



- (51) **International Patent Classification:**
H04N 5/232 (2006.01) *G06F 3/01* (2006.01)
- (21) **International Application Number:**
PCT/US2013/076636
- (22) **International Filing Date:**
19 December 2013 (19.12.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
13/723,075 20 December 2012 (20.12.2012) US
- (71) **Applicant:** GOOGLE INC. [US/US]; 1600 Amphitheater Parkway, Mountain View, CA 94043 (US).
- (72) **Inventor:** HOROWITZ, Bradley; 1600 Amphitheater Parkway, Mountain View, CA 94043 (US).
- (74) **Agents:** GAMBHIR, Ajay et al.; IP Spring, 180 N. LaSalle, Suite 3700, Chicago, IL 60601 (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, 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, LT, 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, 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,

[Continued on next page]

(54) **Title:** CAPTURING PHOTOS WITHOUT A CAMERA

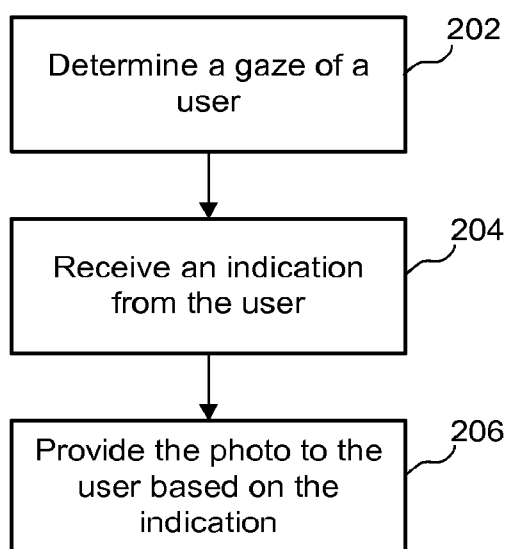


FIG. 2

(57) **Abstract:** Implementations generally relate to capturing photos without a camera. In some implementations, a method includes determining a gaze of a user using a device that is operable to track the gaze. The method also includes receiving an indication from the user that the user desires a photo corresponding to the gaze. The method also includes providing the photo to the user based on the indication.



WO 2014/100449 A1



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, **Published:**
KM, ML, MR, NE, SN, TD, TG).

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

CAPTURING PHOTOS WITHOUT A CAMERA

CROSS REFERENCE TO RELATED APPLICATIONS

[001] The present application claims priority to U.S. Patent Application No. 13/723,075, filed on 20 December 2012, which is hereby incorporated by reference in its entirety.

BACKGROUND

[002] Digital cameras have made taking photos easy and convenient, and various applications have made sharing photos easy and convenient. For example, some applications enable a person to instantly upload photos to a social network system as photos are captured. The social network system may then enable people to share photos with each other.

SUMMARY

[003] Implementations generally relate to capturing photos without a camera. In some implementations, a method includes determining a gaze of a user using a device that is operable to track the gaze. The method further includes receiving an indication from the user that the user desires a photo corresponding to the gaze. The method also includes providing the photo to the user based on the indication.

[004] With further regard to the method, in some implementations, the user provides the indication by making a gesture. In some implementations, the user provides the indication by making a gesture, and where the gesture frames one or more objects. In some implementations, the user provides the indication by making a gesture, and where the method further includes recognizing the gesture by using the device that is operable to track the gaze of the user. In some implementations, the method further includes determining a field of view based on the indication. In some implementations, the method further includes obtaining the photo from a collection of photos based on the gaze. In some implementations, the method further includes obtaining the photo from a collection of photos based on the indication. In some implementations, the method further includes providing a navigator that enables a user to view photos with parameter values that match parameter values of the gaze at different times. In some implementations, the method further includes: collecting attention information associated with the objects; generating an attention map based on the attention information; and enabling the user to select one or more of the photos based on the attention map. In some implementations, the method further includes aggregating collected photos

into a pool of photos; and enabling the user to access the pool of photos based on the indication.

[005] In some implementations, a method includes determining a gaze of a user using a device that is operable to track the gaze, and receiving an indication from the user that the user desires a photo corresponding to the gaze, where the user provides the indication by making a gesture, and where the gesture frames one or more objects. The method further includes recognizing the gesture, and providing a navigator that enables the user to view and select the photo. The method further includes enabling the user to select the photo, and providing the photo to the user based on the selection.

[006] In some implementations, a system includes one or more processors, and logic encoded in one or more tangible media for execution by the one or more processors. When executed, the logic is operable to perform operations including: determining a gaze of a user using a device that is operable to track the gaze; receiving an indication from the user that the user desires a photo corresponding to the gaze; and providing the photo to the user based on the indication.

[007] With further regard to the system, in some implementations, the user provides the indication by making a gesture. In some implementations, the user provides the indication by making a gesture, and where the gesture frames one or more objects. In some implementations, the user provides the indication by making a gesture, and where the logic when executed is further operable to perform operations including recognizing the gesture by using the device that is operable to track the gaze of the user. In some implementations, the logic when executed is further operable to perform operations including determining a field of view based on the indication. In some implementations, the logic when executed is further operable to perform operations including obtaining the photo from a collection of photos based on the gaze. In some implementations, the logic when executed is further operable to perform operations including obtaining the photo from a collection of photos based on the indication. In some implementations, the logic when executed is further operable to perform operations including providing a navigator that enables a user to view photos with parameter values that match parameter values of the gaze at different times. In some implementations, the logic when executed is further operable to perform operations including: collecting attention information associated with the objects; generating an attention map based on the attention information; and enabling the user to select one or more of the photos based on the attention map.

BRIEF DESCRIPTION OF THE DRAWINGS

- [0008] FIG. 1 illustrates a block diagram of an example network environment, which may be used to implement the implementations described herein.
- [0009] FIG. 2 illustrates an example simplified flow diagram for capturing photos without a camera, according to some implementations.
- [0010] FIG. 3 illustrates an example diagram of an eye tracking device that tracks the gaze of a user, according to some implementations.
- [0011] FIG. 4A illustrates an example gesture, according to some implementations.
- [0012] FIG. 4B illustrates another example gesture, according to some implementations.
- [0013] FIG. 5 illustrates an example simplified flow diagram for providing one or more photos to a user, according to some implementations
- [0014] FIGS. 6A, 6B, 6C, and 6D illustrate example simplified photos of objects from different vantage points.
- [0015]
- [0016] FIG. 7 illustrates an example simplified flow diagram for generating an attention map, according to some implementations.
- [0017] FIG. 8 illustrates an example diagram of an attention map, according to some implementations.
- [0018] FIG. 9 illustrates an example simplified flow diagram for generating a model of a location, according to some implementations.
- [0019] FIG. 10 illustrates an example simplified photo composite, according to some implementations.
- [0020] FIG. 11 illustrates an example simplified model of a location, according to some implementations.
- [0021] FIG. 12 illustrates a block diagram of an example server device, which may be used to implement the implementations described herein.

DETAILED DESCRIPTION

[0022] A method for providing photos to a user is described. In various implementations, a system determines a gaze of a user using a device that is operable to track the gaze. The system then receives an indication from the user that the user desires a photo corresponding to the gaze. For example, the user may gaze at a monument and then indicate that the user

desires a photo of the monument. In some implementations, the user may provide the indication by making a gesture, where the gesture frames the photo. In some implementations, the system may recognize the gesture using the device that is operable to track the gaze of the user.

[0023] In some implementations, the system may provide a navigator that enables the user to view photos with parameter values that match parameter values of the gaze at different times. For example, the user may gaze at various monuments throughout an afternoon. The system may then search a pool of photos to identify photos of those monuments, where the identified photos may have parameters (e.g., pitch, rotation, yaw, field of view, etc.) that match corresponding parameters of the gaze of the user during those moments. In some implementations, the system may associate one or more objects with the gaze; collect attention information associated with the objects; generate an attention map based on the attention information; and then enable the user to select one or more of the photos based on the attention map. The system then provides the photo to the user based on the selection.

[0024] FIG. 1 illustrates a block diagram of an example network environment 100, which may be used to implement the implementations described herein. In some implementations, network environment 100 includes a system 102, which includes a server device 104 and a social network database 106. The term system 102 and phrase “social network system” may be used interchangeably. Network environment 100 also includes client devices 110, 120, 130, and 140, which may communicate with each other via system 102 and a network 150.

[0025] For ease of illustration, FIG. 1 shows one block for each of system 102, server device 104, and social network database 106, and shows four blocks for client devices 110, 120, 130, and 140. Blocks 102, 104, and 106 may represent multiple systems, server devices, and social network databases. Also, there may be any number of client devices. In other implementations, network environment 100 may not have all of the components shown and/or may have other elements including other types of elements instead of, or in addition to, those shown herein.

