



(51) International Patent Classification:

H04N 21/218 (2011.01) *H04N 21/414* (2011.01)
G06F 3/0346 (2013.01) *H04N 21/431* (2011.01)
G06T 19/00 (2011.01) *H04N 21/81* (2011.01)
G06V 20/20 (2022.01)

(21) International Application Number:

PCT/US2023/084805

(22) International Filing Date:

19 December 2023 (19.12.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

18/084,103 19 December 2022 (19.12.2022) US

(71) Applicant: **QUINTAR, INC.** [US/US]; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US).

(72) Inventors: **HEIDMANN, Timothy P.**; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US). **JAYARAM, Sankar**; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US). **COCHRAN, Wayne O.**; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US). **SCOTT, John Buddy**; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US).

HARRISON, John; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US). **MATHUR, Durga Raj**; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US). **BAILEY, Richard St Clair**; 2811 Mission College Blvd., Floor 7, Santa Clara, California 95054 (US).

(74) Agent: **MAGEN, Burt**; Vierra Magen Marcus LLP, 4104 24th Street, Suite 390, San Francisco, California 94114-3615 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,

(54) Title: USE OF MULTIPLE REGISTRATIONS FOR AUGMENTED REALITY SYSTEM FOR VIEWING AN EVENT

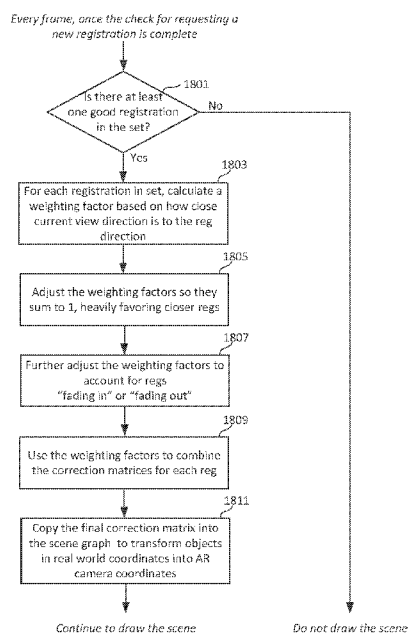


Figure 18

(57) Abstract: Augmented reality systems provide graphics over views from a mobile device for both in-venue and remote viewing of a sporting or other event. A server system can provide a transformation between the coordinate system of a mobile device (smart phone, tablet computer, head mounted display) and a real world coordinate system. Requested graphics for the event are displayed over a view of an event.

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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

USE OF MULTIPLE REGISTRATIONS FOR AUGMENTED REALITY SYSTEM FOR VIEWING AN EVENT

BACKGROUND

[0001] The present technology relates to the use of augmented reality (AR).

[0002] When viewing a sporting event or other activity/event, whether at the actual venue or remotely (such as on television), the activity may be difficult to follow or even see. Although broadcasters sometimes insert graphics into broadcast images or provide alternate views, these are selected by the broadcaster and may not correspond to what individual viewers would like to see. Additionally, when a viewer is watching an event at the venue, such added content may not be available to that viewer at the venue and, even when it is, would not correspond to different viewpoints of different individuals at the event.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] Figures 1 and 2 illustrate examples of the presentation of AR graphics and added content at an outdoor venue and an indoor venue.

[0004] Figure 3 is block diagram of elements for an embodiment of a system to register a user's mobile device and provide augmented reality content to the user's mobile device.

[0005] Figure 4 is a high-level block diagram of one embodiment of a general computing system that can be used to implement various embodiments of the registration processor, registration server and/or content server.

[0006] Figure 5 is a block diagram of a mobile device that can be used for displaying graphics of a view at a venue.

[0007] Figure 6 is a flowchart of one embodiment of a process for operation of an AR system to provide content to viewers at a venue.

[0008] Figure 7A illustrates the collection of survey images by a survey camera at a venue.

[0009] Figure 7B is a block diagram of an embodiment of a camera rig that can be used for taking the survey images.

[0010] Figure 8 illustrates the collection of fiducials at a venue.

[0011] Figure 9 is a flowchart of one embodiment of a process for preparing a venue for a survey.

[0012] Figure 10 is a flowchart of one embodiment of a process for collecting survey images.

[0013] Figure 11 is a high level flowchart of one embodiment of a process for processing imagery.

[0014] Figure 12 illustrates embodiments for registration processing based on a three columned architecture.

[0015] Figures 13A and 13B are flowcharts for embodiments of the registration and tracking process by the mobile device and of the registration process by the registration server.

[0016] Figure 14A is a block diagram of an embodiment for the registration/content server.

[0017] Figures 14B-14D illustrate embodiments for the timing of the different parts of the registration process.

[0018] Figure 15 illustrates weight values for three different registrations as a function of the viewing angle of a mobile device.

[0019] Figure 16 is a flowchart for determining whether the mobile device is in a normal application mode or waiting to reinitialize when using multiple registration for a mobile device.

[0020] Figure 17 is a flowchart of a multiple registration embodiment for deciding whether to initiate a new registration for a mobile device.

[0021] Figure 18 is a flowchart of a multiple registration embodiment for calculating a camera correction from the current set of registrations.

[0022] Figure 19 is a flowchart of an embodiment for a mobile device to handle a response to a registration request on receipt of a reply to a registration request from a registration server.

[0023] Figure 20 illustrates the use of multiple mobile devices with the registration server and content server.

[0024] Figure 21 is a block diagram of an embodiment for supplying content to one or more user's mobile devices.

[0025] Figure 22 is a flowchart for one embodiment of a process for requesting and receiving graphics by a registered mobile device.

[0026] Figures 23 and 24 respectively illustrate examples of a tabletop embodiment for events at a golf course venue and a basketball venue, corresponding to the at-venue embodiments of Figures 1 and 2.

[0027] Figure 25 is a block diagram for a tabletop embodiment.

[0028] Figure 26 is flowchart for the operation of tabletop embodiment.

DETAILED DESCRIPTION

[0029] The following presents techniques for enhancing live sports action and other events for fans who attend events at the venue or to augment their watching experience remote from the venue using augmented reality (AR) with mobile telephones, headsets, glasses, smart televisions, or other devices. At an event's venue, live viewing can enhance the live viewing process, such as by providing individual viewers accurate real time playing surface registration, and allowing live dynamic event data visualization synchronized to the playing surface action so that the entire venue becomes the canvas with accurate wayfinding and location based proposals. At home or other remote viewing locations (such as a sports bar), live tabletop AR streaming can provide dynamic event data visualization synchronized to tabletop streaming and live dynamic event data visualization synchronized to live TV. The techniques can also provide gamification, whether through institutional gaming, friend-to-friend wagering, or similar play for fun.

[0030] To be able provide AR content to users that corresponds to their individual points of view, the users' individual positions and orientations have to be precisely determined

relative to the real world. For example, if the user is at a venue and is viewing the event on a smart phone, the position and orientation of the smart phone and its camera's images will have an internal set of coordinates that need to be correlated with the real world coordinates so that content based on real world coordinates can be accurately displayed on the camera's images. Similarly, when viewing an event on a television, the camera supplying an image will have its coordinate system correlated with the real world coordinate system.

[0031] One way to track a moving camera is through use of simple optical flow techniques to latch on to simple ephemeral patterns in an image and track them frame-to-frame; however, to relate this to the real world, there needs to be a separate process that identifies unique features in the image that have been surveyed and their real world locations used to accurately locate to the viewer. A traditional computer vision approach detects visual features in a reference image, creates a numeric descriptor for that feature, and save numeric descriptor in a database, along with real world location determined by some surveying technique. For a new image, features are then detected in the image, their descriptors computed and found in the database, and the corresponding spatial information in the database is used to determine a viewer's position and orientation. This approach has a number of limitations. In many sports venues, for example, fields of view are made up of organic, non-2-D shapes (for example, trees along a fairway of a golf course) that vary widely with viewing direction and are difficult to uniquely identify. Additionally, the images will often have large areas of features that should be ignored, like moving crowds, changing scoreboards, and moving shadows, for example. Other difficulties include changing lighting conditions that change the appearance of features and many detectable features that are not distinctive enough to be uniquely identified (such as tree trunks or repeating fence posts).

[0032] To improve upon this situation, the following discussion presents a number of novel techniques. By detecting specific kinds of features in an image (e.g., the ridge line and edges of a tent, trunks of trees, location of the peaks of the trees) that can be surveyed, the same details can be identified in an image, and, using starting estimates of view position and orientation (such as from smart phone's GPS, compass, and gravitometer), a correspondence can be established between what a user can see and what has been surveyed in a database. The system can optimize the match between a 2D image of expected features based on the database and position estimates versus the smart phone's 2D camera image. More specifically, rather than use every example of a visual feature, only certain examples of features are used, with iterative

techniques applied to accurately identify those features by their 3D spatial location, even though each feature is not distinctive in itself. Employing multiple feature types together can provide a robust, flexible solution, so that rather than develop an ad-hoc solution for every different viewing environment, the system can create a framework to support detecting different specific features and using them all to solve location problems and add new kinds of features to support different environments.

[0033] Examples of different kinds of features that might be used include straight-line edges of man-made structures and the corners at which they meet, where these might have specific constraints such as one side of the edge is white and a certain number of pixels widths. For outdoor venues, an example can include tree trunks, where these might comprise the 3D points of the bottom and top of a clearly identifiable segment, plus its diameter. In a golf course example, an outline of a green against the rough, the outline of a sand bunker, or a cart path against grass can provide a curving line of points in 3D space. The outline of a tree, or tops of individual trees, against the sky can be a useful reference if it can provide a clean outline and the tree is far away. For any of the features, repeatability of detections regardless of light changes and moving shadows is an important characteristics. To survey the features, the 3D location of features can be measured using multiple views from different positions with instrumented cameras (e.g., cameras with sensors that measure location and/or orientation).

[0034] As used here, surveying a venue is the process of building a collection of features, represented by their logical description along with their 3D position information, in a spatially-organized database. For example, the locations of points could be measured directly, for example, by using a total station (theodolite) survey device, which can accurately measure azimuth, elevation, and distance to a point from a surveyed location and direction. These typically use laser range finding, but might also use multiple view paths, like a stadimeter. On a golf course, for example, sprinkler head locations are useful reference points with accurately surveyed locations. The surveying process may use cameras to collect video or still imagery from multiple locations for the venue. In some embodiments, these survey images can include crowd sourced images. These images are then registered to a real world coordinate system, typically by one or both of accurately measuring the location of the camera using GPS, or by use compass and inertial measurement unit (IMU). This may require special techniques like establishing a reference GPS base station to get sufficient accuracy. Fiducials (visual reference objects) can be placed in well-surveyed positions such that there can be several in the field of

view of any image. The fiducials can also be used to infer the location of other distinctive points within the images. Based on the fiducials and the located distinctive points, the process can register other images that may not contain enough fiducials. In some embodiments, a path of images can be digitized, with features being registered from one image to the next without surveying fiducials and then use post-processing to optimize estimates of the position of those points to match surveyed reference points: For example, a fiducial in the first and last frame of a sequence of images may be enough to accurately position corresponding points across the sequence of images, or these may be determined by structure from motion techniques.

[0035] As used here, registration is the process of establishing a correspondence between the visual frames of reference. For example, registration may include establishing a correspondence between the visual frames of reference that the mobile viewing device establishes on the fly (the coordinates of the mobile device's frame of reference) and a coordinate system of a real world frame of reference. In many situations, an accurate orientation registration may be more important than position registration. Accuracy is determined by how much pixel error there is in, for example, placing a virtual graphic (e.g., image) at a specific location in a real world scene. In one set of embodiments, based on the internal coordinates for a frame of reference of a view-tracking app on a user's device (e.g., ARKit on an iPhone) for a particular image, this can provide information on how 3D rays to several points in the image from the user's mobile device can be used to establish a transformation between the user's mobile device and its real world location so that virtual objects can be accurately drawn atop the video of the scene every frame. Depending on the embodiment, registration for a mobile device can be performed periodically and/or by relying on the mobile device's frame-by-frame tracking ability once a registration is in place. How much of the registration process is performed on the individual user's mobile device versus how much is performed on a remote server can vary with the embodiment and depend on factors such as the nature and complexity of detection of features, database lookup, and solution calibration.

[0036] Figures 1 and 2 illustrate some of the examples of the presentation of AR graphics and added AR content at an outdoor venue and an indoor venue, respectively. Figure 1 illustrates a golf course venue during an event, where the green 120 (extending out from an isthmus into a lake) and an island 110 are marked out for later reference. Figure 1 shows the venue during play with spectators present and a user viewing the scene with enhanced content

such as 3D AR graphics on the display of a mobile device 121, where the depicted mobile device is smart phone but could also be an AR headset, tablet, or other mobile device.

[0037] Some examples of the graphs that can be displayed on a viewer's mobile device are also represented on the main image. These include graphics such as player information and ball location 101 for a player on the green 120, concentric circles indicating distances 103 to the hole, ball trajectories 105 with player information 107 on the tee location, and a grid 109 indicating contours and elevation for the surface of the green. Examples of data related to course conditions include the wind indication graphic 111.

[0038] The graphics can be overlaid on the image as generated by the mobile device. The user can make selections based on a touchscreen or by indicating within the image as captured by the mobile device, such as pointing in front of the device in its camera's field of view to indication a position within the image. For example, the viewer could have a zoomed view 130 displayed on the mobile device. The zoomed view 130 can again display graphics such as player info and ball location 131, concentric distances to the holes 133, and a contour grid 139. The viewer could also rotate the zoom view, such as indicated by the arrows. Also indicated in relation to the zoom image are wager markers 141 as could be done by different viewers on mobile devices on a player-to-player basis, along with an indicator of betting result information 143.

[0039] Figure 2 illustrates the indoor venue example of a basketball game, with a viewer with a mobile device 221 providing 3D AR graphics over the image of the mobile device 221. On the image of the game are shown some example AR graphics, such as player information 251, ball trajectories 253, current ball location 255, and player position and path 257. Other examples of content include a venue model 260, player statistics 261, and a player path 263 in the court.

[0040] Figure 3 is block diagram of one embodiment of a system to register a user's mobile device and provide AR content to the user's mobile device. Figure 3 only illustrates a single mobile device 321, but, as discussed in more detail below, there can be many (e.g., thousands) such devices operating with the system concurrently. In an example where the user is at a venue, the mobile device 321 could be a cell phone, tablet, glasses, or a head mounted display, for example, and, in the case of multiple users, their respective mobile devices can be of

different types. Note that in some embodiments, some of the components of Figure 3 can be combined.

[0041] AR content to display on the mobile device 321, such as on the 2D camera image of a smart phone as illustrated in the examples of Figure 1 and 2, can be provided by a content server 323, where the content can be retrieved from a content database 327 or from a live source, such as in-venue cameras 325. Content database 327 can be one or both of a local database or a cloud database. Examples of content stored in the database can include things such as 3D terrain contours (i.e., elevations of a green for a golf course) or other venue data that can be acquired prior to the event or provided by venue. The content can also include live data about the event, such as scoring, performance related statistics, environmental data (e.g., weather) and other information. Other content can include live image data from cameras 325 that can supplement a user's point of view, such as through a "binocular view" to give a closer point of view or to fill in a user's occlusions, or other live material, such as ball trajectories. The content can be provided from the content server 323 automatically, such as based on previous setting, or directly in response to a request from the mobile device. For example, the user could indicate requested information by touching the display or manually indicating a position such as by placing a finger with the mobile device's field of view. As the content from the content server 323 is referenced to a real world coordinate system, the mobile device 321 will need a transformation between the real world coordinate system and the mobile device's coordinate system.

[0042] The transformation between the mobile device's coordinate system and the real world coordinate system is provided to the mobile device 321 by registration server 311. From the mobile device 321, the registration server 311 receives images and corresponding image metadata. For example, the image metadata can include information associated with the image such as camera pose data (i.e., position and orientation), GPS data, compass information, inertial measurement unit (IMU) data, or some combination of these and other metadata. In some embodiments, this metadata can be generated by an app on the mobile device, such as ARKit running on an iPhone (or other mobile device). Using this data from the mobile device 321 and data in a registration feature database 309, the registration server 311 determines a transform between the coordinate system of the mobile device 321 and a real world coordinate system. In one set of embodiments, the device to real world coordinate transform can be a set of matrices (e.g., transformation matrices) to specify a rotation, translation, and scale dilation

between the real world coordinate system and that of the mobile device. Once that mobile device 321 receives the transformation matrices (or other equivalent data), as the mobile device moves or is oriented differently (a change of pose), the mobile device 321 can track the changes so that the transformation between the mobile device's coordinate system and the real world coordinate system stays current, rather than needing to regularly receive an updated transformation between the mobile device's coordinate system and the real world coordinate system from the registration server 311. The mobile device 321 can monitor the accuracy of its tracking and, if needed, request an updated transformation between the mobile device's coordinate system and the real world coordinate system.

[0043] Registration server 311 is connected to a feature database 309, which can be one or a combination of local databases and cloud databases, that receives content from registration processing 307, which can be a computer system of one or more processors, that receives input from a number of data sources. The inputs for registration processing 307 includes survey images of multiple views from different positions from one or more survey image sources 301, such as one or more instrumented cameras. Embodiments can also include coordinates for fiducial points as inputs for the registration processing 307, where the fiducial points are points with the fields of view of the survey images and that have their coordinates values in the real word coordinate system by use of fiducial coordinate source devices 303, such as GPS or other device that can provide highly accurate real world coordinate values. In some embodiments, a 3D survey data set can also be used as an input for registration processing 307, where the 3D survey data can be generated by 3D surveying device 305 and, for many venues, will have previously been generated and can be provided by the venue or other source.

[0044] To be able to draw 3D graphics accurately over mobile device's 2D picture of the real world, the registration server 311 needs to know the viewer's/mobile device 231 position, the way that it is looking (its pose orientation), and camera details such as field of view and distortion. A process for accurately locating the mobile device and generating accurately aligned camera or other mobile device imagery can be broken down into three steps: First, prior to the event, assembling a database of visible features that will be visible from the range of viewer locations; second, when a viewer initially starts using the app, the location of the viewer's mobile device is determined, and a set of visual features in the mobile device's field of view is established so that the system can accurately register the graphics as presented on the mobile device to the real world; and third, as the viewer continues to use the app, the mobile

device is re-oriented to look at different parts of a scene, tracking features in field of view (such as on a frame-by-frame basis) to maintain an accurate lock between the real world and the augmented reality graphics.

[0045] To build the registration feature database 309, survey data is collected for the venue and assembled into a single reference map to serve as a model for the venue. Within the reference map, viewing areas can be identified and planning can be made for the location of temporary structures such as viewing stands, tents, or signage. Reference markers for use as fiducials are also identified. Note that the reference map may not be a literal map, but a collection of data representing the relevant set of features (as described herein).

[0046] At the venue, prior to event, photos are taken along the line of viewing areas, such as at every 10 feet or 3 meters (or other intervals or distances), and corresponding metadata, such as camera location and orientation, is accurately measured. Multiple cameras can be used, such as three cameras with one looking horizontally in the viewing direction, one camera 45° to the left, and one camera 45° to the right. The photos are taken with high resolution (e.g., 8 megapixel each) and can be saved with high quality JPEG compression, with the imagery and metadata transferred to a central server (e.g., registration processing 307, registration server 311 or another computing device). The cameras can be connected to a very accurate GPS receiver, compass, inclinometer, and gyroscope, so that the camera locations can be known to within a few inches and their orientation to within a few hundredth of a degree. For improved accuracy, the focal length and distortion for each camera can be pre-measured on an optical bench. To more easily move the camera rig 301 around a venue it could be mounted on a golf cart or a drone, for example.

[0047] Once the survey images and their metadata are gathered, they are stored on a computer (e.g., registration processing 307, registration server 311 or another computing device). Surveyed reference points, such as sprinkler locations or visible fiducials placed on reference points, are located prior to taking the photos. The pixel location of fiducial markers can be identified in a variety of the survey images and their 3D coordinates determined via triangulation using the camera parameters, such as discovered from a Structure from Motion (SfM) process. In the processing, these fiducial points are used to refine the measured camera positions and orientations, so that the coordinate system of the photos can be aligned to the real world coordinate system. As described in more detail in the following discussion, given the real world coordinates of the fiducial markers and the SfM coordinates, a transformation is

found that maps between the coordinate system of the individual mobile devices and the real world coordinate system. Figures 7A and 8 respectively illustrate the collection of photos and the use of fiducials, and Figures 9 and 10 respectively present flowcharts for survey preparation and image collection.

[0048] Figure 4 is a high-level block diagram of one embodiment of a more general computing system 401 that can be used to implement various embodiments of the registration processing 307, registration server 311 and/or content server 323. Specific devices may utilize all of the components shown, or only a subset of the components, and levels of integration may vary from device to device. Furthermore, a device may contain multiple instances of a component, such as multiple processing units, processors, memories, transmitters, receivers, etc.

[0049] In Figure 3, the registration server 311 and the content server 323 are represented as separate blocks based on their different uses, but it will be understood that these functions can be implemented within the same server and that each of these blocks can be implemented by multiple servers. Consequently, depending on the embodiment, the registration server 311 and the content server 323 can be implemented as a single server or as a system of multiple servers. The components depicted in Figure 4 include those typically found in servers suitable for use with the technology described herein, and are intended to represent a broad category of such servers that are well known in the art.

[0050] The computing system 401 may be equipped with one or more input/output devices, such as network interfaces, storage interfaces, and the like. The computing system 401 may include one or more microprocessors such as a central processing unit (CPU) 410, a graphic processing unit (GPU), or other microprocessor, a memory 420, a mass storage 430, and an I/O interface 460 connected to a bus 470. The computing system 401 is configured to connect to various input and output devices (keyboards, displays, etc.) through the I/O interface 460. The bus 470 may be one or more of any type of several bus architectures including a memory bus or memory controller, a peripheral bus or the like. The microprocessor 410 may comprise any type of electronic data processor. The microprocessor 410 may be configured to implement registration processing using any one or combination of elements described in the embodiments. The memory 420 may comprise any type of system memory such as static random access memory (SRAM), dynamic random access memory (DRAM), synchronous DRAM (SDRAM), read-only memory (ROM), a combination thereof, or the like. In an

embodiment, the memory 420 may include ROM for use at boot-up, and DRAM for program and data storage for use while executing programs.

[0051] The mass storage 430 may comprise any type of storage device configured to store data, programs, and other information and to make the data, programs, and other information accessible via the bus 470. The mass storage 430 may comprise, for example, one or more of a solid-state drive, hard disk drive, a magnetic disk drive, an optical disk drive, or the like.

[0052] The computing system 401 also includes one or more network interfaces 450, which may comprise wired links, such as an Ethernet cable or the like, and/or wireless links to access nodes or one or more networks 480. The network interface 450 allows the computing system 401 to communicate with remote units via the network 480. For example, the network interface 450 may provide wireless communication via one or more transmitters/transmit antennas and one or more receivers/receive antennas. In an embodiment, the computing system 401 is coupled to a local-area network or a wide-area network for data processing and communications with remote devices, such as other processing units, the Internet, remote storage facilities, or the like. In one embodiment, the network interface 450 may be used to receive and/or transmit interest packets and/or data packets in an ICN. Herein, the term “network interface” will be understood to include a port.

[0053] The components depicted in the computing system of Figure 4 are those typically found in computing systems suitable for use with the technology described herein, and are intended to represent a broad category of such computer components that are well known in the art. Many different bus configurations, network platforms, and operating systems can be used.

[0054] Figure 5 is a high-level block diagram of an embodiment of a mobile device 321 that can be used for displaying graphics of a view at a venue, such as described above. Embodiments of the mobile device can include a smart phone, tablet computer, laptop computer, or other device in which the view of the venue is presented on a display 503, such as a screen with the graphics content also represented on the display. Other embodiments can include head mounted displays, such as AR headsets or AR glasses, that display the graphics over the view of the venue as watched through the head mounted display. The multiple mobile devices that can be used concurrently with the systems presented here can be various combinations of these different varieties of mobile devices. Figure 5 explicitly includes

elements of the mobile device 321 relevant to the discussion presented here, but will typically also include additional elements, but that do not enter into the current discussion and are not shown.

[0055] The embodiment of Figure 5 includes a camera 501 and one or more sensors 507 that respectively provide image data and metadata for the image data that can be used in the registration process described above. Mobile devices 321 such as smart phones typically include a camera 501, such as based on charge coupled devices or other technology, that can provide the image data and also the image of the venue on the mobile device's display screen, while for a head mounted display, the camera 501 would provide the image data, although it may not be displayed directly to the viewer. The sensors 507 can include devices such as GPS receivers, a compass, and an inertial measurement unit (e.g., accelerometer). The metadata from the sensors 507 can provide information on the pose (location and orientation) of the camera 501 when capturing the image data, but will be within the mobile device's internal coordinate system that may only loosely be aligned with the real world coordinate system.

[0056] The mobile device 321 also includes one or more interfaces 505 through which the mobile device 321 can communicate with the registration server 311 and content server 323. The interface 505 can use various standards and protocols (Bluetooth, Wi-Fi, etc.) for communicating with the servers, including communicating with the registration server 311 for the registration process and with the content server 323 to request and receive graphics and other content. The cellular transceiver 511 can also be used to communicate with the registration server 311 and content server 323, as well as for telephony.

