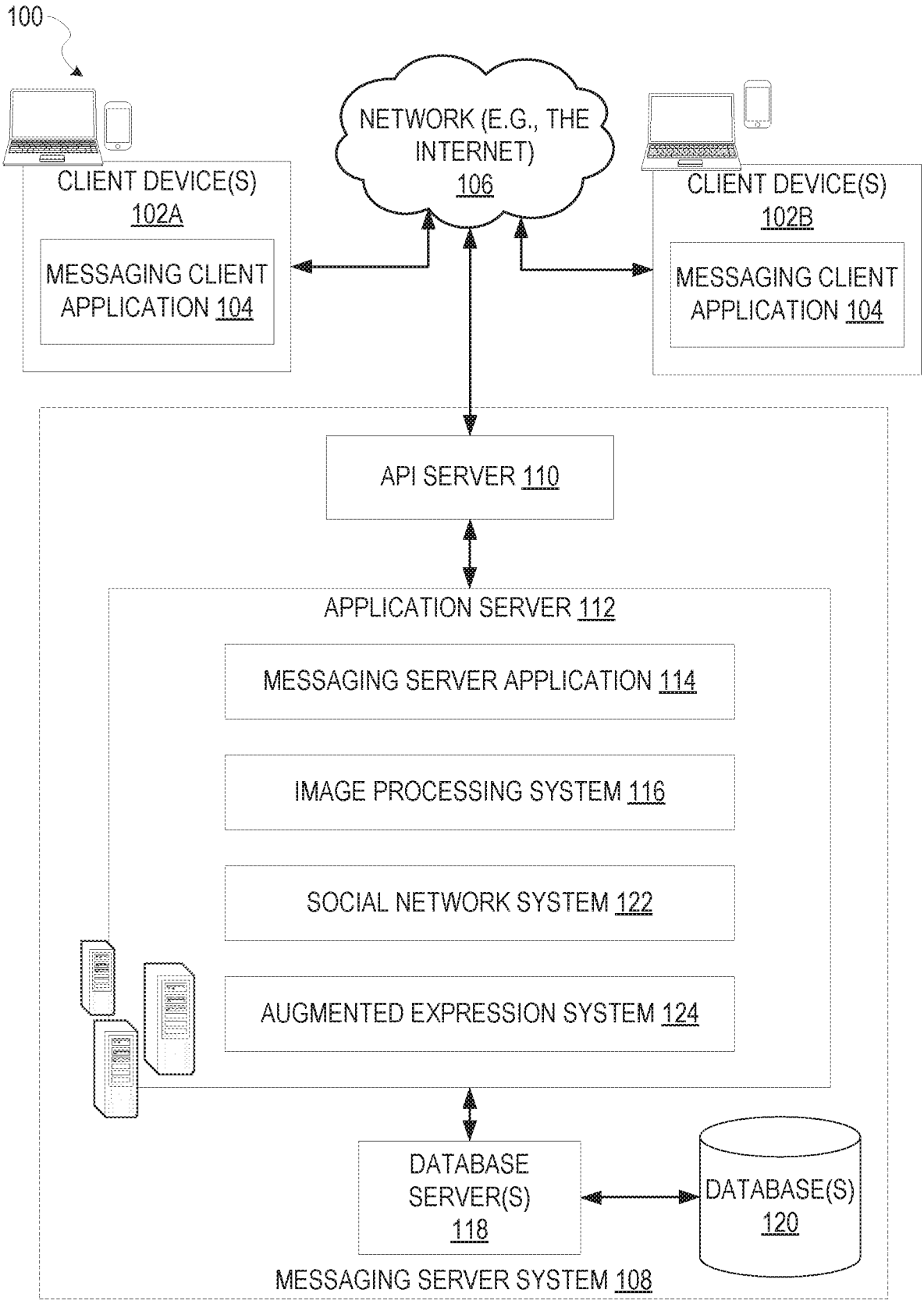

100
NETWORK (E.G., THE INTERNET) 106
CLIENT DEVICE(S) 102A
MESSAGING CLIENT APPLICATION 104
CLIENT DEVICE(S) 102B
MESSAGING CLIENT APPLICATION 104
API SERVER 110
APPLICATION SERVER 112
MESSAGING SERVER APPLICATION 114
IMAGE PROCESSING SYSTEM 116
SOCIAL NETWORK SYSTEM 122
AUGMENTED EXPRESSION SYSTEM 124
DATABASE SERVER(S) 118
DATABASE(S) 120
MESSAGING SERVER SYSTEM 108