[0026] In various implementations, users U1, U2, U3, and U4 may communicate with each other using respective client devices 110, 120, 130, and 140. Users U1, U2, U3, and U4 may also use respective client devices 110, 120, 130, and 140 to take photos. In various implementations, client devices 110, 120, 130, and 140 may include any types of electronic devices such as mobile phones (e.g., smart phones), tablets, notebook computers, desktop

computers, digital cameras, etc. Such client devices 110, 120, 130, and 140 that are not dedicated digital cameras may include integrated digital cameras.

[0027] In various implementations, system 102 may utilize an eye tracking device for collecting attention information, where the eye tracking device may be used in conjunction with a camera device, which may be a dedicated digital camera or a digital camera integrated with an electronic device (e.g., any of client devices 110, 120, 130, 140, etc.). The eye tracking device may itself be integrated with any one or more of client devices 110, 120, 130, 140, etc. As described in more detail below, such an eye tracking device may be any suitable eye tracking device that measures eye positions such as the point of gaze (e.g., the user's line of sight) and/or measures eye movement.

[0028] In some implementations, client devices 110, 120, 130, and 140 may include wearable computing devices, including any hands-free devices. For example, in some implementations, one or more client devices may include devices that operate with a head-mounted camera, head-mounted eye tracking device, and/or head-mounted display.

[0029] FIG. 2 illustrates an example simplified flow diagram for capturing photos without a camera, according to some implementations. Referring to both FIGS. 1 and 2, a method is initiated in block 202, where system 102 determines a gaze of a user using a device that is operable to track the user's gaze.

[0030] FIG. 3 illustrates an example diagram of an eye tracking device 300 that tracks the gaze of a user, according to some implementations. In some implementations, eye tracking device 300 may be positioned in the head area of the user. For example, as shown, eye tracking device 300 may be positioned relatively close to an eye of the user.

[0031] In various implementations, eye tracking device 300 may use any suitable eye tracking technologies, including any suitable eye-tracking hardware components and algorithms, to measure eye positions and eye movement. For example, eye tracking device 300 may use any suitable eye tracking technologies to measure the gaze of the user (e.g., the user's line of sight) or the motion of an eye relative to the head. For example, eye tracking device 300 may use a laser and laser technology to measure eye positions and eye movement relative to objects in the environment.

[0032] In some implementations, eye tracking device 300 may track the gaze of the user by tracking one or more parameters such as pitch, yaw, roll, line of sight, field of view, etc. FIG. 3 shows a pitch axis, yaw axis, and roll axis to illustrate how eye tracking device 300 may move depending on eye movement and/or head movement of the user, as the attention

that the user places on particular objects would influence and correlate to both eye movement and head movement of the user.

[0033] FIG. 3 shows example objects 302, 304, and 306, where object 302 is a tree, object 304 is a large pyramid, and object 306 is a small pyramid. As shown, objects 302, 304, and 306 are in user's line of site 308 and in the user's field of view 310. For ease of illustration, FIG. 3 shows three objects 302, 304, and 306. Any given scene in the user's field of view 310 may include any number of objects, including people.

[0034] In some implementations, system 102 may receive gaze information (e.g., parameter values associated with tracked parameters) directly from eye tracking device 300 or from any other one or more suitable storage locations. For example, in some implementations, eye tracking device 300 may send gaze information to system 102 as the user gazes at particular objects. In various implementations, eye tracking device 300 may send gaze information to system 102 and/or store gaze information local to the user's client device.

[0035] Referring to again to FIG. 2, in block 204, system 102 receives an indication from the user that the user desires a photo corresponding to the gaze. In various implementations, the user provides the indication by making a gesture, where the gesture frames one or more objects to be captured in the photo.

[0036] FIG. 4A illustrates an example gesture 400, according to some implementations. As shown, hands 402 and 404 form a frame 406 around an object (e.g., tree 302), which the user intends to capture in a photo.

[0037] In various implementations, system 102 recognizes the gesture by using a device that is operable to track the gaze of the user. For example, system 102 may recognize gesture 400 and then determine frame 406 based on gesture 400. In some implementations, system 102 determines a field of view 310 based on the indication or gesture. In some implementations, system 102 may recognize various different types of gestures.

[0038] FIG. 4B illustrates another example gesture 410, according to some implementations. As shown, hands 402 and 404 form a frame 412 around objects (e.g., people U1 and U2 in the foreground, and pyramids 304 and a portion of 306 in the background), which the user intends to capture in a photo.

[0039] Referring again to FIG. 2, in block 206, system 102 provides the photo to the user based on the indication. In some implementations, system 102 obtains the photo from a collection of photos (that have already been captured and stored) based on the gaze and/or

based on the indication (e.g., the gesture). In various implementations, system 102 collects photos from multiple people, aggregates the collected photos into a pool of photos, and enables the user to access the pool of photos based on the indication (e.g., a gesture).

[0040] In various implementations, system 102 may cause the photo to be displayed in any suitable display screen. System 102 may then enable the user to select photos for processing (e.g., accessing, downloading, modifying, sharing, etc.). Example implementations directed to providing photos to a user is described in more detail below in connection with FIG. 5.

[0041] FIG. 5 illustrates an example simplified flow diagram for providing one or more photos, according to some implementations. Referring to both FIGS. 1 and 5, a method is initiated in block 502, where system 102 collects photos associated with objects in at least one physical location. In various implementations, system 102 collects the photos from multiple users. For example, system 102 may collect photos of the same object from many different people, from many different vantage points.

[0042] FIGS. 6A, 6B, 6C, and 6D illustrate example simplified photos of objects in a location. FIG. 6A shows a tree 302. FIG. 6B shows tree 302 and a pyramid 304. FIG. 6B also shows a portion of a pyramid 306. FIG. 6C shows pyramid 304 and a portion of pyramid 306. FIG. 6D shows pyramids 304 and 306. For ease of illustration, FIGS. 6A, 6B, 6C, and 6D show three objects from different vantage points. A given location or scene may have many more objects, depending on the particular scenario. Furthermore, system 102 may collect numerous photos of a given object from many different vantage points.

[0043] In block 504, system 102 collects attention information associated with one or more of the objects. The one or more objects may be set in any given visual context. For example, a given object may be a pyramid in a desert setting, as shown in the examples of FIGS. 6A, 6B, 6C, and 6D. Other examples may include buildings in a city, objects in a natural landscape, etc.

[0044] In various implementations, attention information characterizes the attention (e.g., focus, fixation points, etc.) of a given user with respect to a particular object during any given point in time. In various implementations, attention information may include any combination of attention values, attention parameter values, and gaze information, which are described in more detail below.

[0045] In some implementations, gaze information may include various gaze parameter values associated with pitch, yaw, roll, line of sight, field of view, etc. In various

implementations, one or more of the photos are captured using one or more respective devices that are operable to track the gaze (e.g., line of sight) of a user. Devices that are operable to track a gaze facilitate system 102 in determining values of gaze parameters such as pitch, yaw, roll, line of sight, field of view, etc. For example, a given device while capturing photos may also capture gaze information. Such an eye tracking device may be any suitable eye tracking device that measures the point of gaze (e.g., the user's line of sight).

[0046] For ease of illustration, some implementations are described herein in the context of the gaze of a single user. These implementations and others also apply to gazes of multiple users. For example, for a given object (e.g., a pyramid), system 102 may track, log, and aggregate the gaze information of multiple users. As such, in some implementations, attention information is based on tracking the gazes of one or more users. For example, system 102 may determine which objects at the location are receiving attention from the users. Example implementations involving attention information are described in detail below in connection with FIG. 7.

[0047] In block 506, system 102 generates an attention map based on the attention information. The attention map shows how much attention particular objects have received relative to other objects. In various implementations, the attention map is based on various aspects of attention information. For example, the attention map may be based on tracking gazes of one or more of the users relative to one or more objects (e.g., gaze information), and/or based on any other aspect of attention information such as attention values, attention parameter values, etc.). Example implementations directed to generating an attention map and attention information are described in more detail below in connection with FIG. 7.

[0048] In block 508, system 102 enables the user to select one or more of the photos based on the attention map. For example, in some implementations, system 102 may match objects framed by the user's gestures to corresponding objects captured in existing photos. In some implementations, system 102 may match parameter values of the gaze when the indication was received to parameter values associated with the photos from the collection of photos. As such, system 102 may present such photos to the user for optional selections.

[0049] FIG. 7 illustrates an example simplified flow diagram for generating an attention map, according to some implementations. In some implementations, a method is initiated in block 702, where system 102 receives gaze information. In some implementations, the gaze information associated with a given user may also be referred to as the gaze pattern of the user, or the gaze of the user. As indicated above, the gaze information may include one or

more parameter values associated with gaze parameters such as pitch, yaw, roll, line of sight, field of view, etc.

