



- (51) International Patent Classification:
G06K 9/00 (2006.01)
- (21) International Application Number:
PCT/US2015/050908
- (22) International Filing Date:
18 September 2015 (18.09.2015)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:

| | | |
|------------|--------------------------------|----|
| 62/057,219 | 29 September 2014 (29.09.2014) | US |
| 62/057,223 | 29 September 2014 (29.09.2014) | US |
| 14/565,204 | 9 December 2014 (09.12.2014) | US |
| 14/565,236 | 9 December 2014 (09.12.2014) | US |

(74) Agents: **KRATZ, Rudy** et al.; Fitch, Even, Tabin & Flannery, LLP, 120 S. LaSalle Street, Suite 1600, Chicago, Illinois 60603 (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, 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.

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 14/565,236 (CON)
Filed on 9 December 2014 (09.12.2014)

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(71) Applicant: **SONY COMPUTER ENTERTAINMENT INC.** [JP/JP]; 1-7-1 Konan, Minato-ku, Tokyo 108-0075 (JP).

(72) Inventor; and

(71) Applicant (for US only): **BLACK, Glenn** [US/US]; c/o Sony Computer Entertainment America LLC, 2207 Bridgepointe Pkwy., San Mateo, California 94404 (US).

Published:

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

(54) Title: SCHEMES FOR RETRIEVING AND ASSOCIATING CONTENT ITEMS WITH REAL-WORLD OBJECTS USING AUGMENTED REALITY AND OBJECT RECOGNITION

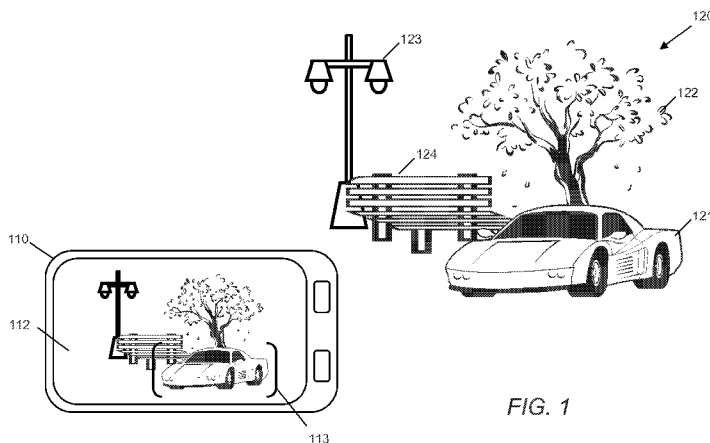


FIG. 1

(57) Abstract: A method includes identifying a real-world object in a scene viewed by a camera of a user device, matching the real-world object with a tagged object based at least in part on image recognition and a sharing setting of the tagged object, the tagged object having been tagged with a content item, providing a notification to a user of the user device that the content item is associated with the real-world object, receiving a request from the user for the content item, and providing the content item to the user. A computer readable storage medium stores one or more computer programs, and an apparatus includes a processor-based device.



SCHEMES FOR RETRIEVING AND ASSOCIATING CONTENT ITEMS WITH
REAL-WORLD OBJECTS USING AUGMENTED REALITY AND OBJECT
RECOGNITION

5

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of and claims benefit from United States Patent Application No. 14/565,236, filed on December 9, 2014, entitled "SCHEMES FOR RETRIEVING AND ASSOCIATING CONTENT ITEMS WITH REAL-
10 WORLD OBJECTS USING AUGMENTED REALITY AND OBJECT RECOGNITION," which claims the benefit of United States Provisional Patent Application No. 62/057,219, filed on September 29, 2014, entitled "SCHEMES FOR RETRIEVING AND ASSOCIATING CONTENT ITEMS WITH REAL-WORLD OBJECTS USING AUGMENTED REALITY AND OBJECT RECOGNITION," and
15 which also claims the benefit of United States Provisional Patent Application No. 62/057,223, filed on September 29, 2014, entitled "METHOD AND APPARATUS FOR RECOGNITION AND MATCHING OF OBJECTS DEPICTED IN IMAGES," the entire content and disclosures of which are all hereby fully incorporated by reference herein in their entireties.

20 This application is related to United States Patent Application No. 14/565,204, filed on December 9, 2014, entitled "METHOD AND APPARATUS FOR RECOGNITION AND MATCHING OF OBJECTS DEPICTED IN IMAGES," the entire content and disclosure of which is hereby fully incorporated by reference herein in its entirety.

25

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates generally to computer software applications, and more specifically to augmented reality apparatus and software.

30

2. Discussion of the Related Art

Augmented Reality (AR) is the general concept of modifying a view of reality with computer generated sensory input. While AR systems exist, they are mostly limited in application and content. For the most part, AR applications are single

purpose programs that function to provide developer generated and controlled content to the user.

SUMMARY OF THE INVENTION

5 One embodiment provides a method, comprising: identifying a real-world object in a scene viewed by a camera of a user device; matching the real-world object with a tagged object based at least in part on image recognition and a sharing setting of the tagged object, the tagged object having been tagged with a content item; providing a notification to a user of the user device that the content item is associated
10 with the real-world object; receiving a request from the user for the content item; and providing the content item to the user.

 Another embodiment provides an apparatus comprising: a processor-based device; and a non-transitory storage medium storing a set of computer readable instructions configured to cause the processor-based device to perform steps
15 comprising: identifying a real-world object in a scene viewed by a camera of a user device; matching the real-world object with a tagged object based at least in part on image recognition and a sharing setting of the tagged object, the tagged object having been tagged with a content item; providing a notification to a user of the user device that the content item is associated with the real-world object; receiving a request from
20 the user for the content item; and providing the content item to the user.

 Another embodiment provides a method, comprising: receiving a selection of a real-world object as viewed by a camera of a first user device from a user; receiving a content item to tag to the real-world object; capturing one or more images of the real-world object with the camera; receiving a sharing setting from the user, wherein
25 the sharing setting comprises whether the real-world object will be matched only with an image of the real-world object or with an image of any object sharing one or more common attributes with the real-world object; and storing the content item, the one or more images of the real-world object, and the sharing setting on a server, wherein the content item is configured to be retrieved by a second user device viewing an object
30 matching the real-world object.

 Another embodiment provides an apparatus comprising: a processor-based device; and a non-transitory storage medium storing a set of computer readable instructions configured to cause the processor-based device to perform steps

comprising: receiving a selection of a real-world object as viewed by a camera of a first user device from a user; receiving a content item to tag to the real-world object; capturing one or more images of the real-world object with the camera; receiving a sharing setting from the user, wherein the sharing setting comprises whether the real-world object will be matched only with an image of the real-world object or with an image of any object sharing one or more common attributes with the real-world object; and storing the content item, the one or more images of the real-world object, and the sharing setting on a server, wherein the content item is configured to be retrieved by a second user device viewing an object matching the real-world object.

A better understanding of the features and advantages of various embodiments of the present invention will be obtained by reference to the following detailed description and accompanying drawings which set forth an illustrative embodiment in which principles of embodiments of the invention are utilized.