FIG. 1

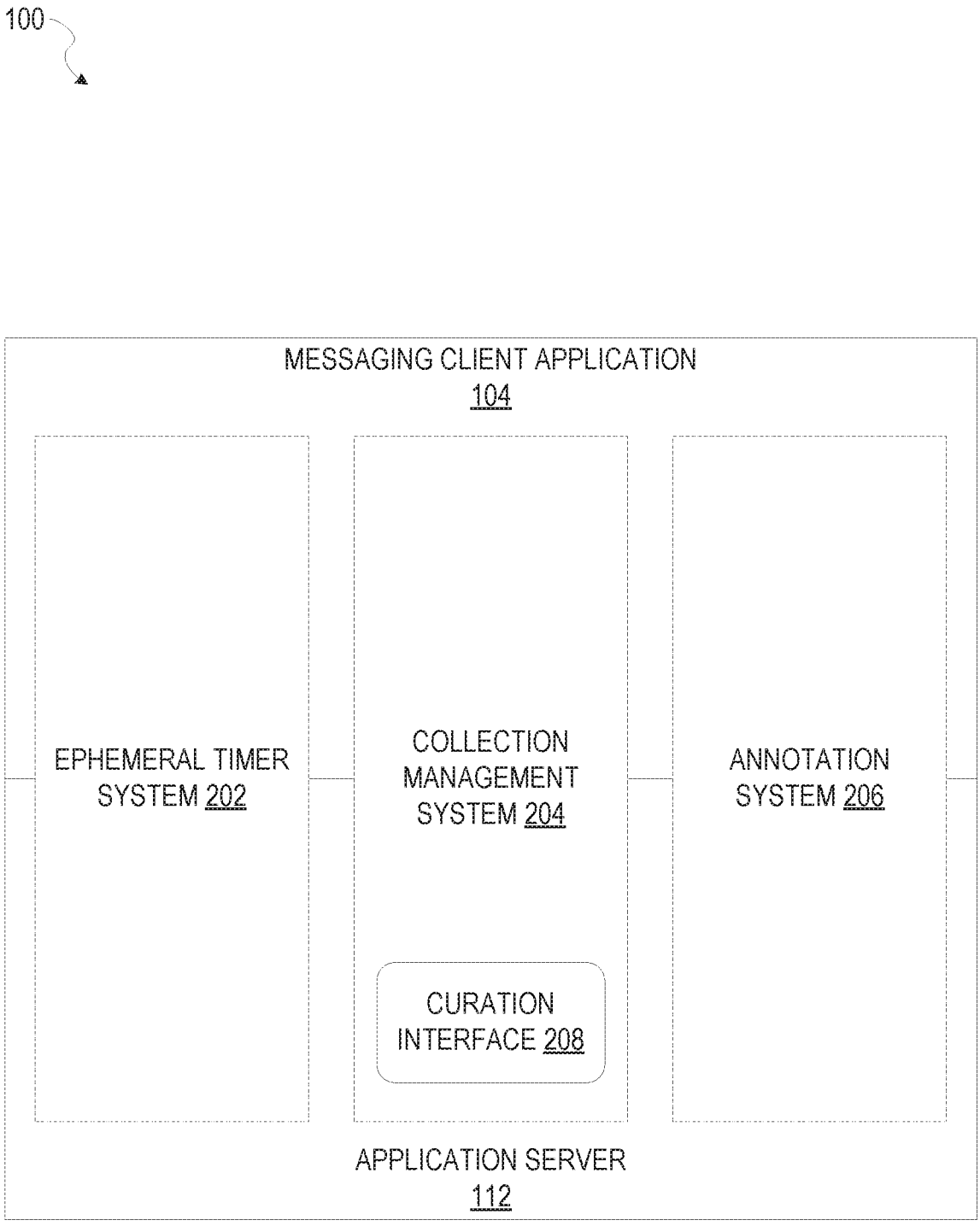

100
MESSAGING CLIENT APPLICATION
104
EPHEMERAL TIMER SYSTEM 202
COLLECTION MANAGEMENT SYSTEM 204
CURATION INTERFACE 208
ANNOTATION SYSTEM 206
APPLICATION SERVER
112

FIG. 2

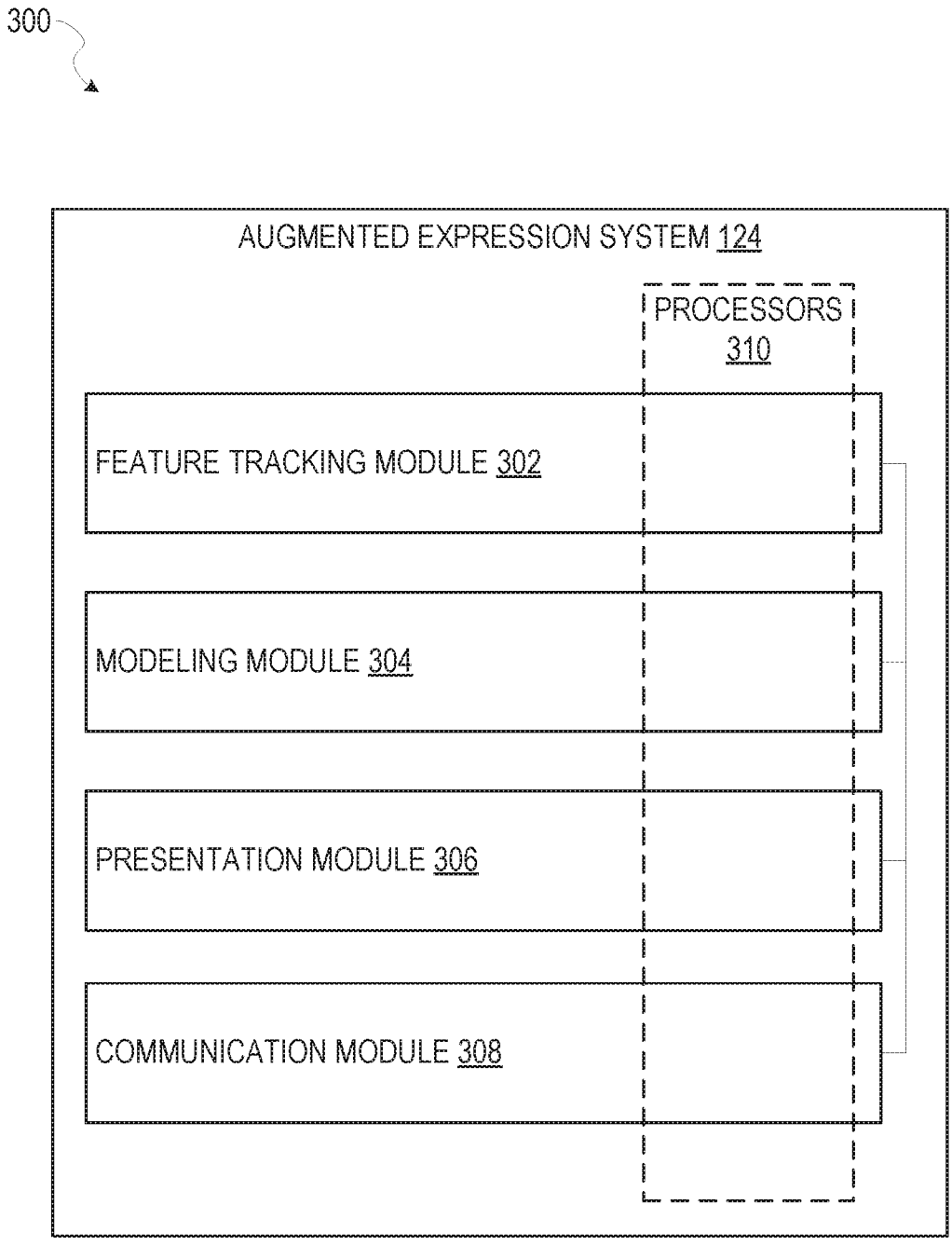

300
AUGMENTED EXPRESSION SYSTEM 124
PROCESSORS 310
FEATURE TRACKING MODULE 302
MODELING MODULE 304
PRESENTATION MODULE 306
COMMUNICATION MODULE 308