[0050] In some implementations, where photos are provided by users and where the photos are associated with the gaze information, system 102 may also associate geotags with both the photos and gaze information. System 102 may include such geotags in the gaze parameters. In some implementations, where photos are not provided, system 102 may include geographical coordinates from gaze information, and include such geographical coordinates with the gaze parameters

[0051] In block 704, system 102 identifies objects based on the gaze information. In some implementations, system 102 may receive one or more photos of one or more objects. System 102 may then identify objects in each photo via any suitable object identification algorithm.

[0052] In some implementations, system 102 may also recognize the identified objects. For example, after system 102 identifies an object such as a monument or face, system 102 may then apply a suitable recognition algorithm to recognize an identity associated with the particular object (e.g., monument or particular person). Example implementations directed to object recognition are described in more detail below.

[0053] In some implementations, system 102 may receive gaze information independently of photos being captured. In some implementations, system 102 may also identify a given object based on the gaze information even before the user captures a photo of the object or even if the user does not ultimately capture a photo of the object. Accordingly, system 102 may capture gaze information even if the user is not concurrently capturing photos. In various implementations, system 102 may track only gaze information without recognizing which user is providing the gaze information. In other words, gaze information may be anonymous. In some implementations, a given user may want system 102 to recognizing the given user as the user providing the gaze information in order to enable system 102 in providing individual and/or customized services for the user. As described in more detail below, system 102 enables users to specify and/or consent to the use of personal information, including gaze information.

[0054] In block 706, system 102 associates the gaze information with each of the one or more objects receiving the attention of the user. For example, in some implementations, the gaze information (e.g., the combination of gaze parameter values) characterizes the gaze of

the user. System 102 may determine, from the gaze information, fixation points on one or more objects in a given photo, including fixation points at particular portions of such objects.

[0055] Note that a photo need not necessarily be provided by the user associated with the gaze information. System 102 may determine an appropriate photo based on geolocation information, other photos provided by the user associated with the gaze information, etc. In some implementations, where system 102 aggregates gazes from multiple users, system 102 may associate the gaze information from the different users with the same one or more objects. For example, referring to FIG. 4, if multiple users gaze at object 304, system 102 may associate the gaze information associated with all such users with the same object 304.

[0056] In block 708, system 102 determines an attention value for each of the one or more objects based on the gaze information. In some implementations, the attention value may be based on one or more attention parameter values. Such attention parameters may include, for example, the amount of time a given user gazed at a given object. In some implementations, such attention parameters have corresponding attention subvalues that system 102 may aggregate in order to derive a given attention value. In various implementations, the phrase “attention parameter value” may be use interchangeably with the phrase “attention subvalue.”

[0057] In some implementations, system 102 may assign an attention subvalue that is proportional to the total amount of time that the user gazed at the object. For example, system 102 may assign a higher attention subvalue if the user gazed at the object for 10 minutes versus only 2 minutes.

[0058] In some implementations, system 102 may assign an attention subvalue that is proportional to the total number of times that the user gazed at a given object. For example, system 102 may assign a higher attention subvalue if the user gazed at the object 5 different times versus a single time.

[0059] In some implementations, system 102 may assign an attention subvalue that is proportional to the total size and/or percentage of a given object at which the user gazed. For example, system 102 may assign a higher attention subvalue if the user gazed at 75% of the object versus 25% of the object.

[0060] In some implementations, system 102 may assign an attention subvalue that is proportional to the total number of people who gazed at a given object. For example, system 102 may assign a higher attention subvalue if 1,000 people gazed at the object versus 5

people. Other attention parameters are possible, and the particular number of attention parameters and the types of parameters will depend on the particular implementation.

[0061] In block 710, system 102 generates an attention map based on the attention values. As indicated above, the attention map may be based on attention information in that the attention information may in turn be based on attention values. An example attention map is described in more detail below in connection with FIG. 8.

[0062] FIG. 8 illustrates an example diagram of an attention map 800, according to some implementations. As indicated above, an attention map shows how much attention particular objects have received relative to other objects. As shown, objects 302, 304, and 306 are shown with Xs overlaid, where the number of Xs is proportional to the attention value. In some implementations, objects 302, 304, and 306 may each be shown with an actual attention value. For example, as shown in this example implementation, object 302 has an attention value of 9, object 304 has an attention value of 97, and object 306 has an attention value of 85. The range of attention values may vary (e.g., 0 to 1.0; 0 to 100; 1 to 1,000, etc.), and the particular range and/or numbering scheme will depend on the particular implementation.

[0063] In some implementations, system 102 may assign a color to each object, where the particular color may correspond to the size of the attention value. For example, an object associated with yellow may have a relatively higher attention value than an object associated with blue; an object associated with orange may have a relatively higher attention value than an object associated with yellow; an object associated with red may have a relatively higher attention value than an object associated with orange. The particular color scheme will vary, depending on the particular implementation.

[0064] FIG. 9 illustrates an example simplified flow diagram for generating a model of a location, according to some implementations. As shown, a method is initiated in block 902, where system 102 identifies the objects in the attention map. For example, system 102 may identify objects (e.g., monuments, buildings, trees, etc.) and/or portions of objects that are receiving attention from people (e.g., receiving high attention from tourists).

[0065] In block 904, system 102 selects photos for each of the objects based on one or more predetermined selection criteria. For example, in various implementations, the one or more predetermined selection criteria may be based on image resolution, quality, field of view, level of attention, etc.

[0066] With regard to resolution, system 102 may rank photos of a particular object based on resolution, and then select photos that meet a predetermined resolution threshold. With

regard to quality, system 102 may rank photos of a particular object based on quality, and then select photos that meet a predetermined quality threshold. With regard to field of view, system 102 may rank photos of a particular object based on field of view, and then select photos that meet a predetermined field of view threshold.

[0067] With regard to level of attention, system 102 may search a corpus of photos for photos that include objects that match objects identified in the attention map. For example, referring to attention map 800 of FIG. 8, system 102 may identify pyramid 304 as receiving a significant amount of attention. System 102 may then search a database for photos of pyramid 304. Similarly, system 102 may search the database for photos of objects that correspond to other objects receiving attention according to attention map 800.

[0068] With further regard to level of attention, system 102 may apply different standard levels to particular criteria (e.g., resolution, quality, field of view, etc.) based on level of attention. For example, system 102 may apply a higher standard to objects that receive a higher level of attention (e.g., more popular objects such as monuments). Conversely, system 102 may apply a lower standard to objects that receive a lower level of attention (e.g., less popular objects such as lawns, clear skies, etc.).

[0069] In some implementations, to determine a standard level for photos of a given object, system 102 may determine one or more attention thresholds, where system 102 may associate predetermined attention thresholds proportionally with levels of attention (e.g., levels of popularity) among users. For example, system 102 may associate a relatively lower attention threshold with a relatively lower level of attention. Conversely, system 102 may associate a relatively higher attention threshold with a relatively higher level of attention. The actual number of predetermined thresholds may vary, and will depend on the particular implementation.

[0070] In some implementations, system 102 may determine an attention value for each of the one or more objects; compare the attention value for each object against one or more predetermined attention thresholds; and then determine if the attention value for each object meets one or more of the predetermined attention thresholds. In some implementations, an attention value meets a given attention threshold if the attention value is greater than the attention threshold. In some implementations, an attention value meets a given attention threshold if the attention value is greater than or equal to the attention threshold. Also, a given attention value may meet multiple attention thresholds. In some implementations, system 102 may associate the attention value with the highest attention threshold met for the

purposes of grouping photos for sharing among users. The particular criteria for meeting an attention threshold may vary and will depend on the particular implementation.

[0071] While criteria based on resolution, quality, field of view, level of attention, are described, other criteria are also possible, depending on the specific implementation. For example, predetermined selection criteria may be based on geolocation. In some implementations, system 102 may ensure that the geolocation of a particular object (e.g., pyramid 304) matches and/or corresponds to the geotag associated with retrieved photos of that object.

[0072] In block 906, system 102 combines the photos into a panorama of the location. In some implementations, system 102 combines photos by registering images of the objects in the photos; calibrating the images of the objects; and then blending the images.

[0073] FIG. 10 illustrates an example simplified photo composite 1000, according to some implementations. As shown, photo composite 1000 includes a combination of photos 1002, 1004, 1006, and 1008, which may correspond to the photos of FIGS. 3A, 3B, 3C, and 3D, respectively. Photo composite 1000 shows tree 302, pyramid 304, and pyramid 306. As shown, the photos are registered and calibrated into photo composite 1000.

[0074] FIG. 11 illustrates an example simplified model 1100 of a location, according to some implementations. As shown, the photos are blended into a single model 1100 of the location. As shown, model 1100 includes tree 302, pyramid 304, and pyramid 306, which were originally from different photos.

