



- (51) **International Patent Classification:**
H04N 5/232 (2006.01)
- (21) **International Application Number:**
PCT/US2013/076288
- (22) **International Filing Date:**
18 December 2013 (18.12.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
13/723,065 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. La-salle, 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:** COLLECTING PHOTOS

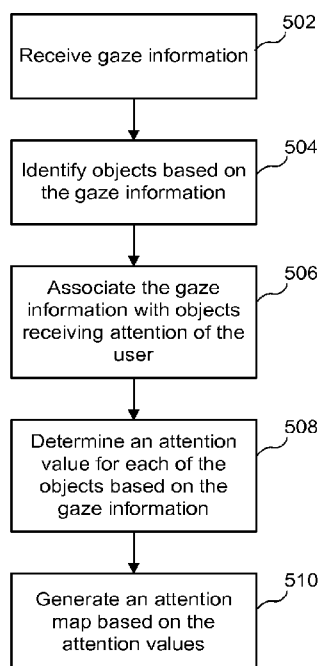


FIG. 5

(57) **Abstract:** Implementations generally relate to collecting photos. In some implementations, a method includes collecting photos associated with one or more objects, where the photos are collected from a plurality of users. The method also includes collecting attention information associated with the one or more objects. The method also includes generating an attention map based on the attention information. The method also includes guiding users to capture photos of particular objects based on the attention map.



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

*before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))*

Published:

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

COLLECTING PHOTOS

CROSS REFERENCE TO RELATED APPLICATIONS

[001] The present application claims priority to U.S. Patent Application No. 13/723,065, 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. Many people capturing photos of the same object are often positioned in the same or similar location. Accordingly, in many situations, numerous photos of the same object are captured, where the photos are substantially the same or are slight variations of each other.

SUMMARY

[003] Implementations generally relate to collecting photos. In some implementations, a method includes collecting photos associated with one or more objects, where the photos are collected from a plurality of users. The method also includes collecting attention information associated with the one or more objects. The method also includes generating an attention map based on the attention information. The method also includes guiding users to capture photos of particular objects based on the attention map.

[004] With further regard to the method, one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user. In some implementations, the method further includes aggregating the collected photos into a pool of photos. In some implementations, the method further includes aggregating the collected photos into a pool of photos; and enabling the plurality of users to have access to the pool of photos. In some implementations, the attention information is based on tracking gazes of one or more of the plurality of users. In some implementations, the attention map is based on tracking gazes of one or more of the plurality of users. In some implementations, to generate the attention map, the method includes: receiving gaze information; identifying the one or more objects; associating the gaze information with each of the one or more objects; and determining an attention value for each of the one or more objects based on the gaze information. In some implementations, the guiding of users includes guiding users to capture

photos of particular objects that receive a level of attention that meets a predetermined attention threshold. In some implementations, the guiding of users includes guiding users to capture photos of particular objects that receive a level of attention that does not meet a predetermined attention threshold. In some implementations, the guiding of users to capture photos of particular objects further includes informing users of optimal vantage points when capturing photos.

[005] In some implementations, a method includes collecting photos associated with one or more objects, where the photos are collected from a plurality of users, and where one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user. The method further includes collecting attention information associated with the one or more objects, where the attention information is based on tracking gazes of one or more of the plurality of users. The method further includes generating an attention map based on the attention information. In some implementations, the generating of the attention map includes: receiving gaze information; identifying the one or more objects; associating the gaze information with each of the one or more objects; and determining an attention value for each of the one or more objects based on the gaze information. The method further includes guiding users to capture photos of particular objects based on the attention map.

[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: collecting photos associated with one or more objects, where the photos are collected from a plurality of users; collecting attention information associated with the one or more objects; generating an attention map based on the attention information; and guiding users to capture photos of particular objects based on the attention map.

[007] With further regard to the system, one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user. In some implementations, the logic when executed is further operable to perform operations including aggregating the collected photos into a pool of photos. In some implementations, the logic when executed is further operable to perform operations including aggregating the collected photos into a pool of photos; and enabling the plurality of users to have access to the pool of photos. In some implementations, the attention information is based on tracking gazes of one or more of the plurality of users. In some implementations, the attention map is based on

tracking gazes of one or more of the plurality of users. In some implementations, to generate the attention map, the logic when executed is further operable to perform operations including: receiving gaze information; identifying the one or more objects; associating the gaze information with each of the one or more objects; and determining an attention value for each of the one or more objects based on the gaze information. In some implementations, the logic when executed is further operable to perform operations including guiding users to capture photos of particular objects that receive a level of attention that meets a predetermined attention threshold. In some implementations, the logic when executed is further operable to perform operations including guiding users to capture photos of particular objects that receive a level of attention that does not meet a predetermined attention threshold.

BRIEF DESCRIPTION OF THE DRAWINGS

[008] FIG. 1 illustrates a block diagram of an example network environment, which may be used to implement the implementations described herein.