[0057] A mobile device 321 also includes one or more processors 509, with associated memory, that are configured to convert the graphics from the content server 323 into the mobile device's coordinate system based on the transformation between the mobile device's coordinate system and the real world coordinate system as received from the registration server 311. The processor(s) 509 can be implemented as ASICs, for example, and be implemented through various combinations of hardware, software, and firmware. The processor or processors 509 can also implement the other functionalities of the mobile device not related to the operations describe here, as well as other more relevant functions, such as monitoring latencies in communications with the servers and adapting the amount of processing for the registration and display of graphics done on the mobile device 321, relative to the servers, based on such latencies.

[0058] The display 503 is configured to present the graphics over the view of the venue. In the case of device where the display 503 is a screen (such as a smart phone or tablet), the view of the venue can be generated by the camera 501, with the graphics also displayed on the screen. In this case, user input (such as related to gamification or requesting specific graphics) can be input by a viewer using the display and/or, in some embodiments, by indicating within the view of the venue from the camera 501, such as by finding the user's fingertip within the image and projecting a ray to this location to, for example, touch where a ball will land or to touch an object to place a bet. In a head mounted display 503, such as AR goggles or glasses, the graphics or other content can be presented over the view of the venue through the mobile device 321, where the user can make indications within the view.

[0059] Figure 6 is a flowchart describing one embodiment for the operation of an AR system for providing viewers with AR graphics over views of an event. Beginning at step 601, the venue is prepared for a survey to collect image mages and fiducial points' coordinates that are supplied to the registration processing 307. Step 601 is discussed in more detail with respect to Figure 9. The survey images are then collected in step 603, which is described in more detail with respect to Figure 10. From the data collected is steps 601 and 603, the registration processing 307 builds a model of the venue, as described further with respect to Figure 11. Steps 601, 603, and 605 are typically performed before the event, although data can also be collected during an event, such as through crowd sourced image data, to refine the model.

[0060] Before the event, mobile devices 321 are registered with a server system including a registration server 311 at step 607. This is done by each mobile device 321 sending the registration server 311 image data and metadata, that will be in the coordinate system of the mobile device, to the registration server. For each mobile device 321, the registration server can then build a transformation for converting positions/locations between the mobile device's coordinate system to a real world coordinate system. The registration server 311 also sends each mobile device 321 template images with a set of tracking points within each of the template images at step 609. The template images with tracking points allow for each of the mobile devices 321 to maintain an accurate transformation between the mobile device's coordinate system and the real world coordinate system as the mobile device changes its pose (i.e., location and orientation). Registration and tracking is described in more detail with respect to Figures 13A and 13B. At step 611 a registered mobile device 321 can then request

and receive AR content, such as graphics to display of views of an event at a venue, from the content sever 323. More details about step 611 are provided below with respect to Figure 22.

[0061] Figure 7A illustrates the collection of survey images by a survey camera at a venue. In this example, the venue is the same as illustrated in Figure 1, but shown as a point cloud 700 generated from features within the venue prior to the event and without spectators. For comparison to Figure 1, the island 710 and green 720 are given reference numbers corresponding to reference numbers 110 and 120 in Figure 1. The individual points of the point cloud 700 correspond to features for use in the registration process as described below. One of the data inputs to the process is the survey data as generated by a survey camera rig 301.

[0062] Figure 7A illustrates the collection of multiple images from multiple locations at the venue, where Figures 9 and 10 describes an embodiment for the process to collect these survey images. In Figure 7A, several dozen sets of images collected at specific points, where several of these image collections (701, 757, 759, 799) at some of these locations are explicitly numbered. The actual process can include additional collections of images, such as in the upper portions of the image, but these are not included in Figure 7A to avoid the Figure becoming overly complicated. The number of such locations and the number of photos taken will vary based on the specifics of the venue and the event, but as described below, these will typically be collected at positions where viewers are likely to be located and with sufficient density be able to perform an accurate registration process.

[0063] In the lower portion of Figure 7A is an expanded view of the collection of images 759 to illustrate the collection more clearly. At the center is the location of the survey camera rig 301 used to collect a set of images, where the survey camera rig 301 can include a single camera or multiple cameras along with equipment to determine the camera location and orientation. The images are represented by a set of N frustums (e.g., truncated pyramids), where a first frustum 759-1 and an Nth frustum 759-N are labeled. The wider base of a frustum (the darker, labelled rectangles) correspond to the 2D image as seen by the camera from its pose when the image is taken and narrow base of a frustum corresponds to the 2D plane of the image collection surface for the camera. The images taken at a given position are taken to overlap and to cover the directions of likely fields of view for users of the mobile devices during the event.

[0064] Figure 7B is a block diagram of an embodiment of a multi-camera survey camera rig 301 that can be used for taking the survey images. In one embodiment, three cameras with a center camera (711a) looking horizontally in the viewing direction, one camera (711b) angled 45° to the left, and one camera (711c) angled 45° to the right. The cameras can have high resolution (e.g., 8 megapixel each) and can use high quality JPEG compression, with the imagery and metadata transferred over interface 715 to a central server. Depending on the embodiment, the images can be processed on the individual cameras (711a, 711b, 711c) or by a separate processing/memory section 713 incorporated into the survey camera rig 301. The survey camera rig 301 can also include instrumentation 717 to determine the metadata for the orientation and location of the cameras' images. The instrumentation can include a GPS receiver, compass, IMU, and gyroscope, for example, so that the camera locations can be known to within a few inches and their orientation to within a few hundredth of a degree.

[0065] Figure 8 illustrates the collection of fiducials at a venue. The venue of Figure 8 is the same as for Figures 1 and 7A and again shows the same point cloud 700 and reference features of the island 710 and green 720, but with the image collections (e.g., 701, 757, 759, 799) not shown. The fiducials will be placed prior to, and included in, the collection of survey images, but the image collections are not shown in Figure 8 for purposes of explanation. The placement and collection of fiducials are described in more detail with respect Figures 9 and 10.

[0066] Figure 8 shows a number of fiducials within the point cloud 700, where several examples of the fiducials (801, 857, 859, 899) are explicitly labelled. As described below, the number and placement of the fiducial will depend on the venue, type of event, and where the survey images are to be collected. The position of the fiducials are determined so that their points' coordinates in the real world coordinate system is well known. This can be done by placing the fiducials at locations with well-known coordinates, such as is often the case for features in the venue (e.g., sprinkler locations of a golf course), by accurately measuring the locations of fiducials by a GPS or other positioning device, or a combination of these.

[0067] Figure 9 is a flowchart of one embodiment of a process for preparing a venue for a survey, providing more detail for step 601 of Figure 6. To organize the collection of survey data, a preliminary model is assembled for the environment of the venue at step 901, where this can be a 2D or 3D model and can often be based on information available from the venue or bases on a rough survey. Based on this model, regions where viewers will be located during

event are identified at step 903. For example, if the venue is a golf course, viewing arrays are typically around the tee, around the green, and along portions of the fairway. In an indoor venue, such as for a basketball game, the viewing arrays correspond to locations in the stands. At step 905, the identified viewer locations can be used to plan a path and spacing for points at which to collect the survey images.

[0068] In step 907, locations that will be within the images are identified as location for fiducials, where these can be objects in known locations that will be visible in the survey images and which can be used to infer the location and orientation of the survey camera location with high accuracy (i.e., down to fractions of inches and degrees). In the example of a golf course, one choice of fiducial locations can be sprinkler head locations, as these are plentiful, easy to find, and their locations are often carefully surveyed by the venue. To make fiducials easier to locate within the survey image, these can be marked by, for example a white or florescent yellow sphere a few inches in diameter mounted on a stand that lets it be located as a specified height (e.g., an inch above a sprinkler head). In some cases, to improve accuracy, a reference GPS base station in communication with the survey camera rig can be set up at step 909.

[0069] Figure 10 is a flow of one embodiment of a process to collect survey images following the preparation of Described with respect to Figure 9 and provides more detail for step 603 of Figure 6. Starting at step 1001, any wanted fiducial marker are placed for a section of the survey path. Depending on the implementation, this can be all of the fiducial markers for the entire survey or for a section of the survey, with the marked moved from views already photographed to subsequent views as the survey camera rig 301 is moved along the survey path. As discussed above, the survey camera rig 301 can be part of rig of multiple cameras along equipment to determine corresponding metadata for the images. The survey camera rig 301 is moved along the path, such as the planned path from step 905, collecting images in step 1003. In the case of a fixed rig of several cameras, at each location the rig can collect a set of images looking in several directions and at different focal lengths, which can be fixed. In terms of instrumentation, the survey camera rig 301 can include an accurate GPS receiver, where this can be referenced to a base station in some embodiments. The GPS receiver can also be integrated with an initial measurement unit, or IMU, with linear and rotational rate sensors, and additionally be integrated with a magnetic compass. Step 1005 records the GPS position and orientation metadata for each of the images. As the images and their metadata are accumulated,

the image quality and metadata accuracy can be monitored at step 1007. Once the images are collected, the fiducial markers can be recovered at step 1009 and the survey imagery and corresponding metadata copied to a server at step 1011.

[0070] In some embodiments, the survey images can be augmented by or based on crowd-sourced survey images from viewers' mobile devices 321. For example, users could be instructed to provide images of a venue before or even during an event, taking photos with several orientations from their viewing positions. This can be particularly useful when an event is not held in a relatively compact venue, such as a bicycling race in which the course may extend a great distance, making a formal survey difficult, but where the course is lined with many spectators who could supply survey image data. In some instances, as viewers provide crowd-sourced survey images, the registration process can be updated during an event. For embodiments where crowd-sourced survey images are provided prior to the event, these crowd-sourced images can be used along with, and in the same manner as, the survey images collected prior to the event by the camera rig 301. When the crowd-sourced survey images are provided during the event, they can be combined with the initial survey data to refine the registration process. For example, based on the pre-event survey images, an initial model of the venue can be built, but as supplemental crowd-sourced survey images are received during an event, the feature database 309 and registration process can be made more accurate through use of the augmented set of survey images and the model of the venue refined. This sort of refinement can be useful if the views of a venue change over the course of the event so that previously used survey images or fiducial points become unreliable.

[0071] In some embodiments, for venues or portions of venues where survey images and fiducials are sparse or absent (e.g., a cycling race), the crowd-sourced survey images and their metadata can be used without the survey images from a camera rig 301 or fiducial point data. The crowd-sourced survey images and their corresponding metadata alone can be used in the same manner as described for the survey images generated by a camera rig 301 and the lack of fiducials from a survey can be replaced by extracting some degree of fiducial point data from the crowd-sourced survey images and their metadata. The model can be generated using crowd-sourced images in combination with survey images, using survey images only, or using crowd-sourced images only. The images are crowd-sourced images as they are provided from the public at large (e.g., those at the venue) and function to divide work between participants to achieve a cumulative result (e.g., generate the model). In some embodiments, the identify

and/or number of the plurality of mobile devices used to provide the crowd sourced images are not known in advance prior to the event at the venue.

[0072] To have accurately generated real world coordinate data for the fiducials, as part of the survey process these locations can be determined by a GPS receiver or other fiducial coordinate source device 303. In some cases, the venue may already have quite accurate location data for some or all of the fiducial points so that these previously determined values can be used if of sufficient accuracy.

[0073] In some embodiments, 3D survey data and similar data can also be used as a source data. For example, this can be established through use of survey equipment such as by a total station or other survey device 305. Many venues will already have such data that they can supply. For example, a golf course will often have contour maps and other survey type data that can be used for both the registration process and also to generate content such as 3D graphics like contour lines.

[0074] Once the source data is generated, this can be used by the registration processing 307 to generate the feature database 309. The processing finds detectable visual features in the images, for those that can be detected automatically. The better features are kept for each image (such as, for example, the best N features for some value N), while keeping a good distribution across the frame of an image. For each image, a descriptor is extracted and entered into a database of features and per-image feature location. Post-processing can merge features with closely matching descriptors from multiple images of the same region, using image metadata to infer 3D locations of a feature and then enter it into the feature database 309. By spatially organizing the database, it can be known what is expected to be seen from a position in direction. Although one feature provides some information about position and orientation, the more features that are available, the more accurate the result will be. When a venue is a constructed environment, such as a football stadium or a baseball park, there will typically be enough known fiducials to determine position and orientation. In more open venues, such as golf course fairway with primarily organic shapes such as trees and paths, additional reference points may need to be collected.

[0075] Non-distinctive features in the images, such as a tree trunk, edge of a cart path, or the silhouette of trees against the sky, can be correlated across adjacent views to solve for 3D locations and then entered into the feature database 309. Such features can typically be

detected, but often not identified uniquely. However, if where the image is looking is roughly known, it is also roughly known where to expect the features to be located. This allows for their arrangement in space to be used to accurately identify them and to accurately determine a location, orientation, and camera details. The process can also collect distinctive information extracted from the features, such as width of a tree trunk or size of a rock, to help identify the objects and include these in the database.

[0076] Once the images have been registered, they can be used in conjunction with a 2D venue map to identify spectator areas as 3D volumes. The tracking and registration process can ignore these volumes and not attempt to use features within them as they will likely be obscured. Other problem areas (large waving flags, changing displays, vehicle traffic areas) can similarly be ignored. In some cases, it can be useful to perform a supplemental survey shortly before an event to include added temporary structures that may be useful for registration and also reacquire any imagery that can be used to correct problems found in building the initial feature database 309. The feature database 309 can also be pruned to keep the better features that provide the best descriptor correlation, are found in a high number of images, and that provide a good distribution across fields of view.

[0077] Figure 11 is a flow chart describing one embodiment for processing the imagery in registration processing 307 to generate the data for the feature database 309 from the survey images, fiducial points' coordinates, and 3D survey data. The process of Figure 11 is an example implementation of step 605 of Figure 6. The processing can be done offline, with manual operations performed by several people in parallel, and with a mix of automated and manual effort. For the individual collected images, at step 1101 fiducials within the image are identified and the position metadata fine-tuned. Also, within the individual images, at step 1103 various types of macro features (i.e., large scale features identifiable visually be a person) that can be used for registration are identified. At step 1105 the GPS position and orientation metadata for the images are recorded, where the positions can be stored in cartesian coordinates as appropriate for the venue, for example. In addition to camera position and orientation, the metadata can also include camera intrinsic parameters such as focal distance, optical center, and lens distortion properties. Step 1107 looks at adjacent sets of images and identifies features present in multiple images and solves for their 3D location. The feature database 309 is assembled at step 1109, where this can be organized by viewing location and view direction,

so that the registration server 311 can easily retrieve features that should be visible from an arbitrary location and view direction.

[0078] Figure 12 is a more detailed flowchart of the process for an embodiment for operation of the registration processing 307 based on a three columned architecture and illustrating how the steps of Figure 11 fit into this architecture. Other embodiments may not include all of the columns, such as by not using the third column. In Figure 12, the left most column uses the survey images, possibly including supplemental crowd-sourced survey images to generate descriptors and coordinate data for features. The middle column uses a combination of survey images and fiducial points' coordinates to generate macro feature coordinate data. The right column uses 3D survey data to generate 3D contours.

[0079] In terms the elements of Figure 4, the inputs (the survey images, fiducial points coordinates, 3D survey dataset) can be received through the network interfaces 450 and the outputs (feature descriptor coordinate data, macro coordinate data, 3D contours) transmitted to the feature database or databases 309 by the network interfaces 450. The processing steps of Figure 12 (e.g., 1201, 1215, 1221, 1225) can be performed by the microprocessor 410, with the resultant data (e.g., 1213, 1217, 1219, 1223, 1229) stored in the memory 420 or mass storage 430, depending on how the microprocessor stores it for subsequent access. For process operations that may require some degree of manual operation, such 1211, 1227, or 1231, these can also be performed by microprocessor 410 with manual input by way of the I/O interface 460.

[0080] Considering the left most column, the survey images can be acquired as described above with respect to the flows of Figures 9 and 10 and also, in some embodiments, incorporate crowd-sourced images. In some embodiments, Structure-from-Motion (SfM) techniques can be applied to process the images in block 1201, where SfM is a photogrammetric range imaging technique that can estimate 3D structures from a sequence of images. For example, the COLMAP SfM pipeline or SfM techniques can be used. The resultant output is a set of descriptors and coordinate data for the extracted features. For example, this can be in the form of scale-invariant feature transform (SIFT) descriptors that can be stored in the feature database 309. The SIFT descriptors can be, for example, in the form of a vector of 128 floating points values that allows for features to be tracked and matched by descriptors that are robust under varying viewing conditions and are not dependent on the features illumination or scale. The

output of the structure-from-motion can also include camera pose data from the images for use in the second column of Figure 12.

[0081] The second column of Figure 12 includes inputs of the same survey images as the left column, both directly and through the camera pose data (i.e., position and orientation metadata) 1217, and of the fiducial points' coordinates. The fiducials within the survey images are labelled in block 1211, where this can include both automated and manual labelling as described above. The result of the labelling are the fiducial 2D coordinates within the images at block 1213.

[0082] The camera pose data obtained from structure-from-motion 1217 will be referenced to a coordinate system, but this is a free floating coordinate system used for the structure-from-motion process and not that of the real world. As the 3D graphics and other content that will be provided to the mobile device 321 needs to be in the same coordinate system as the images, the coordinate system of the camera pose data of structure-from-motion 1217 needs to be reconciled with a real world coordinate system. This is performed in the processing of structure-from-motion to real world solver 1215. The data inputs to the structure-from-motion to real world solver 1215 are the camera pose data of structure-from-motion 1217, the fiducial 2D coordinates data 1213, and the fiducial points' coordinates. The resultant output generated by the structure-from-motion to real world solver is a structure to real world transform 1219. In some embodiments, operations corresponding to some or all of the additional elements of the middle column of Figure 12 can be moved to the registration server 311. For example, the elements 1221, 1223, and 1225 or their equivalents could be performed on the registration server 311, in which case the structure-from-motion transformation between the mobile device's coordinate system and the real world coordinate system would be stored in the feature database 309. As represented in Figure 12, the additional elements of 1221, 1223, and 1225 are performed prior to the storage of data in the feature database 309.

[0083] Considering the structure-from-motion to real world transform 1219 in more detail, structure-from-motion is performed in a normalized coordinate system appropriate for numeric purposes and the camera extrinsic data is expressed in this coordinate system. The transform 1219 is a similarity transformation that maps points from the SfM coordinate system into the target, real world coordinate system. The cameras' coordinate system can be converted to a real world coordinate system based on a combination of a rotation and translation and a scale,

rotation, and translation operation. The combination of these can be used to generate a transform matrix between the two coordinates systems.

[0084] As shown in the embodiment of Figure 12, the registration processing 307 continues on to a transform pose process 1221 to transform the camera poses (their locations and orientations) used during the survey process to the real world coordinate system based on the camera pose from the structure-from-motion 1217 and the structure-from-motion to world transform 1219. The resultant data output is the camera pose to real world coordinate transformation 1223, allowing the camera pose in the camera's coordinate system to be changed into the camera's pose in the real world coordinate system.

[0085] The system also performs bundle adjustment 1225 based on the camera pose to world coordinate transformation 1223 data labeled macro 2D feature data 1229 as an input. The labeled macro 2D feature data 1229 is generated by a label macro features process 1227 to assign labels to the large scale macro features, where this can be a manual process, an automated process, or a combination of these, where this is often based on the types of features. Bundle adjustment is a process of, given a set of images depicting a number of 3D points from different viewpoints, simultaneously refining the 3D coordinates describing the scene geometry, the parameters of the relative motion, and the optical characteristics of the cameras employed to acquire the images. The bundle adjustment 1225 can be an optimization process for minimizing the amount of error between differing projections of the images, resulting in the output data of the macro features' coordinate data for storage in the feature database 309.

[0086] In embodiments including the third column of Figure 12, a set of 3D contour data is generated from the 3D survey dataset by extracting and name contours process 1231. This can be a manual process, an automated process, or a combination of these. As noted above, the 3D survey dataset can include existing data provided by the event venue as well as data newly generated for the registration process.

[0087] As described above with respect to Figure 3, the data from registration processing 307 are features' descriptor and coordinate data, macro-feature coordinate data, and 3D contour data. This data is stored in the feature database 309, from which the registration server 311 can retrieve these as point feature data, large scale feature data, and shape feature data for use in the registration process.

[0088] To register a viewer's mobile device 321, the registration server 311 receives the position, orientation, and field of view (or pos/orient/fov) data from the mobile device 321, such as from an API on phone or other mobile device 321. Prior to sending this data, which serves as metadata for the image data from the mobile device 321, the GPS and compass on the mobile device will calibrate themselves, this may include prompting the user to get a clearer view of the sky or perhaps move the mobile device through a figure-eight pattern, for example. Typically, this can provide a position within about 5 meters, an orientation within about 10 degrees, and a field of view withing about 5 degrees. The camera or other mobile device 321 can grab images, every 5 seconds for example, and perform basic validity checks, and send the image data and image metadata to the server.

[0089] Once the image data and metadata are at the registration server 311, the registration server 311 finds distinctive and non-distinctive features within the image and, using image metadata for position and orientation, compares this to expected features in the feature database 309. For example, the registration server 311 can use distinctive features to refine the position and orientation values, then use this location to identify the non-distinctive features to further solve for the position, orientation, and field of view of the mobile device 321 within the real world coordinate system. On the registration server 311, the solving problem identifies alignment errors for each feature, where these errors can be accumulated across multiple viewers and used to improve the 3D location estimation of the feature.

[0090] In some embodiments, the registration server 311 can prompt the user to do a pan left-right for the mobile device 321. The images from the pan can be captured and used to build up a simple panorama on the registration server 311. The registration server 311 can then build a pyramid of panorama images at a range of resolution values, find likely tracking points and reference, or "template", images including the likely tracking points, and sends these to the mobile device 321. Based on the tracking points and template images, the mobile device 321 can locate, find, and match reference points in image frames quickly on a frame-by-frame basis to get an accurate orientation value for the mobile device 321.

[0091] Once the mobile device 321 is registered, it can track the images, maintaining a model (such as a Kalman-filtered model) of the mobile device's camera's orientation, where this can be driven by the IMU of the mobile device 321 and tracking results from previous frames. This can be used by the mobile device 321 to estimate the camera parameters for the current frame. The mobile device can access the current set of simple features at their predicted

location with a current image, such as by a simple template matching, to refine the estimate. Typically, it is expected that a mobile device 321 may have its orientation changed frequently, but that its location will change to a lesser amount, so that the orientation of the mobile device 321 is the more important value for maintaining graphics and other content locked on the imagery with the real world coordinate system.

[0092] The active set of simple features can be updated so that the area of view is covered, with simple features being discarded or updated based upon which simple features can be readily found and factors such as lighting changes. In some embodiments, the features can be reacquired periodically and re-solved for location and orientation to account for a viewer moving or due to a drifting of fast tracking values, for example. This could be done on a periodic basis (e.g., every minute or so), in response to the mobile device's GPS or IMU indicating that the viewer has moved, or in response to the matching of local reference features starting to indicate difficulties for this process. If the mobile device is unable to locate template features within the current image, a more detailed match against the panorama images can be performed, where this can start with the lower resolution images, to reacquire an orientation for the mobile device 321 or determine that the view is obstructed. In response to being unable to locate template features within the current image, the AR graphics and other content may be hidden or, alternately, continued to be displayed using a best guess for the mobile device's orientation. In some embodiments, the mobile device 321 can provide the user with a visual indication of the level of accuracy for the tracking, so that the user can be trained to pan smoothly and with a consistent camera orientation (i.e., mostly upward), and maintain a view of the scene in which obstructions are minimized.

[0093] Figures 13A and 13B are flowcharts describing embodiments of the registration and tracking process of step 607 and 609 of Figure 6. Figure 13A describes the process performed by the mobile device 321 and Figure 13B describes the registration process performed by the registration server 311. Once a user is at the venue, the user's phone or other mobile device 321 obtains one or more frames of image data containing from camera 501 along with the image's corresponding camera position and orientation metadata from the sensors 507, as described in the preceding paragraphs. Step 1301 of Figure 13A is the capturing of the one or more images by the mobile device and step 1303 includes the accumulation of the corresponding metadata at the mobile device. Once accumulated and stored in the processors/memory 509, the image and image metadata can then be sent from the mobile device

321 to the registration server 311 at step 1305 over the interfaces 505 or cellular transceiver 511.

[0094] At steps 1307 and 1309, the mobile device 321 receives the transformation between the mobile device's coordinate system and the real world coordinate system and the tracking points and template images from the registration server 311. Before going to steps 1307 in Figure 13A, however, Figure 13B is discussed as it describes how the received information at steps 1307 and 1309 is generated on the registration server.