[0075] For ease of illustration, model 1100 of FIG. 11 shows a simple panoramic view. Depending on the number of objects captured, photos available, points of view, etc., models of locations may provide various types panoramas of varying complexities (e.g., 360 degree panoramas, cylindrical panoramas, spherical panoramas, cubical panoramas, etc.). The specific panoramas provided by a given model of a location will depend on the specific implementation.

[0076] In block 908, system 102 iteratively replaces photos in the panorama with replacement photos based on one or more predetermined replacement criteria. In various implementations, the one or more predetermined replacement criteria may be the same or similar to the predetermined selection criteria described above. For example, in various implementations, the one or more predetermined selection criteria may be based on resolution, quality, field of view, level of attention, geolocation, etc.

[0077] In various implementations, system 102 selects replacement photos continuously in order to improve a given model of a location. In many situations, numerous photos of the same object are substantially the same, yet are slight variations of each other. Such variations in photos may be due to gaze differences (e.g., differences in height among people, etc.), camera differences (e.g., differences in quality, settings, etc.), lighting differences (e.g., differences in time of day, weather, etc.), background and environmental differences (e.g., variations in the sky, clouds, airplanes, etc.), foreground differences (e.g., people, friends, family, etc.), etc. In various implementations, system 102 factors in such differences when initially selecting photos for a given model of a location. Similarly, system 102 factors in such differences when replacing photos of the model. Such a replacement process improves the model of the location over time.

[0078] In various implementations, by utilizing an attention map for selecting photos for objects, system 102 allocates appropriate resources for building and refining a given model of a location. In other words, system 102 concentrates various resources on objects of a model of a given location that are most meaningful (e.g., receives the most attention from people).

[0079] In various implementations, system 102 may cause the generated model of a given location to be displayed on a display screen and/or navigator screen to enable a user to view the model.

[0080] In some implementations, system 102 provides a navigator that enables a user to view photos with parameter values that match parameter values of the gaze at different times. In various implementations, system 102 frames the matching objects in the photos based on gestures indicated by the user. Via the navigator, system 102 enables the user to select one or more photos.

[0081] In some implementations, system 102 enables users to manually modify and improve a given model. For example, in some implementations, system 102 enables a user to overlay one or more user-selected photos over one or more portions of the panorama. This allows one or more users to improve a given model of a location as desired.

[0082] In addition to enabling users to capture photos without a camera, models provided by implementations described herein may have various other applications. For example, in some implementations, a model may provide a high-resolution texture map for users. In some implementations, a model may enable a user to enhance, augment, or improve new photos captured. In some implementations, a model may guide a user to different aspects of a location, which may inform the user of different photo opportunities. Such photo

opportunities may benefit the user when capturing new photos in that the user may notice objects via a model that the user might not otherwise notice. Such photo opportunities may also improve a given model and/or benefit a community of people in that users may contribute more photos to be incorporated into the given model.

[0083] Although the steps, operations, or computations described herein may be presented in a specific order, the order may be changed in particular implementations. Other orderings of the steps are possible, depending on the particular implementation. In some particular implementations, multiple steps shown as sequential in this specification may be performed at the same time. Also, some implementations may not have all of the steps shown and/or may have other steps instead of, or in addition to, those shown herein.

[0084] While system 102 is described as performing the steps as described in the implementations herein, any suitable component or combination of components of system 102 or any suitable processor or processors associated with system 102 may perform the steps described.

[0085] In various implementations, system 102 may utilize a variety of recognition algorithms to recognize faces, landmarks, objects, etc. in photos. Such recognition algorithms may be integral to system 102. System 102 may also access recognition algorithms provided by software that is external to system 102 and that system 102 accesses.

[0086] In various implementations, system 102 enables users to specify and/or consent to the use of personal information. Use of personal information may include, for example, system 102 using their faces in photos or using their identity information in recognizing people identified in photos, and may also include system 102 tracking the gaze of a user even when the user is not capturing photos. In some implementations, system 102 may provide users with multiple selections directed to specifying and/or consenting to the use of personal information. For example, selections with regard to specifying and/or consenting may be associated with individual photos, all photos, individual photo albums, all photo albums, etc. The selections may be implemented in a variety of ways. For example, system 102 may cause buttons or check boxes to be displayed next to various selections. In some implementations, system 102 enables users to specify and/or consent to the use of using their photos for facial recognition in general. Example implementations for recognizing faces and other objects are described in more detail below.

[0087] In various implementations, system 102 obtains reference images of users, where each reference image includes an image of a face that is associated with a known user. The

user is known, in that system 102 has the user's identity information such as the user's name and other profile information. In some implementations, a reference image may be, for example, a profile image that the user has uploaded. In some implementations, a reference image may be based on a composite of a group of reference images.

[0088] In some implementations, to recognize a face in a photo, system 102 may compare the face (i.e., image of the face) and match the face to reference images of users. Note that the term “face” and the phrase “image of the face” are used interchangeably. For ease of illustration, the recognition of one face is described in some of the example implementations described herein. These implementations may also apply to each face of multiple faces to be recognized.

[0089] In some implementations, system 102 may search reference images in order to identify any one or more reference images that are similar to the face in the photo. In some implementations, for a given reference image, system 102 may extract features from the image of the face in a photo for analysis, and then compare those features to those of one or more reference images. For example, system 102 may analyze the relative position, size, and/or shape of facial features such as eyes, nose, cheekbones, mouth, jaw, etc. In some implementations, system 102 may use data gathered from the analysis to match the face in the photo to one more reference images with matching or similar features. In some implementations, system 102 may normalize multiple reference images, and compress face data from those images into a composite representation having information (e.g., facial feature data), and then compare the face in the photo to the composite representation for facial recognition.

[0090] In some scenarios, the face in the photo may be similar to multiple reference images associated with the same user. As such, there would be a high probability that the person associated with the face in the photo is the same person associated with the reference images.

[0091] In some scenarios, the face in the photo may be similar to multiple reference images associated with different users. As such, there would be a moderately high yet decreased probability that the person in the photo matches any given person associated with the reference images. To handle such a situation, system 102 may use various types of facial recognition algorithms to narrow the possibilities, ideally down to one best candidate.

[0092] For example, in some implementations, to facilitate in facial recognition, system 102 may use geometric facial recognition algorithms, which are based on feature

discrimination. System 102 may also use photometric algorithms, which are based on a statistical approach that distills a facial feature into values for comparison. A combination of the geometric and photometric approaches could also be used when comparing the face in the photo to one or more references.

[0093] Other facial recognition algorithms may be used. For example, system 102 may use facial recognition algorithms that use one or more of principal component analysis, linear discriminate analysis, elastic bunch graph matching, hidden Markov models, and dynamic link matching. It will be appreciated that system 102 may use other known or later developed facial recognition algorithms, techniques, and/or systems.

[0094] In some implementations, system 102 may generate an output indicating a likelihood (or probability) that the face in the photo matches a given reference image. In some implementations, the output may be represented as a metric (or numerical value) such as a percentage associated with the confidence that the face in the photo matches a given reference image. For example, a value of 1.0 may represent 100% confidence of a match. This could occur, for example, when compared images are identical or nearly identical. The value could be lower, for example 0.5 when there is a 50% chance of a match. Other types of outputs are possible. For example, in some implementations, the output may be a confidence score for matching.

[0095] For ease of illustration, some example implementations described herein have been described in the context of a facial recognition algorithm. Other similar recognition algorithms and/or visual search systems may be used to recognize objects such as landmarks, logos, entities, events, etc. in order to implement implementations described herein.

[0096] Implementations described herein provide various benefits. For example, implementations described herein synthesize static scenes based on photos captured by different people. Implementations continually improve such static scenes as different people contribute new photos to the scenes.

[0097] FIG. 12 illustrates a block diagram of an example server device 1200, which may be used to implement the implementations described herein. For example, server device 1200 may be used to implement server device 104 of FIG. 1, as well as to perform the method implementations described herein. In some implementations, server device 1200 includes a processor 1202, an operating system 1204, a memory 1206, and an input/output (I/O) interface 1208. Server device 1200 also includes a social network engine 1210 and a media application 1212, which may be stored in memory 1206 or on any other suitable

storage location or computer-readable medium. Media application 1212 provides instructions that enable processor 1202 to perform the functions described herein and other functions.

[0098] For ease of illustration, FIG. 12 shows one block for each of processor 1202, operating system 1204, memory 1206, I/O interface 1208, social network engine 1210, and media application 1212. These blocks 1202, 1204, 1206, 1208, 1210, and 1212 may represent multiple processors, operating systems, memories, I/O interfaces, social network engines, and media applications. In other implementations, server device 1200 may not have all of the components shown and/or may have other elements including other types of elements instead of, or in addition to, those shown herein.

[0099] Although the description has been described with respect to particular embodiments thereof, these particular embodiments are merely illustrative, and not restrictive. Concepts illustrated in the examples may be applied to other examples and implementations.