[009] FIG. 2 illustrates an example simplified flow diagram for collecting photos from multiple users, according to some implementations.

[0010] FIGS. 3A, 3B, 3C, and 3D illustrate example simplified photos of objects from different vantage points.

[0011] FIG. 4 illustrates an example diagram of an eye tracking device that tracks the gaze of a user, according to some implementations.

[0012] FIG. 5 illustrates an example simplified flow diagram for generating an attention map, according to some implementations.

[0013] FIG. 6 illustrates an example diagram of an attention map, according to some implementations.

[0014] FIG. 7 illustrates an example simplified flow diagram for guiding users to capture photos of particular objects, according to some implementations.

[0015] FIG. 8 illustrates a block diagram of an example server device, which may be used to implement the implementations described herein.

DETAILED DESCRIPTION

[0016] Implementations described herein collect photos from multiple people. In various implementations, a system collects photos associated with one or more objects by guiding

users to capture photos of particular objects based on an attention map. For example, the system may guide people to more popular objects for photos as well as less popular, yet potentially interesting objects. As a result, the system may collect photos of the same objects from many different people, from many different vantage points.

[0017] In some implementations, the system collects photos from multiple users, where one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user. The system then collects attention information associated with the one or more objects, where the attention information is based on tracking gazes of one or more of the users. The system then generates an attention map based on the attention information. In various implementations, the attention map shows how much attention particular objects have received relative to other objects.

[0018] As indicated above, the system guides users to capture photos of particular objects based on the attention map. More specifically, in some implementations, the system may guide a user to capture a photo of a particular object based on whether the particular object receives a level of attention that meets a predetermined attention threshold. For example, the system may guide a given user to an object (e.g., a monument) that is popular among tourists, yet the given user has not noticed. In some implementations, the system may inform users of optimal vantage points when capturing photos.

[0019] 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.

[0020] 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.

[0021] 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.

[0022] 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.

[0023] 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.

[0024] FIG. 2 illustrates an example simplified flow diagram for collecting photos from multiple users, according to some implementations. Referring to both FIGS. 1 and 2, a method is initiated in block 202, where system 102 collects photos associated with one or more objects. In various implementations, the photos are collected from multiple users. For example, system 102 may collect photos of the same object from many different people, from many different vantage points.

[0025] FIGS. 3A, 3B, 3C, and 3D illustrate example simplified photos of objects from different vantage points. FIG. 3A shows two pyramids 304 and 306 in a desert setting. FIG. 3B shows the same two pyramids 304 and 306 from a further distance than shown in FIG. 3A. FIG. 3B also shows a tree 302. FIGS. 3C and 3D show two pyramids 304 and 306 from different vantage points.

[0026] For ease of illustration, FIGS. 3A, 3B, 3C, and 3D show four different vantage points, and these vantage points are significantly distinct. Many other vantage points are possible, depending on the particular scenario. Furthermore, many people capturing photos of the same object are often positioned in the same or similar location. For example, two people could be standing next to each other while capturing photos of the same object. In

another example, two people could take turns standing in the same spot to capture photos of the same object. Accordingly, in many situations, numerous photos of the same object are substantially the same or 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 difference (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.

[0027] In block 204, system 102 collects attention information associated with the one or more 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. 3A, 3B, 3C, and 3D.. Other examples may include a statue in a plaza, a particular person's face in a group of people, etc.

[0028] 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.

[0029] 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 point of gaze (e.g., the user's line of sight). Example implementations of a device that is operable to track the gaze of a user are described in more detail in below in connection with FIG 4.

[0030] 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 gazes information of multiple users. As such, in some implementations, attention information is based on tracking gazes of one or more users. Example

implementations involving attention information are described in detail below in connection with FIG. 5.

[0031] In block 206, 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, because 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 are described in more detail below in connection with FIG. 5.

[0032] In block 208, system 102 guides users to capture photos of particular objects. In various implementations, system 102 guides users based on the attention map. For example, system 102 may guide a given user to an object (e.g., a monument) or portion of an object that is popular among tourists (e.g., receiving high attention from different people), yet the given user has not noticed. Example implementations directed to guiding users to capture photos of particular objects are described in more detail below in connection with FIG. 7.

[0033] FIG. 4 illustrates an example diagram of an eye tracking device 400 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.

[0034] In various implementations, eye tracking device 400 may use 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 400 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 400 may use a laser and laser technology to measure eye positions and eye movement relative to objects in the environment.

[0035] In some implementations, eye tracking device 400 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. 4 shows a pitch axis, yaw axis, and roll axis to illustrate how eye tracking device 400 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.

[0036] FIG. 4 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 408 and in the user's field of view 410. For ease of illustration, FIG. 4 shows three objects 302, 304, and 306. Any given scene in the user's field of view 410 may include any number of objects, including people.