FIG. 3

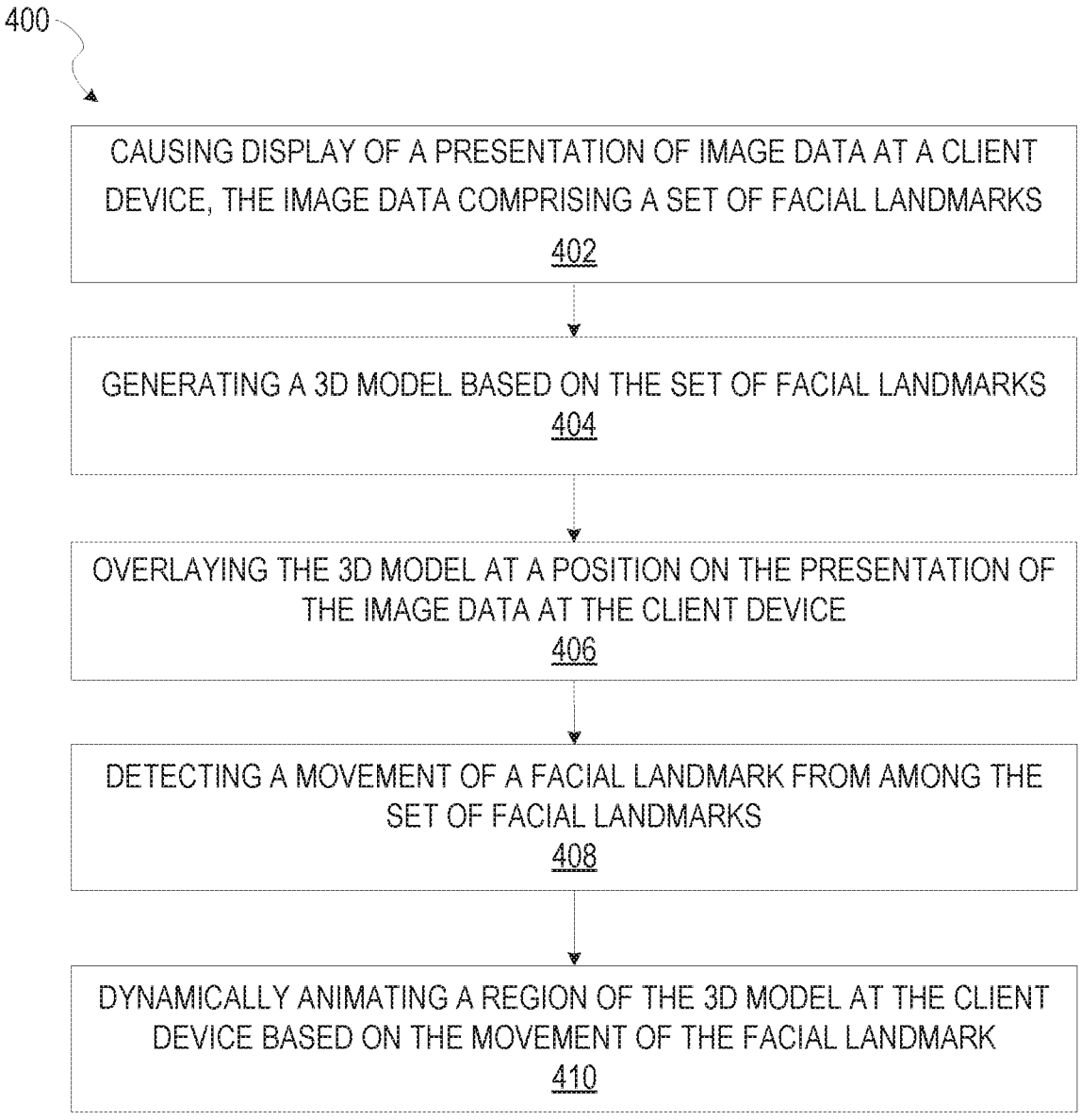

400
CAUSING DISPLAY OF A PRESENTATION OF IMAGE DATA AT A CLIENT DEVICE, THE IMAGE DATA COMPRISING A SET OF FACIAL LANDMARKS
402
GENERATING A 3D MODEL BASED ON THE SET OF FACIAL LANDMARKS
404
OVERLAYING THE 3D MODEL AT A POSITION ON THE PRESENTATION OF THE IMAGE DATA AT THE CLIENT DEVICE
406
DETECTING A MOVEMENT OF A FACIAL LANDMARK FROM AMONG THE SET OF FACIAL LANDMARKS
408
DYNAMICALLY ANIMATING A REGION OF THE 3D MODEL AT THE CLIENT DEVICE BASED ON THE MOVEMENT OF THE FACIAL LANDMARK
410