[0095] More specifically, Figure 13B describes how the data sent from the mobile device 321 at step 1105 is used by the registration server 311 to generate the data received back the mobile device in steps 1307 and 1309. Starting at step 1351, the registration server 311 receives the image and image metadata from the mobile device 321 over the network interfaces 450. Based on the images' metadata, the registration server 311 retrieves the descriptors of expected features at step 1353 from feature database 309 over the network interfaces 450, where this data can be stored in the memory 420 or mass storage 430. Starting from the expected positions and shapes of the features in the images, and given the corresponding metadata (position, orientation, field of view, distortion), at step 1355 the registration server 311 locates, to the extent possible, the actual features. From the located features, at step 1357 registration server can adjust the initial measurement of the mobile device's metadata (camera position, orientation, focal length, distortion) and determine an optimal alignment. The tracked real world position and orientation of the mobile device 321 are then used by the microprocessor 410 of the registration server 311 to calculate the transformation between the mobile device's coordinate system and the real world coordinate system at step 1359. The registration server also calculates tracking points and template images for the individual mobile devices 321 at step 1361, where, as described in more detail below, the tracking points and template images are used by the mobile device to update its transformation between the mobile device's coordinate system and the real world coordinate system as the mobile device 321 changes pose. The transformation between the mobile device's coordinate system and the real world coordinate system can be in the form of a set of matrices for a combination of a rotation, translation, and scale dilation to transform between the coordinate system of the mobile device 321 and the real world coordinates. The calculated transformation between the mobile device's coordinate system and the real world coordinate system and tracking points/template images

are respectively sent from the registration server 311 over the network interfaces 450 to the mobile device 321 at steps 1363 and 1365.

[0096] Returning now to Figure 13A and the flow as seen by the mobile device, the mobile device 321 receives the transformation between the mobile device's coordinate system and the real world coordinate system (step 1307) and the tracking points and template images (step 1309). Once the registration is complete and the information of steps 1307 and 1309 received, by using this data by the processors/memory 509 the mobile device 321 can operate largely autonomously without further interaction from the registration server as long the tracking is sufficiently accurate, with the internal tracking of the mobile device 321 continuing to operate and generate tracking data such as, for example, on a frame-by-frame basis.

[0097] At step 1311, the mobile device 321 aligns its coordinate system with the real world coordinate system based on the transformation between the mobile device's coordinate system and the real world coordinate system. This can include retrieving, for each frame of the images, tracking position and orientation, converting these to real world coordinates, and drawing 3D graphics content from the content server over the images. This correction can be implemented as an explicit transformation in the 3D graphics scene hierarchy, moving 3D shapes into the tracking frame of reference so that it appears in the correct location when composited with over the mobile devices images.

[0098] Using the tracking points and template images, the alignment of the device to real world coordinate systems is tracked at step 1313 and the accuracy of the tracking checked at step 1315. For example, every frame or every few frames, the basic features supplied by the registration process at step 1309 are detected in the mobile device's camera 501 and verified that they are in the expected location. If the tracking is accurate, the flow loops back to step 1313 to continue tracking. If the reference features cannot be found, or if they are not within a margin of their expected location, the registration process can be initiated again at step 1317 by sending updated image data and metadata to the registration server 311. Additionally, the mobile device 321 can periodically report usage and accuracy statistics back to the registration server 311.

[0099] Although Figure 3 explicitly illustrates only a single mobile device 321, and the flows of Figures 13A and 13B are described in terms of only a single mobile device, in operation the system will typically include multiple (e.g., thousands) such mobile devices and

the flows of Figures 13A and 13B can be performed in parallel for each such mobile device. Additionally, the distribution of the amount of processing performed the mobile device relative to the amount of processing performed on the servers can vary based on the embodiment and, within an embodiment, may vary with the situation, such as by the mobile devices or registration servers could monitor the communication speed in real time. For example, if a latency in communications between a mobile device and the servers exceed a threshold value, more processing may be shifted to the mobile devices, while if transmission rates are high additional processing could be transferred to servers to make use of their greater processing power.

[00100] Figure 14A is a more detailed flowchart of an embodiment for the operation of registration server 311. The registration server 311 retrieves the output of the three columns from registration processing 307 from the feature database 309 and combines these with the image data and metadata from a mobile device 321 to determine the transformation between the mobile device's coordinate system and the real world coordinate system. In terms of Figure 4, the inputs (image data and image metadata from the mobile devices 321 and point features, large scale features, and shape features from the feature database 309) can be received through the network interfaces 450 and the outputs (the coordinate transformations and tracking points and template images) transmitted to the mobile device 321 by the network interfaces 450. The processing steps of Figure 14A (e.g., 1411, 1415, 1419, 1421, 1425, 1433) can be performed by the microprocessor 410, with the resultant data (e.g., 1413, 1417, 1423, 1431) stored in the memory 420 or mass storage 430, depending on how the microprocessor stores it for subsequent access.

[00101] The point features from the database 309, such as in the form a descriptor and 3D real world coordinates in the form of scale invariant feature transformation (SIFT) features, and the mobile device image data and image metadata are supplied to processing block 1411 to determine 2D feature transformations, with the resultant output data of 2D and 3D feature transformation pairs 1413, which can again be presented in a SIFT format. The processing of to find 2D macro features 1415 matches the mobile device's 2D image data to the 3D large scale features. To find the 2D macro features from the mobile device's image data, the inputs are the 2D image data and corresponding image metadata from the mobile device 321 and the large scale feature data (macro features and their 3D coordinate data) from the feature database 309. The processing to find 2D macro features 1415 from the mobile device's images can

implemented as a convolutional neural network (CNN), for example, and generates matches as 2D plus 3D transformation pairs 1417 data for the large scale macro features of the venue.

[00102] For embodiments that use the 3D survey dataset, shape features extracted from the 3D survey data are combined with the image data and image metadata from the mobile device 321. The mobile device's image data and image metadata undergo image segmentation 1421 to generate 2D contours 1423 for the 2D images as output data. The image segmentation can be implemented on the registration server 311 as a convolutional neural network, for example. The 2D contour data 1423 can then be combined with the 3D contour data from the feature database 309 in processing to render the 3D contours to match the 2D contours within the images from the mobile device 321.

[00103] A camera pose solver 1419 generates the camera pose for mobile device 321 in real world coordinates 1431 as output data. The camera pose solver 1419 input data are the image data and image data from the mobile device 321, the 2D plus 3D feature transformation pairs 1413 data, and the macro 2D plus 3D transformation pairs 1417 data. The camera pose solver 1419 can also interact with the rendering of 3D contours and matching with 2D contour processing 1425. Based on these inputs, the output data is the camera pose of mobile device 321 in the real world coordinates 1431, which are then used to determine the transform so that the mobile device 321 can align its coordinate system to real world. The processing to calculate the pose offset transform 1433 uses the camera pose in real world coordinates 1431 and the image data and image metadata from mobile device 321. The device to real world coordinate transform can be a matrix of parameters for a translation to align the origins of the two coordinate systems, a rotation to align the coordinate axes, and a dilation, or scale factor, as distances may be measured differently in the two coordinate systems (e.g., meters in the mobile device 321 whereas measurement for a venue are given in feet). The device to real world coordinate transform can then be sent from the registration server 311 to the mobile device 321 along a set of tracking points and template images. Although described in terms of a single mobile device 321, this process can be performed concurrently for multiple mobile devices by the registration server.

[00104] Figures 14B-14D illustrate implementations for the registration of a mobile augmented reality device 321 with a central registration server or servers 311. In the embodiment of Figure 14A, the implementation sequentially performs each of the elements the registration process where the mobile device 321 sends image data and image metadata to a

central registration server 311, extracts features from the images data, matches features against the feature database, solves for the pose of the mobile device 321, and sends a device/real world coordinate transformation (either for an initial transformation to align the coordinate systems or to correct/update the transformation) back to the device. As the speed of the response of the registration server 311 can be a factor in a positive user experience, alternate implementations can be used to provide a quicker response time, such as the quick/detailed implementation of Figure 14C or the pipelined approach of Figure 14D. The presentation of Figures 14B-14D present the process in terms of three steps (extract features, match features, and solve for pose), it will be understood that alternate embodiments can use additional or different steps.

[00105] In the approach of Figure 14C, an initial correction is returned to the mobile device 321 followed by a more detail solution for solving the mobile device's pose. As represented in Figure 14C, the determination and return of an initial correction is shown in the upper sequence, with the more detailed solution in the lower sequence. The upper sequence is similar to Figure 14B and begins with the mobile device 321 sending image data and image metadata to the registration sever 311, but now only a subset of features is extracted from the image data by the registration server 311. As the number of extracted features is reduced, the determination of an initial correction can be performed more quickly than for the full process of Figure 14B. After the subset of features are extracted, the subset is matched against the feature database 309 to determine a quick solve for the mobile device's pose, with this initial correction then sent from the registration server 311 to the mobile device 321. The mobile device can then begin an initial alignment of coordinate systems based on the initial correction data. To provide a more detailed solve for the pose of the mobile device 321, the registration server 311 extracts the remaining features from the image data, matches these against the feature database 309, and then can refine the quick solve to generate a more detailed solve for the pose of the mobile device 321. The more detailed correction can then be used by the mobile device 321 to refine the quick result. Although Figure 14C illustrates the rough solution being determined and sent prior to starting the full registration process, in some embodiments these can overlap, such as beginning to extract the remaining features while the subset of features is being matched against the database.

[00106] Figure 14D illustrates an extension of the process of Figure 14C to a pipelined approach, incrementally returning better results as the registration server 311 repeatedly extracts features from the image data, matches each set of extracted features against the feature

database 309, repeatedly solves for the pose of the mobile device 321, and returns the updated corrections to the mobile device 321 from the registration server 311. How many features that are found and matched by the registration server 311 before solving and returning an initial solution to the mobile device 321 can be a tunable parameter, as can also be the solution accuracy requirements. For example, the system can adjust the thresholds for the number of features found, matched, and included in the pose solution before returning a solution based on the system's load to adapt to the number of devices undergoing the registration process. The approach of Figures 14C and 14D provide an early or partial result that may be of lower accuracy than that of Figure 14, but still be sufficient to start operating without the user wait that would result in waiting for the full quality result of the arrangement of Figure 14B.

[00107] The accuracy of the registration process can be increased through use of multiple concurrent registrations of a single mobile device. As described in the above discussion, to do effective augmented reality, the apparent position of virtual objects must correspond to real world locations in the video atop which they are displayed with a very high degree of accuracy. To do this, the mobile device can continually track its camera position and orientation relative to some arbitrary reference frame, then the registration process determines the relationship of that arbitrary frame to the real-world coordinate system of the venue. This lets the system draw something at a specific world location in its corresponding location in the arbitrary frame of the camera. As used above and in the following, a frame is a coordinate system with an origin, and orientation, and scale (e.g., meters or feet); and the combination of a camera's position and orientation is its pose. The correspondence of frames is determined by recording an image from the camera at a known pose, finding of visual features in that image, and then comparing the apparent location of those features to a database of known visual features and their 3D real-world locations. This provides the real-world pose of the camera, from which the system can calculate how to convert between the arbitrary frame and the real-world frame. This process of taking a camera image along with its pose in the arbitrary frame, finding features, and solving for the real-world pose and the correspondence between the frames is the registration for the mobile device. This correspondence can be represented by a 4x4 transformation matrix, a correction matrix.

[00108] The correction matrix obtained from a registration works extremely well when the mobile device is looking in the same direction as it was looking when the registration was performed. There is usually some error in the inferred camera pose, but the combination of

position error in one direction and orientation or focal length error in the other makes the visual features line up well. However, when the mobile device looks in a different direction, say 180 degrees away from the direction used in the registration, the errors can combine to cause a poor alignment between virtual and real world objects. This harms the illusion of the virtual objects actually being in the real world. There is also some inherent error in the mobile device's gyroscopes' ability to measure absolute orientation, and a 180 pan may be reported as a 175 degree pan, causing additional mis-alignment between the virtual and the real world. This error is not easily measurable, but it tends to be somewhat repeatable.

[00109] To address this source of error, multiple registrations can be performed for the same mobile device at several different camera orientations, then the mobile device can use the correspondence from whichever registration whose direction most closely matches the mobile device's current viewing direction. This results in a good alignment in whichever direction the mobile device is looking, whether the misalignment was caused by registration errors or gyroscope inaccuracies. By creating a smooth weighting function based on the current viewing direction and the directions of all the registrations, the mobile device can transition through the range of viewing directions, calculating a correction matrix that changes smoothly, resulting in an excellent visual result.

[00110] The alignment between the virtual and the real world diminishes over time because of normal drift due to tracking inaccuracies. To maintain good alignment over time, the system can periodically perform new registrations, typically determined by the age of the registration. In addition, whenever the mobile device looks in a new direction, a new registration can be performed to maintain accuracy throughout the entire viewing range. As a result, the system can continually maintain a set of active registrations, adding new ones and removing old ones. To keep the alignment between the virtual and the real smooth, rather than suddenly change the set of registrations used, new registrations can gradually "fade in" after being added to the mobile device's set of registrations and old registrations can "fade out" before being deleted from the set. This can be done by adjusting weights determined from the mobile device's view direction function.

[00111] Considering the multiple registration, or multi-reg, process in further detail, the view-tracking app (e.g., ARKit on an iPhone) on a user's mobile device 321 does not accurately measure rotation over large ranges of pan or tilt. If the system registers the camera looking in some direction, then pans away 180 degrees, the alignment between camera imagery and virtual

graphics will be off, typically by a few degrees. This error is somewhat consistent, in that if the mobile device 321 pans back to the original direction, the alignment will again be fairly accurate, and if panned those 180 degrees again, the error will be about the same amount. Additionally, if a user continues to pan another 180 degrees in the same direction to complete a 360 degree circle, ending up looking in the original direction, the alignment error will be about twice the 180 degree error. Panning and tilting the camera while looking at distant objects also causes some drift as the alignment varies over time. Additionally, any camera position error in a registration looking in a given direction will show up as an alignment error when panning 180 degrees away because a position error causes an orientation error that works against the position error to align correctly, and that orientation error adds to the position error when looking 180 degrees away. To reduce such alignment errors when panning a large amount from an original registration, multiple registrations corresponding to different viewing directions are remembered and the one that best corresponds to the current viewing direction is selected.

[00112] A mobile device 321 (or, more accurately, its view-tracking application such as ARKit on an iPhone) establishes its coordinate system when the app starts, with, in a common orientation, the origin at the location of the camera, +Y axis up, +X to the right and +Z pointing toward the back of the viewer based on whichever direction the viewer is facing at the time in a common embodiment. In this orientation, [0,0,-1] is the vector in mobile device's world space describing the direction the camera is looking. If the mobile device pan's 90 degrees to the right, the view direction will be [1,0,0], for example. The camera parameter to the app's render callback contains the camera-to-world transform as a 4x4 correction matrix. In general the columns of that matrix describe the current right, up, and behind vectors for the current orientation, in the coordinates of the mobile device's coordinate system. If the camera is in the exact position and orientation as when the app started, that matrix will be the identity.

[00113] An important quantity is the direction that the camera of the mobile device 321 is facing, which, as the Z direction, is the 3rd of the 4 columns of the correction matrix and (in the definition of the last paragraph) describes the behind vector in view-tracking app coordinates. As such, the current viewing direction is the negative value and should be set as a unit vector. To transform it to the real world coordinate system of a venue, it would be multiplied it by the inverse of the correction matrix returned by the registration process. If the correction matrix is \mathbf{M}_{cor} , \mathbf{Pt}_{rw} is a point vector in the real world coordinate system, and \mathbf{Pt}_{md}

is a point vector in the coordinate system of the mobile device 321, these are related as in equation (1):

$$\mathbf{Pt}_{md} = \mathbf{M}_{cor} \times \mathbf{Pt}_{rw}, \quad \text{eq. (1)}$$

where \times is a matrix multiplication. If \mathbf{Dir}_{rw} is the corresponding direction vector in the real world coordinate system and \mathbf{Dir}_{md} is the corresponding point vector in the view-tracking application coordinate system of the mobile device 321, these are related as in equation (2):

$$[\mathbf{M}_{cor}]^{-1} \times \mathbf{Dir}_{md} = \mathbf{Dir}_{rw}, \quad \text{eq. (2)}$$

where $[\mathbf{M}_{cor}]^{-1}$ is the inverse matrix of \mathbf{M}_{cor} . In this representation, the points are represented by a 4-element column vector with a 1 in the fourth element and the direction vectors are a 4-element column vector with a 0 in the fourth element. These are affected by orientation and scale, but not by translation.

[00114] Once a registration for the mobile device 321 is performed, the system knows where the mobile device's camera is located in the world's coordinates. If the location of an object in the venue (e.g., the location of a tee or a hole in a golf course), the system can calculate the direction from a viewer to the object as a unit vector with a 0 in the 4th element, the "target direction". A dot product (inner product multiplication) between the viewing direction vector and the target direction vector can provide information on how the close to the target is to the viewing direction, with a "1" corresponding to looking directly at it and a "-1" corresponding to looking 180 away from it.

[00115] Figures 16-19 are flow charts that present the multi-registration process in more detail, but first considering the process at a higher level, the process begins with a registration to get the real world to mobile device correction matrix. According to equation (1), the viewer's location in the real world coordinate system is:

$$[\mathbf{M}_{cor}]^{-1} \times \mathbf{Pt}_{md} = \mathbf{Pt}_{rw}$$

where \mathbf{Pt}_{md} is the viewer's location in ARKit space, or the column vector $[0, 0, 0, 1]$, which is the same as the last column of the inverse correction matrix. The mobile device remembers this correction matrix and the corresponding view direction in the venue's real world coordinates. The mobile device can then use that correction matrix to draw graphics over images of the venue on its display.

[00116] The mobile device continues to calculate the dot product between the registration view direction and the current view direction, both in the venue's real-world coordinates.

When the viewer is looking in a different direction and holding the camera still such that the view direction is not changing much over time, the system can start another registration. The mobile device will also remember the additional correction matrix and view direction in the mobile device's view-tracking app coordinates.

[00117] With multiple such registrations saved on the mobile device, at each frame the mobile device calculates the dot product the current view direction vector and each registration's view direction vector. The mobile device can then calculate a weighted sum of the different correction matrices based on how close the current view direction is to each one, then use that matrix to draw the AR graphics over the view. The weighted correction matrix, $M_{corTotal}$, can be expressed as:

$$M_{corTotal} = weight_A * M_{corA} + weight_B * M_{corB} + \dots, \quad eq. (3)$$

where $weight_i$ is the weight factor for registration I, M_{cori} is the correction matrix for registration I, and the sum runs over all the current registrations and the sum of the weights is 1. A number of embodiments are possible for the determination of the weights. For example, in one embodiment to calculate weight values, for each registration i a $factor_i$ can be computed as:

$$factor_i = 1/(1.01-dot_i),$$

where dot_i is a dot product between the current view vector and the view factor corresponding registration i. As the dot product can range from -1 to 1, in this example the $factor_i$ values can range from about 0.5 to 100. The weights can then be calculated as:

$$weight_i = factor_i / (\text{sum of all factors}),$$

so that they are then normalized to sum to one. Consequently, a $weight_i$ will be almost 1 when the current viewing direction is near the direction in which the registration i was performed, and otherwise be between 0 and 1. The weights should also vary smoothly as view direction pans around. Figure 15 illustrates an example. Having multiple registrations available to the mobile device 321 can be used to improve its position estimation. If multiple registrations are close in time, so that the effect of tracking drift is minimal, the mobile device 321 can average the position estimates to provide a better position fix. The mobile device can use the onboard tracking to estimate how much the phone has moved between registrations and factor that in to the determination. An estimate of accuracy for the position fix can be included as part of the registration process, with those accuracy estimates used to better weight position estimates more in the average.

[00118] Figure 15 illustrates weight values for three different registrations as a function of the viewing angle of the mobile device. In the example, the three registrations corresponding to a view direction of -90° , 0° , and 90° and the corresponding weight values are respectively 1505, 1501, and 1503. As illustrated in Figure 15, the weight 1501 corresponding to an azimuth of 0° is 1 at 0° and smoothly decrease to 0 as the view angle go to -90° or $+90^\circ$. Similarly, the weight 1503 corresponding to a view angle of 90° is 1 at this view angle and varies smoothly and is 0 at 0° and -90° , and the weight 1505 corresponding to a view angle of -90° is 1 at this view angle and varies smoothly and is 0 at 0° and $+90^\circ$.

[00119] In some embodiments, a confidence value can be assigned to registrations and be taken into account when determining the weights. For example, the confidence value can be a function of time, decreasing as the registration ages and, if it falls below a threshold value, eventually discarded. Other factors can also be used to determine confidence values at the time of registration, such how stable the mobile was during the registration, the quality or number of reference features used in the registrations, or other factors that could affect the accuracy of the registration.

[00120] When a new registration is added to a set of a mobile device's set of registrations, its weight value can gradually increase from 0 to its final value over some time interval (e.g., one second or so) to avoid jumps in the positioning of displayed AR content. Similarly, as registrations age, their weights can be gradually decreased to 0 before being deleted from the mobile device's set of registrations. For example, after calculating the weights for each registration of every frame, the mobile device can also multiply a weight by a fade in/fade out factor of 0 to, then normalize all the weights so they still add to 1.

[00121] With respect to embodiments for determining when and if to perform an additional registration, the mobile device can consider the current view direction relative to the viewing directions for the existing registrations by computing the dot product for the existing registrations, which corresponds to the cosine of the angle between the directions. If the cosine is less than, for example, 0.5 (an angle of 60°), the mobile device can request the system to initiate a new registration for this view angle. A determination of whether the mobile device is relatively stable, a calculation of the how fast the mobile device is panning or tilting can be made by subtracting the current view vector of a frame from the view vector of a previous frame and if the amount of change in direction per frame is below a threshold, a registration can be initiated.

[00122] Figures 16-19 provide more detail on embodiments for multiple registration of a mobile device 321, providing additional detail to Figures 13A and 13B when using multiple registrations. Figures 15-18 can be performed for every frame (typically 60 times a second) as the mobile device reports its tracking results, but before it updates the display with video and AR graphics. Figure 19 is performed asynchronously after making a registration request of the registration server. The following discussion is presented in the context of an embodiment where the registrations come from the registration server 311, but in other embodiments some or all of the multiple registrations can be based on discrete registrations done on the mobile device 321 itself.

[00123] Figure 16 is a flowchart of a multiple registration embodiment for determining whether the mobile device 321 is in a normal application mode or waiting to reinitialize. This can be performed for every frame once camera tracking is complete for the mobile device. Starting at 1601, the mobile device checks its tracking state and checks for status such signal drops, orientation, and camera exposure and can also update status and other timer values, such as those mentioned in later steps of the flow. Step 1603 then determines whether the mobile device is in its normal operating mode and, if not, the flow continues on to step 1605 to determine whether the tracking state has been good for greater than a threshold time of over, for example, some number $\langle n \rangle$ seconds. (Here, $\langle n \rangle$ seconds is a threshold depending on the embodiments and the time threshold of other parameters (e.g., step 1609 or 1613 in Figure 16 and steps in later figures) using the same notation may be the same or independent values.) If the mobile device's tracking has not been good for over the threshold time (No path from step 1605), the flow continues on to step 1607 and the mobile device clears out the current set of registrations, after which no further processing is performed for the frame and the AR graphics are not displayed.

[00124] Backing up to step 1603, if the mobile device 321 is in normal operation mode, the flow goes to step 1609 for the series of decisions down the right-hand column of Figure 16. In step 1609 the mobile device 321 determines whether the tracking state has been bad for over a threshold value (e.g., a number of $\langle n \rangle$ seconds) and, if so, goes to step 1607; and, if not continues on to step 1611. In step 1611 the mobile device determines whether it has been dropped (i.e., lost is wireless connection to the servers) and, if so, goes to step 1607; and, if not, continues to the decision of step 1613. Step 1613 determines whether or not the mobile devices has been laying face-up or face down for over some threshold amount of time (e.g.,

<n> seconds), since this could indicate that the mobile device 321 is not currently being actively used and has been, for example, set down on a table. If it has been stationary in this way for some time (Yes path out of step 1613), the flow continues on to step 1607; and, if not, the flow continues to step 1615. Step 1615 looks at whether the exposure from the mobile device's camera have been dark for over some threshold amount of time, such as would be the case if the mobile device's user have placed it in a pocket and was no longer actively viewing: if so, the flow goes to step 1607; and, if not, the process continues on to the subsequent steps of deciding whether to perform a registration process. The various decisions of steps 1609, 1611, 1613, and 1615 are shown in a particular order in Figure 16, but, depending on the embodiment, these can be done in other order or have one of more of these steps done concurrently.

[00125] Figure 17 is a flowchart of a multiple registration embodiment for deciding whether to initiate a new registration for a mobile device 321. The mobile device can perform this process on a frame by frame basis once the mobile device's tracking is complete and state of the state of the view-tracking app is updated as described with respect to Figure 16. Starting at step 1701, the mobile device 321 determines its current orientation and viewing direction. Based on this information, the mobile device can then determine whether the current viewing orientation is acceptable based on its current set of registrations at step 1703 and, if not, the mobile device 321 continues on with using the current registrations and move on the next steps of operation. If the mobile device 321 is an acceptable orientation for performing a registration, the flow then continues on to step 1705 and checks on whether there is already a pending registration near the current viewing direction at step 1705, since, if so, the mobile device 321 can continue on to use the current registration and continuation with normal operation while waiting for the registration server 311 to complete and provide the pending registration. If there is not a pending registration request near the current viewing direction, step 1707 continues on to check on whether the mobile device 321 has an existing registration of high confidence near the current viewing direction: if so, the mobile device can continue to use the current registrations and normal operation; if not, the flow continues on to 1709 to start a request for new registration in the current view direction.

[00126] The registration process for the current viewing direction can then be performed much as described in Figures 13A and 13B, beginning for the mobile device 321 at steps 1709 by capturing image data and metadata, which it can then re-scale and compress. The mobile

device 321 can then send the compressed image and metadata to the registration server 311 as part of a registration request at 1711 and then continue on to the next steps of operating.