[0037] In some implementations, system 102 may receive gaze information (e.g., parameter values associated with tracked parameters) directly from eye tracking device 400 or from any other one or more suitable storage locations. For example, in some implementations, eye tracking device 400 may send gaze information to system 102 as the user gazes at particular objects. In some implementations, when used with a camera device, eye tracking device 400 may send gaze information to system 102 as a camera device sends photos to system 102 (e.g., as the photos are captured). In some implementations, eye tracking device 400 may store gaze information local to the user's client device (e.g., if used with a dedicated digital camera, or if used with a mobile phone or other electronic device that has an integrated digital camera, etc.).

[0038] FIG. 5 illustrates an example simplified flow diagram for generating an attention map, according to some implementations. Referring to FIGS. 1, 4, and 5, a method is initiated in block 502, 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.

[0039] 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.

[0040] In block 504, 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.

[0041] 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.

[0042] 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.

[0043] In block 506, 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 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.

[0044] 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 of such users with the same object 304.

[0045] In block 508, 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 used interchangeably with the phrase “attention subvalue”.

[0046] 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.

[0047] 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.

[0048] 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.

[0049] 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.

[0050] In block 510, 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 be in turn based on attention values. An example attention map is described in more detail below in connection with FIG. 6.

[0051] FIG. 6 illustrates an example diagram of an attention map 600, 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.

[0052] 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.

[0053] FIG. 7 illustrates an example simplified flow diagram for guiding users to capture photos of particular objects, according to some implementations. As indicated above, system 102 guides users to capture photos of particular objects based on the attention map. In some implementations, a method is initiated in block 702, where system 102 determines one or more attention thresholds. System 102 may associate predetermined attention thresholds proportionally with levels of attention. 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.

[0054] In block 704, system 102 determines an attention value for each of the one or more objects. Various implementations directed to attention values are described above in connection with FIG. 5.

[0055] In block 706, system 102 compares the attention value for each object against one or more predetermined attention thresholds.

[0056] In block 708, system 102 determines 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. The particular criteria for meeting an attention threshold may vary and will depend on the particular implementation. 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 guiding users to the associated object.

[0057] In block 710, system 102 guides users to capture photos of particular objects based on the attention map. More specifically, in some implementations, system 102 may draw the attention of users to a particular object based on whether the attention value of the particular object meets a given predetermined attention threshold. For example, if the attention value of a given object meets a given predetermined attention threshold, system 102 may deem the object to have a sufficiently high attention level (e.g., attention value) to warrant a given user's attention, especially if that user has not noticed the object.

[0058] In some implementations, system 102 guides users to capture photos of particular objects that receive a level of attention that does not meet a predetermined attention threshold. For example, in some implementations, system 102 may guide users to objects that have not received a particularly high level of attention, but that users may be interested in nonetheless. For example, a museum may have a new piece that may be potentially interesting, yet new to the museum. In another example, an individual may have captured a new perspective of an old object (e.g., the backside of a museum piece, the underside of a bridge, etc.). System 102 may detect a relatively quick increase in attention level of a given object. Yet, the current attention level may still be considered a relatively low attention level. As such, users may appreciate system 102 guiding them to capture photos of less popular objects.

[0059] In some implementations, when system 102 guides users to capture photos of particular objects, system 102 may also inform users of optimal vantage points when capturing photos. In some implementations, system 102 may associate the same object with multiple attention levels based on different vantage points, where system 102 may associate each vantage point with a unique set of gaze parameters such as pitch, yaw, roll, line of sight, field of view, etc. In an example scenario, a given object (e.g., a monument) may have a lower attention level when viewed up close (e.g., 5 feet away) compared to being viewed from a distance (e.g., 30 feet away) where the entire object may be captured in a photo. In another example scenario, a given object (e.g., the same monument) may have a higher attention level when viewed from across a bridge compared to standing in front of the object. Accordingly, system 102 may deem certain vantage points as optimal vantage points based on a relatively higher attention level (e.g., where people tend to capture photos), and based on associated gaze parameters.

[0060] In some implementations, system 102 may draw the attention of a given user to a particular object by causing an attention indicator to be displayed on a display screen and/or

navigator screen. For example, if the user is viewing a scene via a lens (e.g., a viewing lens), system 102 may cause an attention indicator to be displayed on a display screen that is superimposed over the lens, integrated with the lens, or operates in connection with the lens. For example, in some implementations, system 102 may overlay the attention indication on top of the object as seen through the lens/on the display screen and/or navigator screen.

[0061] In some implementations, system 102 may display an attention map on the display screen and/or navigator screen such that the user may view different objects and associated attention information (e.g., attention indicators, attention values, etc.) in real-time while looking out at the scenery.

[0062] In some implementations, system 102 may create multiple predetermined attention thresholds corresponding with multiple levels of attention (e.g., levels of popularity) among users. As such, system 102 may indicate a level of popularity for the user to consider when the user decides whether to capture a picture of a given object. In some implementations, system 102 may navigate the user (e.g., using global positioning system technology, geotags in photos, geographical coordinates associated with gaze parameters, etc.) to optimal vantage points, where system 102 also indicates an attention level with the attention indication on the display screen.