[00100] Note that the functional blocks, methods, devices, and systems described in the present disclosure may be integrated or divided into different combinations of systems, devices, and functional blocks as would be known to those skilled in the art.

[00101] Any suitable programming languages and programming techniques may be used to implement the routines of particular embodiments. Different programming techniques may be employed such as procedural or object-oriented. The routines may execute on a single processing device or multiple processors. Although the steps, operations, or computations may be presented in a specific order, the order may be changed in different particular embodiments. In some particular embodiments, multiple steps shown as sequential in this specification may be performed at the same time.

[00102] A “processor” includes any suitable hardware and/or software system, mechanism or component that processes data, signals or other information. A processor may include a system with a general-purpose central processing unit, multiple processing units, dedicated circuitry for achieving functionality, or other systems. Processing need not be limited to a geographic location, or have temporal limitations. For example, a processor may perform its functions in “real-time,” “offline,” in a “batch mode,” etc. Portions of processing may be performed at different times and at different locations, by different (or the same) processing systems. A computer may be any processor in communication with a memory. The memory may be any suitable processor-readable storage medium, such as random-access memory

(RAM), read-only memory (ROM), magnetic or optical disk, or other tangible media suitable for storing instructions for execution by the processor.

CLAIMS

What is claimed is:

1. A method comprising:
 - determining a gaze of a user using a device that is operable to track the gaze;
 - receiving an indication from the user that the user desires a photo corresponding to the gaze, wherein the user provides the indication by making a gesture, and wherein the gesture frames one or more objects;
 - recognizing the gesture;
 - providing a navigator that enables the user to view and select the photo;
 - enabling the user to select the photo; and
 - providing the photo to the user based on the selection.
2. A method comprising:
 - determining a gaze of a user using a device that is operable to track the gaze;
 - receiving an indication from the user that the user desires a photo corresponding to the gaze; and
 - providing the photo to the user based on the indication.
3. The method of claim 2, wherein the user provides the indication by making a gesture.
4. The method of claim 2, wherein the user provides the indication by making a gesture, and wherein the gesture frames one or more objects.

5. The method of claim 2, wherein the user provides the indication by making a gesture, and wherein the method further comprises recognizing the gesture by using the device that is operable to track the gaze of the user.
6. The method of claim 2, further comprising determining a field of view based on the indication.
7. The method of claim 2, further comprising obtaining the photo from a collection of photos based on the gaze.
8. The method of claim 2, further comprising obtaining the photo from a collection of photos based on the indication.
9. The method of claim 2, further comprising providing a navigator that enables a user to view photos with parameter values that match parameter values of the gaze at different times.
10. The method of claim 2, further comprising:
 - collecting attention information associated with the objects;
 - generating an attention map based on the attention information; and
 - enabling the user to select one or more of the photos based on the attention map.
11. The method of claim 2, further comprising:
 - aggregating collected photos into a pool of photos; and
 - enabling the user to access the pool of photos based on the indication.

12. A system comprising:
 - one or more processors; and
 - logic encoded in one or more tangible media for execution by the one or more processors and when executed operable to perform operations comprising:
 - determining a gaze of a user using a device that is operable to track the gaze;
 - receiving an indication from the user that the user desires a photo corresponding to the gaze; and
 - providing the photo to the user based on the indication.
13. The system of claim 12, wherein the user provides the indication by making a gesture.
14. The system of claim 12, wherein the user provides the indication by making a gesture, and wherein the gesture frames one or more objects.
15. The system of claim 12, wherein the user provides the indication by making a gesture, and wherein the logic when executed is further operable to perform operations comprising recognizing the gesture by using the device that is operable to track the gaze of the user.
16. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising determining a field of view based on the indication.
17. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising obtaining the photo from a collection of photos based on the gaze.

18. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising obtaining the photo from a collection of photos based on the indication.

19. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising providing a navigator that enables a user to view photos with parameter values that match parameter values of the gaze at different times.

20. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising:

collecting attention information associated with the objects;

generating an attention map based on the attention information; and

enabling the user to select one or more of the photos based on the attention map.