[00127] Figure 18 is a flowchart of a multiple registration embodiment for calculating a camera correction from the current set of registrations, where the mobile device 321 can perform this process on a frame by frame process once checking that a request for a new registration is complete. Step 1801 checks that there is at least one good registration available for the mobile device 321 and, if not, the AR content is not displayed. If there is at least one good registration in the set of registrations, in step 1803 the mobile device calculates weighting factor for each of the good registrations in the set based on how close the current viewing direction align the corresponding direction of each of the registrations. Normalization of the weighting factors follows in step 1805 so that the weight sum to 1 while favoring the more closely aligning registrations, as discussed above with respect to Figure 15. In some embodiments, in step 1807 the mobile device 321 can then further adjust the weights based their age or other confidence factors and on if a registration has been newly introduced to allow fading in and fading out of the registrations. The weights can then be used to combine the correction matrices as described above with respect to equation (3) in step 1809. The weighted sum of step 1809 is then used as the correction matrix in step 1811 to transform AR content supplied from the content server 323 from the venue's real world coordinate system into the mobile device's coordinate system for display of the AR content by the mobile device 321.

[00128] Figure 19 is a flowchart of an embodiment for a mobile device 321 to handle a response to a registration request on receipt of a reply to a registration request from a registration server 311. The process can be performed asynchronously after making a registration request of the registration server. In an initial check at step 1901, the mobile device 321 checks on whether the registration succeeded and, if not, the mobile device continues operating. If the registration was successful, at step 1902 in some embodiments the mobile device 321 can check on whether there are any nearby existing registrations and, if so, the mobile device 321 continues on and decides on whether to use the registration. If there are not any nearby registrations available, the flow can go from step 1902 to step 1911 as described below. In other embodiments, this decision can be incorporated into step 1909. In step 1903 the mobile devices checks for other registrations in the set that have a registration viewing direction close to the that of the new registration's direction, where this can be done by taking a dot product of the corresponding direction vectors and seeing whether it is above some

threshold value. The mobile device 321 checks the confidence level of the registration against the confidence level of these close registrations in step 1905 and if the new registration does not have a higher confidence level, it is not introduced into the current set of registrations and nothing is done with it at this point as far as this new registration. If any existing close registration has lower confidence than the new one, it is then faded out at step 1907 over some number of <n> seconds and then deleted.

[00129] Continuing on to step 1909 the mobile device determines whether the new registration has a higher confidence than any of the existing nearby registrations, not just the close ones, as done above. If not, the new registration is discarded at step 1913 before continuing on. If, instead, the new registration is of a higher confidence than any registration existing nearby registration (or if there is not a nearby existing registration), it is faded in to the set of some number of <n> seconds in step 1911 and the mobile device continues on with providing the AR content based on the set of registrations. In some embodiments, following the addition of the new registration, at step 1915 can be used by the mobile device 321 to improve its estimate of its position. If multiple registrations are close in time (so the effect of tracking drift over time is minimal), the mobile device 321 can average the position estimates to provide a better position fix. Onboard tracking can estimate how much the mobile device 321 has moved between registrations and factor this into the determination. The mobile device 321 can estimate how good its position fix is as part of the registration process, and use those accuracy estimates to weight better position estimates more in the average.

[00130] Although the preceding of discussion of multi-registration has been discussed with respect to a single mobile device 321, when multiple mobile devices are in use (as in the following discussion), it will be understood that multiple registrations can applied to any or all of these mobile devices, although the following discussion will only refer to use of a single registration.

[00131] Figure 20 illustrates the use of multiple mobile devices 321a, 321b, 321c, 321d, and 321e with the registration server 311 and content server 323. The example of Figure 20 shows five mobile devices, but the number can range from a single device to large numbers of such devices used by viewers at an event venue. The mobile device can be of the same type or of different types (smart phone, tablet, or AR headset, for example). Each of the mobile devices 321a, 321b, 321c, 321d, and 321e can independently supply the registration server 311 with image data and image metadata as described above for a single mobile device 321. The

registration server 311 can concurrently and independently perform the registration process for each of the mobile devices, providing them with their corresponding transformation between the mobile device's coordinate system and the real world coordinate system and with their own set of tracking points and reference images. Each of the mobile devices 321a, 321b, 321c, 321d, and 321e can independently request and receive 3D graphics and other content from the content server 323. Although Figure 20 represent the registration server 311 and content server 323 as separate blocks, in an actual implementation each of these can correspond to one or more servers and parts or all of their functions can be combined within a single server.

[00132] In some embodiments some or all of the mobile devices 321a, 321b, 321c, 321d, and 321e can provide crowd-sourced survey images that can be used by registration processing 307 to supplement or, in some cases, replace the survey images from a survey camera rig 301. Depending on the embodiment, the crowd-sourced survey images can be one or both of the image data and image metadata supplied as part of the registration process or image data and image data generated in response to prompts from the system. The crowd-sourced survey images can be provided before or during an event. In some cases, such as extended outdoor venue (a golf course or route for a cycling race), there may be activity at the location of some viewers but not others, so that some of the crowd-sourced survey images could be used for assembling the feature database 309 relevant to a location prior to activity at the location, while other crowd-sourced survey images or other data would be relevant to locations of current activity.

[00133] Once a mobile device 321 has been registered, it can receive 3D graphics and other content for display on the mobile device. Figures 1 and 2 include some example of such content, with Figure 21 presenting a block diagram of the distribution of content to user's mobile devices.

[00134] Figure 21 is a block diagram of an embodiment for supplying content to one or more user's mobile devices. Figure 21 explicitly represents two such mobile devices, 321a and 321b, but at an actual event there could be large numbers of such mobile devices at a venue. The mobile devices 321a and 321b request and receive content from the content server 323. Although the specifics will vary depending on the venue and the type of event, Figure 21 illustrates some examples of content sources, where some examples of content were described above with respect to Figures 1 and 2.

[00135] A content database 327 can be used to supply the content server 323 with information such as 3D graphics and other information that can be determined prior to an event, such as player information, elevation contours, physical distances, and other data that can be determined prior to event. Some of this content, such as 3D contours may also be provided from the registration server and the feature database 309. The content server 323 may also receive live data from the venue to provide as viewer content on things such as player positions, ball positions and trajectories, current venue conditions (temperature, wind speed), and other current information on the event so that live, dynamic event data visualization can be synchronized to the playing surface live action. One or more video cameras 325 at the venue can also provide streamed video content to the mobile devices 321a and 321b: for example, in some embodiments if a user of a mobile device requests a zoomed view or has there is subject to occlusions, the cameras 325 can provide a zoomed view or fill in the blocked view.

[00136] For some embodiments, the different mobile devices 321a and 321b can also exchange content as mediated by the content server 323. For example, the viewers can capture and share content (amplified moments such as watermarked photos) or engage in friend-to-friend betting or other gamification. The viewer can also use the mobile device 321a or 321b to send gamification related requests (such as placing bets on various aspects of the event, success of a shot, final scores, and so on) and responses from the content server 323 to the internet, such as for institutional betting or play for fun applications.

[00137] Figure 22 is a flowchart describing one embodiment of a process for requesting and receiving graphics by a registered mobile device 321, providing more detail for step 611 of Figure 6. At step 2201 the registered mobile devices 321a, 321b, 321c, 321d, 321e of Figure 20 request graphics content from content server 323. (The mobile devices 321a, 321b, 321c, 321d, 321e will have already received the transformation between the mobile device's coordinate system and the real world coordinate system from the registration server 311.) The requests for graphics at step 2201 can be based both on direct user input and on automatic requests by a mobile device 321. For example, as the mobile device has its field of view changed, new graphics can be requested based on the corresponding change in pose, in which case the mobile device can automatically issue a request for graphs appropriate to the new view of the venue. The graphics can also be used based on what is occurring in the view, such as when one set of players in a golf tournament finish a hole and a new set of players start the hole. User input to select graphics can be selected through the display of the mobile device

321, such as by the touch screen of a smart phone or laptop computer, or by pointing within the field of view of the camera for the mobile device. For example, a viewer may indicate a player's position within the view to request graphics of information on the player.

[00138] In step 2203, mobile devices 321a, 321b, 321c, 321d, 321e receive from content server 323 their respective graphics to be displayed by the mobile devices 321a, 321b, 321c, 321d, 321e over a view of the venue, where the graphics are specified by location and orientation in the real world coordinate system. Each of the mobile devices 321a, 321b, 321c, 321d, 321e can then use processor(s) 509 to convert the graphics into the mobile device's coordinate system based on the transformation at step 2205. The transformed graphics are then presented over a view of the venue by display 503 at step 2207.

[00139] The discussion to the point has focused on embodiments of augmented reality system using mobile devices, such as augmented reality enabled devices such as mobile phones, headsets, or glasses that are used to enhance a viewer's experience at an event's venue. The techniques can also be extended for use at remote locations, such as at home or a sport bar, for example, where the event is viewed on a television in conjunction with a smart television as part of "tabletop" embodiment.

[00140] Figures 23 and 24 illustrate examples of a tabletop embodiment for respective events at a golf course venue and a basketball venue, corresponding to the at-venue embodiments of Figures 1 and 2. In a tabletop embodiment, in addition to being able to view the event on a television, the views can also view the event on mobile devices, such as a smart phone, with overlaid graphs and also to view graphics on a model of the venue with graphics.

[00141] Figure 23 illustrates the same event and venue as Figure 1, but viewed at a remote venue on a television 2300. The event can again be viewed on the display of a mobile device 2321a or 2321b with graphics and other AR content displayed along with the view of the event. A tabletop view 2330, similar to the zoomed view 130 of a model of the view in Figure 1 can also be viewed by a head mounted display 2323. The augmented view can also present content, such as player statistics 2301 or course conditions such as the wind indication graphic 2311.

[00142] The tabletop view 2330 can include the graphics as described above for the in-venue view, both on the mobile device 121 and also in the zoomed view 130 of Figure 1. Some examples include player info and ball location 2331, concentric distances to the holes 2333, and a contour grid 2339, as well as gamification graphics such as wager markers 2341.

[00143] Figure 24 illustrates the same event and venue as Figure 2, but viewed at a remote venue on a television 2400. A viewer can again view the event with augmented reality graphics on a mobile device 2421 with a display screen, the same as those presented above for in-venue viewing, or as a tabletop view 2330 presentation when viewed with an augmented reality head mounted display 2423. In the tabletop view 2460, the augmented reality content can again include content such as player statistics 2451 and 2461 described above with respect to Figure 2, along with gamification graphics 2441.

[00144] Figure 25 is a block diagram of elements of a tabletop embodiment. Similar to Figure 3, Figure 25 again illustrates a registration server 2511 and a content server 2523, along with a mobile device 2521 such as a smart phone or other mobile device with a screen display. These elements can operate much as described above for the corresponding elements of Figure 3 and other figures, but where the other elements of Figure 3 are not explicitly shown in Figure 25.

[00145] Figure 25 also includes a television 2551 for remote viewing of the event, where the television may be connected to receive content from one or both of the registration server 2511 and content server 2523, receive content by another channel (cable or internet, for example), or a combination of these. The mobile device 321 may also interact with the television 2551 to receive content or transmit control signals, such as to change views or request content. Figure 25 further includes a head mounted display 2531 such as an AR headset or AR glasses. The display of the head mounted display 2531 can display the tabletop view 2530, along with AR graphics.

[00146] Figure 26 is flowchart for the operation of tabletop embodiment. As with the in-venue flow of Figure 6, prior to an event a model of the venue is built. At step 2601 the venue is prepared for survey, with the survey images collected at step 2603. Steps 2601 and 2603 can be as described above with respect to steps 601 and 603 and can be the same as these steps, with the process for in-venue enhanced viewing and the process for remote viewing being the same process. At step 2605 a tabletop model of the venue is built in much the same way as described with respect to step 605, but additionally the model of the venue is built for a tabletop display. In the tabletop view such as 2330 or 2460, rather than being display over a view of the venue as viewed through a head mounted display of the mobile device or on the display of the mobile device, at a tabletop position at the remote venue a representation of the venue is also presented, with the AR graphics presented over the representation. When viewed with an

augmented reality head mounted display 2323 or 2423, the venue representation with graphics is displayed at a designed location (i.e., a tabletop) within the remote venue.

[00147] At step 2607 the mobile devices 2321/2421 and 2323/2423 are register similarly to step 607 of Figure 6, but now the position of where the tabletop view 2330/2460 is to be located by the head mounted displays is also determined. This position can be determined by input from the views of the head mounted displays 2323/2423 within venue at step 2609. Although the movements at a remote venue will often be more limited than for in-venue viewing, tracking (similar to step 609) is performed at step 2611, both to accurately display the graphics, but also to maintain the laptop model in its location. At step 2613, requested graphics are again provided to the views on their mobile devices.

[00148] According to one set of aspects, a method includes maintaining by a mobile device of a plurality of registrations for the mobile device within a venue, each of registrations corresponding to a different viewing direction within the venue and including a correction between a real world coordinate system for the venue and a coordinate system of the mobile device, and determining by the mobile device of a current viewing direction for a camera of the mobile device. The method also includes forming by the mobile device of a weighted sum of the corrections for the plurality of registrations based upon the current viewing direction, a weight for each of the corrections in the weighted sum depending on how closely the current viewing direction aligns with viewing direction of the corresponding registration of the correction; displaying on a display of the mobile device a view of the venue in the current viewing direction of the mobile device from the camera of the mobile device; receiving by the mobile device of augmented reality (AR) content for the venue in the real world coordinate system of the venue; transforming by the mobile device of the AR content for the venue into the mobile device's coordinate system using the weighted sum of the corrections; and displaying the transformed AR content over the view of the venue on the display of the mobile device.

[00149] In other aspects, a mobile device includes mobile device, comprising: a camera configured to generate image data; a display configured to display the generated image data; memory; and one or more processing circuits. The one or more processing circuits are configured to: maintain in the memory a correction between a real world coordinate system for a venue and a coordinate system of the mobile device for each of a plurality of registrations for the mobile device within a venue, each of registrations corresponding to a different viewing

direction within the venue; determine a current viewing direction within the venue for the camera; form a weighted sum of the corrections for the plurality of registrations based upon the current viewing direction, a weight for each of corrections in the sum depending on how closely the current viewing direction aligns with viewing direction of the corresponding registration of the correction; display on the display a view of the venue in the current viewing direction of from the camera;

[00150] receive augmented reality (AR) content for the venue in the real world coordinate system of the venue; transform the AR content for the venue into the mobile device's coordinate system using the weighted sum of the corrections; and display the transformed AR content over the view of the venue on the display.

[00151] In other aspects, a system includes one or more mobile devices and one or more servers. Each of the mobile devices comprises: one or more mobile devices, each of the mobile devices comprising: a camera configured to generate image data; a display configured to display the generated image data; memory; and one or more processing circuits. The one or more processing circuits are configured to: maintain in the memory a correction between a real world coordinate system for a venue and a coordinate system of the mobile device for each of a plurality of registrations for the mobile device within a venue, each of registrations corresponding to a different viewing direction within the venue; determine a current viewing direction within the venue for the camera; form a weighted sum of the corrections for the plurality of registrations based upon the current viewing direction, a weight for each of corrections in the sum depending on how closely the current viewing direction aligns with viewing direction of the corresponding registration of the correction; display on the display a view of the venue in the current viewing direction of from the camera; receive corresponding augmented reality (AR) content for the venue in the real world coordinate system of the venue; transform the corresponding AR content for the venue into the mobile device's coordinate system using the weighted sum of the corrections; and display the transformed corresponding AR content over the view of the venue on the display. The one or more servers are configured to provide the AR content for the venue in the real world coordinate system of the venue to the corresponding mobile device in response to a request for the corresponding AR content.

[00152] For purposes of this document, reference in the specification to "an embodiment," "one embodiment," "some embodiments," or "another embodiment" may be used to describe different embodiments or the same embodiment.

[00153] For purposes of this document, a connection may be a direct connection or an indirect connection (e.g., via one or more other parts). In some cases, when an element is referred to as being connected or coupled to another element, the element may be directly connected to the other element or indirectly connected to the other element via intervening elements. When an element is referred to as being directly connected to another element, then there are no intervening elements between the element and the other element. Two devices are “in communication” if they are directly or indirectly connected so that they can communicate electronic signals between them.

[00154] For purposes of this document, the term “based on” may be read as “based at least in part on.”

[00155] For purposes of this document, without additional context, use of numerical terms such as a “first” object, a “second” object, and a “third” object may not imply an ordering of objects, but may instead be used for identification purposes to identify different objects.

[00156] For purposes of this document, the term “set” of objects may refer to a “set” of one or more of the objects.

[00157] The foregoing detailed description has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit to the precise form disclosed. Many modifications and variations are possible in light of the above teaching. The described embodiments were chosen in order to best explain the principles of the proposed technology and its practical application, to thereby enable others skilled in the art to best utilize it in various embodiments and with various modifications as are suited to the particular use contemplated. It is intended that the scope be defined by the claims appended hereto.

CLAIMS

We claim:

1. A method, comprising:

maintaining by a mobile device of a plurality of registrations for the mobile device within a venue, each of registrations corresponding to a different viewing direction within the venue and including a correction between a real world coordinate system for the venue and a coordinate system of the mobile device;

determining by the mobile device of a current viewing direction for a camera of the mobile device;

forming by the mobile device of a weighted sum of the corrections for the plurality of registrations based upon the current viewing direction, a weight for each of the corrections in the weighted sum depending on how closely the current viewing direction aligns with viewing direction of the corresponding registration of the correction;

displaying on a display of the mobile device a view of the venue in the current viewing direction of the mobile device from the camera of the mobile device;

receiving by the mobile device of augmented reality (AR) content for the venue in the real world coordinate system of the venue;

transforming by the mobile device of the AR content for the venue into the mobile device's coordinate system using the weighted sum of the corrections; and

displaying the transformed AR content over the view of the venue on the display of the mobile device.

2. The method of claim 1, further comprising:

determining the weight for each of the corrections in the weighted sum based an angle between a vector for the current viewing direction and a vector for the viewing direction of the corresponding registration of the correction.

3. The method of any of claims 1-2, wherein each of the plurality of weights is a smoothly varying function of the current viewing direction.

4. The method of any of claims 1-3, wherein the plurality of weights sum to one and each of the plurality of weights has a value of one when the current viewing direction is the same as the viewing direction of the corresponding registration of the correction.

5. The method of any of claims 1-4, further comprising:

requesting by the mobile device of an additional registration from a registration server, the requesting including transmitting from the mobile device to the registration server of image data and metadata of a corresponding viewing direction for the additional registration;

in response to requesting by the mobile device of the additional registration from a registration server, receiving the additional registration; and

including the additional registration within the plurality of registrations.

6. The method of claim 5, wherein, when including the additional registration in the plurality of registrations, the additional registration is faded in over a time interval.

7. The method of any of claims 5-6, wherein requesting an additional registration includes:

determining whether the mobile device is in an acceptable viewing orientation;

and

determining whether there is not already a high confidence registration near the viewing direction corresponding to the additional registration.

8. The method of any of claims 5-7, further comprising:

determining by the mobile device from one or more of the plurality of registrations a position estimate for the mobile device; and

updating by the mobile device of the position estimate based on the additional registration.

9. The method of any of claims 1-8, wherein the weight for each of the corrections in the weighted sum further depends on a confidence value of the corresponding registration of the correction.

10. The method of claim 9, wherein each of the confidence values is dependent on a quality of the registration.

11. The method of any of claims 9-10, wherein each of the confidence values is dependent on an age of the registration.

12. The method of any of claims 9-11, further comprising:
deleting a first registration from the plurality of registrations in response to the confidence value of the first registration having a low value.

13. A mobile device, comprising:

a camera configured to generate image data;

a display configured to display the generated image data;

memory; and

one or more processing circuits configured to:

maintain in the memory a correction between a real world coordinate system for a venue and a coordinate system of the mobile device for each of a plurality of registrations for the mobile device within a venue, each of registrations corresponding to a different viewing direction within the venue;

determine a current viewing direction within the venue for the camera;

form a weighted sum of the corrections for the plurality of registrations based upon the current viewing direction, a weight for each of corrections in the sum depending on how closely the current viewing direction aligns with viewing direction of the corresponding registration of the correction;

display on the display a view of the venue in the current viewing direction of from the camera;

receive augmented reality (AR) content for the venue in the real world coordinate system of the venue;

transform the AR content for the venue into the mobile device's coordinate system using the weighted sum of the corrections; and

display the transformed AR content over the view of the venue on the display.

14. The mobile device of claim 13, wherein the one or more processing circuits are further configured to:

determine the weight for each of the corrections in the weighted sum based on an angle between a vector for the current viewing direction and a vector for the viewing direction of the corresponding registration of the correction.

15. The mobile device of any of claims 13-14, wherein the one or more processing circuits are further configured to:

request an additional registration from a registration server, the request including transmitting to the registration server of image data and metadata of a corresponding viewing direction for the additional registration;

in response to requesting the additional registration from a registration server, receiving the additional registration; and

include the additional registration within the plurality of registrations.

16. The mobile device of claim 15, wherein the one or more processing circuits are further configured to:

determining by the mobile device from one or more of the plurality of registrations a position estimate for the mobile device; and

updating by the mobile device of the position estimate based on the additional registration.

17. The mobile device of any of claims 13-16, wherein the weight for each of the corrections in the weighted sum further depends on a confidence value of the corresponding registration of the correction.

18. A system, comprising:

one or more mobile devices, each of the mobile devices comprising:

a camera configured to generate image data;

a display configured to display the generated image data;

memory; and

one or more processing circuits configured to:

maintain in the memory a correction between a real world coordinate system for a venue and a coordinate system of the mobile device for each of a plurality of registrations for the mobile device within a venue, each of registrations corresponding to a different viewing direction within the venue;

determine a current viewing direction within the venue for the camera;

form a weighted sum of the corrections for the plurality of registrations based upon the current viewing direction, a weight for each of corrections in the sum depending on how closely the current viewing direction aligns with viewing direction of the corresponding registration of the correction;

display on the display a view of the venue in the current viewing direction of from the camera;

receive corresponding augmented reality (AR) content for the venue in the real world coordinate system of the venue;

transform the corresponding AR content for the venue into the mobile device's coordinate system using the weighted sum of the corrections; and

display the transformed corresponding AR content over the view of the venue on the display;

one or more servers configured to:

provide the AR content for the venue in the real world coordinate system of the venue to the corresponding mobile device in response to a request for the corresponding AR content.

19. The system of claim 18, wherein the one or more processing circuits of each of the mobile devices is further configured to

request an additional registration from the one or more servers, the request including transmitting from the mobile device to the registration server of image data and metadata of a corresponding viewing direction for the additional registration;

receive the additional registration; and

include the additional registration within the plurality of registrations, and wherein the one or more servers further configured to:

supply the additional registration in response to the request the additional registration.

20. The system of any of claims 18-19, wherein the weight for each of the corrections in the weighted sum further depends on a confidence value of the corresponding registration of the correction.