[0063] In various implementations, system 102 receives photos taken by different users and aggregates the collected photos into a pool of photos, and enables the users to have access to the pool of photos. As a result, system 102 enables users to access a large pool of photos of the same object(s) based on various attention levels associated with the object(s). Furthermore, the ever growing and evolving nature of the pool of photos facilitates system 102 in better guiding users in capturing photos.

[0064] 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.

[0065] 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.

[0066] 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.

[0067] 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.

[0068] 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.

[0069] 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.

[0070] 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.

[0071] 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.

[0072] 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.

[0073] 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.

[0074] 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.

[0075] 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.

[0076] 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.

[0077] Implementations described herein provide various benefits. For example, implementations described herein provide users with a large pool of photos for a given object. Implementations also increase the size of the pool and evolve the pool by guiding users to capture particular photos based on an attention map.

[0078] FIG. 8 illustrates a block diagram of an example server device 800, which may be used to implement the implementations described herein. For example, server device 800 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 800 includes a processor 802, an operating system 804, a memory 806, and an input/output (I/O) interface 808. Server device 800 also includes a social network engine 810 and a media application 812, which may be stored in memory 806 or on any other suitable storage location or computer-readable medium. Media application 812 provides instructions that enable processor 802 to perform the functions described herein and other functions.

[0079] For ease of illustration, FIG. 8 shows one block for each of processor 802, operating system 804, memory 806, I/O interface 808, social network engine 810, and media application 812. These blocks 802, 804, 806, 808, 810, and 812 may represent multiple processors, operating systems, memories, I/O interfaces, social network engines, and media applications. In other implementations, server device 800 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.

[0080] 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.

[0081] 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.

[0082] 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.

[0083] 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:

collecting photos associated with one or more objects, wherein the photos are collected from a plurality of users, and wherein one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user;

collecting attention information associated with the one or more objects, wherein the attention information is based on tracking gazes of one or more of the plurality of users;

generating an attention map based on the attention information, wherein the generating of the attention map comprises:

receiving gaze information;

identifying the one or more objects;

associating the gaze information with each of the one or more objects; and

determining an attention value for each of the one or more objects based on the gaze information; and

guiding users to capture photos of particular objects based on the attention map.

2. A method comprising:

collecting photos associated with one or more objects, wherein the photos are collected from a plurality of users;

collecting attention information associated with the one or more objects;

generating an attention map based on the attention information; and

guiding users to capture photos of particular objects based on the attention map.

3. The method of claim 2, wherein one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user.
4. The method of claim 2, further comprising aggregating the collected photos into a pool of photos.
5. The method of claim 2, further comprising:
aggregating the collected photos into a pool of photos; and
enabling the plurality of users to have access to the pool of photos;
6. The method of claim 2, wherein the attention information is based on tracking gazes of one or more of the plurality of users.
7. The method of claim 2, wherein the attention map is based on tracking gazes of one or more of the plurality of users.
8. The method of claim 2, wherein the generating of the attention map comprises:
receiving gaze information;
identifying the one or more objects;
associating the gaze information with each of the one or more objects; and
determining an attention value for each of the one or more objects based on the gaze information.

9. The method of claim 2, wherein the guiding of users comprises guiding users to capture photos of particular objects that receive a level of attention that meets a predetermined attention threshold.
10. The method of claim 2, wherein the guiding of users comprises guiding users to capture photos of particular objects that receive a level of attention that does not meet a predetermined attention threshold.
11. The method of claim 2, wherein the guiding of users to capture photos of particular objects further comprises informing users of optimal vantage points when capturing photos.
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:
collecting photos associated with one or more objects, wherein the photos are collected from a plurality of users;
collecting attention information associated with the one or more objects;
generating an attention map based on the attention information; and
guiding users to capture photos of particular objects based on the attention map.
13. The system of claim 12, wherein one or more of the photos are captured using one or more respective devices that are operable to track a gaze of a user.

14. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising aggregating the collected photos into a pool of photos.
15. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising:
- aggregating the collected photos into a pool of photos; and
 - enabling the plurality of users to have access to the pool of photos;
16. The system of claim 12, wherein the attention information is based on tracking gazes of one or more of the plurality of users.
17. The system of claim 12, wherein the attention map is based on tracking gazes of one or more of the plurality of users.
18. The system of claim 12, wherein, to generate the attention map, the logic when executed is further operable to perform operations comprising:
- receiving gaze information;
 - identifying the one or more objects;
 - associating the gaze information with each of the one or more objects; and
 - determining an attention value for each of the one or more objects based on the gaze information.
19. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising guiding users to capture photos of particular objects that receive a level of attention that meets a predetermined attention threshold.

20. The system of claim 12, wherein the logic when executed is further operable to perform operations comprising guiding users to capture photos of particular objects that receive a level of attention that does not meet a predetermined attention threshold.