1/12

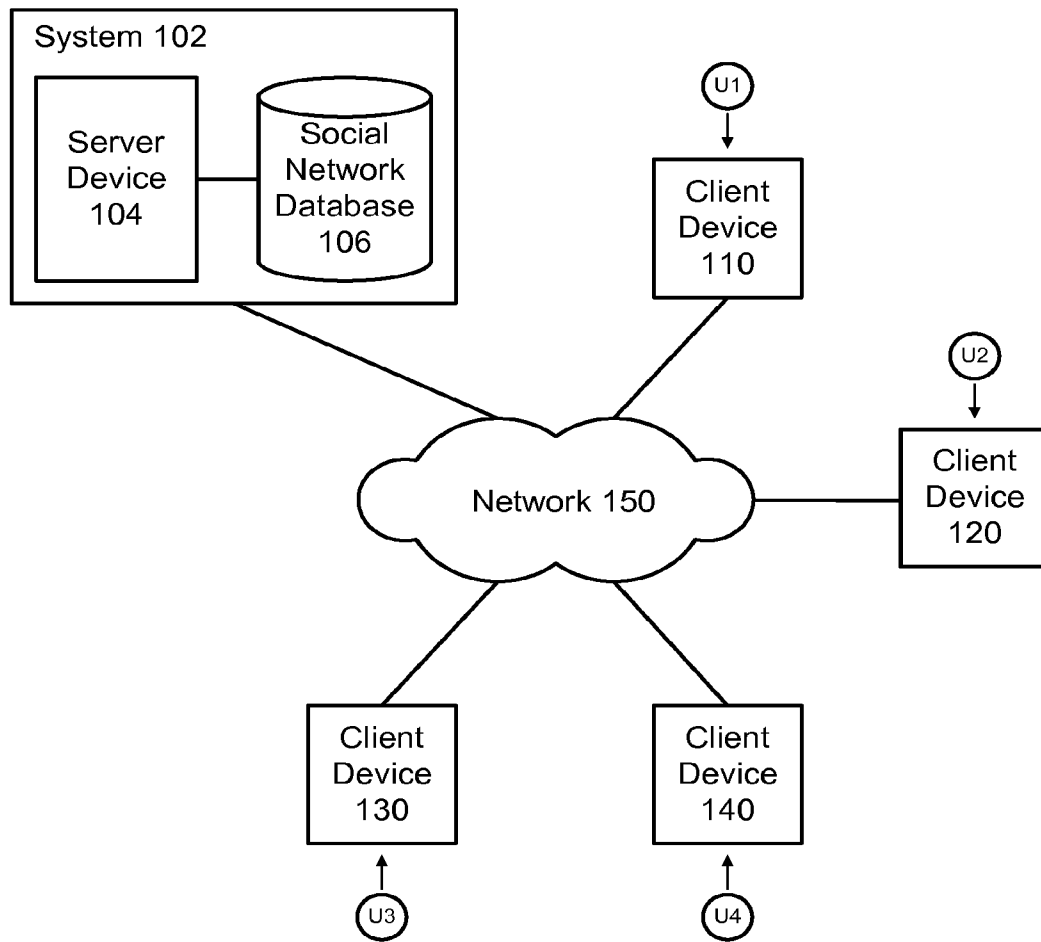
100

FIG. 1

2/12

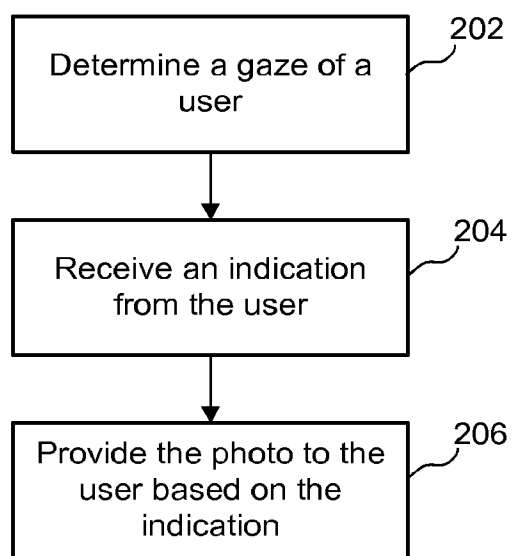


FIG. 2

3/12

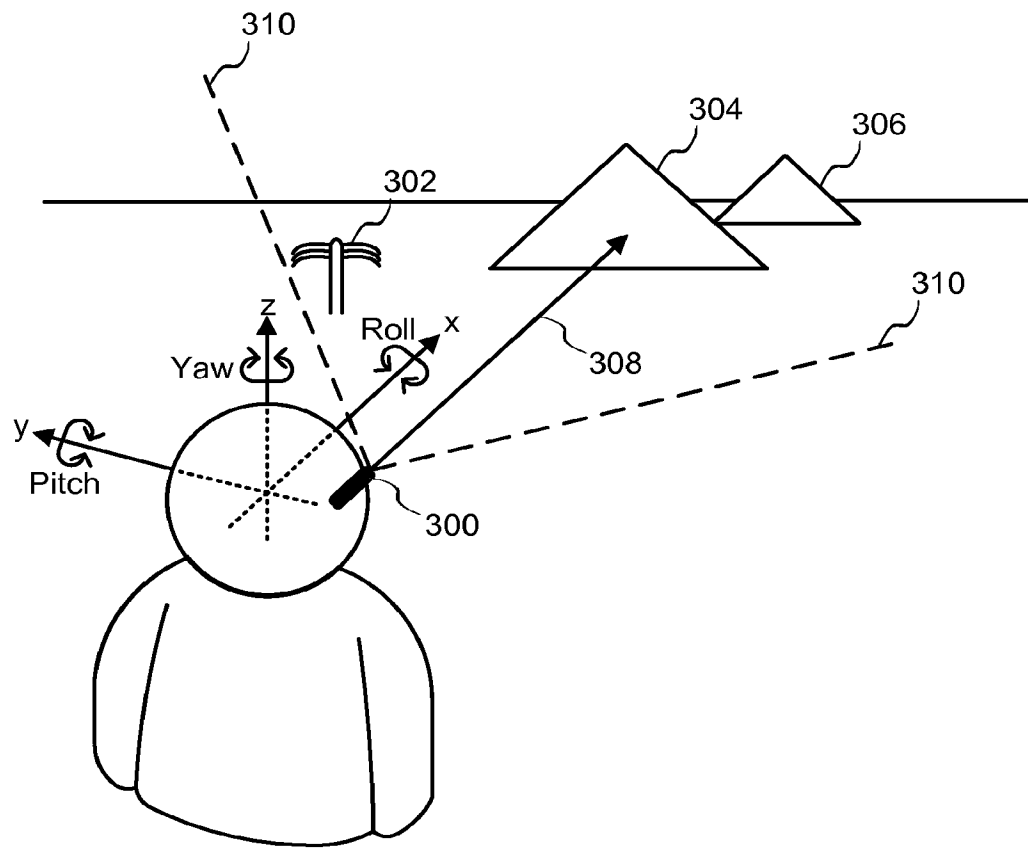
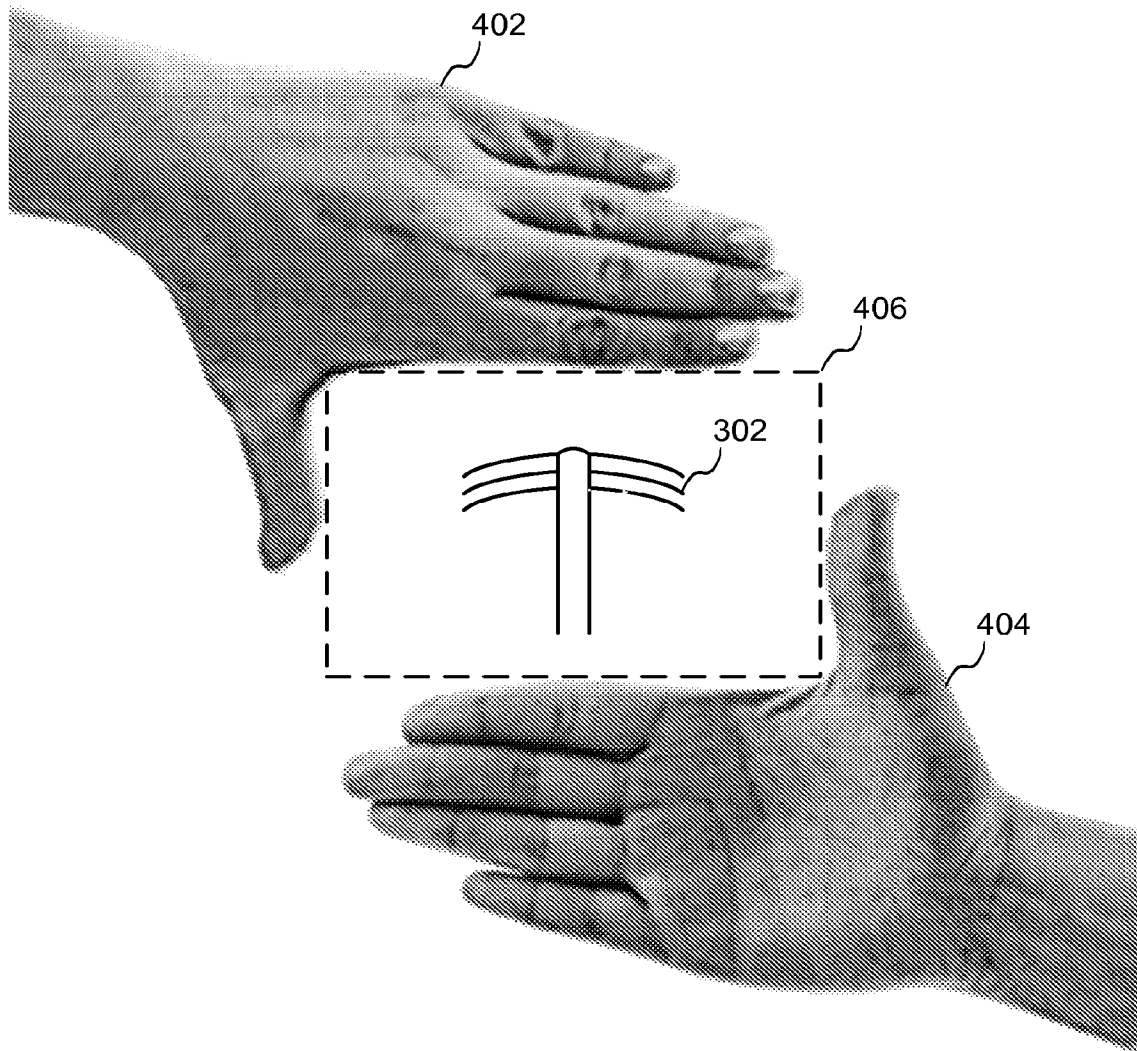