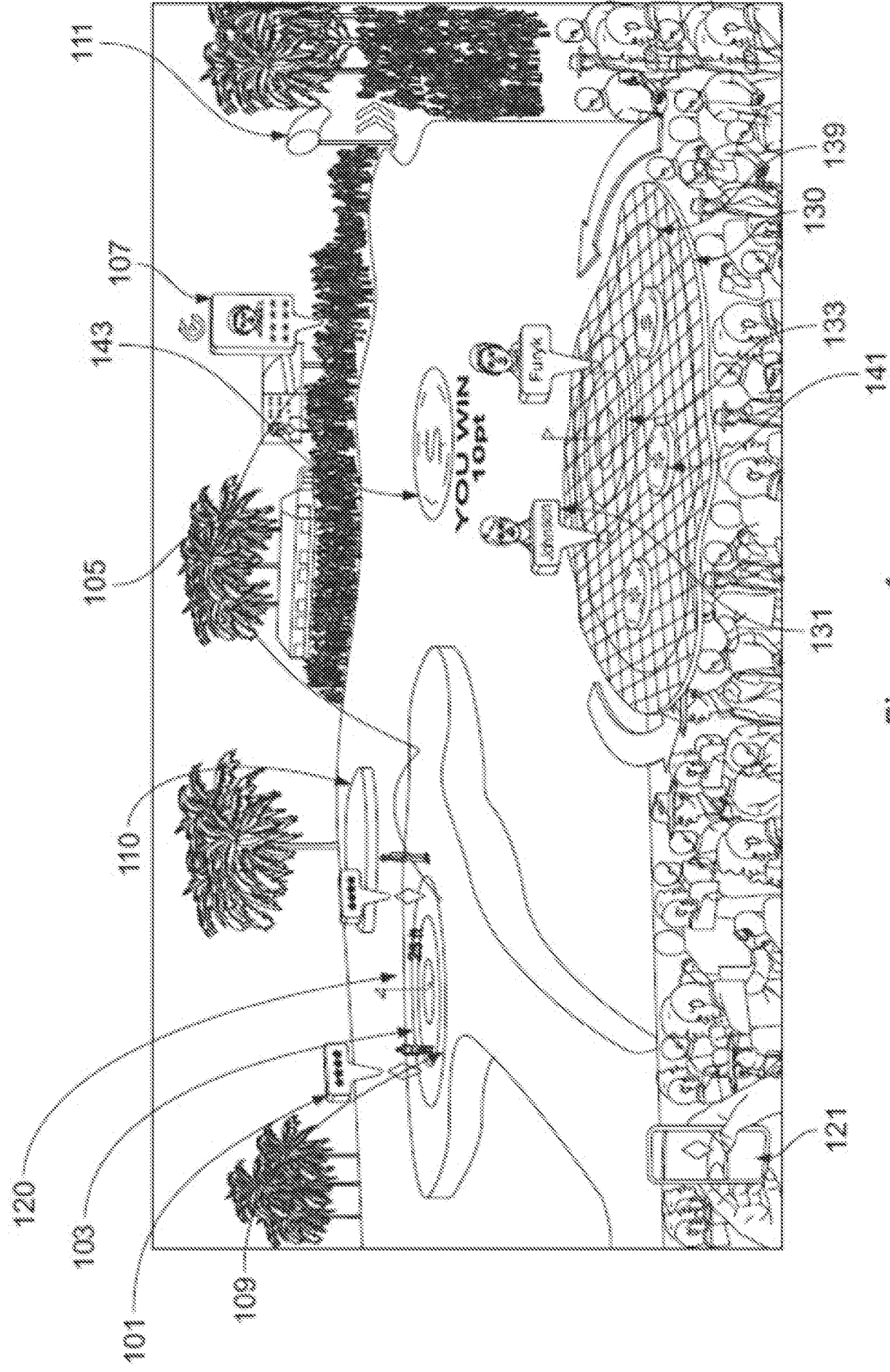


Figure 1

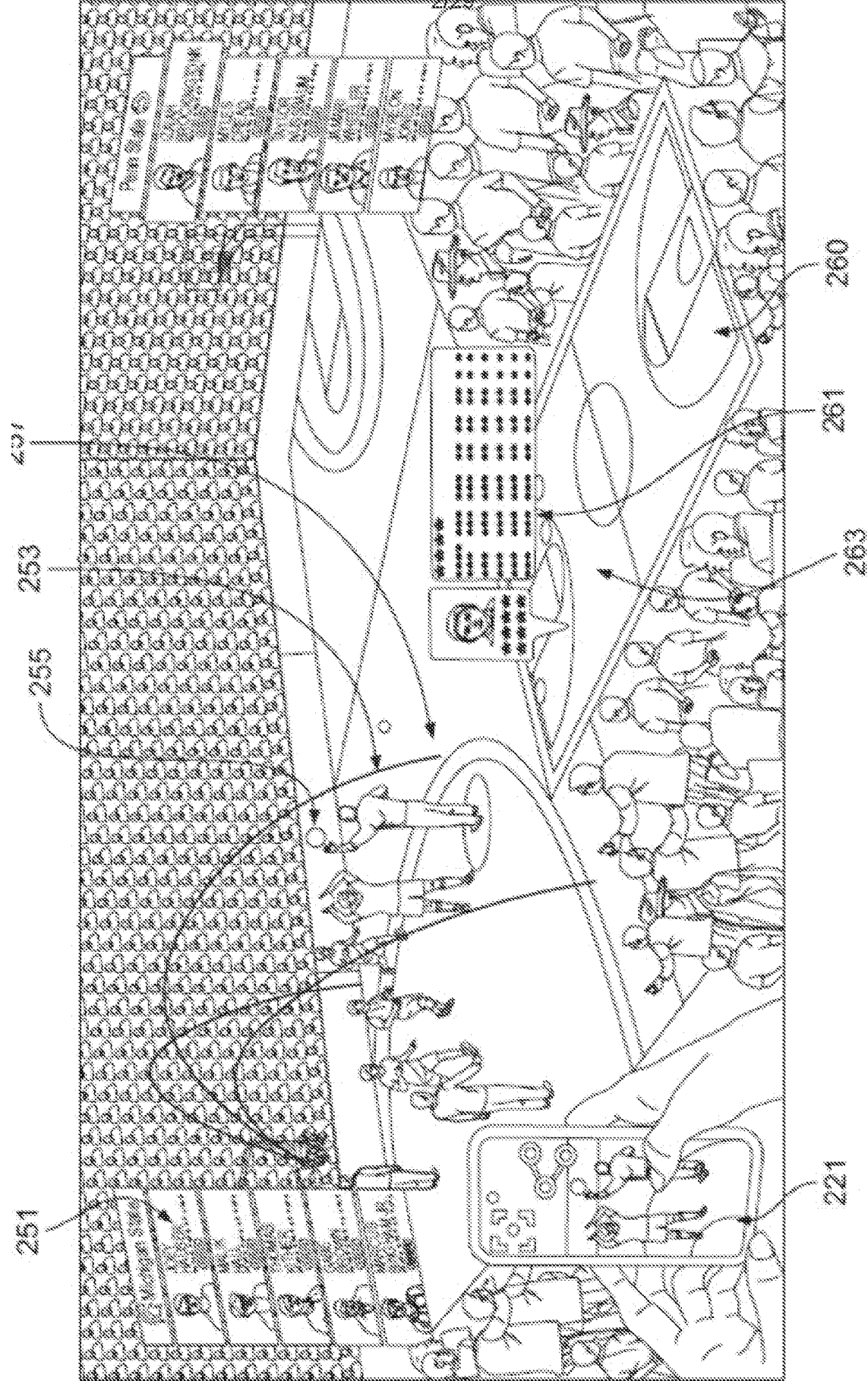


Figure 2

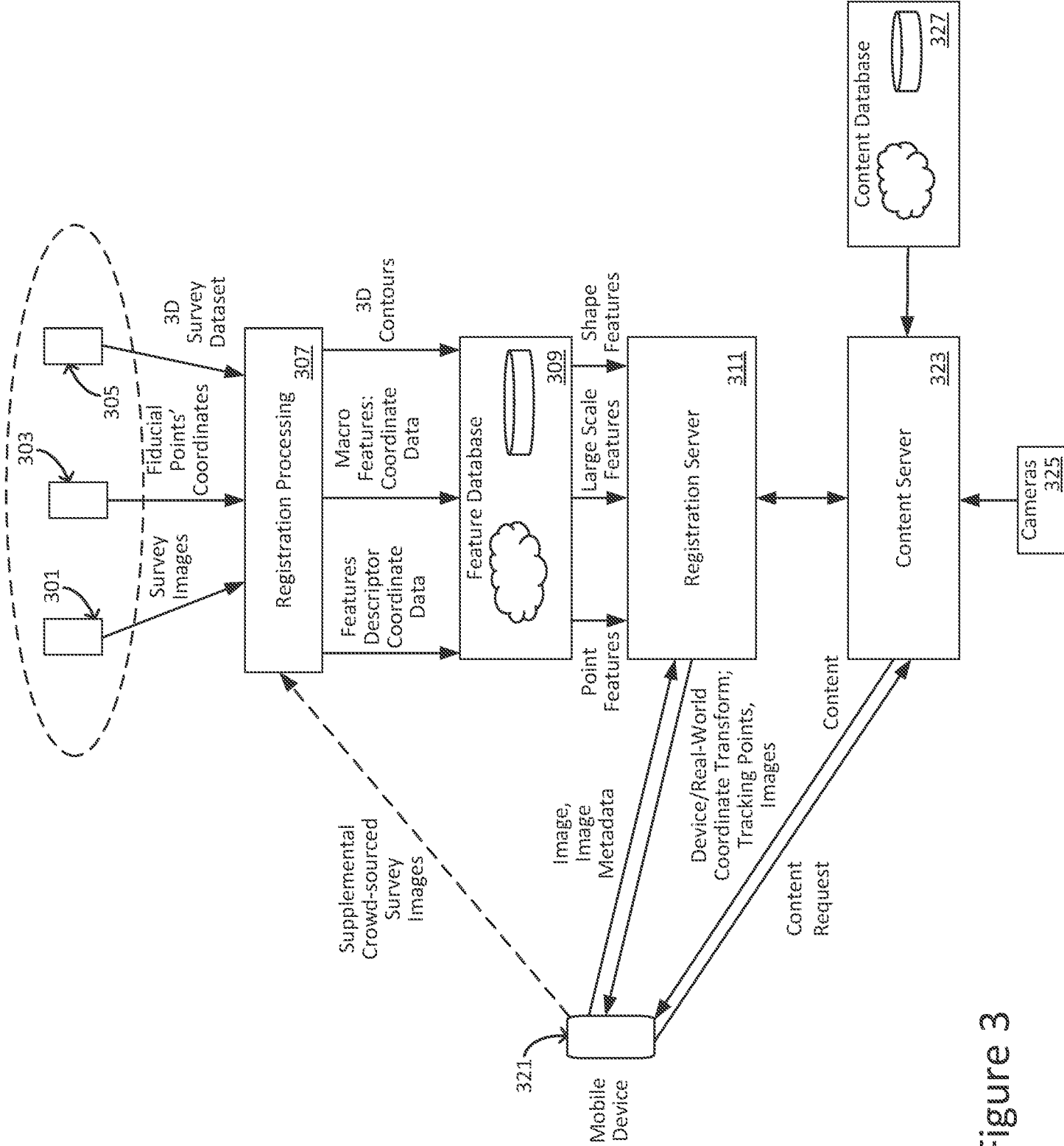


Figure 3

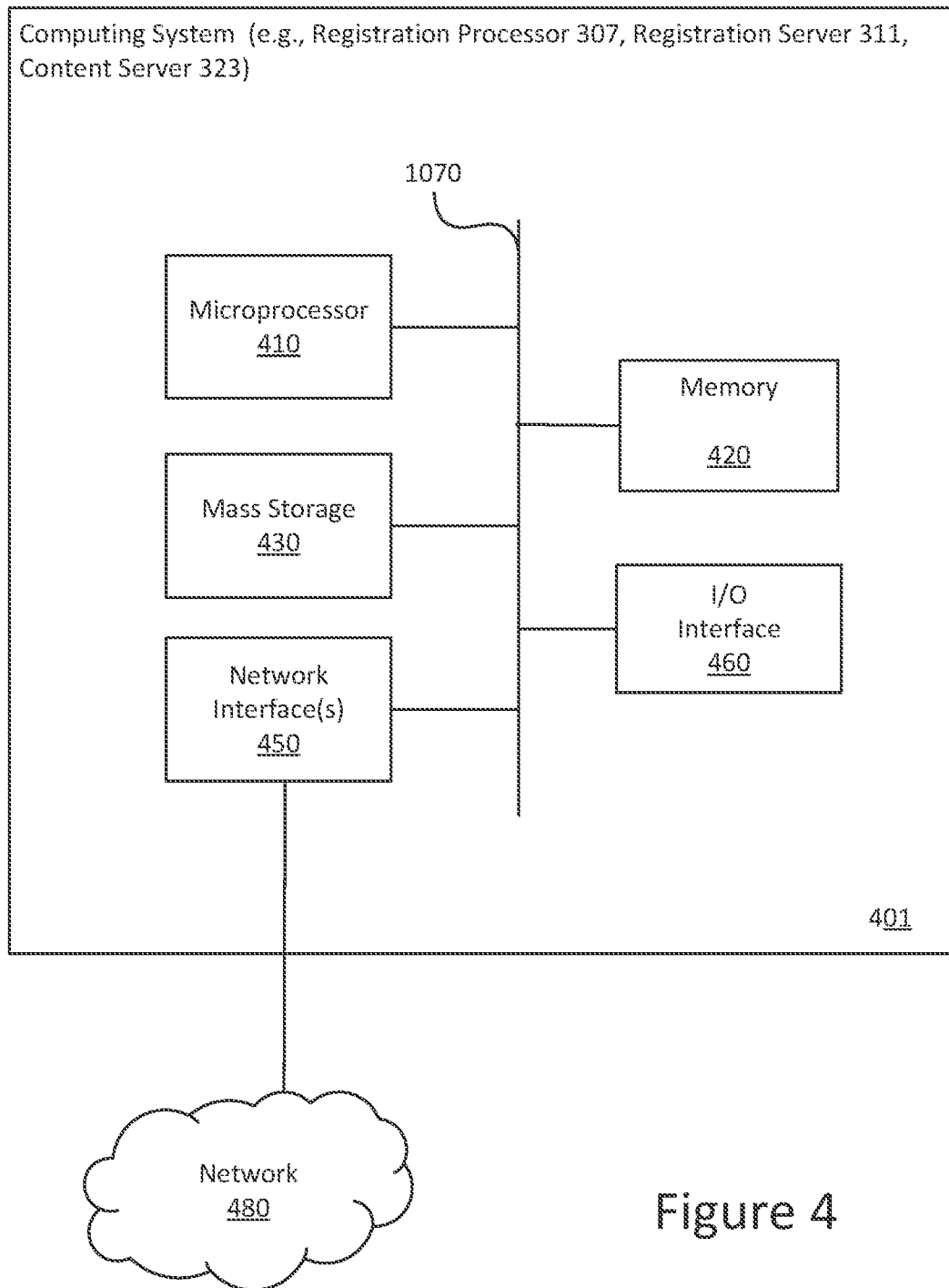


Figure 4

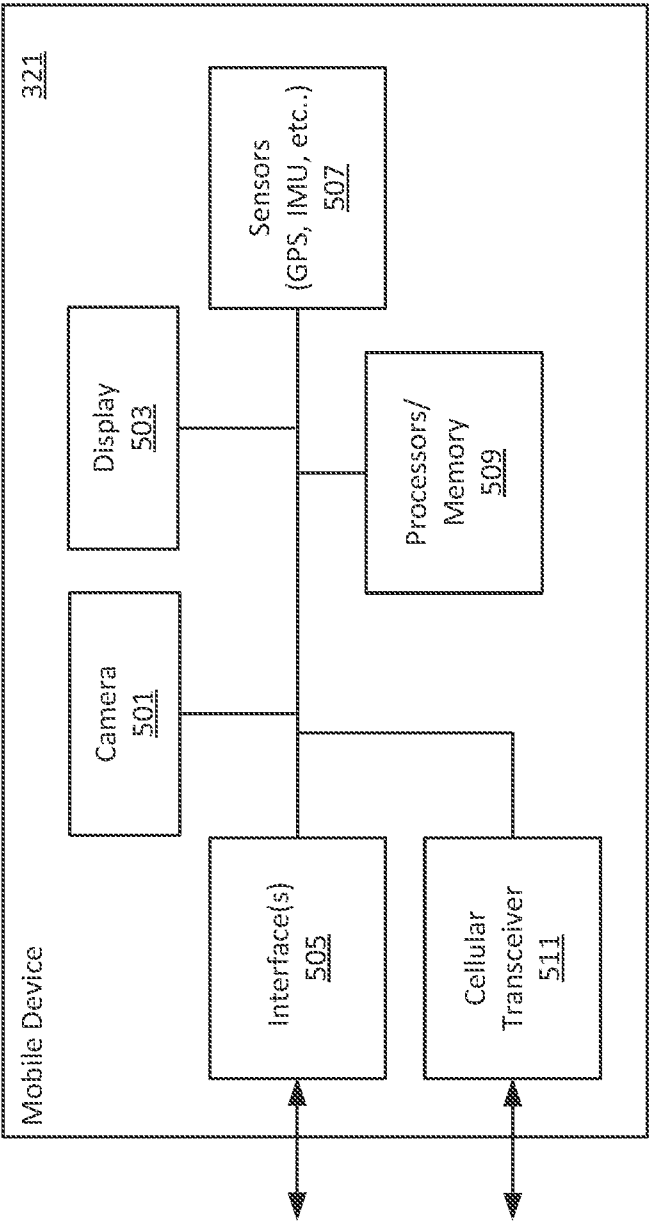


Figure 5

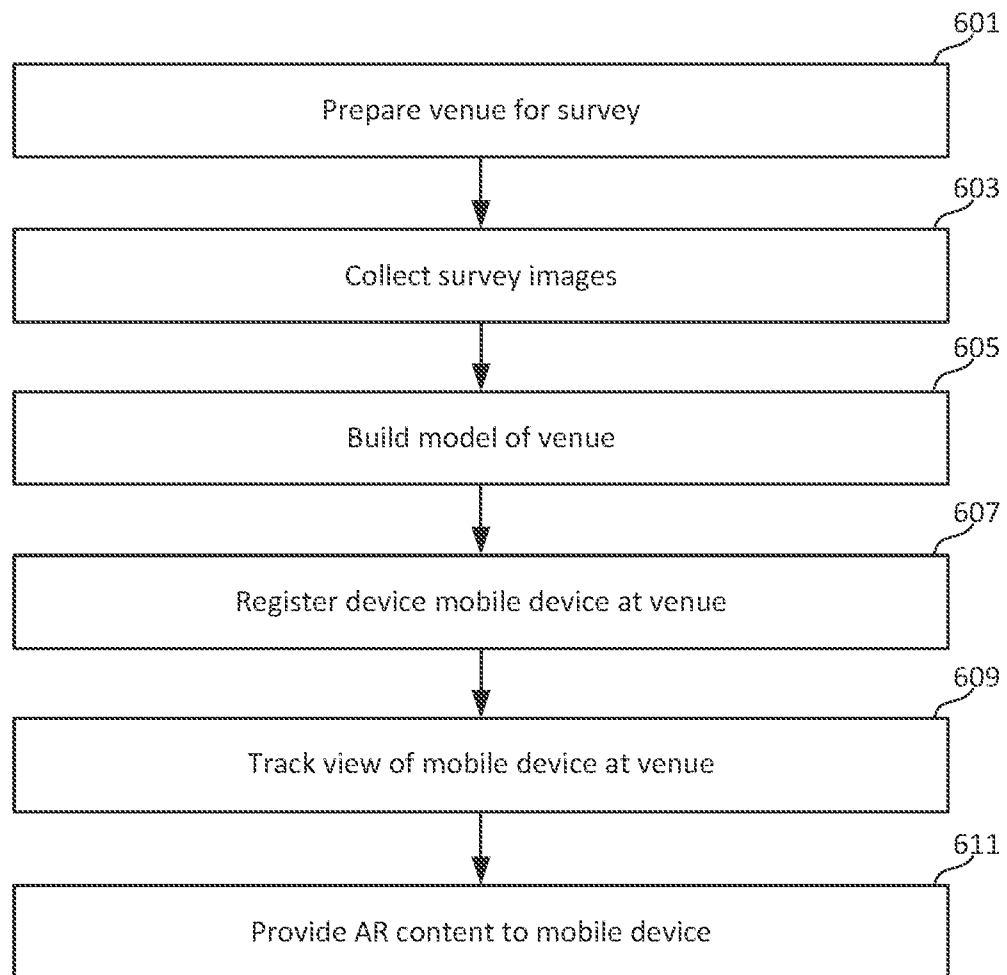


Figure 6

700

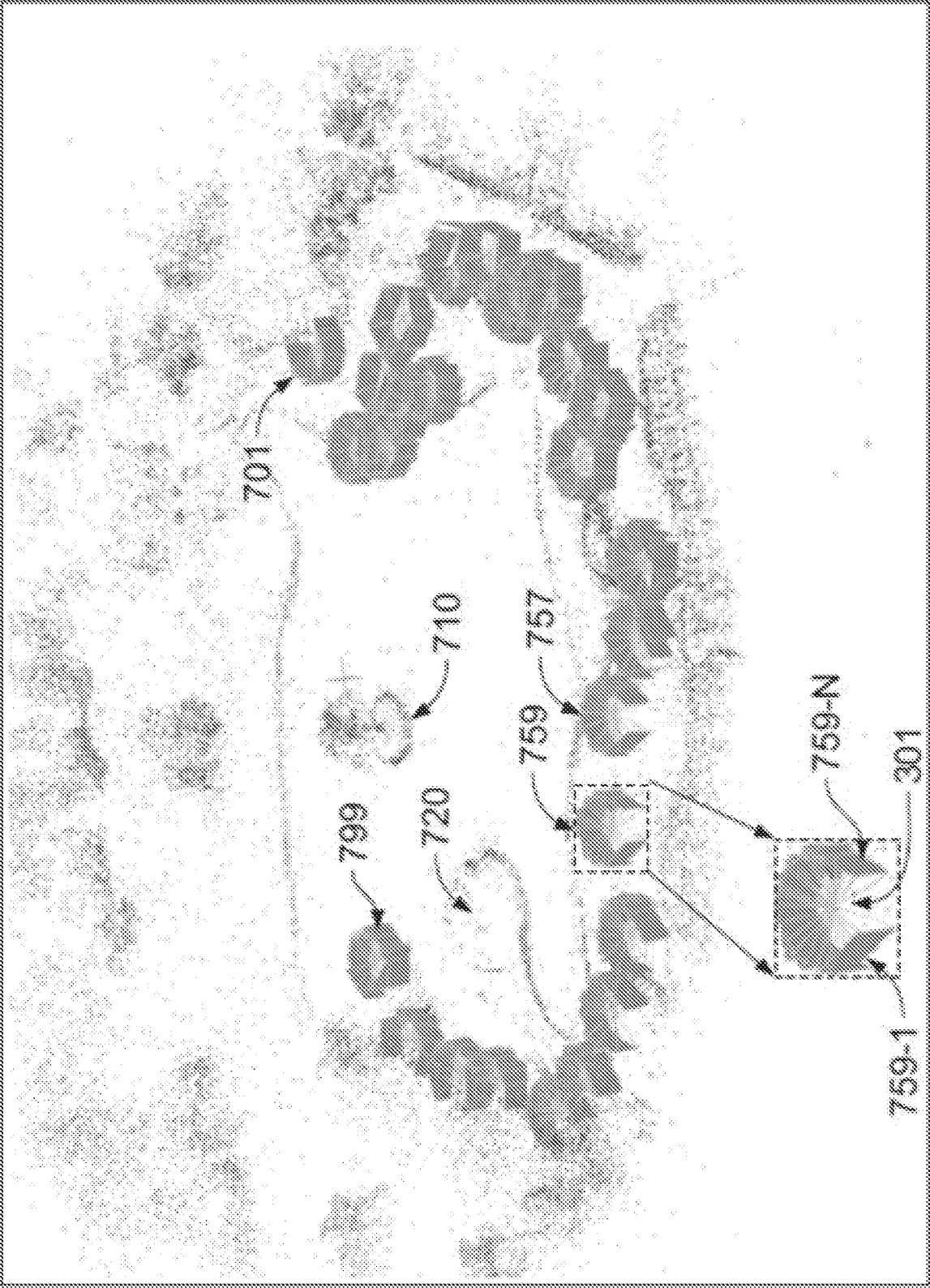


Figure 7A

700



Figure 8

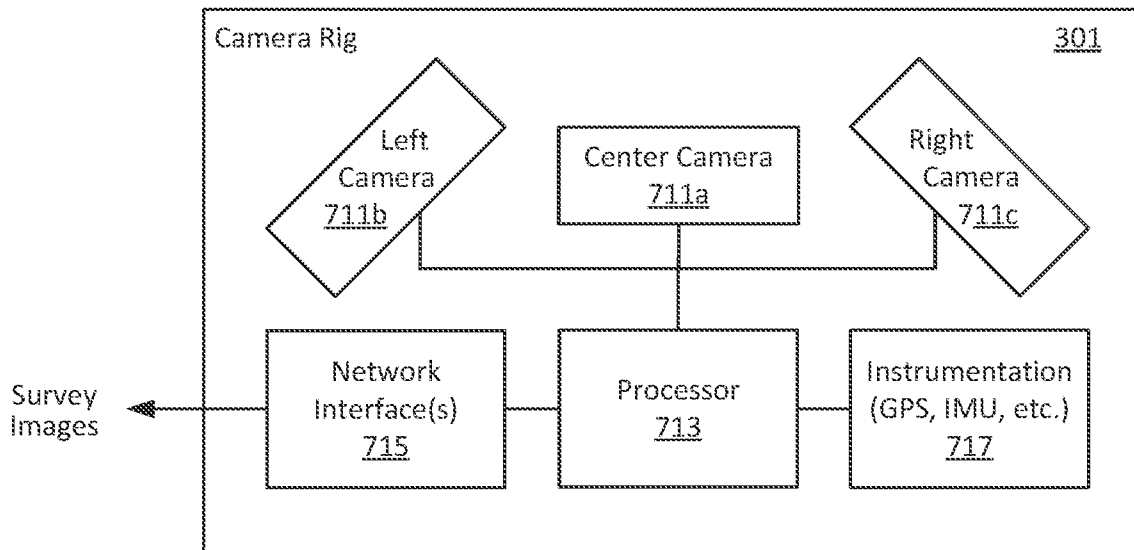