1/8

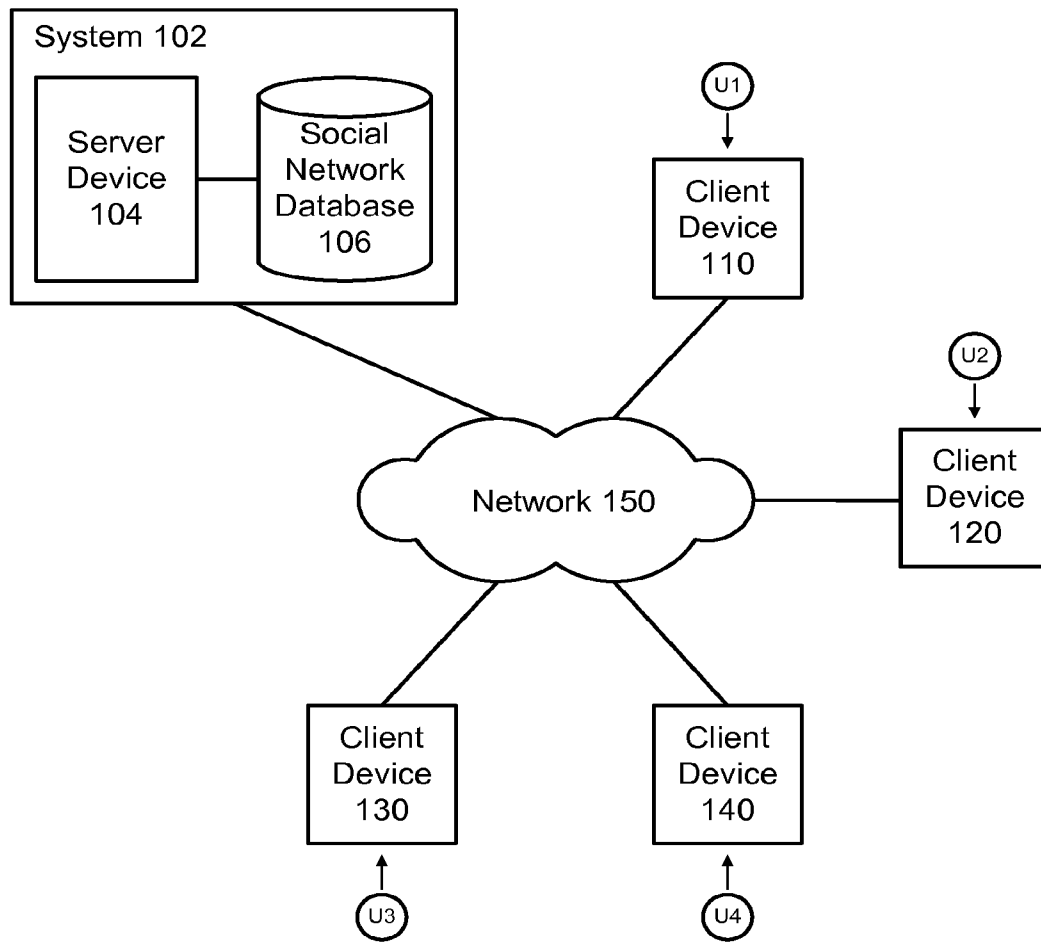
100

FIG. 1

2/8

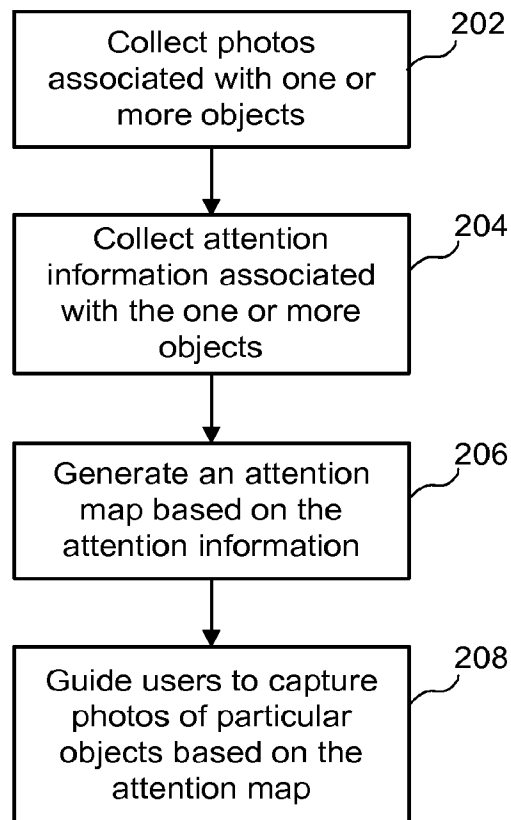


FIG. 2

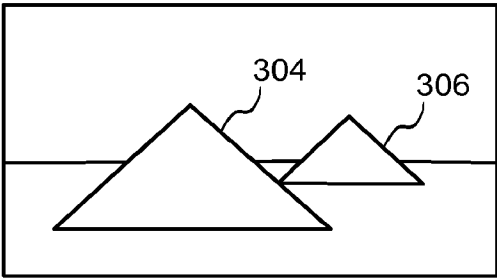