BRIEF DESCRIPTION OF THE DRAWINGS

The above and other aspects, features and advantages of embodiments of the present invention will be more apparent from the following more particular description thereof, presented in conjunction with the following drawings wherein:

FIG. 1 is a diagram illustrating a device viewing a real-world scene in accordance with some embodiments of the present invention;

FIG. 2 is a diagram illustrating a method of virtually tagging a real-world object in accordance with some embodiments of the present invention;

FIG. 3 is a diagram illustrating a method of providing a content item tagged to a real-world object in accordance with some embodiments of the present invention;

FIGS. 4-6 are screenshots illustrating user interfaces in accordance with some embodiments of the present invention;

FIG. 7 is a diagram illustrating a device viewing a real-world scene in accordance with some embodiments of the present invention;

FIG. 8 is a block diagram illustrating a system that may be used to run, implement and/or execute any of the methods and techniques shown and described herein in accordance with the embodiments of the present invention.

DETAILED DESCRIPTION

Augmented Reality (AR) can be used to provide supplemental information about real-world objects for story telling. The following discussion will focus on example embodiments of the present invention that allow users to share stories based on real-world objects via an AR story telling system. In some embodiments, the AR story telling system provides users with a way to publish and share memories and stories about real-world objects with a great degree of customization and control. Specifically, in some embodiments a user can tag a real-world object with content items such as text, photo, video, and audio to share with other users. In some embodiments the system provides a user a notification when a real-world object in his/her surrounding has a content item tagged to or associated with it.

In some embodiments, by using Augmented Reality, users can "tag" objects with personal content such as photos, videos, or voice recordings. In some embodiments, the AR system can use image recognition and/or Global Positioning System (GPS) to detect objects in the physical world that have been "tagged" by other users. When a tagged object is detected, the system can render an alert over the object and allow the wearer or user of the AR device to review the information left in the tag (e.g., watch the video, view the photo, read the note, etc.). Tags can be shared based on "following" people through social networks. Tags can last forever, expire after a user selected amount of time, or be limited to a specific number of views (e.g., only after one person views the tag it is removed).

An example application of an embodiment of the present invention may include a user tagging his favorite baseball cap with a video of himself catching a home run ball. His friends encountering a similar hat could then see that it was tagged, and watch the video. As another example, a user tags the menu of his or her favorite restaurant with a voice note of the user's favorite items. Other users who see the menu can then hear their friend's recommendation. And as another example, a user tags a park bench with a childhood memory of having a picnic there. Other users later walking by could then participate in the nostalgic story.

Some embodiments of the present invention focus on using image recognition to provide a physical reference point for digital data, which allows alerts to be triggered based on a user's interaction with the real world. In some embodiments, to place a tag the user locates the desired object through an AR device, then chooses the type of tag (e.g., photo, video, audio, text, etc.) and composes the "post" (e.g., title,

description, hashtags, etc.). When the AR system detects a tagged object in the real world, it can use any kind image overlay to alert the user and direct the user toward the information. In this way, some embodiments of the present invention can provide social interaction and/or information sharing through augmented reality.

5 Referring to FIG. 1, there is illustrated a device viewing a real-world scene in accordance with some embodiments of the present invention. In FIG. 1, a user device 110 is used to view a real-world scene 120. The real-world scene 120 includes real-world objects such as a car 121, a tree 122, a lamppost 123, and a bench 124. The user device 110 may be a portable electronic device such as a Smartphone, a tablet
10 computer, a mobile computer, a pad-like device, a head-mounted display, and/or a wearable device. The user device 110 includes a display screen 112 and other input/output devices such as a touch screen, a camera, a speaker, a microphone, a network adapter, a GPS sensor, a tilt sensor, etc. The display screen 112 may display an image of the real-world scene 120 as viewed by an image capture device of the
15 user device 110 along with computer generated images to enable story telling via AR. For example, in FIG. 1, a bracket 113 is shown in the display screen 112 around an image of the car 121 to highlight the car 121. The user may select an object in the display screen 112 to interact with virtually. For example, the user may tag one of the objects with a content item or select an object to view content item(s) that has been
20 tagged to it. Further details of methods for providing AR story telling are provided below with reference to FIGS. 2-8.

While FIG. 1 shows a user device 110 with a display screen 112 that reproduces an image of the real-world scene, in some embodiments, the user device
25 110 may include a transparent or translucent lenses such as those used in glasses-type and/or head-mounted display devices. With such devices, a user may view the real-world scene 120 directly through one or more lenses on the device, and the computer generate images are displayed in a way as to overlay the direct view of the real-world scene to produce a combined view similar to those shown in FIGS. 1 and 4-6.

Referring to FIG. 2, there is illustrated an example of a method 300 for
30 virtually tagging a real-world object that operates in accordance with some embodiments of the present invention. In some embodiments, steps of method 300 may be performed by one or more server devices, a user device, or a combination of server and user devices.

In step 210, a user selection of a real-world object in a scene viewed by a camera of a first user device is received. The scene viewed by a camera of a user device may be an image captured in real-time or an image that was captured with a time delay. In some embodiments, a user selects an object by selecting a region of the image corresponding to the object. The system may use the image of the selected region for object recognition. In some embodiments, the user may select an object by zooming or framing the image of the real scene such that the object is the dominate object in the image. In some embodiments, the system may indicate, in the image of the real-world scene, objects that it automatically identifies through object recognition. For example, as shown in FIG. 1, the system may put a bracket around an identified object, in this case, a car 121, that the user can select. In some embodiments, the user may first select to a region of an image, and the system will then determine an object associated with the selected region of an image. For example, the user may touch the region above the image of the car 121, and the system may identify the tree 122 in that region. In some embodiments, the selected real-world object may be previously unknown to the system, and the system may isolate the object in the image by using image processing techniques such as edge detection. For example, the system may not recognize the lamppost 123 as a lamppost but is able to recognize it as a separate object based on edge detection. In some embodiments, the user may be prompted to enter attributes, such as name, model number, object type, object category, etc. to describe the object such that the system may “learn” of a previously unknown object.

In step 220, the system receives a content item to tag to, or associate with, the real-world object. The content item may be one or more of a text comment, an image, an audio clip, a video clip, a hyperlink, etc. The system may provide interfaces for a user to enter, upload, and/or capture the content item. For example, the system may provide a comment box and/or buttons to select a file to upload.

In step 230, one or more images of the real-world object are captured with the camera of the user device. The captured image may be later processed to determine additional information related to the selected object. The real-world object may be identified based on the shape, color, relative size, and other attributes of the object identifiable from an image of the real-world object. For example, the system may identify the car 121 as a car based on its shape and color. In some embodiments, the

system may further identify other attributes of the object. For example, the system may identify the make, model, and model year of the car 121 based comparing the captured image to images in a known objects database. In some embodiments, the captured image may include specific identifying information of an object. For example, the image may include a license plate number of the car 121. In other examples, the captured image may include serial number or model number of electronic devices, unique identifier of limited edition of collectable items, etc.

In some embodiments, step 230 is performed automatically by the system without requiring user input. For example, the system may acquire one or more images of the real-world object when the user selects the object in step 210. In some embodiments, the system may prompt the user to take images of the real-world object. For example, if the object selected in step 210 is not one that is recognized by the system, the system may prompt the user to take images of the object from different angles so that it may be better matched with other images of the object. In some embodiments, after step 240, the system may prompt the user to capture additional images of the object based on the sharing settings. For example, if the user wants to tag the content item specifically to only the car 121 and not any other car that is similar in appearance, the system may ask for an image of unique identifying information, such as the car's license plate, if the license plate is not visible in the current view of the car.

