



(51) International Patent Classification:
G09G 5/00 (2006.01)

(21) International Application Number:
PCT/US2018/043164

(22) International Filing Date:
20 July 2018 (20.07.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
15/658,310 24 July 2017 (24.07.2017) US
15/658,280 24 July 2017 (24.07.2017) US

(71) Applicant: **VISOM TECHNOLOGY, INC.** [US/US];
13939 209th Avenue NE, Woodinville, WA 98077 (US).

(72) Inventors: **KELLOGG, Ryan**; 13939 209th Avenue
NE, Woodinville, WA 98077 (US). **PHILLIPS, Charles**;
13939 209th Avenue NE, Woodinville, WA 98077 (US).
BUCHANAN, Sean; 13939 209th Avenue NE, Wood-
inville, WA 98077 (US).

(74) Agent: **SAAD, Hassan et al.**; Fenwick & West LLP, 801
California Street, Mountain View, California 94041 (US).

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

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

(54) Title: **MARKERLESS AUGMENTED REALITY (AR) SYSTEM**

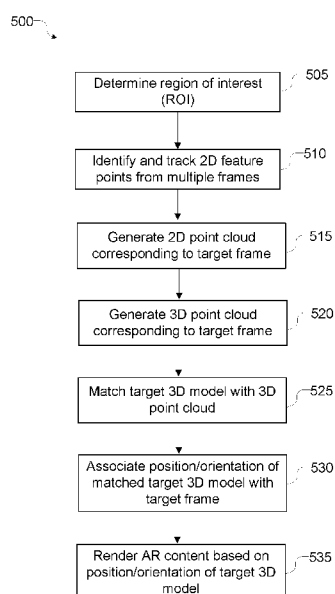


FIG. 5

(57) Abstract: A markerless augmented reality (AR) can track 2D feature points among video frames, generate 2D point clouds and 3D point clouds based thereon, and match a 3D model against 3D point cloud to obtain proper positional information of the model with respect to a frame. The AR system can use the 3D model with the obtained positional information to render and project AR content to a user's view. Additionally, the AR system can maintain associations between frames and 3D model positional information for search and retrieval.

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

MARKERLESS AUGMENTED REALITY (AR) SYSTEM

BACKGROUND

[0001] This application claims priority to U.S. Patent Application No. 15/658,280 filed July 24, 2017, entitled "MARKERLESS AUGMENTED REALITY (AR) SYSTEM," and U.S. Patent Application No. 15/658,310 filed July 24, 2017, entitled "MARKERLESS AUGMENTED REALITY (AR) SYSTEM," both of which applications are incorporated by reference herein in their entireties.

[0002] Augmented reality (AR) generally refers to a live, direct or indirect, view of a physical, real-world environment whose elements are augmented by computer-generated sensory input such as video, graphics, sound, or GPS data. AR devices, such as AR Head-Mounted Display (HMD) devices, may include transparent display elements that enable a user to see virtual content superimposed (e.g., overlaid or projected) over the user's view of the real world. Virtual content that appears to be superimposed over the user's real-world view is commonly referred to as AR content, which may include "holographic" objects as well as other sensory information or data.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] One or more embodiments of the present disclosure are illustrated by way of example and not limitation in the figures of the accompanying drawings, in which like references indicate similar elements.

[0004] Figure 1 illustrates an example of an AR system in accordance with some embodiments of the presently disclosed technology.

[0005] Figure 2 illustrates an example of feature points extraction and tracking from multiple images in accordance with some embodiments of the presently disclosed technology.

[0006] Figure 3 illustrates an example of epipolar geometry based feature point tracking in accordance with some embodiments of the presently disclosed technology.

[0007] Figure 4 illustrates an example of 3D point cloud generated in accordance with some embodiments of the presently disclosed technology.

[0008] Figure 5 is a flowchart illustrating a process for generating a 3D point cloud and matching with a target 3D model in accordance with some embodiments of the presently disclosed technology.

[0009] Figure 6 is a flowchart illustrating a process for pre-computing and matching of 3D point clouds based on likely positions and/or orientation of a target 3D model in accordance with some embodiments of the presently disclosed technology.

[0010] Figure 7 is a flowchart illustrating a process for identifying an existing association between a target 3D model and 2D feature points of a target frame in accordance with some embodiments of the presently disclosed technology.