FIG. 3

4/12



400

FIG. 4A

5/12

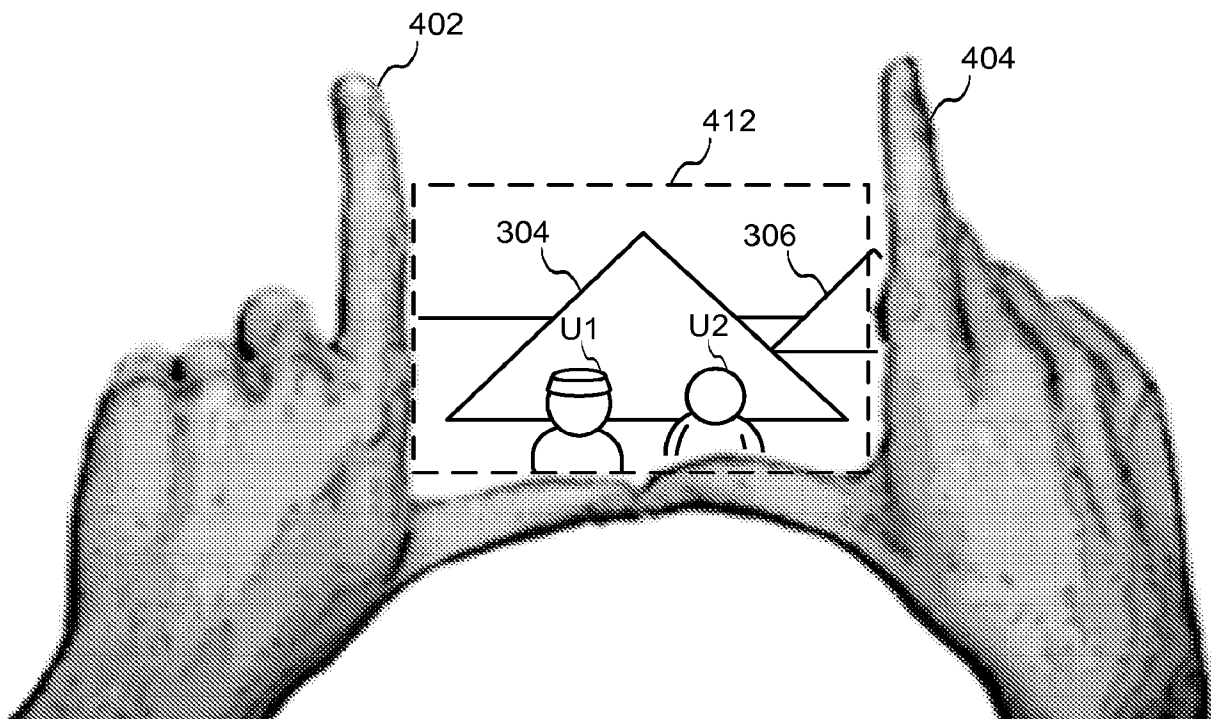
410

FIG. 4B

6/12

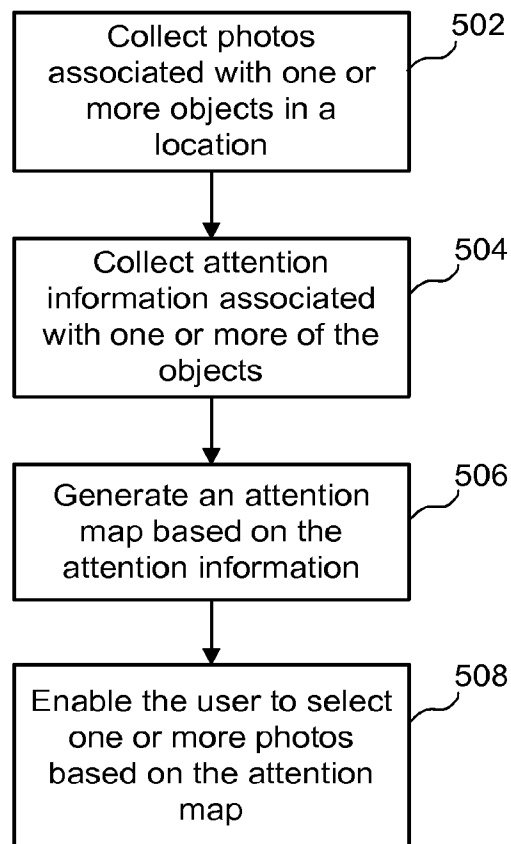


FIG. 5

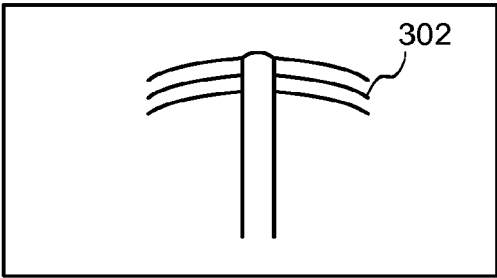


FIG. 6A

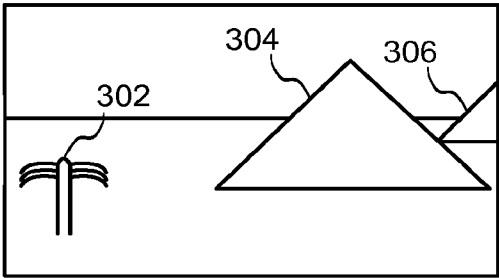


FIG. 6B

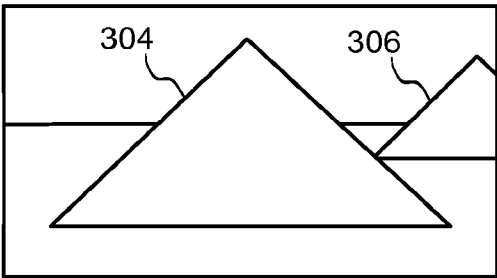


FIG. 6C

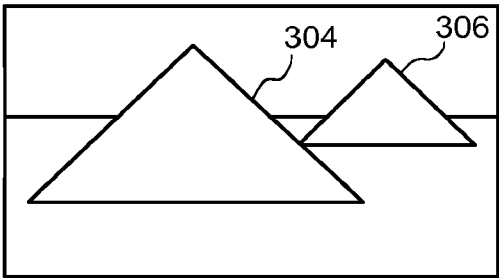


FIG. 6D

8/12

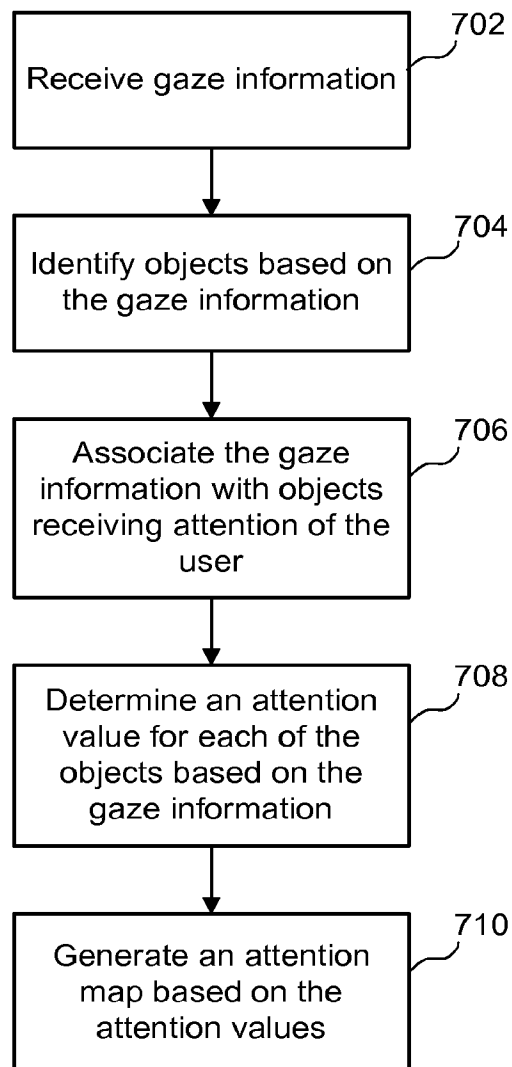
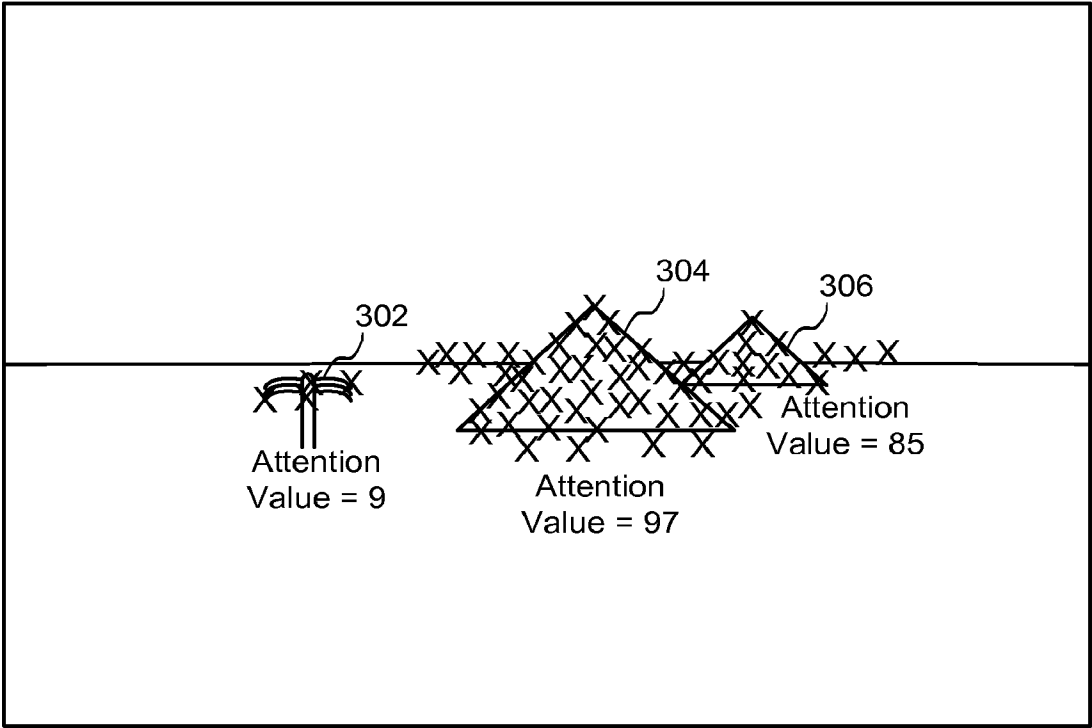