Generally, the system may utilize variously known computer object recognition techniques to recognize objects in the image of the real-world scene. In some embodiments, object recognition may use appearance-based methods which compares an image with reference images of known objects to recognize an object. Examples of appearance-based methods include edge matching, grayscale matching, histograms of receptive field responses, etc. In some embodiments, object recognition may use feature-based methods that rely on matching object features and image features. Examples of feature-based methods include pose clustering, geometric hashing, scale invariant feature transform, interpretation trees, etc. The system may use one or more object recognition methods in combination to enhance the accuracy of object identification.

In step 240, sharing setting is received from the user. In some embodiments, the sharing setting includes whether the real-world object will be matched only with

an image of the real-world object or with an image of any object sharing one or more common attributes with the real-world object. For example, if the user configures the sharing setting such that the real-world object will be matched only with an image of the specific real-world object, the system may determine and/or request the user to enter a unique identifier associated with the real-world image. The unique identifier may be based on a location of the object, a serial number the object, a license plate number, or other identifying text or image of the object. For example, the car 121 may be uniquely identified based on its license plate number and/or painted text or graphic on its body, the bench 124 and lamppost 123 may be uniquely identified based on their location information and/or an inscription, a baseball cap may be uniquely identified based on a signature on the cap or the locations of wear and tear on the cap, a trophy may be uniquely identified by the text on the trophy, etc. In some embodiments, the system may automatically determine these identifying attributes. For example, if the system detects a license a plate, name, or serial number in the image, the system may use that information as the identifying attribute. In another example, if the system recognizes the real-world object as a stationary object, such as the lamppost 123, a building, a statue, etc., the system may use the location of the object or the location of the device that captured the image as the identifying attribute. In some embodiments, the system may prompt the user to designate the identifying attribute. For example, the system may prompt the user to take a picture of or enter the serial number of an electronic device as the identifying attribute.

If the user selects to have the real-world object match with an image of any object having one or more common attributes with the real-world object, the system may prompt the user to select the matching attribute(s). Matching attributes refer to attributes that an object must have to be matched with a tagged object to receive content items tagged to the tagged object. In some embodiments, the matching attributes may be based on the appearance of the object, attributes of the object that are determined using image recognition, and/or the location of the object. For example, the user may configure the car 121 to match with any car, cars with the same color, cars by the same manufacturer, cars of the same model, or cars of the same model and model year, etc. While the color of the car may be determined based on the image alone, attributes such as the manufacturer of the car may be determined using image recognition. For example, the system may match the image of the car 121

with a reference image of a known car to determine the car's make, model, and model year. The common attribute can further be a location-based attribute. For example, the user may configure the bench 124 to match with any bench in the same park, any bench in the same city, any bench in a municipal park, any bench above a certain
5 altitude, etc. Location attribute information of the real-world object may be determined based on the devices' GPS information and/or the location metadata information of the captured image of the real-world object.

In some embodiments, the system may perform image recognition and/or location analysis on the selected real-world object and present the user with a list of
10 identified attributes for selection. The user can configure the sharing setting by selecting attributes used for matching the real-world object to another object from the list. For example, if the user selects the car 121 on the user device 110, the system may generate a list of attributes such as:

Object type: automobile
15 Color: gray
Make: DeLorean
Model: DMC-12
Model year: 1981
Location: Hill Valley, California

20 The user may then select one or more of the identified attributes to configure the matching attributes in the sharing setting. For example, the user can select make and model as the matching attributes, and the content item tagged to the car 121 will only be matched to another image of a DeLorean DMC-12. In some embodiments, the user can manually enter attributes of the object to configure the sharing setting. For
25 example, the user may configure the location-based sharing setting to be any location within 20 miles or 50 miles from Hill Valley, California, etc.

In some embodiments, the sharing settings further includes social network based sharing settings that controls who can view the content item tagged to the real-world object based on the user's social network connections. For example, the user
30 can elect to share the content item to all "followers," only with friends, a selected group of users, or members of an existing group within a social network, etc.

While steps 220, 230, and 240 are shown sequentially in FIG. 2, in some embodiments, these steps can be performed in different orders. For example, in some

embodiments, the user may configure the sharing setting prior to entering the content item to tag to the object. In some embodiments, one or more images of the real-world item may be captured prior to the content item being received or after the sharing setting is entered. In some embodiments, anytime during steps 220-240, the user may
5 return to one of the previously performed steps to edit and/or enter information. In some embodiments, an object tagging user interface is provided to allow the user to perform two or more of steps 220, 230, and 240 in any order. In some embodiments, an image of the real-world object is captured when the user selects the real-world object in step 210. The system then determines whether more images are needed
10 based on the received sharing settings in step 240, and prompt the user to capture one or more images as needed.

In step 250, the content item, the image of the real-world object, and the sharing setting entered in steps 220-240 are stored on a network accessible database such that the content item is configured to be retrieved by a second user device
15 viewing an object matching the real-world object based on the sharing setting. Thus, in some embodiments, the content item is now said to be tagged to, or associated with, the real-world object. In some embodiments, the content item, the image of the real-world object, and the sharing setting are stored as they are entered in steps 220-240. In some embodiments, after step 250, the user may return to edit or delete the
20 information entered in steps 210-240 using the user device and/or another user electronic device by logging into a user profile.

Referring to FIG. 3, there is illustrated an example of a method 300 that operates in accordance with some embodiments of the present invention. In some embodiments, steps of method 300 may be used to provide AR story telling as
25 described herein. In some embodiments, method 300 may be performed by one or more server devices, a user device, or a combination of server and user devices.

In step 310, a real-world object in a scene viewed by a camera of a user device is identified based on image recognition. The system may utilize any known image recognition techniques. The scene viewed by a camera of a user device may be an
30 image captured in real-time or an image that was captured with a time delay. In some embodiments, the user device uploads an image of the real-world scene to a server and the server identifies the real-world object(s) in the image. In some embodiments, the user device identifies attributes of the real-world object (such as color, gradient

map, edge information) and uploads the attributes to the server to reduce network delay. The real-world object may be identified based on the shape, color, relative size, and other attributes of the object identifiable from an image of the real-world object. In some embodiments, object recognition may be based on multiple frames of a video image. In some embodiments, an object may be identified based in part on the GPS location of the user device when the image is captured. For example, the system may be able to distinguish two visually similar benches based on the location of the device. In some embodiments, in step 310, the system will attempt to identify all objects viewed by a camera of the user device.

10 In step 320, a real-world object is matched to a tagged object. The tagged object may be a real-world object tagged through the process shown in FIG. 2. As such, in some embodiments the tagged object has a content item associated with it. The matching in step 320 is based on the image recognition of the real-world object and the sharing setting of the tagged object. In some embodiments, the matching may be based on comparing the attributes of the real-world object with tagged objects in a tagged object database. For example, if the real-world object in step 310 is a red sports car, the real-world object may be matched with a tagged object designated to be matched with red cars, sports cars, etc, but would not be matched with a green sports car having a sharing setting that restricts sharing to only green cars. In some 15 embodiments, the matching is also based on the social network sharing setting of the tagged object. For example, if a tagged object has been designated to only be shared with the author's "friends," and the user using the device to view the real-world object is not connected with the author of the content item in the social network service, then the real-world object would not be matched with that particular tagged object.