[0011] Figure 8 is a block diagram illustrating an example of the architecture for a computer system (or computing device) that can be utilized to implement various portions of the presently disclosed technology.

DETAILED DESCRIPTION

[0012] In order to accurately and promptly superimpose AR content in various contexts (e.g., navigation, gaming, education, entertainment), 3D positional information (e.g., location and orientation in a world coordinate system) of a camera that acquires image or video (e.g., multiple frames of images) is typically required in an accurate and real-time manner. In this regard, marker-based AR relies on the presence of artificial markers in the user's view. These markers, however, may distract from a view of the subject of interest, contribute to certain unnatural feel of the AR experience, or otherwise adversely affect user experience. Also, marker-based AR can simply be inapplicable in many cases because the artificial markers cannot be added to certain real-world scenes.

[0013] The presently disclosed technology is directed to markerless AR systems and methods that enable efficient tracking of feature points among images (e.g., consecutive frames within a video) of natural and/or never-before-seen surroundings, 3D model matching based on the tracked feature points, and

AR content rendering using positional information of the matched 3D model. In contrast with typical AR systems or methods that optimize for the global accuracy of a map of surrounding environment or a pose of the camera, the presently disclosed technology focuses on the local accuracy of feature points relative to the camera in order to accurately align virtual and real objects for AR. Accordingly, the presently disclosed technology can be computationally more efficient for implementation in real-time with relatively limited computational resources (e.g., on a CPU of a mobile device such as an AR-HMD or smartphone).

[0014] Figures 1-8 are provided to illustrate representative embodiments of the presently disclosed technology. Unless provided for otherwise, the drawings are not intended to limit the scope of the claims in the present application.

[0015] Many embodiments of the technology described below may take the form of computer- or controller-executable instructions, including routines executed by a programmable computer or controller. The programmable computer or controller may or may not reside on a corresponding AR device. For example, the programmable computer or controller can be an onboard computer of the AR device, a separate but dedicated computer associated with the AR device, or part of a network or cloud based computing service. Those skilled in the relevant art will appreciate that the technology can be practiced on computer or controller systems other than those shown and described below. The technology can be embodied in a special-purpose computer or data processor that is specifically programmed, configured or constructed to perform one or more of the computer-executable instructions described below. Accordingly, the terms "computer" and "controller" as generally used herein refer to any data processor and can include Internet appliances and mobile devices (including palm-top computers, wearable computing devices, cellular or mobile phones, multi-processor systems, processor-based or programmable consumer electronics, network computers, mini computers and the like). Information handled by these computers and controllers can be presented at any suitable display medium, including an LCD (liquid crystal display) or AR-HMD's transparent display. Instructions for performing computer- or controller-executable tasks can be stored in or on any suitable computer-readable

medium, including hardware, firmware or a combination of hardware and firmware. Instructions can be contained in any suitable memory device, including, for example, a flash drive, USB (universal serial bus) device, and/or other suitable medium.

[0016] Figure 1 illustrates an example of an AR system 100 in accordance with some embodiments of the presently disclosed technology. The AR system 100 can include an AR device 102, a processing system 104, a model data service 106, and an association data service 108 that are communicatively connected with one another via connections 110.

[0017] The AR device 102 can be an AR-HMD, smartphone, or other mobile device that can implement at least some portion of the technology disclosed herein. The AR device 102 can include a head fitting, by which the AR device 102 can be worn on a user's head. The AR device can include one or more transparent AR display devices, each of which can overlay or project holographic images on the user's view of his or her real-world environment, for one or both eyes (e.g., by projecting light into the user's eyes). The AR device 102 can further include one or more eye-tracking cameras for gaze capturing, one or more microphones for voice input, one or more speakers for audio output, and one or more visible-spectrum video cameras for capturing surrounding environment and/or user gestures. Those of art in the skill will understand that the AR device 102 can include other sensors that provide information about the surrounding environment and/or the AR device 102 (e.g., one or more depth sensors for determining distances to nearby objects, GPS or IMU for determining positional information of the AR device 102, or the like). The AR device 102 can also include circuitry to control at least some of the aforementioned elements and perform associated data processing functions (e.g., speech and gesture recognition and display generation). The circuitry may include, for example, one or more processors and one or more memories. Some embodiments may omit some of the aforementioned components and/or may include additional components not mentioned above.