Figure 7B

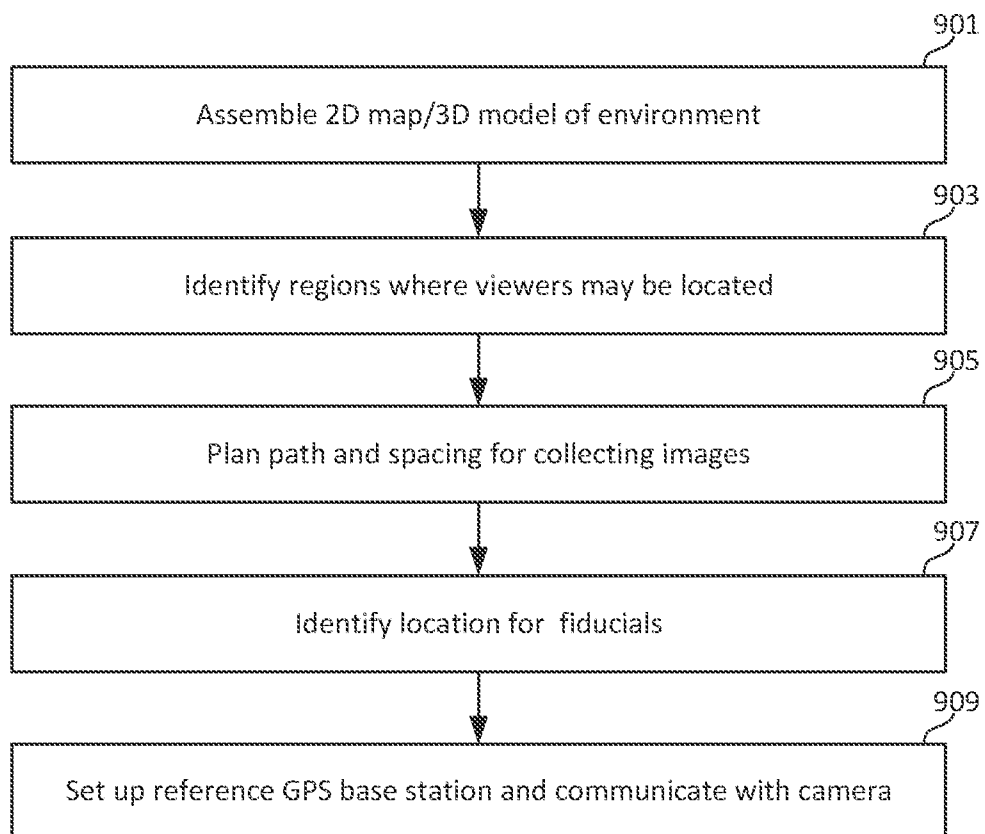


Figure 9

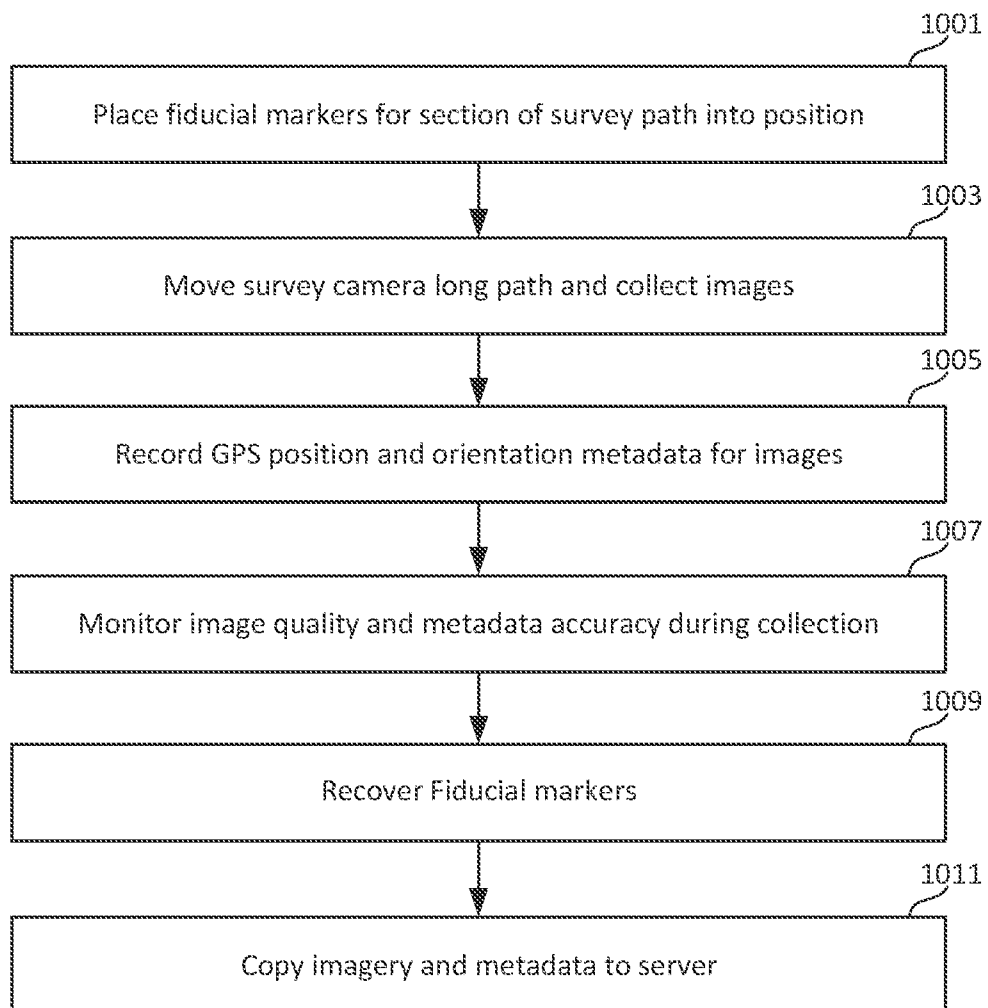


Figure 10

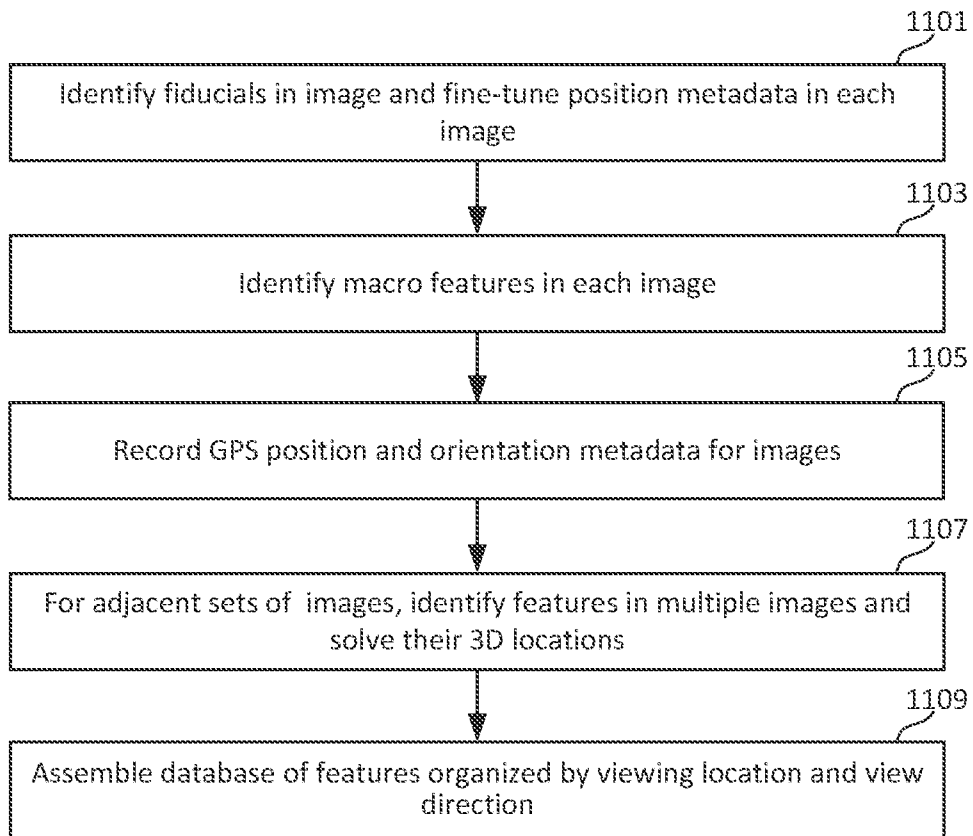


Figure 11

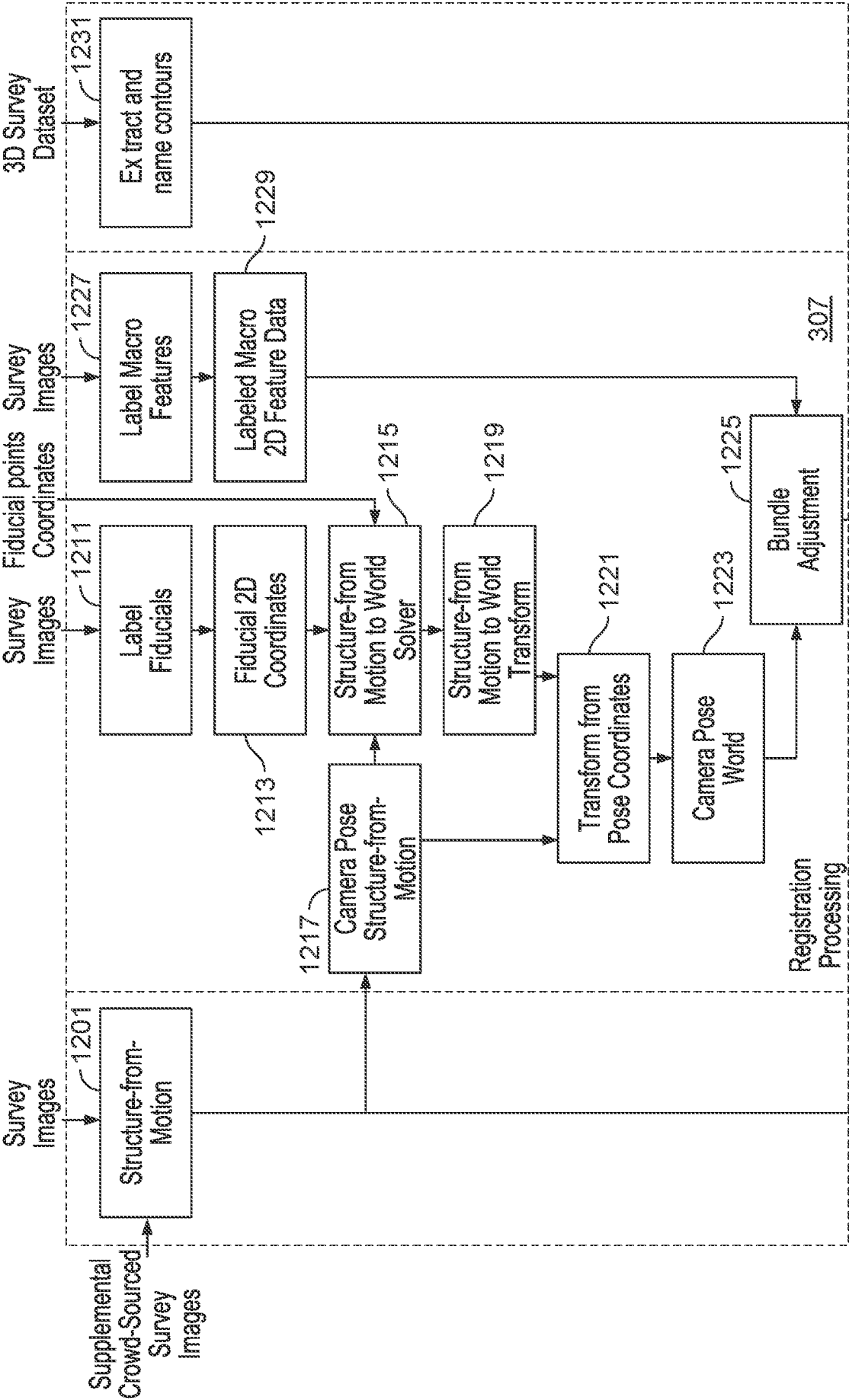


Figure 12

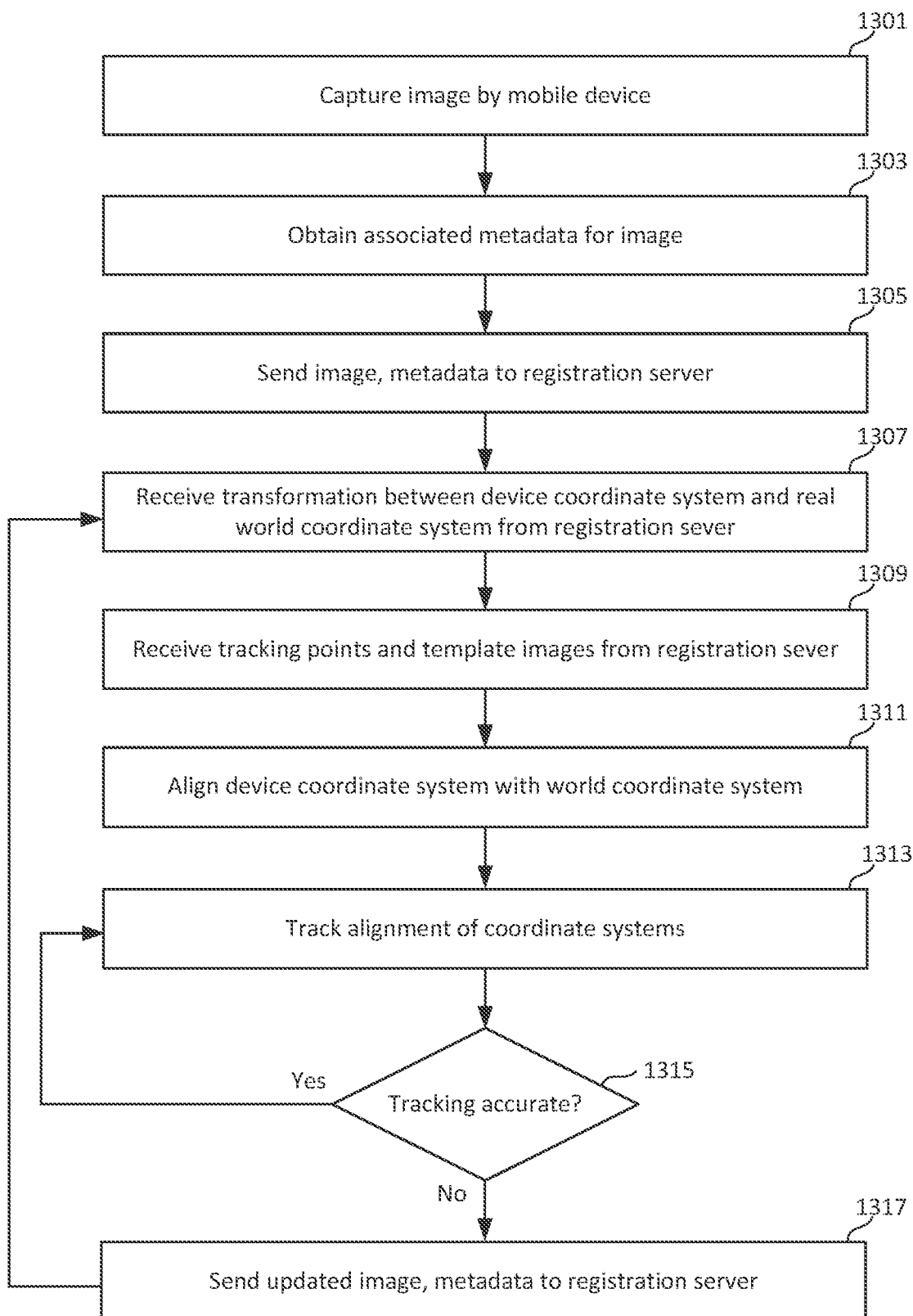


Figure 13A

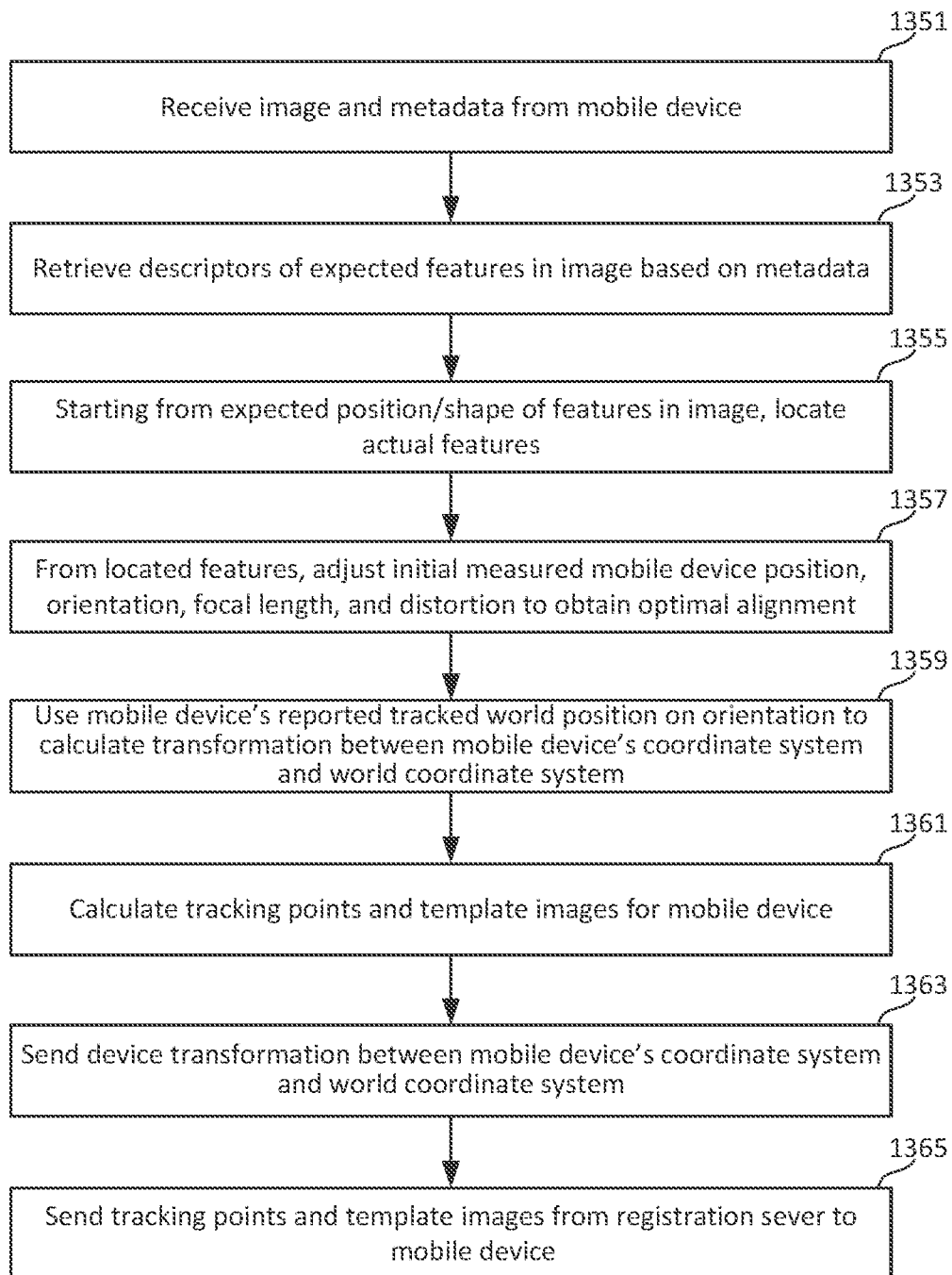


Figure 13B

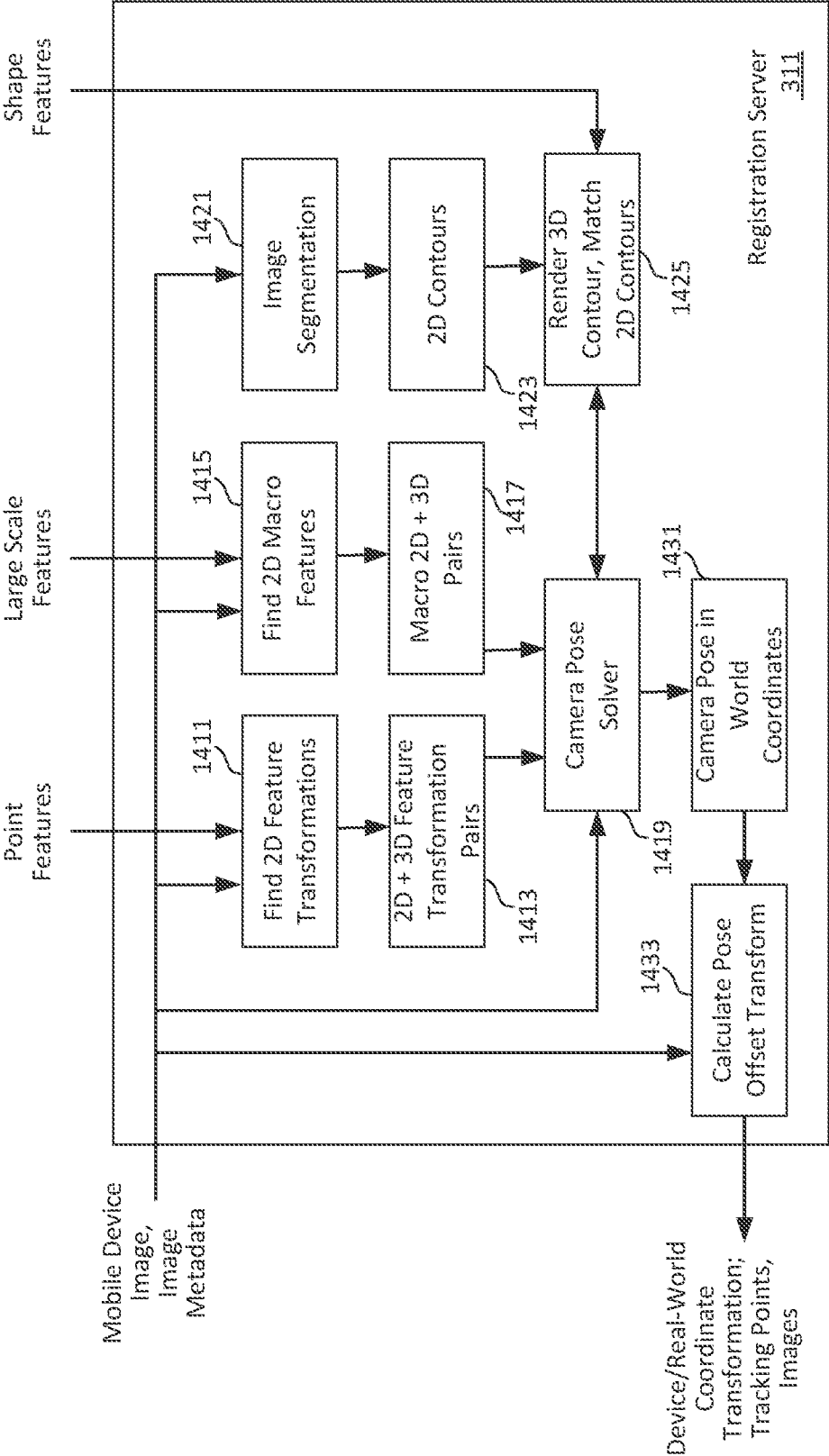


Figure 14A

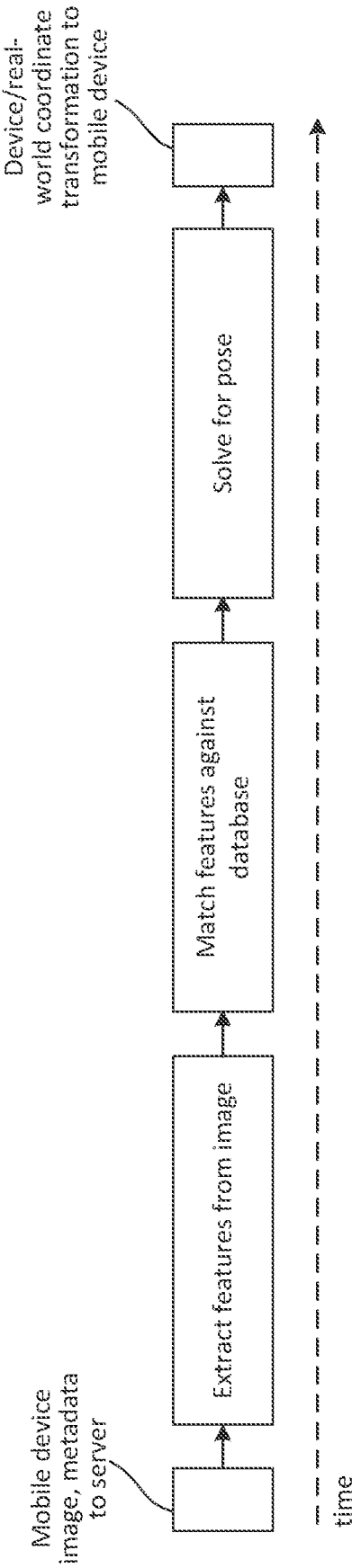