25 In some embodiments, the matching in step 320 may also be based on the location of the device viewing the real-world object. For example, if a tagged object has been configured to only be shared with items within a geographical region (GPS coordinate, address, neighborhood, city, etc.) that tagged object would not be matched with image of real-world objects taken outside of that region. In some embodiments, step 320 is performed for each object in a real-world scene identified in step 310.

30 In step 330, once a matching tagged object is found in step 320, a notification is provided to the user. The notification may be provided by one or more of a sound notification, a pop-up notification, a vibration, and an icon in a display of the scene

viewed by the camera of the user device. For example, the system may cause a graphic indicator to appear on the screen to indicate to the user that an item in the real-world scene has been tagged. The graphic indicator may be an icon overlaying a view of the object, a bracket that surround the object, or a color/brightness change to highlight the object, etc. In some embodiments, if an object has been tagged with more than one content item, the notification may include an indicator for each content item. For example, if an item has been tagged by two different users or tagged with different type of content item, the graphic indicator may also indicate the author and/or content type for each content item.

10 In some embodiments, a notification may be provided based on location information alone, which then prompts the user to search for the tagged object in the surrounding. For example, if the system determines that there is a tagged object nearby based on matching the GPS coordinates of the device and the tagged object, the system may cause the device to vibrate or beep. The user may then point the camera of the device at surrounding objects to find a tagged object. The system may cause the graphic indicator to be displayed on the user device when the tagged object comes within the view of the camera.

In step 340, a request for the content item is received. The request may be generated by the user device when the user selects the real-world object and/or a graphical indicator in the user interface. For example, the user may tap an icon displayed on a touch screen of the user device to select a tagged object or a content item. In some embodiments, if an object has been tagged with multiple content items, the selection may include a selection of a content item and the request may include an indication for the specific content item. In some embodiments, the request may be generated by the user device automatically and the content item may be temporarily stored or cached on the user device. In some embodiments, the request may be required only for certain content item types. For example, the system may transmit text and image content items automatically upon detecting a tagged object, but stream audio and video type content items to the device only upon receiving a request from the user.

30 In step 350, a content item is provided to the user device. In some embodiments, all content items associated with a selected tagged object are provided to the user device. In some embodiments, only the selected content item is provided.

In some embodiments, video and audio type content items are provided in a streaming format.

In some embodiments, after step 350, the viewing user may have the option to comment on the content item and/or the tagged object by providing his/her own
5 content item.

Referring to FIG. 4, there is illustrated example screenshots of user interfaces for providing AR story telling. An object selection user interface 410 displays an image of a real-world scene viewed by a camera of the device, and may be used by a user to tag real-world objects with virtual content items. In some embodiments, the
10 user can select a region of the real-world scene associated with an object. For example, a user may tap an area, draw a line around an area, and/or click and drag a rectangle around an area to select an area. In some embodiments, the system captures an image of the selected region and attempts to identify an object in the selected region. In some embodiments, the individual objects may be identified based on edge
15 detection. In some embodiments, the system identifies one or more objects recognized by the system in the real-world scene to the user. For example, the system may overlay brackets 414 or other graphic indicators on the image of the real world scene. The graphic indicators may be selectable by the user.

After an object has been selected in the real-world scene view, a tagging user
20 interface 420 may be displayed. The tagging user interface 420 may include an image 422 of the object selected to be tagged. The tagging user interface 420 may further include various options to enter and configured the content item. For example, the user may use an “add content” button 424 to enter or attach a text comment, an image, an audio file, and/or a video file. In some embodiments, the user interface also allows
25 the user to record a new image, audio, and/or video to tag to the object without leaving the user interface and/or application. A user may use a “sharing setting” button 426 to configure various sharing settings discussed herein. For example, the use may select matching attributes, configure social network sharing setting, and/or enter location restrictions in the tagging user interface 420.

Referring to FIG. 5, there is illustrated additional example screenshots of user
30 interfaces for providing AR story telling. A tagged object notification user interface 510 may be displayed to a user viewing a real-world scene through the user device. The tagged object notification user interface 510 may include one or more graphic

indicators to identify tagged objects in the view of the real-world scene. For example, icon 512 indicates that the lamppost has been tagged with a text comment, icon 514 indicates that the bench has been tagged with an audio clip, and icon 516 indicates that the car has also been tagged with a text comment. In some embodiments, other
5 types of graphical indicators are used to indicate that an object has a viewable content item. For example, a bracket similar to the bracket 414 in FIG. 4, an arrow, a blinking icon, an icon containing the profile image of the author of the content item, a thumbnail of the content item, etc. may be used. In some embodiments, the object or the surrounding's color and/or shading may be changed to distinguish a tagged object
10 from its surrounding. For example, FIG. 6 shows a car 610 that is highlighted against the shaded background 620.

In some embodiments, the tagged object notification user interface 510 may be the same interface as the tagging user interface 410. For example, when the user opens a program or an application (or "app") on the user device, the user is presented
15 with an image of the real-world scene viewed by the camera of the user device. If any of the objects in the real-world scene are tagged, a notification may be provided. The user may also select any object in the scene to virtually tag a new content item to the object.

When a tagged object or a content item is selected in the tagged object
20 notification user interface 510, a content item viewing user interface 520 is shown. In the content item viewing user interface 520, the content item selected from the tagged object notification user interface 510 is retrieved and displayed to the user. The content may be displayed in full screen or as an overlay 522 of the image of the real world scene as shown in FIG. 5. The content item may include one or more of a text
25 comment, an image, an audio clip, and a video clip. In some embodiments, if more than one content item is associated with a tagged object, multiple content items from different authors may be displayed at the same time.

The object selection user interface 410, the tagging user interface 420, the tagged object notification user interface 510, and the content item viewing user
30 interface 520 may each be part of a program or application running on a user device and/or a server device. In some embodiments, the user interfaces may be parts of an app on a Smartphone that communicates with an AR server to provide the images and the contents of the user interfaces that enables AR storytelling. In some embodiments,

the user interfaces are generated and provided entirely at the server.

Referring to FIG. 7, there is illustrated a diagram showing a glasses-type user device viewing a real-world scene in accordance with some embodiments of the present invention. In some embodiments, the above described user interfaces may also
5 be implemented with glasses-type and/or head-mounted type display devices. With such devices, a user may view the real-world 710 scene directly through a lens 720 on the device, and computer generated images 722 may be projected on the lens and/or wearer's eyes in a way as to overlay the direct view of the real-world scene to produce a combined view as shown in FIG. 7. The overlay images 722 may indicate a selected
10 or selectable object for tagging and/or identify a tagged object as described herein. The user's selection of an object may be detected through eye gaze detection and/or voice command recognition. In some embodiments, the object may be selected through a touch pad on the device. For example, the user may slide over the touch pad to "scroll" through the selectable objects viewed through the lens of the device. In
15 some embodiments, the user may add a text comment or an audio clip content item through spoken command received by a microphone of the head-mounted display. In some embodiments, selections can be made by gesturing in the view of the real-world scene which are captured and interpreted by a camera on the device. While only the frame of one lens is shown in FIG. 7, it is understood that the same principle applies
20 to head-mounted devices with two lenses.