[0018] In the illustrated example, the AR device 102 is configured to communicate with one or more other components of the AR system via one or more connections 110, which can include a wired connection, a wireless connection, or a combination thereof. In some embodiments, however, the AR device 102 can implement all the functionalities of the AR system 100 as disclosed herein and can operate as a standalone device. The connection 110 can be configured to carry any kind of data, such as image data (e.g., still images and/or full-motion video, including 2D and 3D images), audio data (including voice), multimedia, and/or any other type(s) of data. The connection 110 can be, for example, a universal serial bus (USB) connection, Wi-Fi connection, Bluetooth or Bluetooth Low Energy (BLE) connection, Ethernet connection, cable connection, DSL connection, cellular connection (e.g., 3G, LTE/4G or 5G), a local area network (LAN), a wide area network (WAN), an intranet, a metropolitan area network (MAN), the global Internet, or the like, or a combination thereof.

[0019] The processing system 104 can be implemented, for example, on a personal computer, game console, tablet computer, smartphone, or other type of processing device. Alternatively or in addition, the processing system 104 (or at least a portion thereof) can be implemented via a network or cloud based computing service. As discussed above, in some embodiments, the processing system 104 can be implemented, in part or in whole, by the AR device 102. The processing system 104 can receive images, video, audio, or other data collected by one or more sensors of the AR device 102, process the received data in real-time or substantially real-time (e.g., within a threshold of delay) for extracting and/or tracking feature points in 2D and/or 3D, and generating point clouds in 2D and/or 3D in accordance with some embodiments of the presently disclosed technology.

[0020] The processing system 104 can query, search, retrieve, and/or update 3D models from the model data service 106, which can be implemented, for example, on a personal computer, game console, tablet computer, smartphone, or other type of processing device. Alternatively or in addition, the model data service 106 (or at least a portion thereof) can be implemented via a network or

cloud based computing service. As discussed above, in some embodiments, the model data service 106 can be implemented, in part or in whole, by the AR device 102. The model data service 106 may include one or more databases or data stores that maintain one or more 3D models of AR objects (e.g., 3D mesh models or 3D point cloud models). The processing system 104 can match a 3D model selected from the model data service 106 with 3D point cloud(s) generated based on feature points, and determine a proper position and/or orientation for the 3D model. Based on the match, the processing system 104 can render corresponding AR content, and cause the AR device 102 to overlay or otherwise superimpose the AR content on a user's view.

[0021] The processing system 104 can query, search, retrieve, and/or update associations between features of images and positions and/or orientations of 3D models that are maintained by the association data service 108. The association data service 108 can be implemented, for example, on a personal computer, game console, tablet computer, smartphone, or other type of processing device. Alternatively or in addition, the association data service 108 (or at least a portion thereof) can be implemented via a network or cloud based computing service. As discussed above, in some embodiments, the association data service 108 can be implemented, in part or in whole, by the AR device 102. In response to a match between a selected 3D model with one or more 3D point clouds generated based on feature points, the processing system 104 can associate the position and/or orientation of the matched 3D model with one or more images (e.g., video frames, or feature points derived from video frames) that provided basis for the matched 3D point cloud(s). Such associations can be transmitted to, stored, or otherwise maintained by, for example, one or more databases or data stores of the association data service 108.

[0022] In some embodiments, the processing system 104, the model data service 106, or the association data service 108 can pre-compute 3D point clouds and/or feature point patterns based on various pre-determined or predicted positions and/or orientations of a selected 3D model. The pre-computation results and their associated position and/or orientation of the 3D model can also be

maintained by the association data service 108, for example, to supplement, reinforce, and/or verify other associations that have been determined based on actual image or video data received from the AR device 102.

[0023] In some embodiments, the processing system 104 can identify an applicable association from the association data service 108 for any incoming image(s), for example, based on feature point patterns derived from the incoming image. Based on the position and/or orientation of a selected 3D model as indicated by the identified association, the processing system 104 can render AR content, and cause the AR device 102 to properly overlay or otherwise superimpose the AR content on the user's view.