FIG. 3A

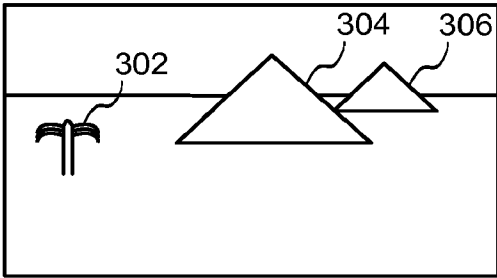


FIG. 3B

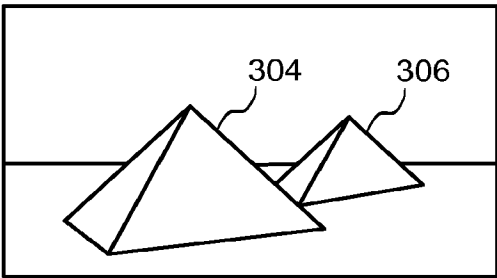


FIG. 3C

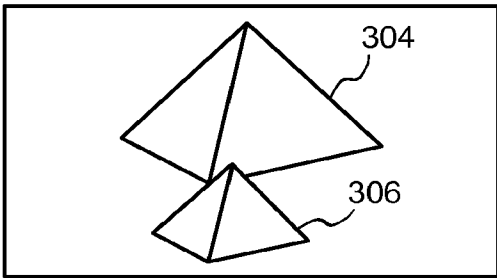