FIG. 8 is a block diagram illustrating a system that may be used to run, implement and/or execute any of the methods and techniques shown and described herein in accordance with some embodiments of the present invention. The system 800 includes user devices 810, 820, and 830, an AR server 840, and a third
25 party/social networking server 870 communicating over a network 805.

The user device 810 may be any portable user devices such as a Smartphone, a tablet computer, mobile computer or device, pad-like device, a head-mounted display, and/or a wearable device. The user device 810 may include a processor 811, a memory 812, a network interface 813, a camera 814, a display 815, one or more other
30 input/output devices 816, and a GPS receiver 817. The processor 811 is configured to execute computer readable instructions stored in the memory 812 to facilitate one of more steps of the methods described with reference to FIGS. 2-3. The memory 812 may include RAM and/or a hard drive memory device. The network interface 813 is

configured to transmit and receive data to and from at least the AR server 840 through the network 805. The camera 814 may be any image capture device. The display 815 may be a liquid crystal display (LCD), a light-emitting diode (LED) display, an organic light emitting diode (OLED) display, a liquid crystal on silicon (LCoS) display, a LED illuminated display, and the like. In some embodiments, the display 815 is a touch display that is configured to receive input from the user. In some embodiments, the display 815 is a head-mounted display that projects images to a lens or into the wear's eyes. The display may be used to display variously user interfaces described with reference to FIGS. 1-7 above. The GPS receiver 817 may be configured to detect GPS signals to determine a coordinate. The determined GPS coordinate can be used by the system 800 to determine the location of a device, an image, and/or an object. The user device 810 may further include other input/output devices 816 such as a microphone, a speaker, a tilt sensor, a compass, a USB port, a secondary camera, a graphic processor, etc. The input/output devices 816 may be used to enter and playback content items. In some embodiments, the tilt sensor and the compass are used by the system to better position the computer generated graphics over the image of the real-world scene, in addition image recognition.

User devices 820 and 830 may be user devices similar to user device 810 and may be operated by one or more other users to access the AR story telling system 800. It is understood that while three user devices are shown in FIG. 8, the system 800 may be accessed by any number of users and user devices.

The AR server 840 includes a processor 842, a memory 841, and a network interface 844. The processor 842 is configured to execute computer readable instructions stored in the memory 841 to perform one or more steps described herein with reference to FIGS. 2-3. The AR server 840 may be connected to or include one or more of an object database 860 and a content item database 850. The content item database 850 may store information relating to each of the tagged real-world objects, content item(s) tagged to each object, and the sharing settings associated with the content item and/or tagged object. In some embodiments, step 250 in FIG. 2 is performed by storing the information into the content item database 850. The object database 860 may include a database of known objects used by the object recognition algorithm of the system to identify one or more real-world objects and their attributes. In some embodiments, image(s) of the real-world object taken by the camera 814 of

the user device 810 is compared to one or more images of objects in the object database 860. When there is a match, the object database 860 may provide further information relating to the object, such as object name, object type, object model, etc.

In some embodiments, one or more of the object database 860 and the content item database 850 may be part of the AR server 840. In some embodiments, the object database 860 and the content item database 850 may be implemented as a single database. In some embodiments, the AR server 840 communicates with one or more of the object database 860 and the content item database 850 also via the network 805. In some embodiments, the object database 860 may be maintained and controlled by a third party.

In some embodiments, the object database 860 may “learn” new objects by receiving user provided images and attributes, and adding the user provided information to its database. For example, when a user takes a picture of a car and enters its make and model, the object database may subsequently be able to recognize another image of the same car and be able to provide information on its make and model. While only one AR server 840 is shown, it is understood that the AR server 840, the object database 860, and content item database 850 may be implemented with one or more physical devices connected through a network.

The social networking server 870 provides social networking functions for users to connect to each other and build social networks and groups. The social networking server 870 may be part of the AR server 840 or a third party service. The connections and groups built in the social networking service may be used to configure the sharing setting discussed herein. For example, if a content item is configured to be shared only with “friends” of the author, the AR server 840 may query the social networking server 870 for the social networking status between two users to determine whether a content item should be provided to a second user. In some embodiments, when a user is configuring sharing settings, information can be retrieved from the social networking server 870 such that the user can select among his/her friends and/or social groups to share the content item.

In some embodiments, one or more of the embodiments, methods, approaches, and/or techniques described above may be implemented in one or more computer programs or software applications executable by a processor based apparatus or system. By way of example, such processor based system may comprise the

processor based apparatus or system 800, or a computer, entertainment system, game console, graphics workstation, server, client, portable device, pad-like device, etc. Such computer program(s) may be used for executing various steps and/or features of the above-described methods and/or techniques. That is, the computer program(s) may be adapted to cause or configure a processor based apparatus or system to execute and achieve the functions described above. For example, such computer program(s) may be used for implementing any embodiment of the above-described methods, steps, techniques, or features. As another example, such computer program(s) may be used for implementing any type of tool or similar utility that uses any one or more of the above described embodiments, methods, approaches, and/or techniques. In some embodiments, program code macros, modules, loops, subroutines, calls, etc., within or without the computer program(s) may be used for executing various steps and/or features of the above-described methods and/or techniques. In some embodiments, the computer program(s) may be stored or embodied on a computer readable storage or recording medium or media, such as any of the computer readable storage or recording medium or media described herein.

Therefore, in some embodiments the present invention provides a computer program product comprising a medium for embodying a computer program for input to a computer and a computer program embodied in the medium for causing the computer to perform or execute steps comprising any one or more of the steps involved in any one or more of the embodiments, methods, approaches, and/or techniques described herein. For example, in some embodiments the present invention provides one or more non-transitory computer readable storage mediums storing one or more computer programs adapted to cause a processor based apparatus or system to execute steps comprising: identifying a real-world object in a scene viewed by a camera of a user device; matching the real-world object with a tagged object based at least in part on image recognition and a sharing setting of the tagged object, the tagged object having been tagged with a content item; providing a notification to a user of the user device that the content item is associated with the real-world object; receiving a request from the user for the content item; and providing the content item to the user. In another example, in some embodiments the present invention provides one or more non-transitory computer readable storage mediums storing one or more computer programs adapted to cause a processor based

apparatus or system to execute steps comprising: receiving a selection of a real-world object as viewed by a camera of a first user device from a user; receiving a content item to tag to the real-world object; capturing one or more images of the real-world object with the camera; receiving a sharing setting from the user, wherein the sharing
5 setting comprises whether the real-world object will be matched only with an image of the real-world object or with an image of any object sharing one or more common attributes with the real-world object; and storing the content item, the one or more images of the real-world object, and the sharing setting on a server, wherein the content item is configured to be retrieved by a second user device viewing an object
10 matching the real-world object.

While the invention herein disclosed has been described by means of specific embodiments and applications thereof, numerous modifications and variations could be made thereto by those skilled in the art without departing from the scope of the invention set forth in the claims.