FIG. 4

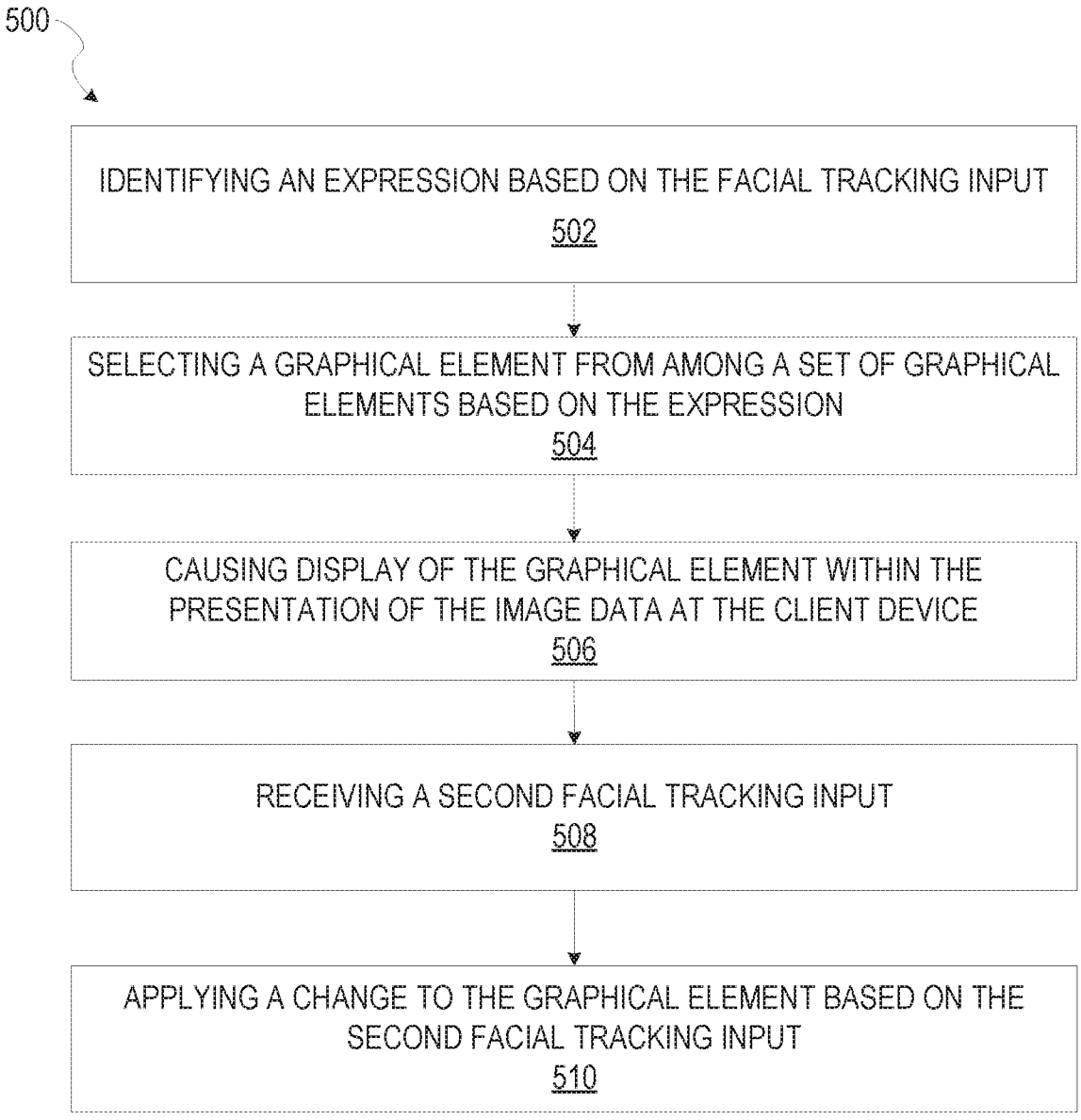

500
IDENTIFYING AN EXPRESSION BASED ON THE FACIAL TRACKING INPUT
502
SELECTING A GRAPHICAL ELEMENT FROM AMONG A SET OF GRAPHICAL ELEMENTS BASED ON THE EXPRESSION
504
CAUSING DISPLAY OF THE GRAPHICAL ELEMENT WITHIN THE PRESENTATION OF THE IMAGE DATA AT THE CLIENT DEVICE
506
RECEIVING A SECOND FACIAL TRACKING INPUT
508
APPLYING A CHANGE TO THE GRAPHICAL ELEMENT BASED ON THE SECOND FACIAL TRACKING INPUT
510

FIG. 5

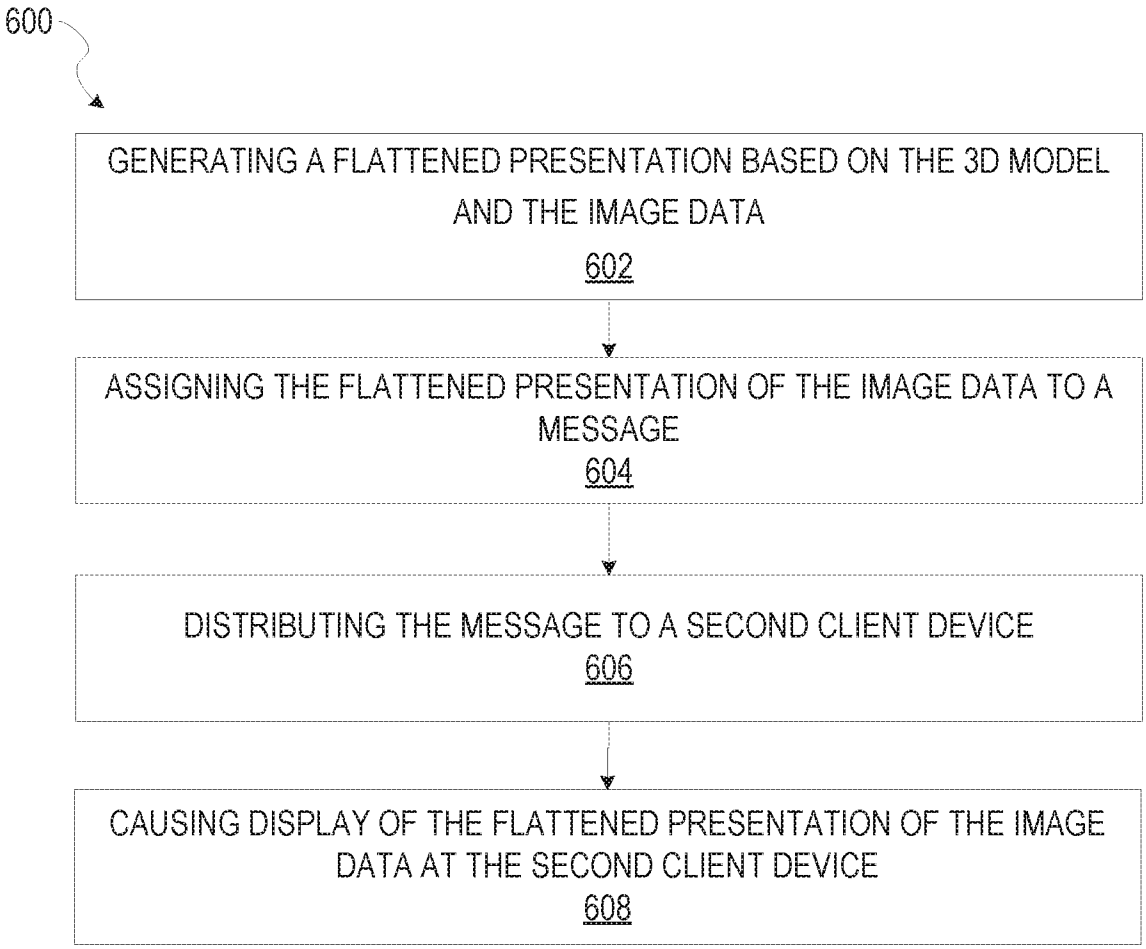

600
GENERATING A FLATTENED PRESENTATION BASED ON THE 3D MODEL AND THE IMAGE DATA
602
ASSIGNING THE FLATTENED PRESENTATION OF THE IMAGE DATA TO A MESSAGE
604
DISTRIBUTING THE MESSAGE TO A SECOND CLIENT DEVICE
606
CAUSING DISPLAY OF THE FLATTENED PRESENTATION OF THE IMAGE DATA AT THE SECOND CLIENT DEVICE
608