FIG. 3D

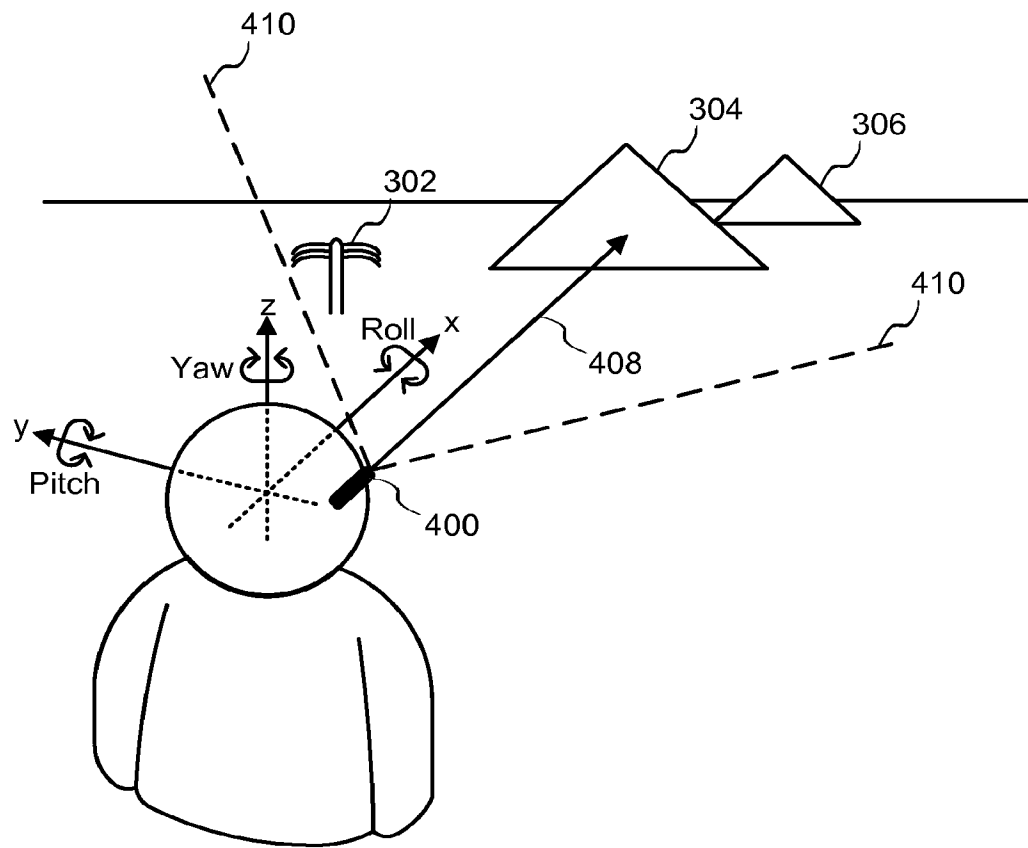


FIG. 4

5/8

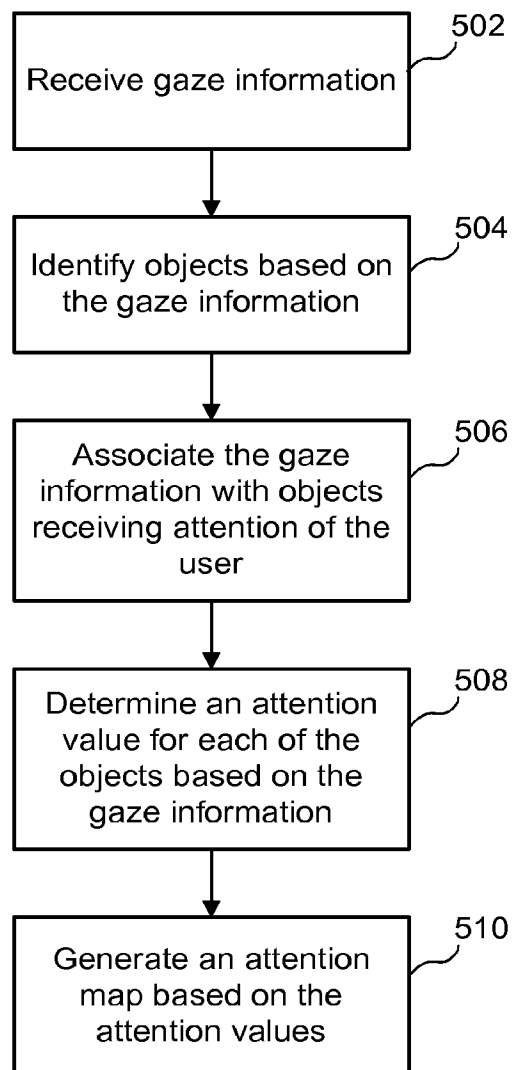
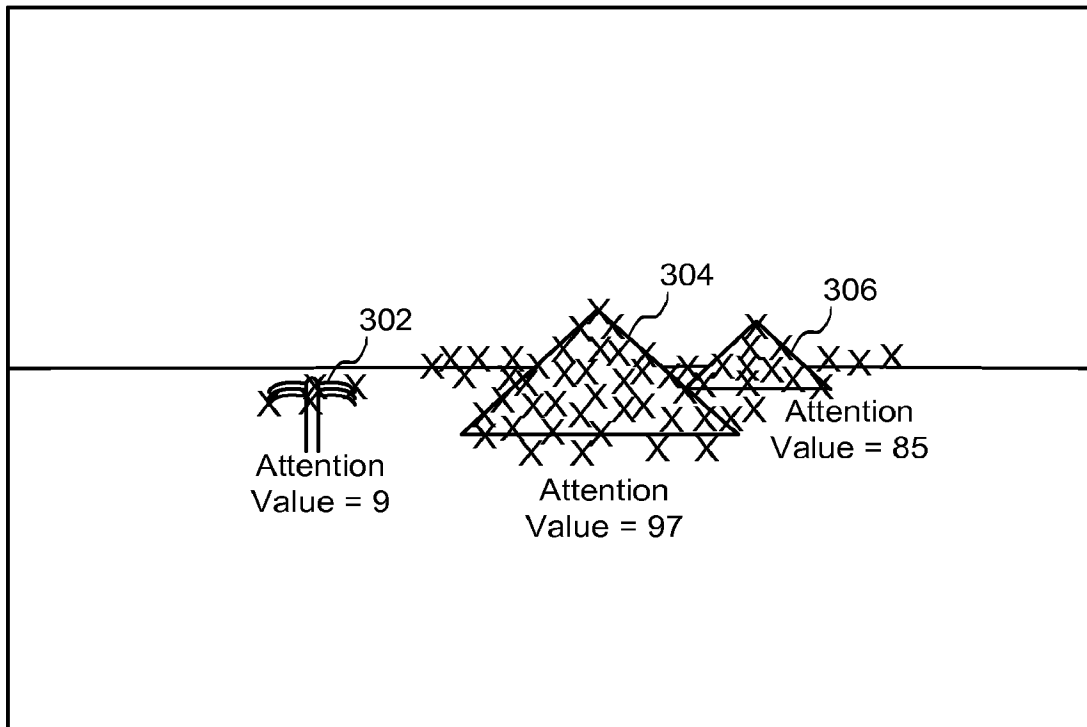