15

CLAIMS

What is claimed is:

1. A method comprising:
 - identifying a real-world object in a scene viewed by a camera of a user device;
 - matching the real-world object with a tagged object based at least in part on image recognition and a sharing setting of the tagged object, the tagged object having been tagged with a content item;
 - providing a notification to a user of the user device that the content item is associated with the real-world object;
 - receiving a request from the user for the content item; and
 - providing the content item to the user.
2. The method of claim 1, wherein the sharing setting comprises whether the tagged object will be matched with only with an image of the tagged object or with an image of any object sharing one or more common attributes with the tagged object.
3. The method of claim 1 or 2, wherein the matching of the real-world object with the tagged object is further based on one or more of a location of the user device and a social network connection between the user and an author of the content item.
4. The method of claim 1 or 2, wherein the notification comprises one or more of a sound notification, a pop-up notification, a vibration, and a icon in a display of the scene viewed by the camera of the user device.
5. The method of claim 1 or 2, wherein the content item comprises one or more of a text comment, an image, an audio clip, and a video clip.
6. An apparatus comprising:
 - a processor-based device; and
 - a non-transitory storage medium storing a set of computer readable instructions configured to cause the processor-based device to perform steps

comprising:

identifying a real-world object in a scene viewed by a camera of a user device;
matching the real-world object with a tagged object based at least in part on image recognition and a sharing setting of the tagged object, the tagged object having been tagged with a content item;
providing a notification to a user of the user device that the content item is associated with the real-world object;
receiving a request from the user for the content item; and
providing the content item to the user.

7. The apparatus of claim 6, wherein the sharing setting comprises whether the tagged object will be matched with only with an image of the tagged object or with an image of any object sharing one or more common attributes with the tagged object.

8. The apparatus of claim 6 or 7, wherein the matching of the identified real-world object with the tagged object is further based on one or more of a location of the user device and a social network connection between the user and an author of the content item.

9. The apparatus of claim 6 or 7, wherein the notification comprises one or more of a sound notification, a pop-up notification, a vibration, and a icon in a display of the scene viewed by the camera of the user device.

10. The apparatus of claim 6 or 7, wherein the content item comprises one or more of a text comment, an image, an audio clip, and a video clip.

11. A method comprising:
receiving a selection of a real-world object as viewed by a camera of a first user device from a user;
receiving a content item to tag to the real-world object;
capturing one or more images of the real-world object with the camera;
receiving a sharing setting from the user, wherein the sharing setting comprises whether the real-world object will be matched only with an image of the

real-world object or with an image of any object sharing one or more common attributes with the real-world object; and

storing the content item, the one or more images of the real-world object, and the sharing setting on a server, wherein the content item is configured to be retrieved by a second user device viewing an object matching the real-world object.

12. The method of claim 11, wherein the content item comprises one or more of a text comment, an image, an audio clip, and a video clip.

13. The method of claim 11, wherein the sharing setting further comprises restrictions on who can view the content item.

14. The method of any one of claims 11-13, wherein the sharing setting is entered by the user selecting from a list of attributes corresponding to the real-world object.

15. The method of claim 14, wherein the list of attributes is generated through object recognition performed on the one or more images of the real-world object.

16. An apparatus comprising:
a processor-based device; and
a non-transitory storage medium storing a set of computer readable instructions configured to cause the processor-based device to perform steps comprising:

receiving a selection of a real-world object as viewed by a camera of a first user device from a user;

receiving a content item to tag to the real-world object;

capturing one or more images of the real-world object with the camera;

receiving a sharing setting from the user, wherein the sharing setting comprises whether the real-world object will be matched only with an image of the real-world object or with an image of any object sharing one or more common attributes with the real-world object; and

storing the content item, the one or more images of the real-world object, and

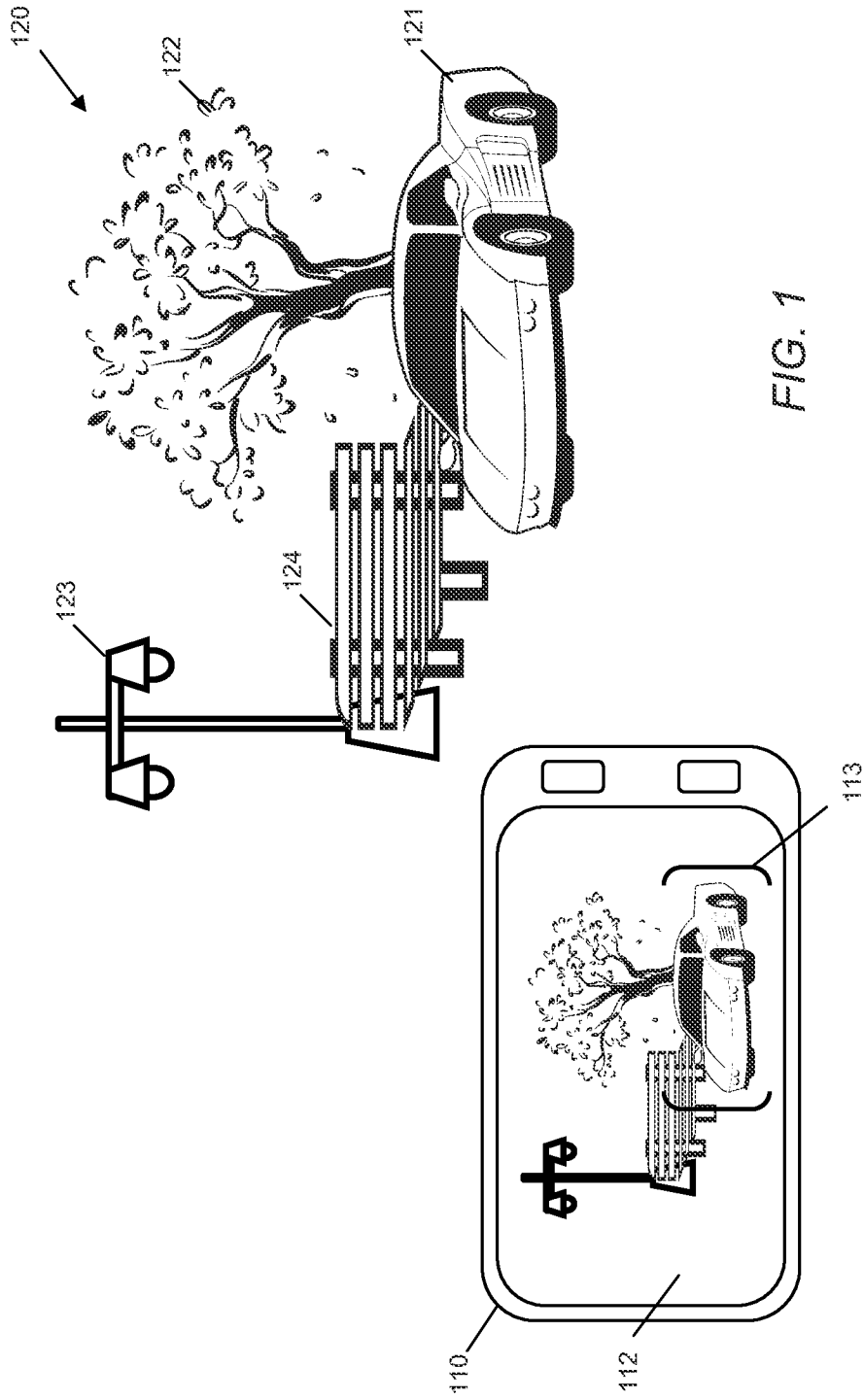
the sharing setting on a server, wherein the content item is configured to be retrieved by a second user device viewing an object matching the real-world object.