FIG. 6

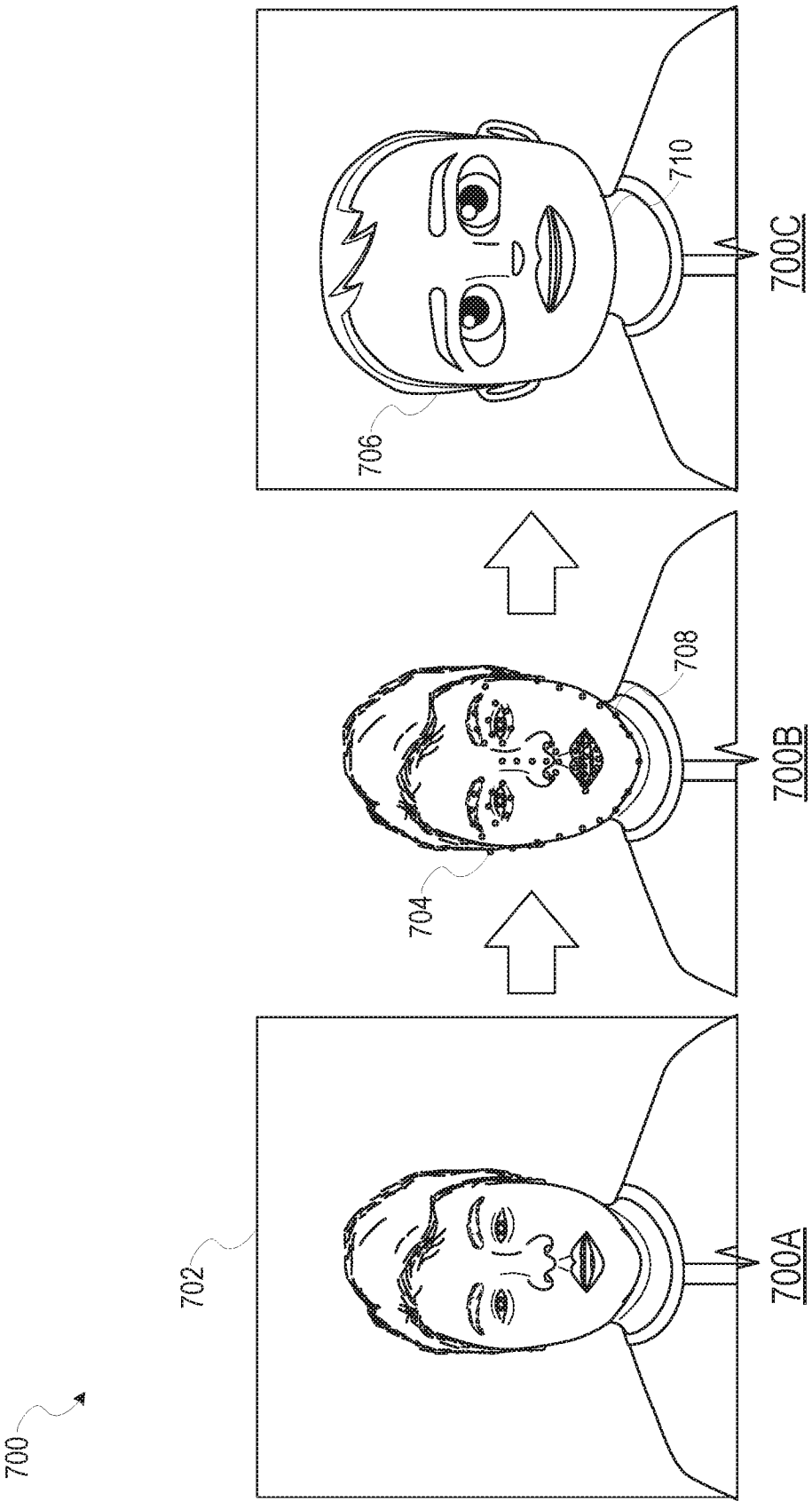


FIG. 7

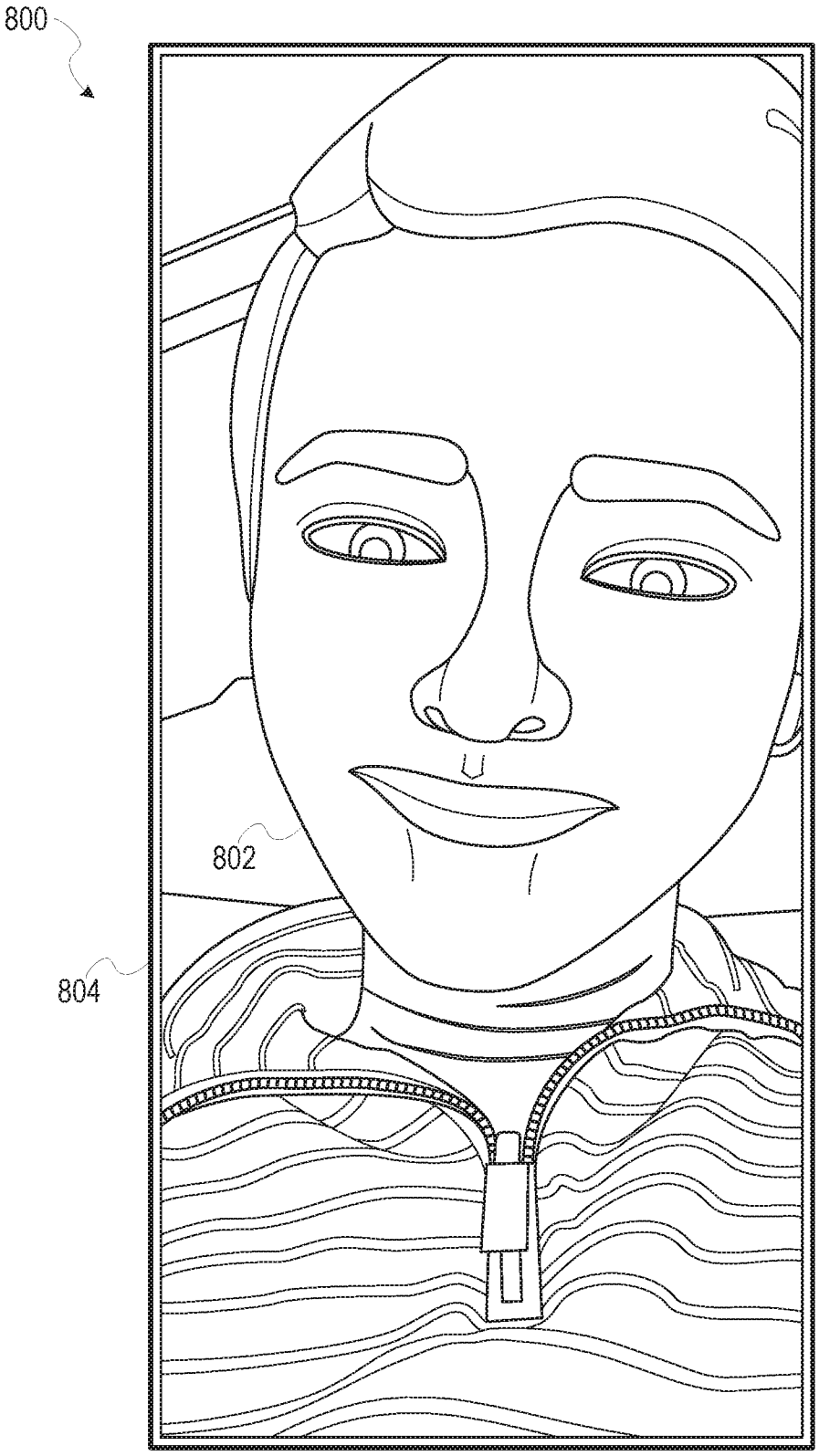


FIG. 8

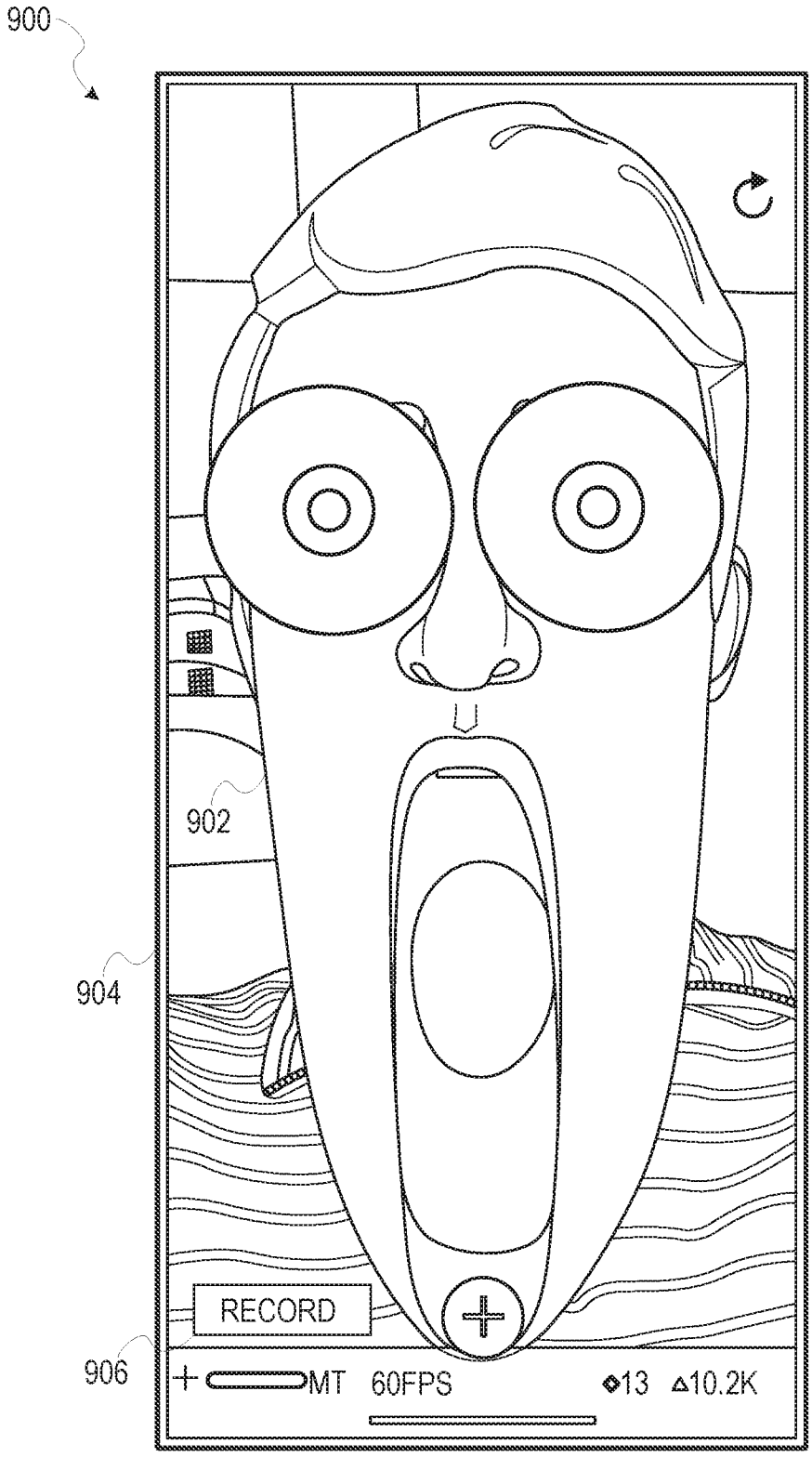


FIG. 9

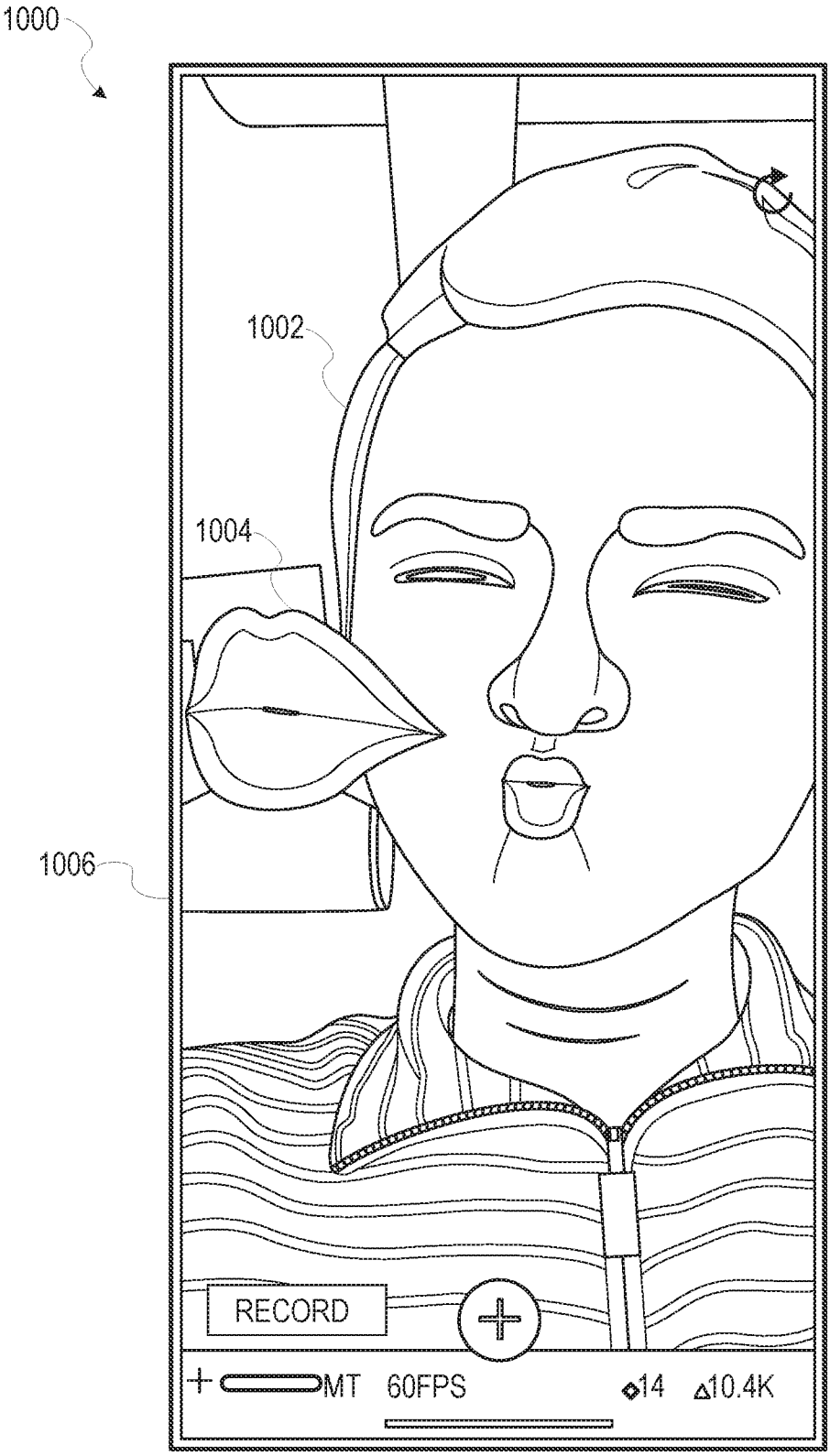


FIG. 10