Figure 14B

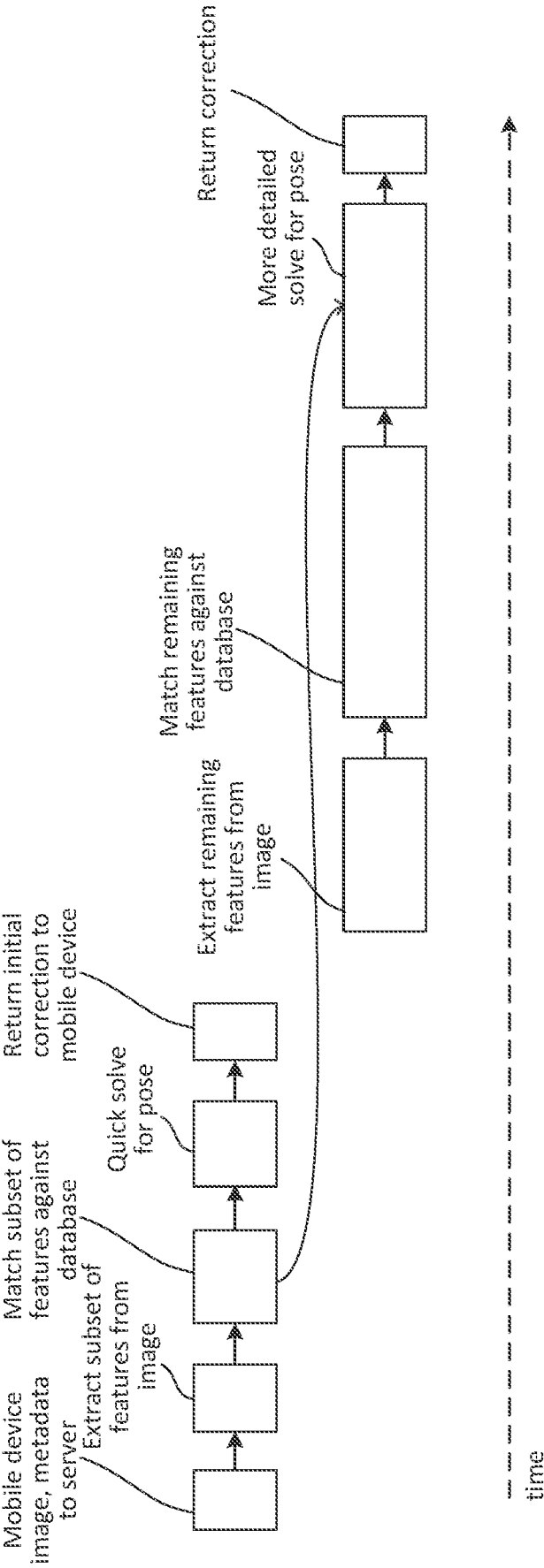


Figure 14C

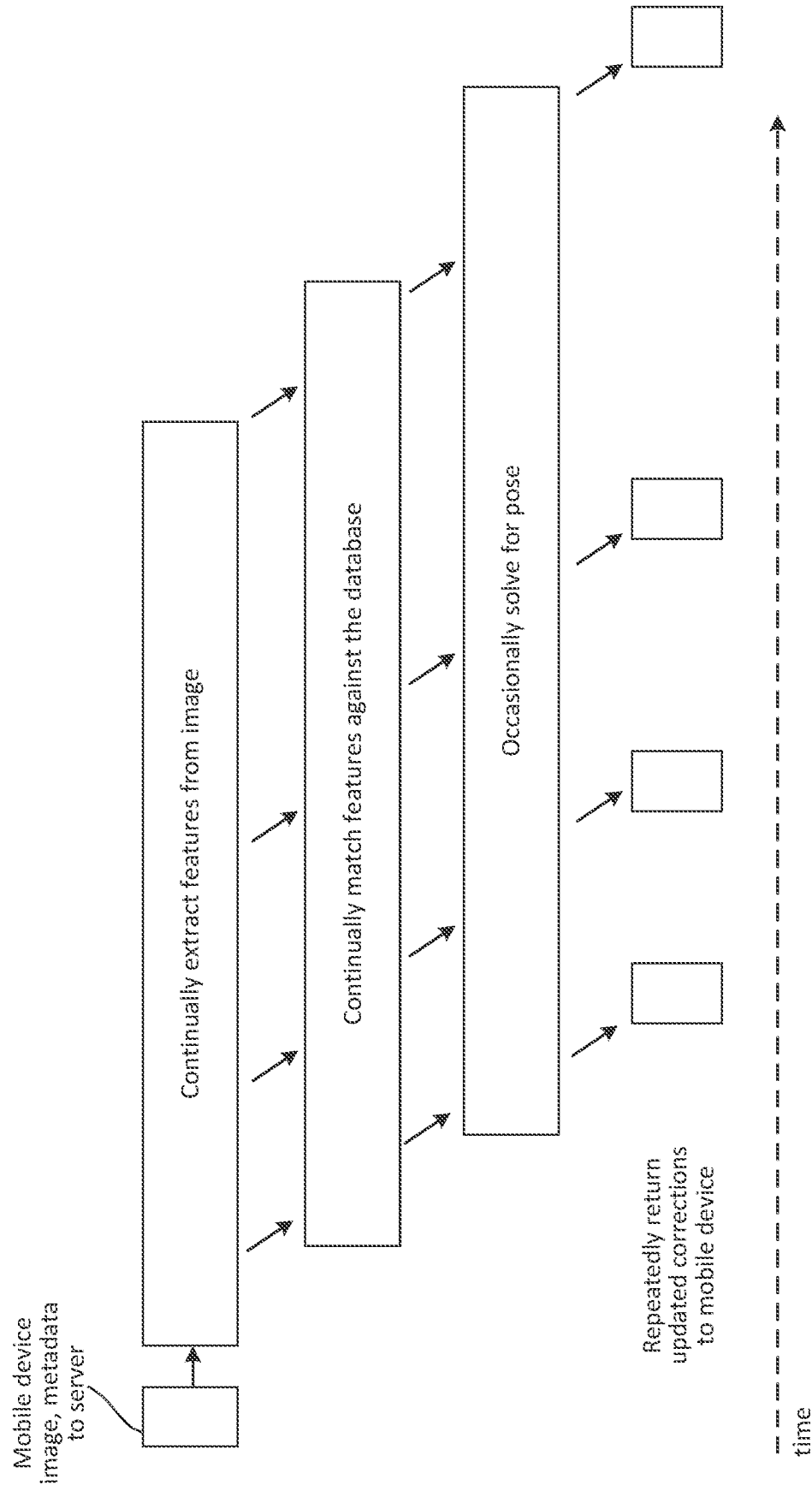


Figure 14D

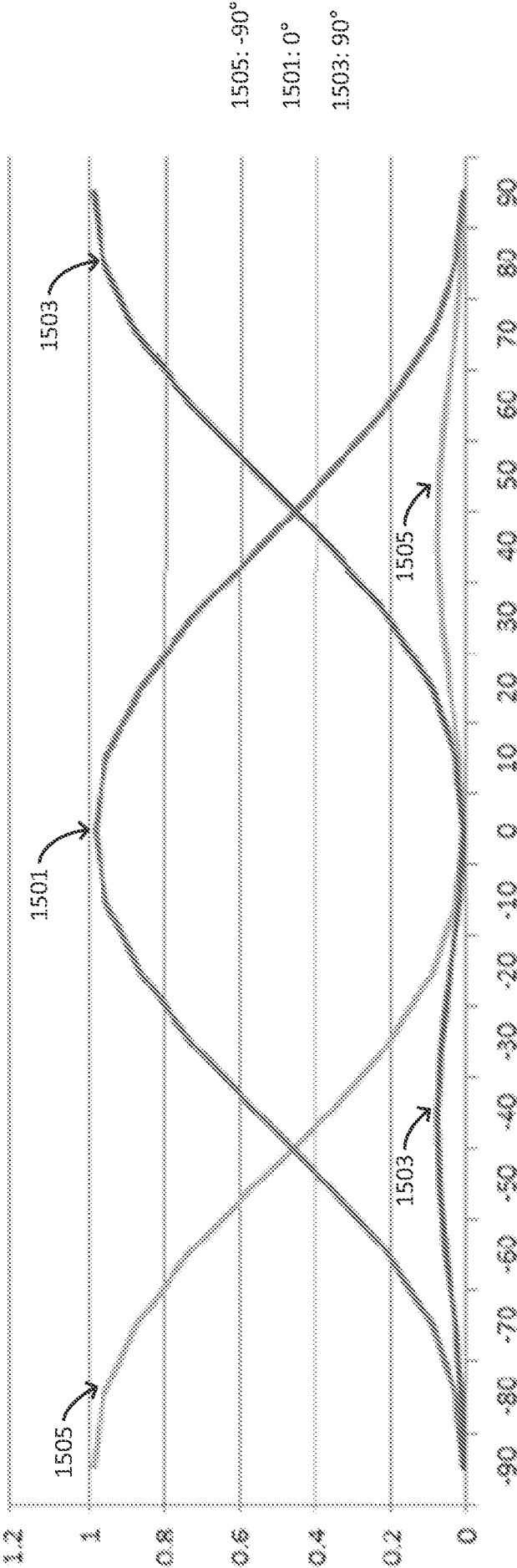


Figure 15

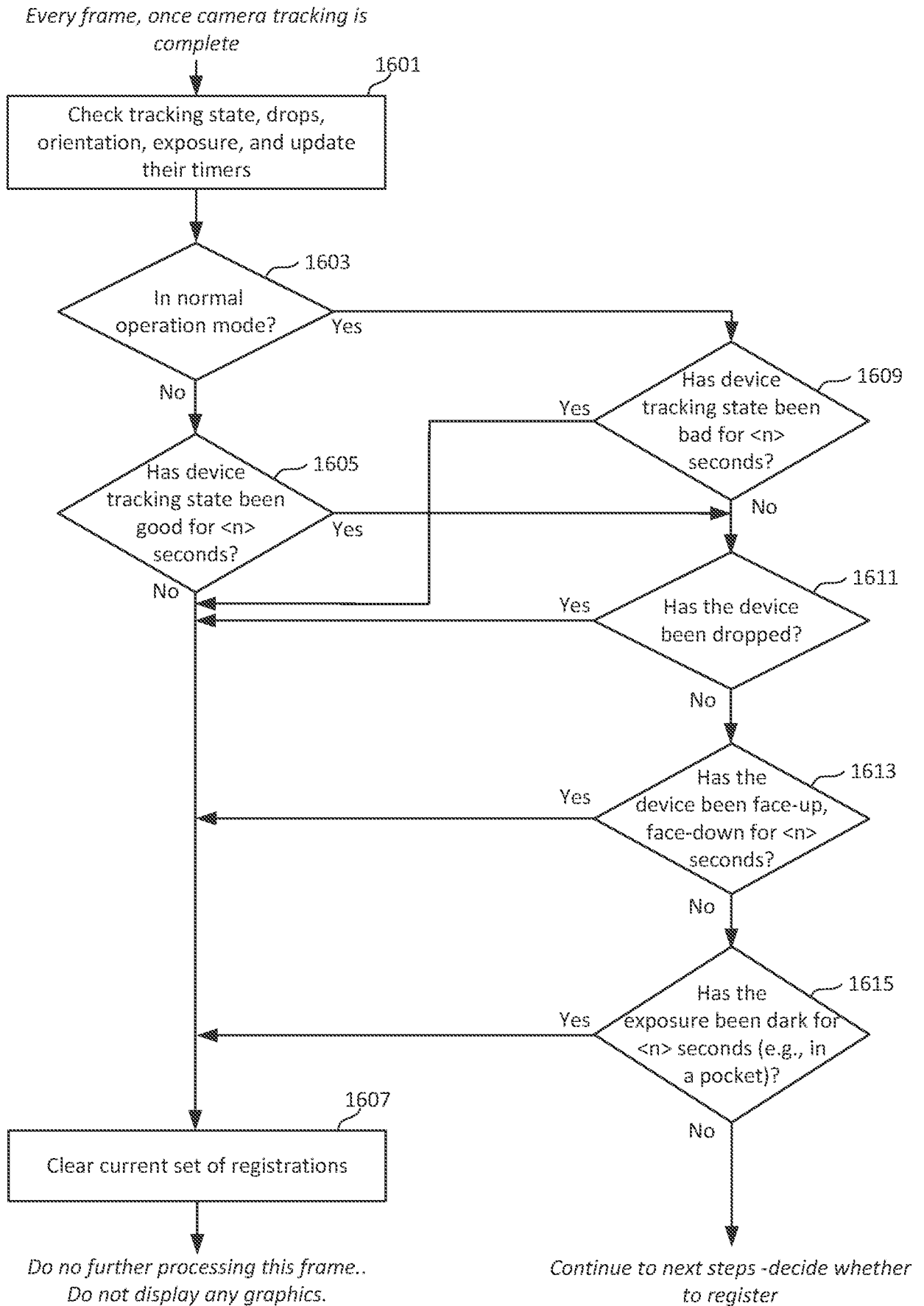


Figure 16

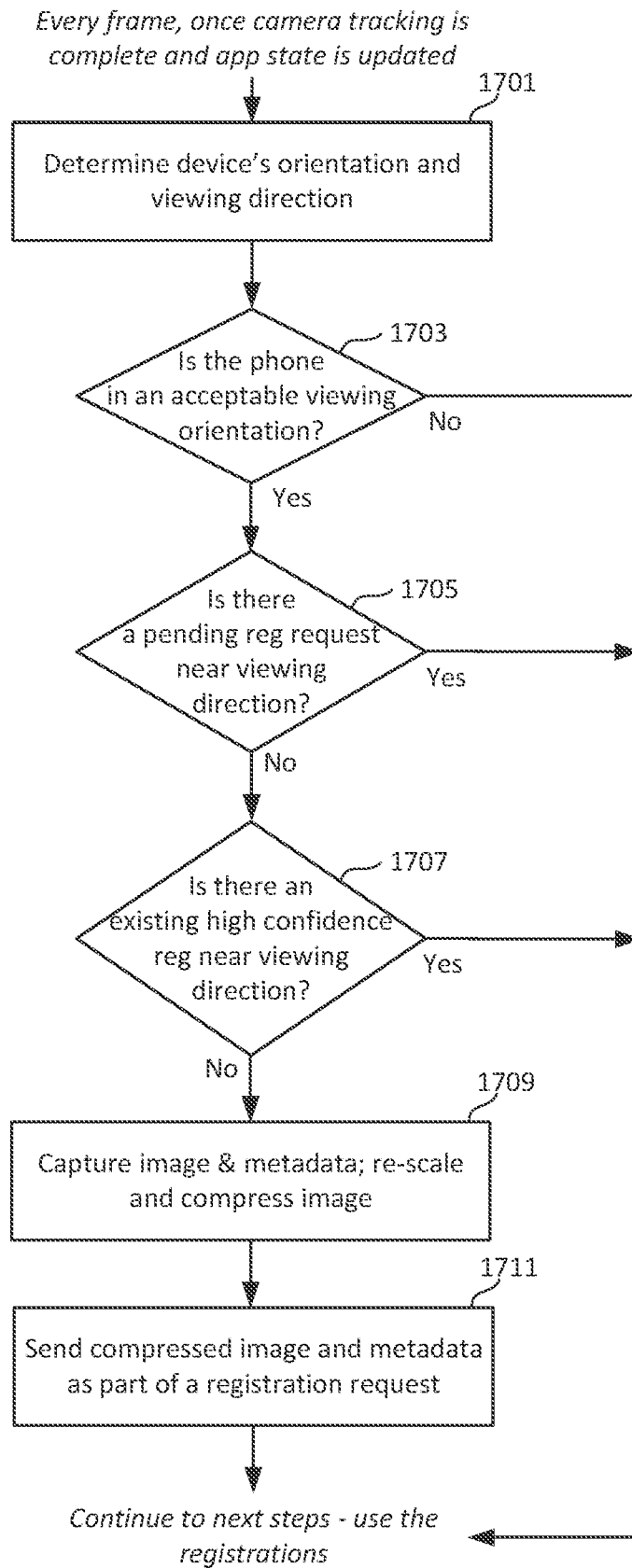


Figure 17

Every frame, once the check for requesting a new registration is complete

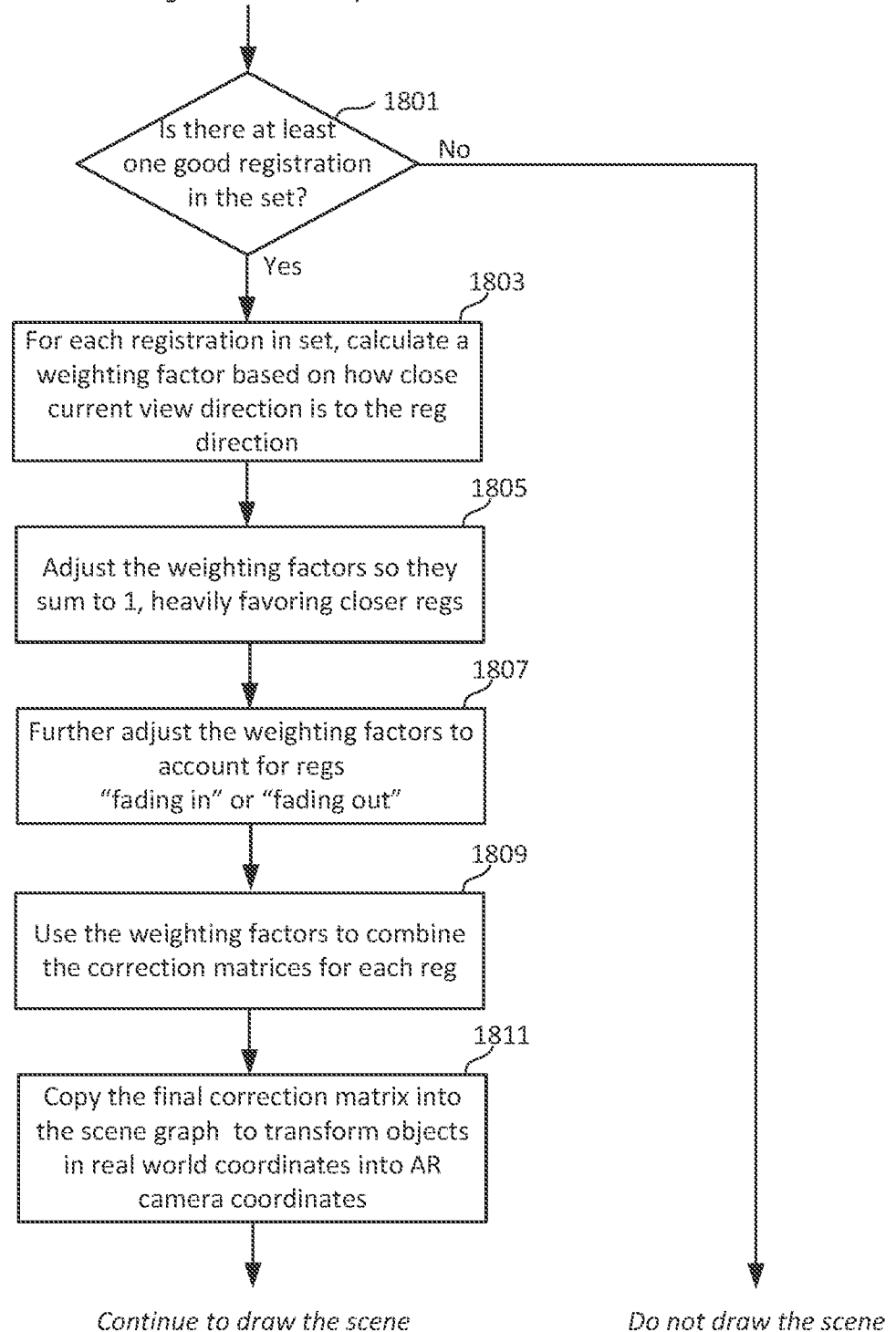
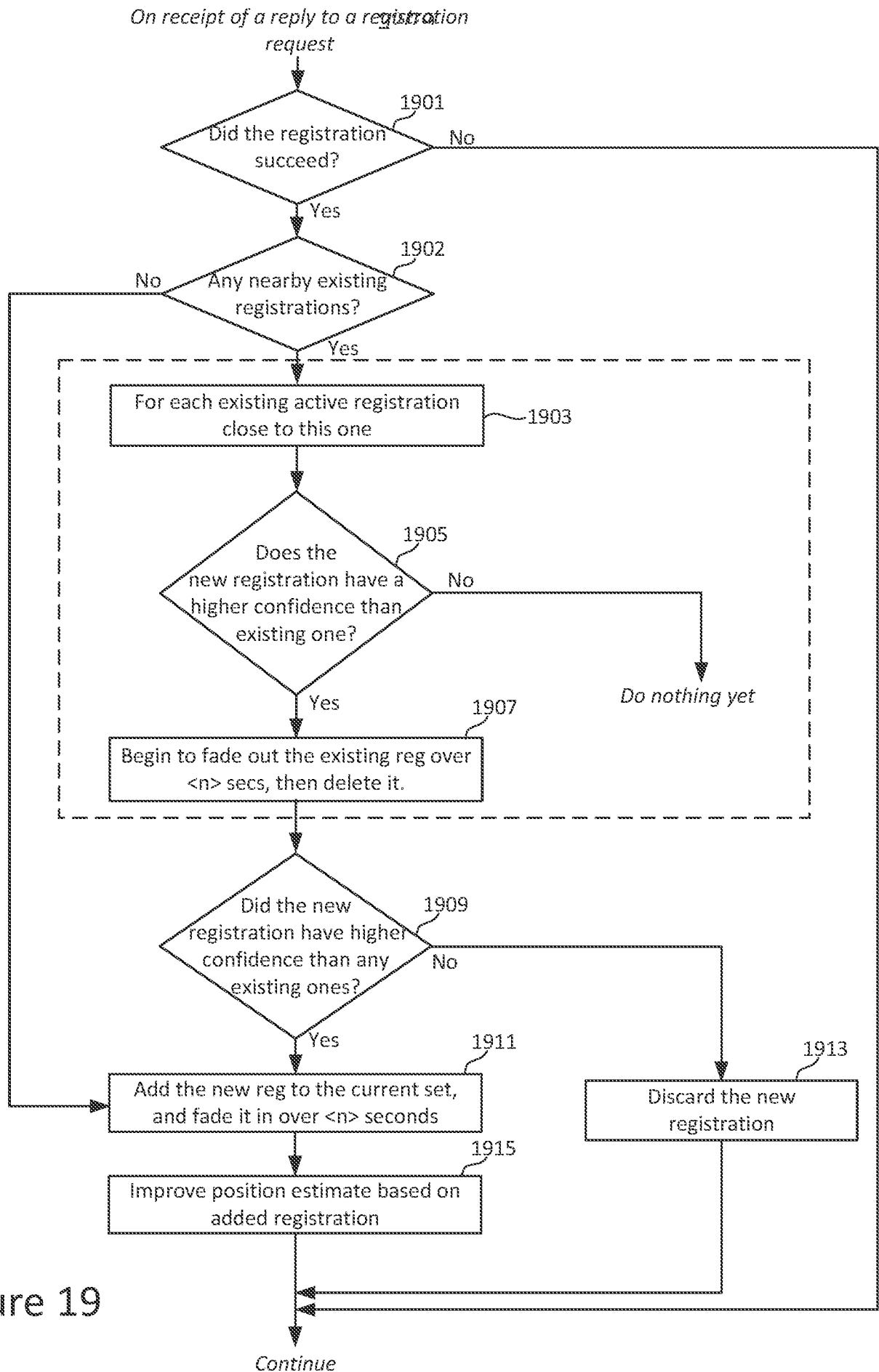