17. The apparatus of claim 16, wherein the content item comprises one or more of a text comment, an image, an audio clip, and a video clip.

18. The apparatus of claim 16, wherein the sharing setting further comprises restrictions on who can view the content item.

19. The apparatus of any one of claims 16-18, wherein the sharing setting is entered by the user selecting from a list of attributes corresponding to the real-world object.

20. The apparatus of claim 19, wherein the list of attributes is generated through automatic object recognition performed on the one or more images of the real-world object.



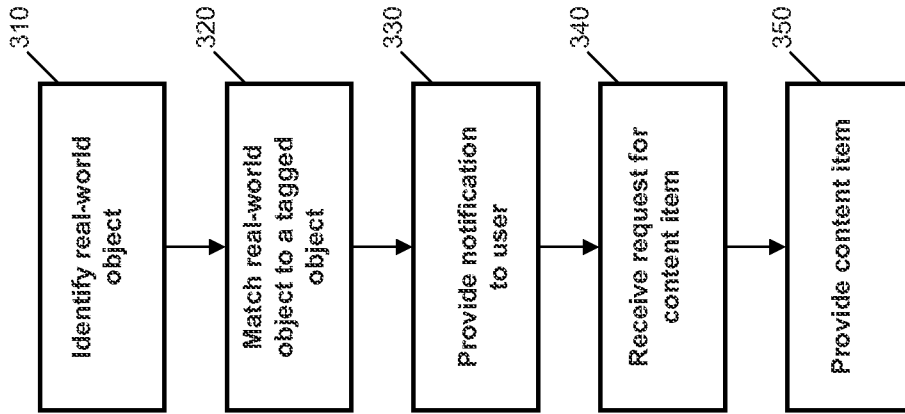


FIG. 3

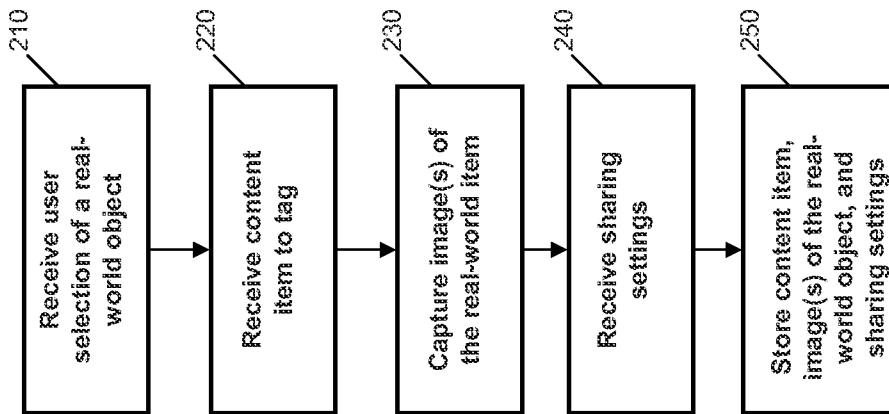


FIG. 2

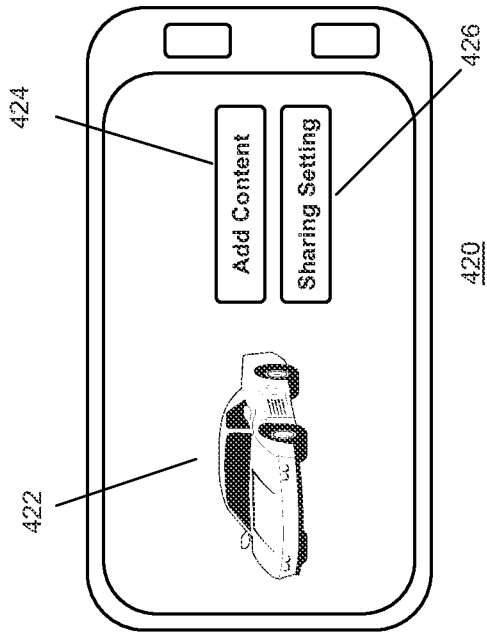


FIG. 4

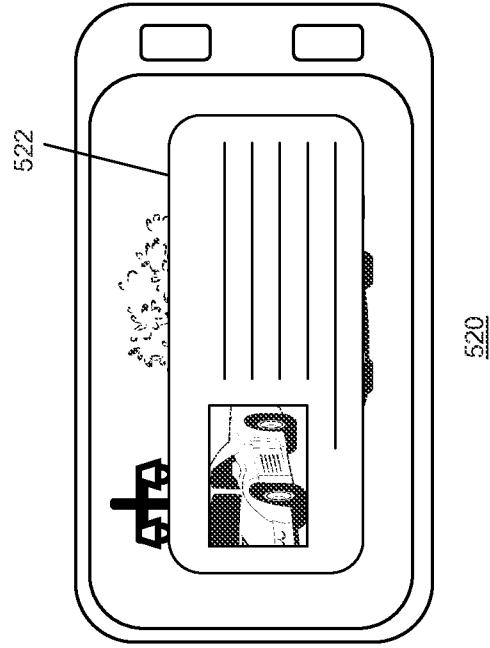
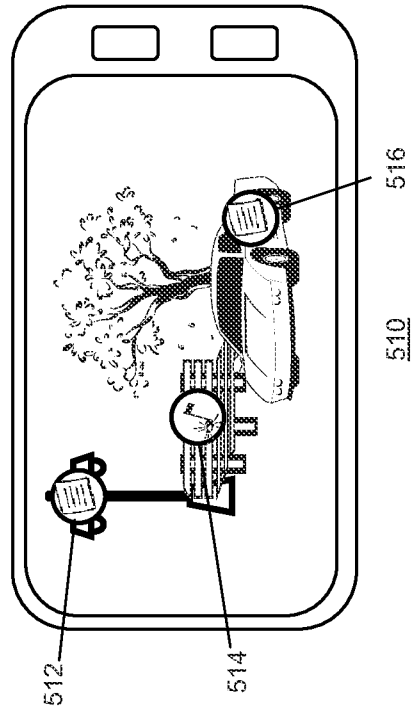
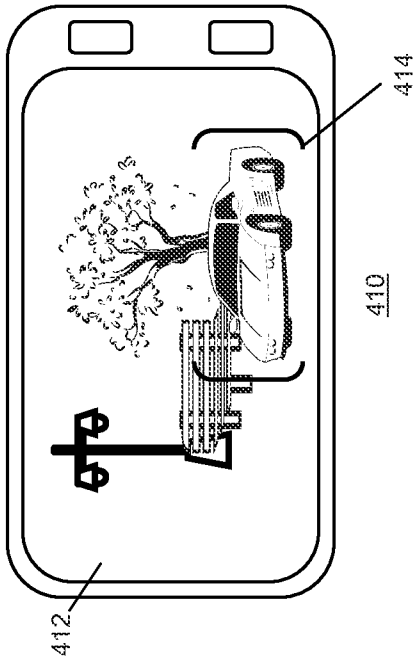