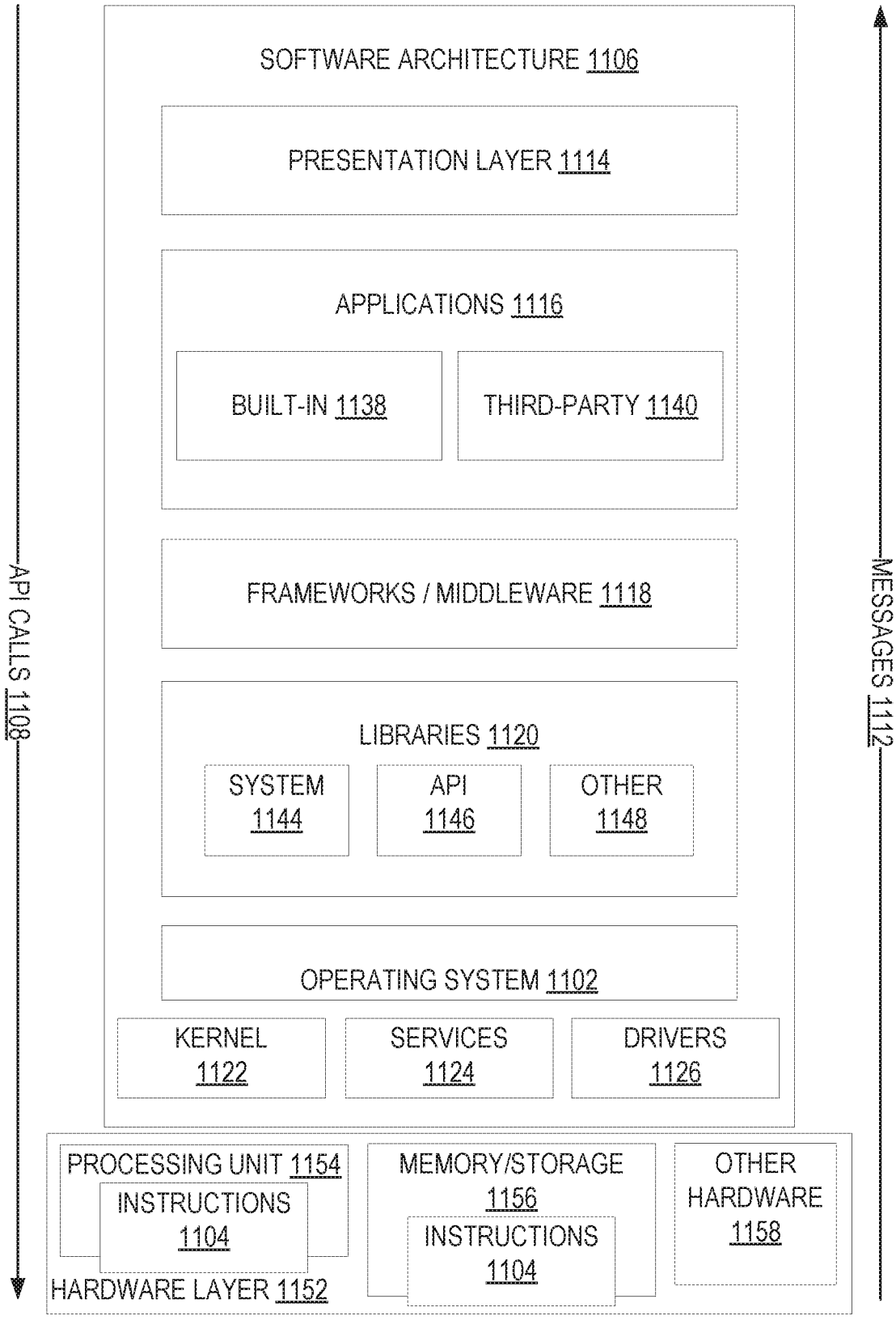

SOFTWARE ARCHITECTURE 1106
PRESENTATION LAYER 1114
APPLICATIONS 1116
BUILT-IN 1138
THIRD-PARTY 1140
FRAMEWORKS / MIDDLEWARE 1118
LIBRARIES 1120
SYSTEM 1144
API 1146
OTHER 1148
OPERATING SYSTEM 1102
KERNEL 1122
SERVICES 1124
DRIVERS 1126
PROCESSING UNIT 1154
INSTRUCTIONS 1104
HARDWARE LAYER 1152
MEMORY/STORAGE 1156
INSTRUCTIONS 1104
OTHER HARDWARE 1158
API CALLS 1108
MESSAGES 1112