Figure 18



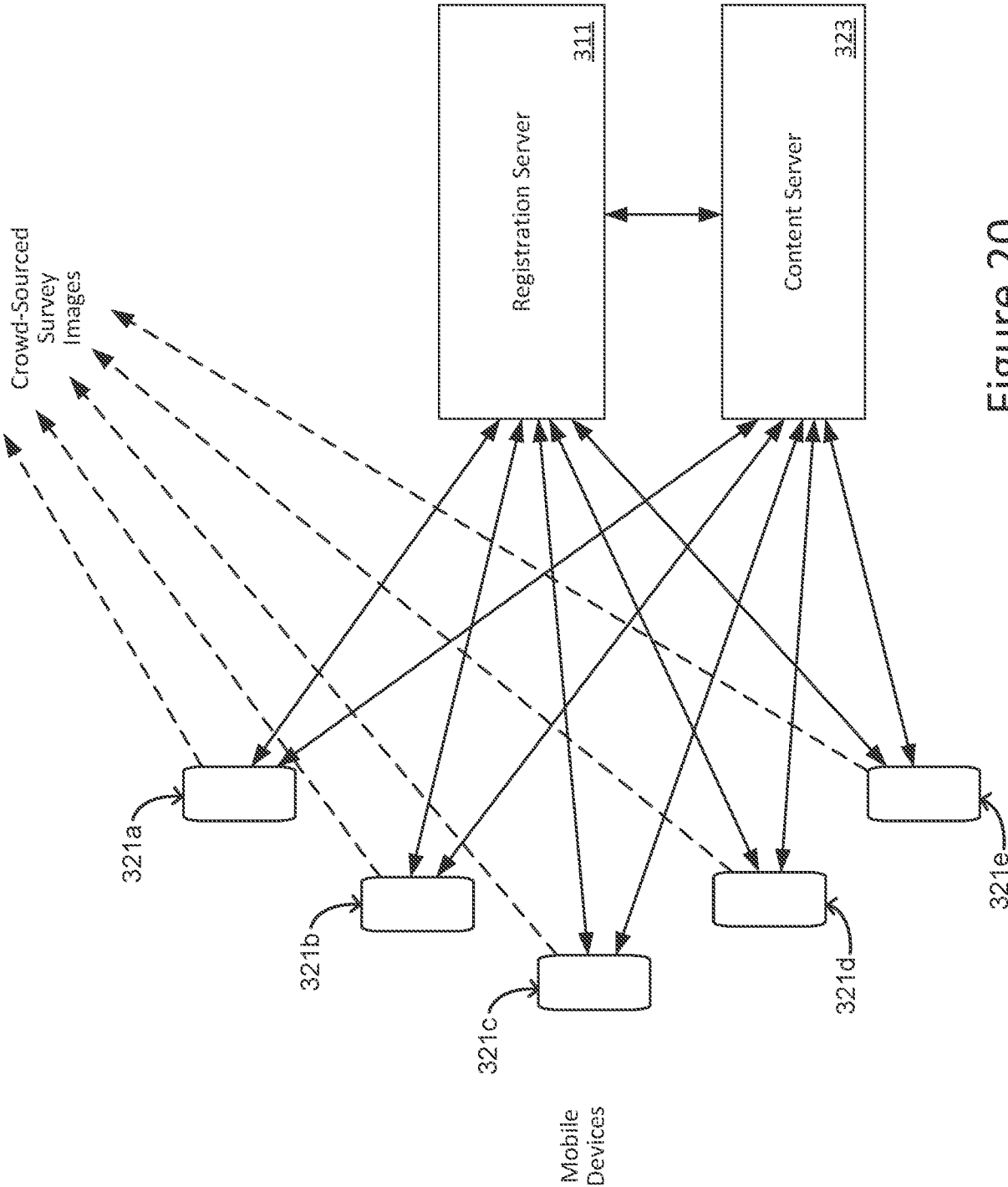


Figure 20

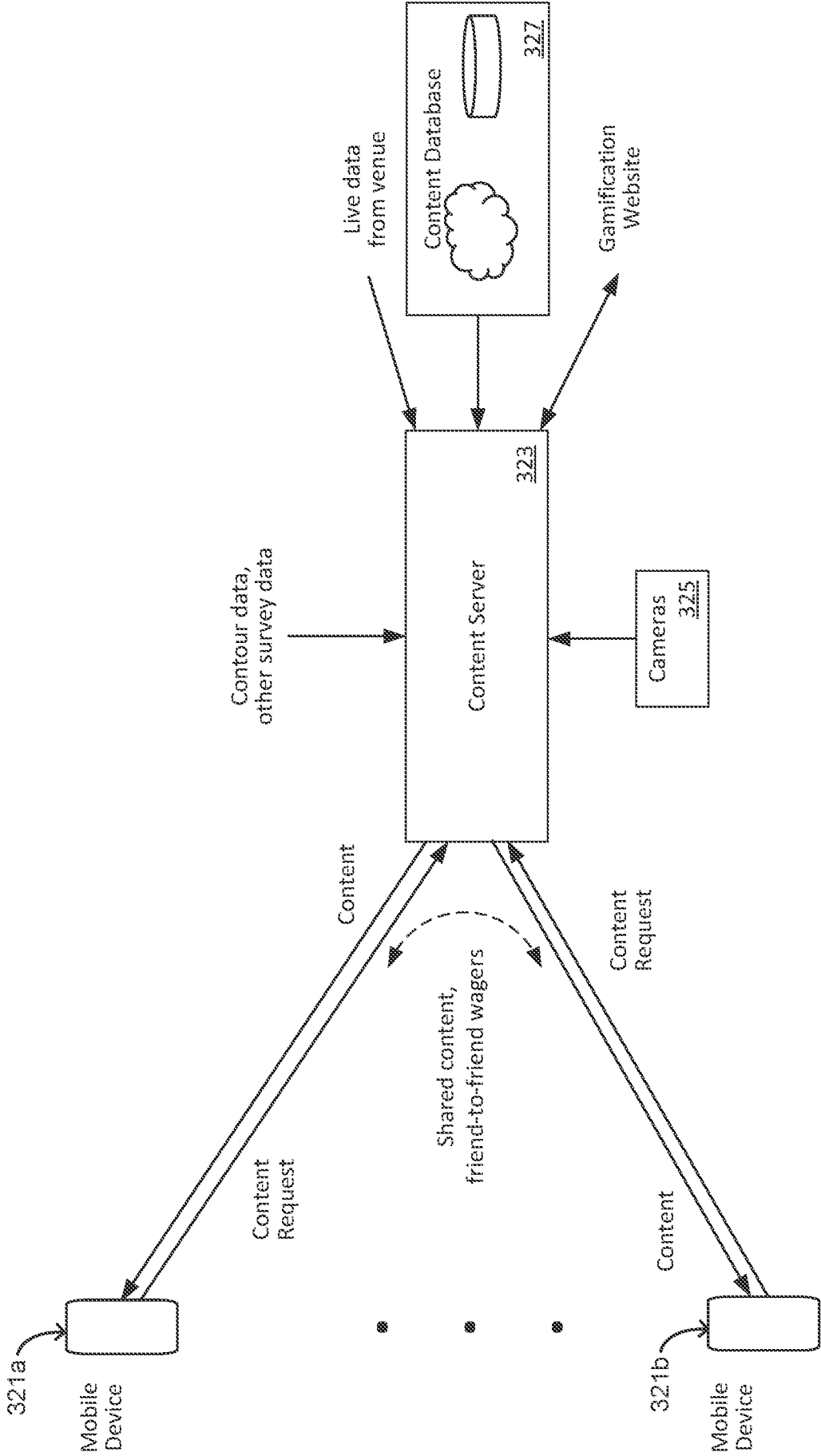


Figure 21

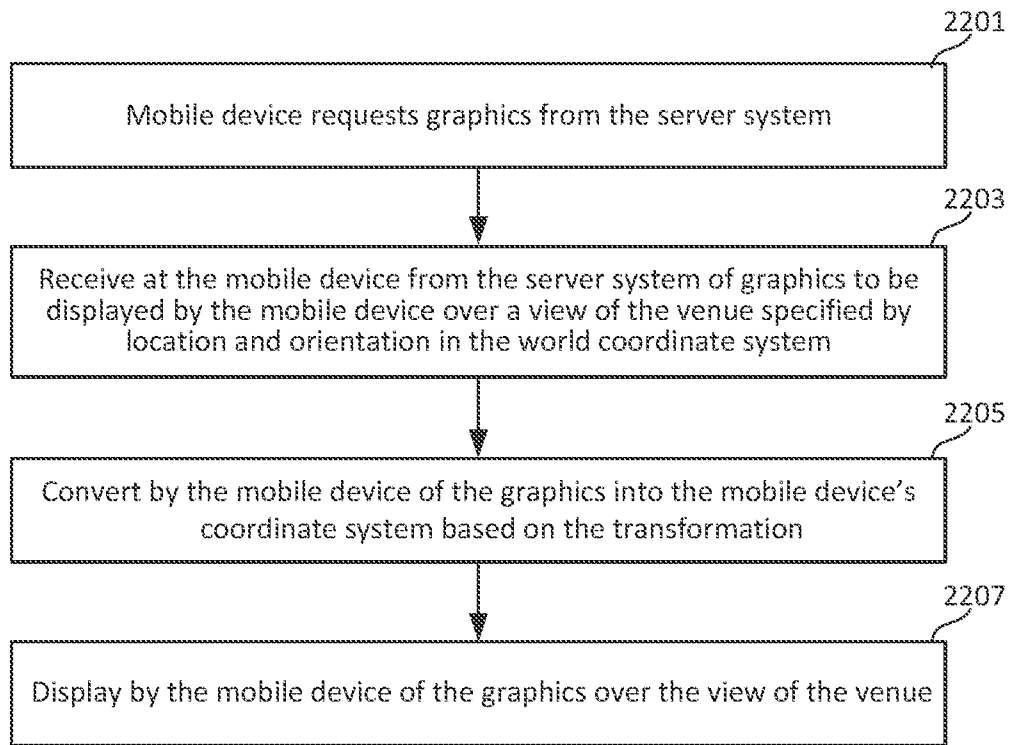
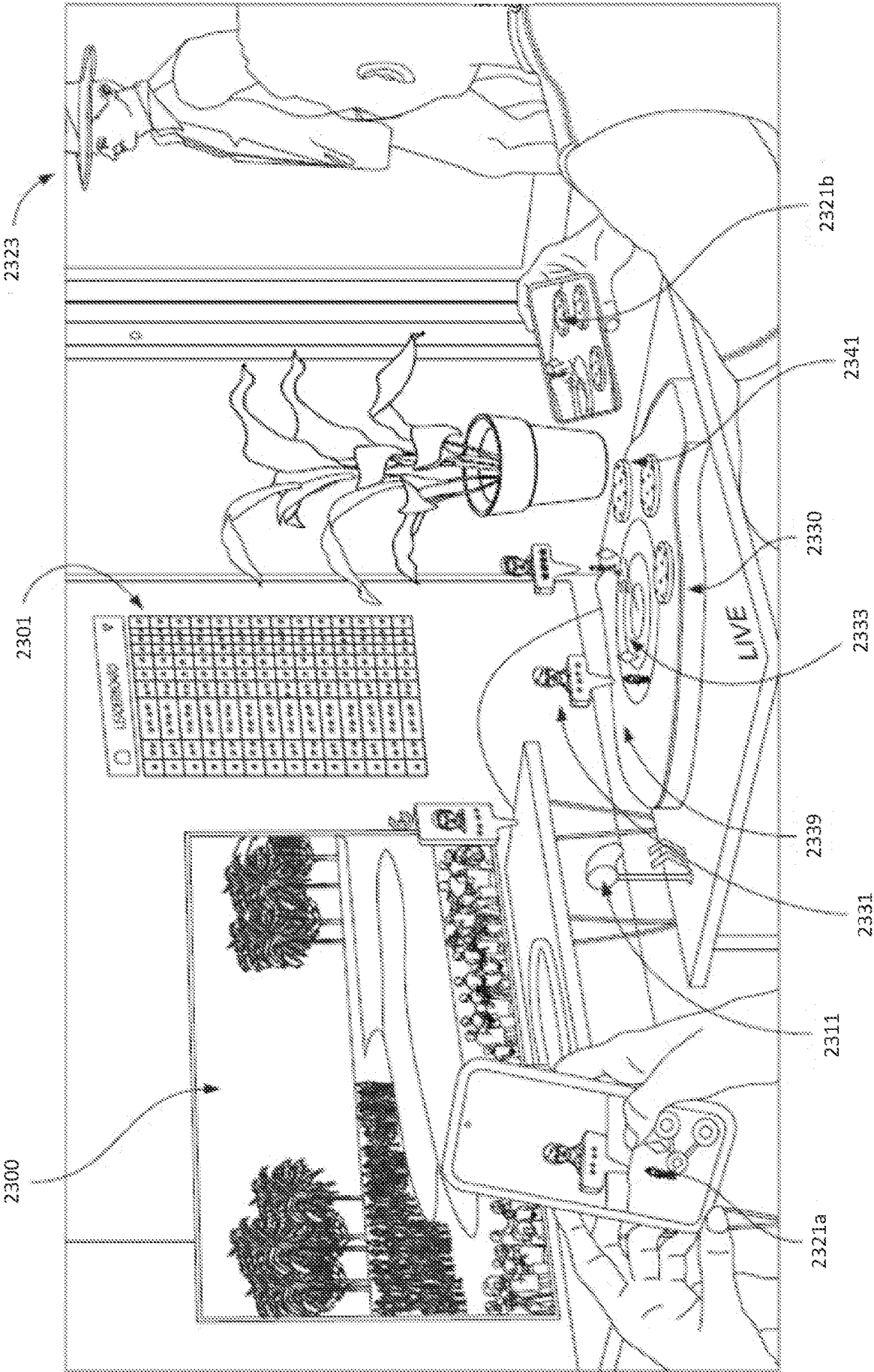


Figure 22



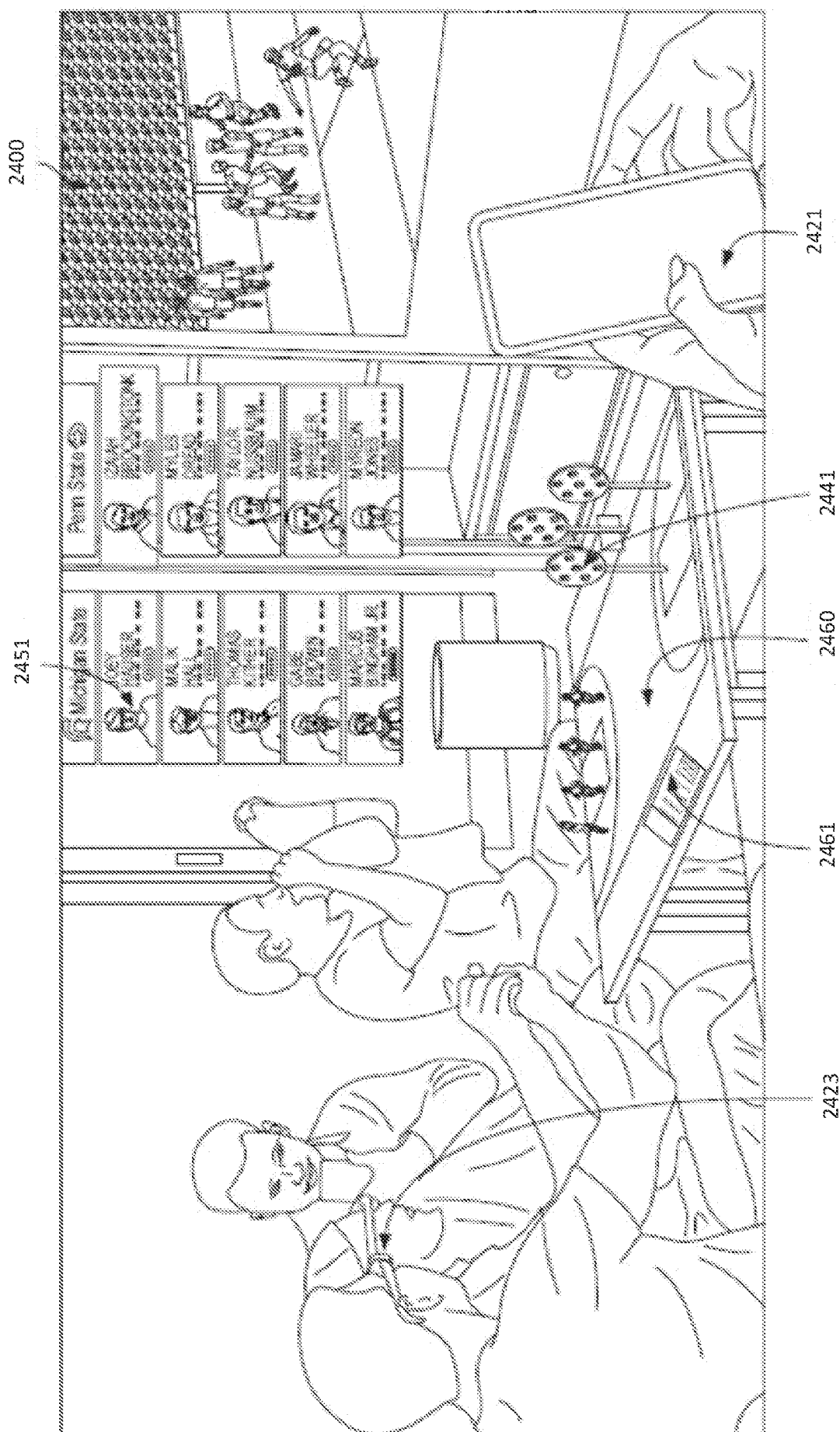


Figure 24

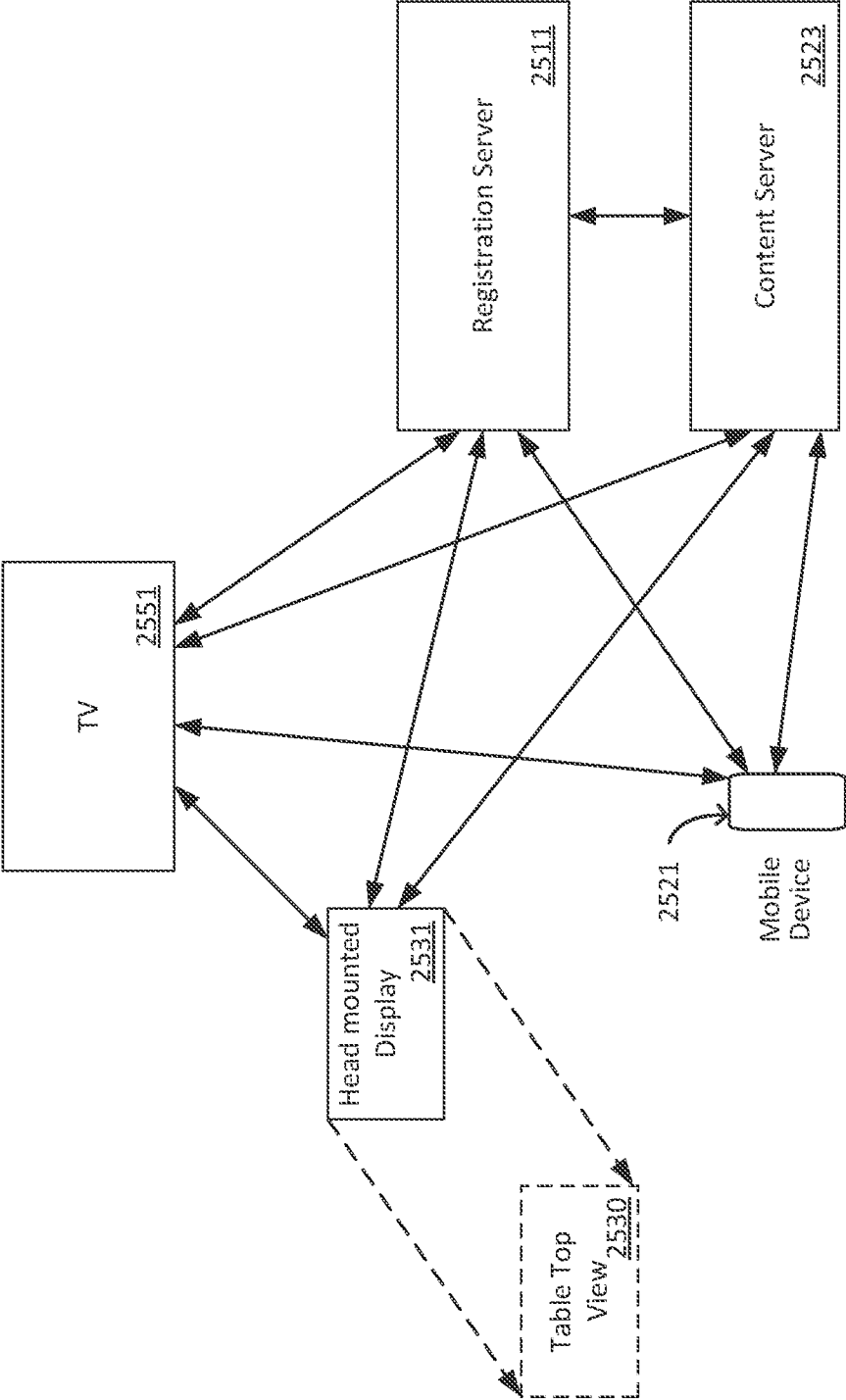


Figure 25

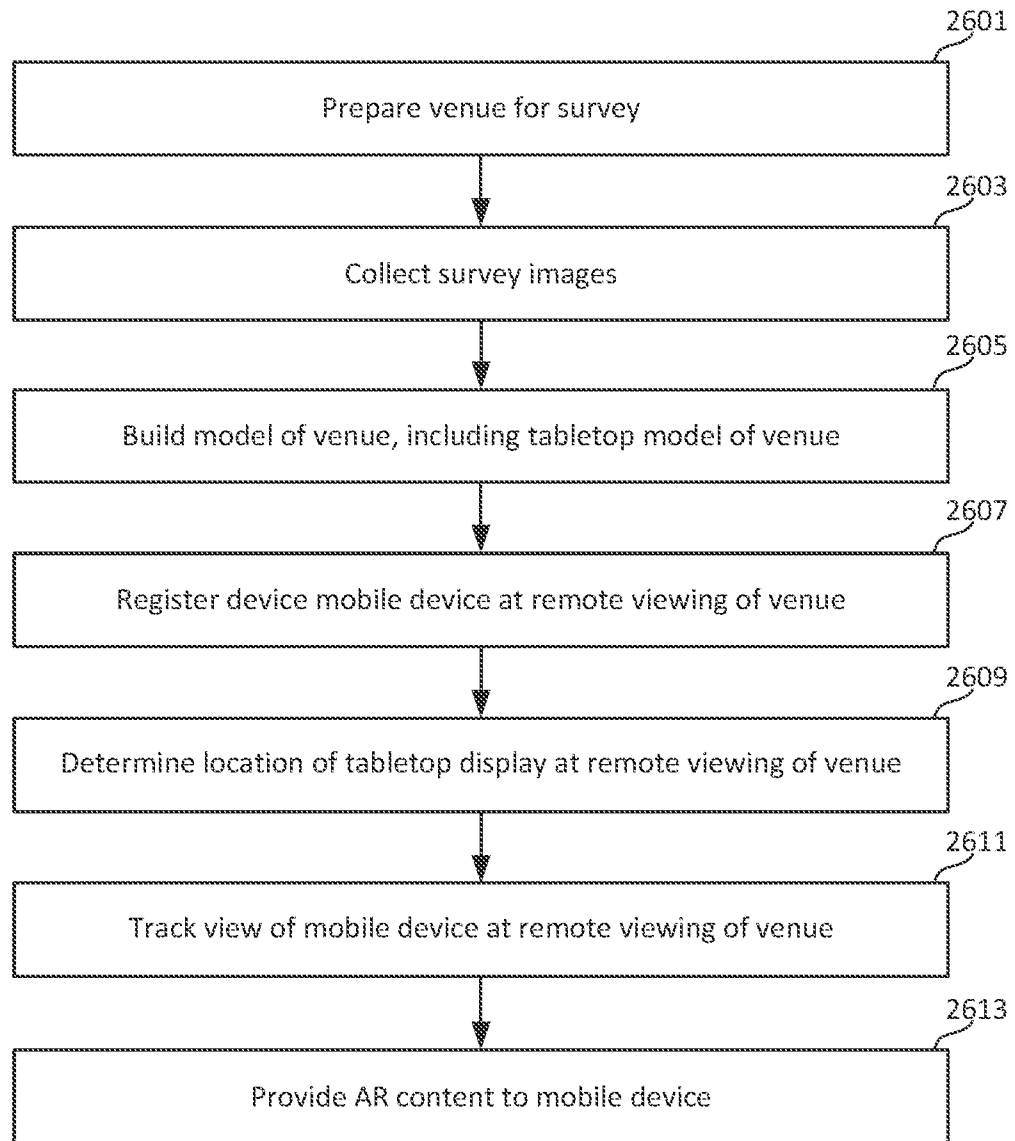


Figure 26

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/084805

| A. CLASSIFICATION OF SUBJECT MATTER INV. H04N21/218 G06F3/0346 G06T19/00 G06V20/20 H04N21/414 H04N21/431 H04N21/81 ADD. According to International Patent Classification (IPC) or to both national classification and IPC | | |
|---|---|---|
| B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) H04N G06F G06V G06T Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| X | US 10 832 417 B1 (TZUR YOCHAY [IL]) 10 November 2020 (2020-11-10) | 1, 3-7, 11-13, 15, 18, 19 |
| Y | figures 1-4 column 2, line 8 - column 6, line 4 column 8, line 50 - line 60 ----- | 2, 8-10, 14, 16, 17, 20 |
| Y | US 2022/028095 A1 (BLEYER MICHAEL [US] ET AL) 27 January 2022 (2022-01-27) | 2, 9, 10, 14, 20 |
| A | figures 1, 23, 32 paragraph [0011] paragraph [0063] - paragraph [0080] paragraph [0124] - paragraph [0154] paragraph [0185] - paragraph [0205] ----- -/- | 3 |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 12 March 2024 | | Date of mailing of the international search report 21/03/2024 |
| 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 Döttling, Martin |

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2023/084805

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|--|-----------------------|
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| Y | US 2022/295040 A1 (JAYARAM SANKAR [US] ET AL) 15 September 2022 (2022-09-15) figures 1-3, 12, 13 paragraph [0036] - paragraph [0066] paragraph [0100] - paragraph [0106] ----- | 1-20 |
| Y | YOON JONG-HYUN ET AL: "Increasing Camera Pose Estimation Accuracy Using Multiple Markers", 29 November 2006 (2006-11-29), SAT 2015 18TH INTERNATIONAL CONFERENCE, AUSTIN, TX, USA, SEPTEMBER 24-27, 2015; [LECTURE NOTES IN COMPUTER SCIENCE; LECT.NOTES COMPUTER], SPRINGER, BERLIN, HEIDELBERG, PAGE(S) 239 - 248, XP047402137, ISBN: 978-3-540-74549-5 abstract page 239 - page 245 ----- | 1-20 |

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2023/084805

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---------------------|----------------------------|-----------------------------|
| US 10832417 | B1 | 10-11-2020 | NONE |
| US 2022028095 | A1 | 27-01-2022 | EP 4186029 A1 31-05-2023 |
| | | | US 2022028095 A1 27-01-2022 |
| | | | WO 2022019977 A1 27-01-2022 |
| US 2022295040 | A1 | 15-09-2022 | EP 4305598 A1 17-01-2024 |
| | | | US 2022295040 A1 15-09-2022 |
| | | | WO 2022192166 A1 15-09-2022 |