FIG. 7



800

FIG. 8

10/12

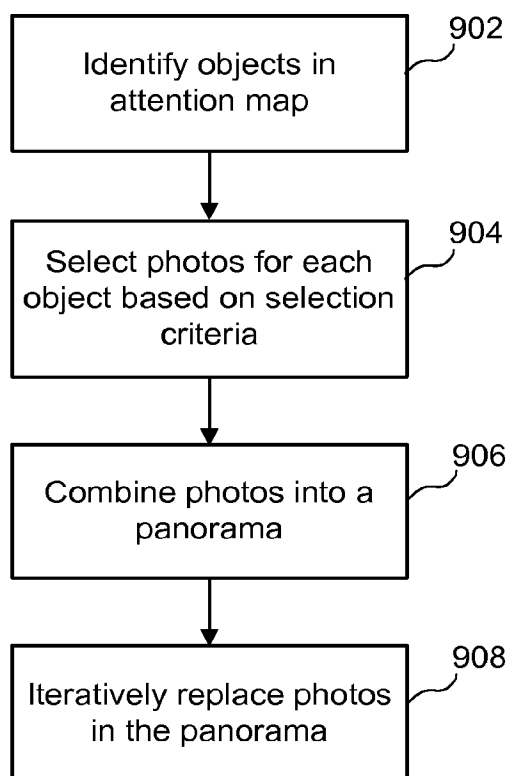


FIG. 9

11/12

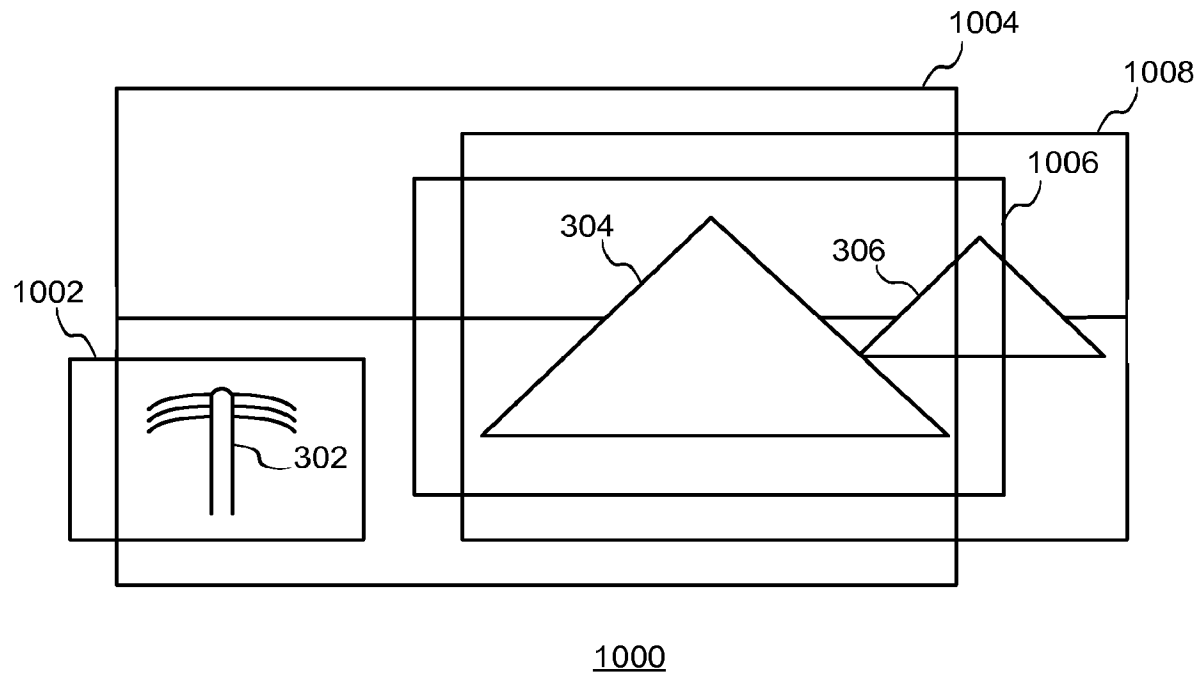


FIG. 10

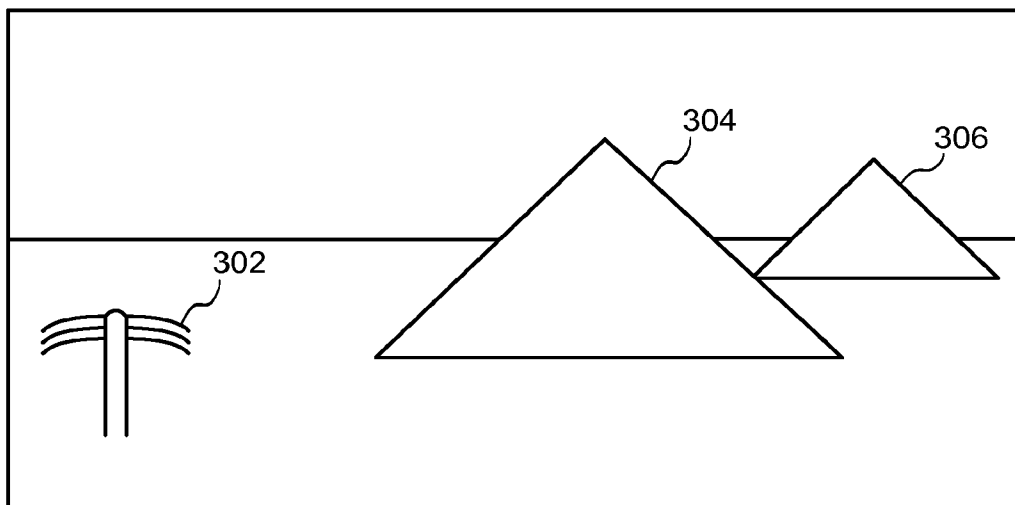


FIG. 11

12/12

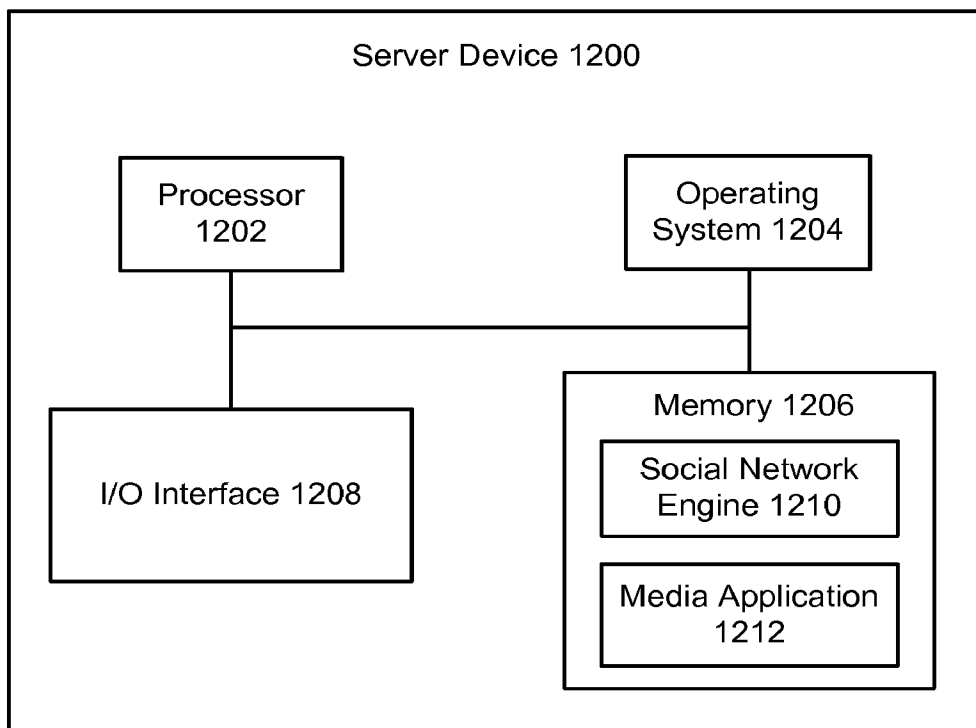


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/076636

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N5/232 G06F3/01
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)
H04N G06F

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)

EP0-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	US 2012/236030 A1 (BORDER JOHN N [US] ET AL) 20 September 2012 (2012-09-20) paragraph [0370] paragraph [0372] paragraph [0394] paragraph [0706] - paragraph [0707]	1-20
A	Anonymous: "Ubi-Camera - Take Photos With Your Hands - DigInfo TV", 28 March 2012 (2012-03-28), XP055109990, Retrieved from the Internet: URL:http://www.diginfo.tv/v/12-0047-r-en.php [retrieved on 2014-03-26] the whole document ----- -/--	4,6,14, 16

☒ 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

28 March 2014

Date of mailing of the international search report

03/04/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Ogor, Pascal

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/076636

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 2 410 733 A2 (SONY ERICSSON MOBILE COMM AB [SE]) 25 January 2012 (2012-01-25) paragraph [0027]	10,20
A	----- US 2009/109795 A1 (MARTI STEFAN [US]) 30 April 2009 (2009-04-30) abstract figures 2-4 -----	1,2,12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2013/076636

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012236030	A1	20-09-2012	NONE
EP 2410733	A2	25-01-2012	EP 2410733 A2 25-01-2012
		US 2012019703 A1	26-01-2012
US 2009109795	A1	30-04-2009	NONE