FIG. 11

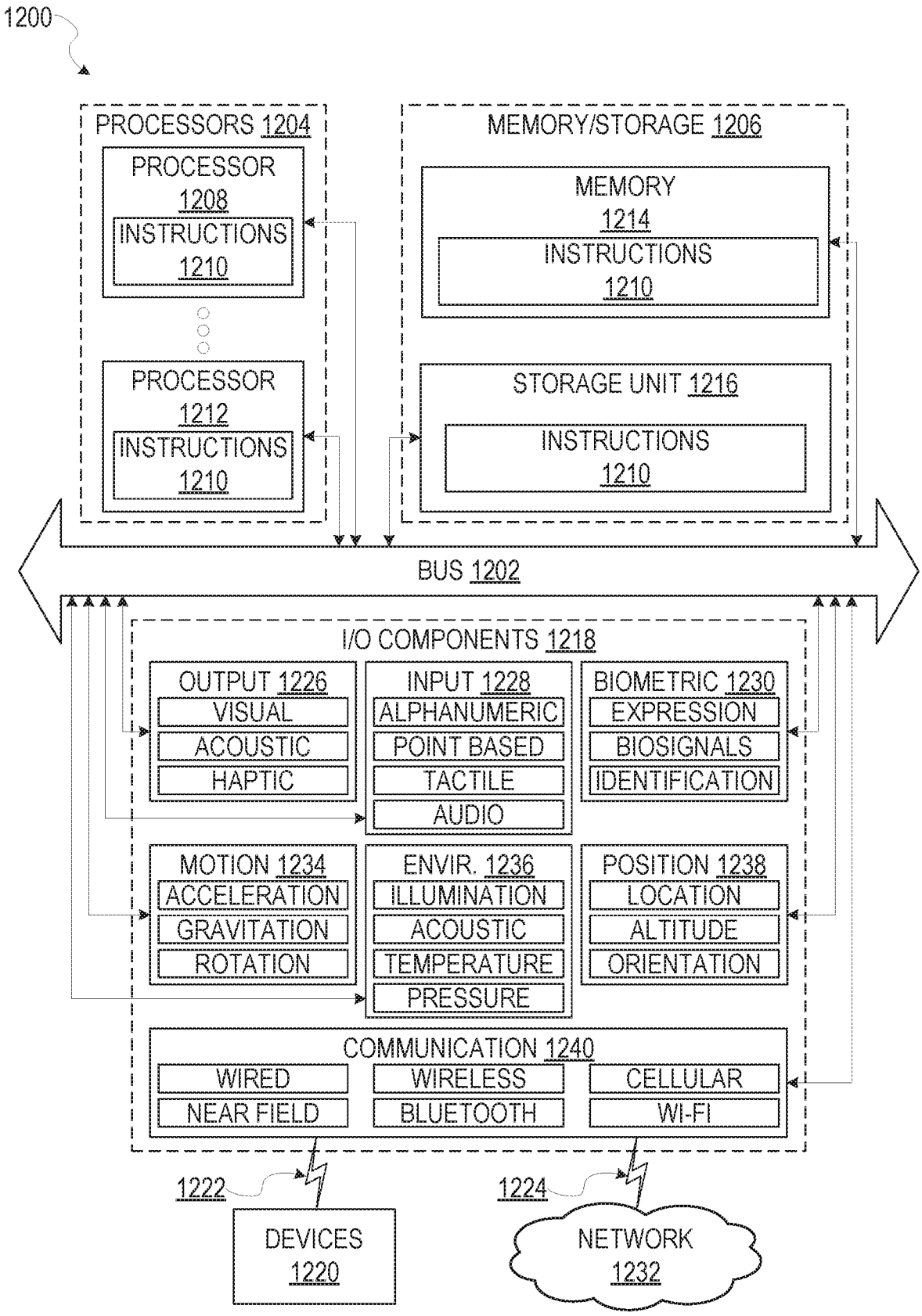

1200
PROCESSORS 1204
PROCESSOR 1208
INSTRUCTIONS 1210
PROCESSOR 1212
INSTRUCTIONS 1210
MEMORY/STORAGE 1206
MEMORY 1214
INSTRUCTIONS 1210
STORAGE UNIT 1216
INSTRUCTIONS 1210
BUS 1202
I/O COMPONENTS 1218
OUTPUT 1226
VISUAL
ACOUSTIC
HAPTIC
INPUT 1228
ALPHANUMERIC
POINT BASED
TACTILE
AUDIO
BIOMETRIC 1230
EXPRESSION
BIOSIGNALS
IDENTIFICATION
MOTION 1234
ACCELERATION
GRAVITATION
ROTATION
ENVIR. 1236
ILLUMINATION
ACOUSTIC
TEMPERATURE
PRESSURE
POSITION 1238
LOCATION
ALTITUDE
ORIENTATION
COMMUNICATION 1240
WIRED
NEAR FIELD
WIRELESS
BLUETOOTH
CELLULAR
WI-FI
1222
DEVICES 1220
1224
NETWORK 1232

FIG. 12

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/027900

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06T13/40 G06T7/20 G06T19/00 G06K9/00
ADD.

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)
G06T G06K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2013/027893 A1 (KANG JUN-KYU [KR]) 28 February 2013 (2013-02-28) figures 1-4, 6, 12, 14, 15 abstract paragraph [0034] - paragraph [0042] paragraph [0070] - paragraph [0080] paragraph [0118] - paragraph [0126] -----	1-20
X	US 2013/235045 A1 (CORAZZA STEFANO [US] ET AL) 12 September 2013 (2013-09-12) figures 1, 3-5, 8A-8C paragraphs [0006], [0014], [0046] paragraphs [0047], [0052], [0055] paragraphs [0061], [0067], [0073] paragraph [0082] ----- -/--	1-20

[X] Further documents are listed in the continuation of Box C.

[X] 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	Date of mailing of the international search report
1 July 2019	08/07/2019

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Höller, Helmut

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/027900

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2014/194439 A1 (INTEL CORP [US]; LI WENLONG [CN]; DU YANGZHOU [CN]; TONG XIAOFENG [CN]) 11 December 2014 (2014-12-11) figures 1, 3A-3E, 8, 9A-9C, 11, 12 page 11, line 16 - page 15, line 2 page 20, line 11 - page 24, line 2 page 27, line 9 - page 29, line 17 page 36, line 24 - page 37, line 8	1-20
X	US 2018/005420 A1 (BONDICH ARTEM [RU] ET AL) 4 January 2018 (2018-01-04) figures 1, 4, 5, 11-13 paragraph [0020] - paragraph [0023] paragraph [0033] paragraph [0043] - paragraph [0051] paragraphs [0074], [0099], [0100]	1-20
A	US 2012/130717 A1 (XU NING [CN] ET AL) 24 May 2012 (2012-05-24) figures 1, 3, 5, 8 paragraph [0041] - paragraph [0050] paragraph [0072]	1-20
A	WO 2008/096099 A1 (AMEGOWORLD LTD [GB]; STOREY JOHN [GB]; CROSS ROBERT [GB]) 14 August 2008 (2008-08-14) the whole document	1-20
A	WO 2016/090605 A1 (INTEL CORP [US]; ZHANG LIDAN [CN] ET AL.) 16 June 2016 (2016-06-16) the whole document	1-20

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2019/027900

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013027893 A1	28-02-2013	KR 20130022434 A	07-03-2013
		WO 2013027893 A1	28-02-2013
US 2013235045 A1	12-09-2013	US 2013235045 A1	12-09-2013
		US 2016163084 A1	09-06-2016
WO 2014194439 A1	11-12-2014	CN 105190700 A	23-12-2015
		US 2015213604 A1	30-07-2015
		US 2018025506 A1	25-01-2018
		WO 2014194439 A1	11-12-2014
US 2018005420 A1	04-01-2018	CN 109643370 A	16-04-2019
		EP 3479292 A1	08-05-2019
		KR 20190022811 A	06-03-2019
		US 2018005420 A1	04-01-2018
		WO 2018006053 A1	04-01-2018
US 2012130717 A1	24-05-2012	CN 102568023 A	11-07-2012
		US 2012130717 A1	24-05-2012
WO 2008096099 A1	14-08-2008	AU 2007346312 A1	14-08-2008
		CA 2677051 A1	14-08-2008
		CN 101669352 A	10-03-2010
		EP 2127341 A1	02-12-2009
		GB 2459073 A	14-10-2009
		JP 2010519791 A	03-06-2010
		RU 2009133354 A	20-03-2011
		US 2010141662 A1	10-06-2010
		WO 2008096099 A1	14-08-2008
WO 2016090605 A1	16-06-2016	CN 107077750 A	18-08-2017
		EP 3238176 A1	01-11-2017
		JP 2018505462 A	22-02-2018
		KR 20170095817 A	23-08-2017
		US 2016361653 A1	15-12-2016
		WO 2016090605 A1	16-06-2016