FIG. 5

6/8



600

FIG. 6

7/8

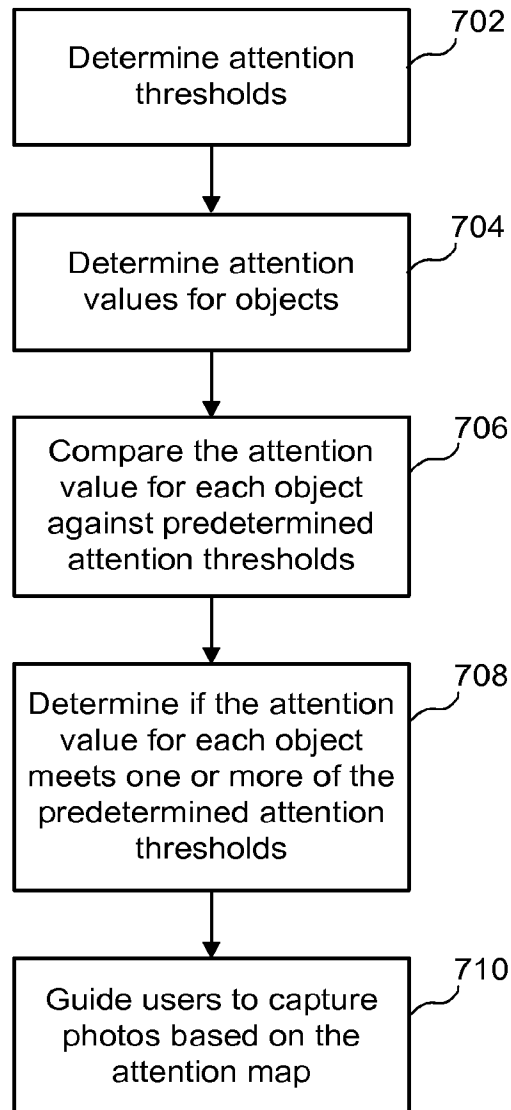


FIG. 7

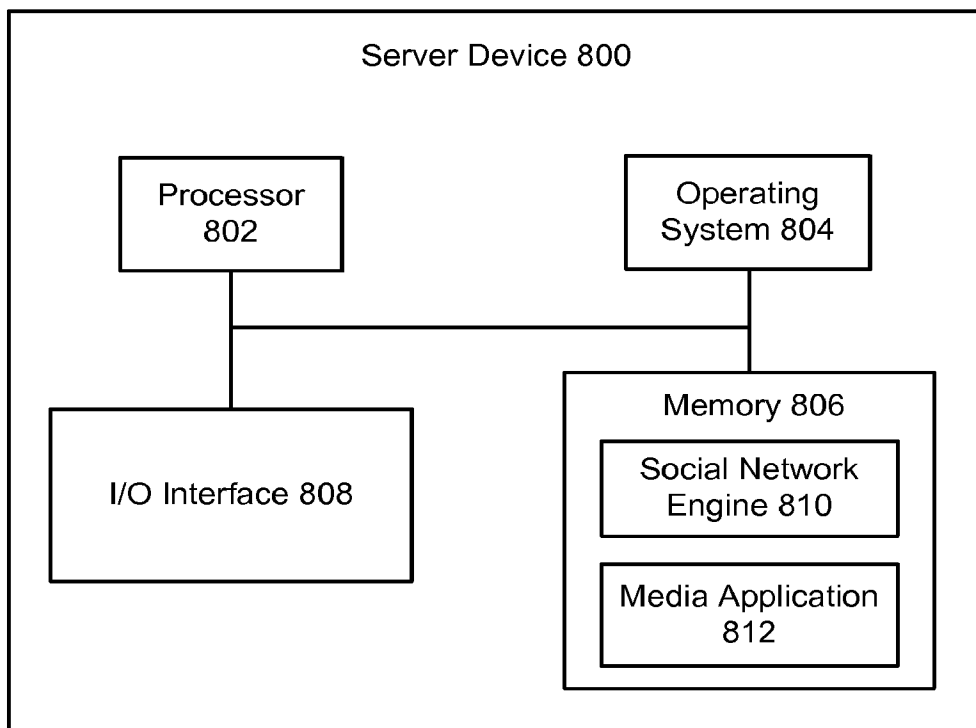


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/076288

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N5/232
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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 1 422 923 A2 (EASTMAN KODAK CO [US]) 26 May 2004 (2004-05-26) paragraphs [0005], [0007], [0010] - [0014], [0025], [0042], [0056] - [0059], [0084] - [0089]; claim 1; figures 1,13-15	1-20
A	----- US 2004/212712 A1 (STAVELY DONALD J [US] ET AL) 28 October 2004 (2004-10-28) abstract paragraphs [0021], [0022], [0032], [0036], [0037]; figures 1b,3 ----- -/-	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

9 April 2014

Date of mailing of the international search report

22/04/2014

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

Fournier, Christophe

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2013/076288

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	O Spakov ET AL: "Visualization of eye gaze data using heat maps", MEDICINE TECHNOLOGY, 2007, pages 55-58, XP055113222, Retrieved from the Internet: URL:http://vddb.library.lt/fedora/get/LT-eLABa-0001:J.04~2007~ISSN_1392-1215.N_2_74. PG_55-58/DS.002.1.01.ARTIC the whole document -----	1,2,12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2013/076288

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
EP 1422923	A2	26-05-2004	EP 1422923 A2	26-05-2004
			JP 4511821 B2	28-07-2010
			JP 2004181233 A	02-07-2004
			US 2004103111 A1	27-05-2004

US 2004212712	A1	28-10-2004	JP 2005033768 A	03-02-2005
			US 2004212712 A1	28-10-2004