FIG. 5



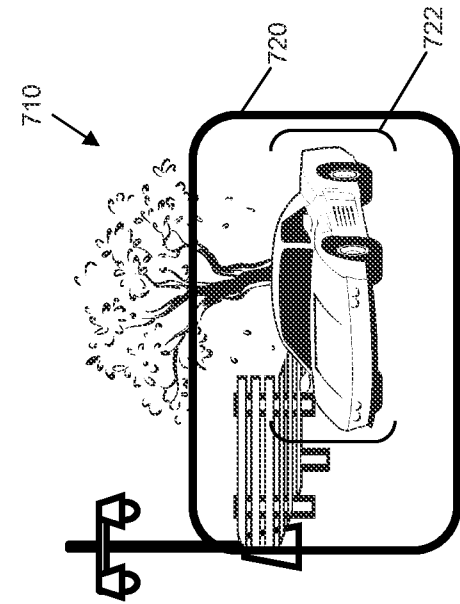


FIG. 6

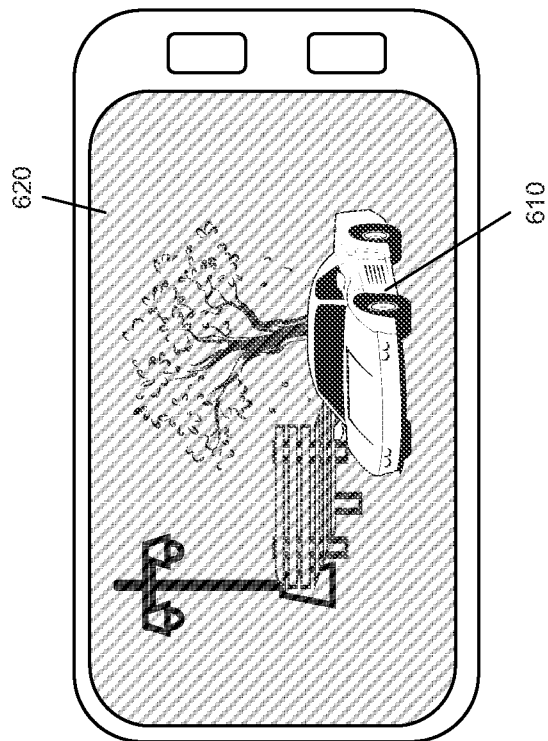


FIG. 7

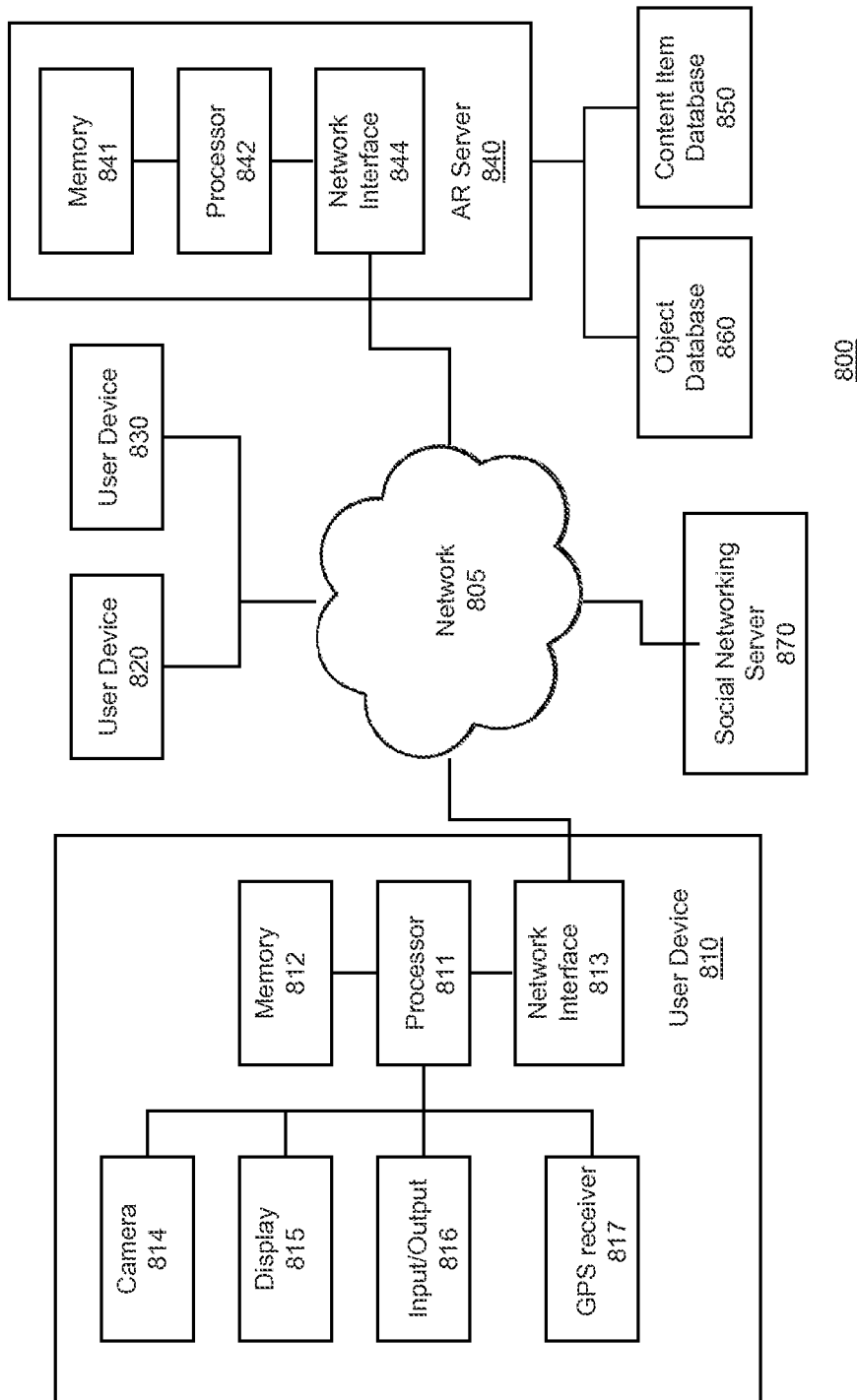


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 15/50908

A. CLASSIFICATION OF SUBJECT MATTER
 IPC(8) - G06K 9/00 (2015.01)
 CPC - G06K 9/00 G06K 9/2054 G06K 2209/01
 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)
 CPC: G06K 9/00 G06K 9/2054 G06K 2209/01; IPC(8): G06K 9/00 (2015.01); USPC: 382/181

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 USPC: 382/195
 CPC: G 06K 9/00469, G 06K 9/222 (keyword limited, see below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 PatBase, Google Web, Google Patens
 Search terms: object identification, object tagging, image recognition, content tagging

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| X | US 2013/0129142 A1 (Miranda-Steiner) 23 May 2013 (23.05.2013), entire document, especially abstract, para [0026], [0028], [0055], [0016], [0063], [0034], [0052], [0039], [0020], [0005], [0040]. | 1-20 |

Further documents are listed in the continuation of Box C.

| | |
|---|--|
| * Special categories of cited documents: | "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 |
| "A" document defining the general state of the art which is not considered to be of particular relevance | "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 |
| "E" earlier application or patent but published on or after the international filing date | "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 |
| "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) | "&" document member of the same patent family |
| "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 | |

| | |
|--|--|
| Date of the actual completion of the international search 04 November 2015 (04.11.2015) | Date of mailing of the international search report 18 DEC 2015 |
|--|--|

| | |
|---|--|
| Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300 | Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774 |
|---|--|