[0024] Figure 2 illustrates an example of feature points extraction and tracking from multiple images in accordance with some embodiments of the presently disclosed technology. For a period of time between two points (e.g., t_i and t_{i+k}), an AR device 202 (e.g., corresponding to the AR device 102 of Figure 1) may capture multiple images 220a, 220b, 220c, ..., 220d (e.g., consecutive video frames) that correspond to a user's real-world view at different time within the time period. The images can be captured by a camera 212 of the AR device 202 at different positions and/or orientations. The presently disclosed technology can examine each image (e.g., processing relevant pixels within the image) in real-time or substantially real-time as it is captured and identify 2D feature points (e.g., feature pixels) from the image. While 2D feature points are being identified, the presently disclosed technology can track respective 2D feature points across multiple images captured within a time window based on correlations between or among the multiple images. In some embodiments, the feature points extraction and tracking process does not require additional knowledge of the user's real-world view other than the two dimensional data captured by the images. In some embodiments, the feature points extraction and tracking process does not require extraction of any 2D plane within a 3D reference system based on the feature points.

[0025] Figure 3 illustrates an example of epipolar geometry based feature point tracking in accordance with some embodiments of the presently disclosed

technology. Figure 3 shows two images 320a and 320b which are captured by a camera of an AR device at two different positions and/or orientations. O and O' represent the camera centers that correspond to the images 320a and 320b, respectively. The projection of O' on image 320a corresponds to an epipole point, e . Similarly, the projection of O on image 320b corresponds to an epipole point, e' . In other words, epipoles are the points of intersection of a line through camera centers and the image planes.

[0026] For a 2D feature point x extracted from image 320a, determining its corresponding 3D point X in a real-world environment is infeasible without additional information, because every point on the line OX projects to the same point x on the image plane of image 320a. But different points on the line OX project to different points x' on the image plane of image 320b.

[0027] The projection of the different points on OX form an epilines l' (corresponding to 2D feature point x) on the image plane of image 320b. To efficiently track a 2D feature point x' on image 320b that corresponds to the 2D feature point x on image 320a, the currently disclosed technology may examine pixels on the epilines l' (or within a threshold thereof) without processing other portions of the image 320b. Therefore, any 2D feature points extracted on one image can possibly be tracked using their corresponding epilines on one or more other images. Such a tracking mechanism provides better computational performance and accuracy.

[0028] With feature points tracked across two or more images, the presently disclosed technology can triangulate the location of corresponding 3D feature points (e.g., point X) in a real-world environment and track the 3D feature points across multiple images as well.

[0029] Figure 4 illustrates an example of 3D point cloud generated in accordance with some embodiments of the presently disclosed technology. As discussed above, the presently disclosed technology can extract and track 2D feature points from multiple images. The correlation of 2D feature points between images (e.g., based on epipolar geometry) can be used to determine relative rotations and/or translations between 2D reference systems (e.g., 2D coordinate

system) of the respective images. Accordingly, the presently disclosed technology can transform 2D feature points extracted from multiple images (each associated with a respective 2D reference system) captured during a sliding time window into corresponding 2D feature points that reside within a 2D reference system associated with a target image (e.g., image captured at the beginning, mid-point, end, or other point within the sliding time window). These transformed 2D feature points can be combined with 2D feature points extracted from the target image itself to form a 2D point cloud.

[0030] The presently disclosed technology can convert the 2D point cloud into a 3D point cloud (e.g., 3D point cloud 412 as shown in Figure 4) in a 3D reference system (e.g., 3D reference system 410) associated with the real-world environment or associated with the AR device or its camera. As discussed above, locations of 3D feature points that correspond to extracted 2D feature points can be determined, for example, based on triangulation. In some embodiments, the conversion from 2D point cloud to 3D point cloud can be implemented based on any other suitable technique known to those of skill in the art. For example, Perspective-*n*-Point (PnP) based methods can be used. In some embodiments, outliers of the feature points can be removed, for example, using bundle adjustment based methods.

[0031] Figure 5 is a flowchart illustrating a process 500 that can be implemented by the AR system 100 for generating a 3D point cloud and matching with a target 3D model in accordance with some embodiments of the presently disclosed technology. At block 505, the AR system 100 determines one or more regions of interest (ROI) for superimposing AR content. In some embodiments, the AR system 100 can present a user interface that enables a user to manually select one or more ROIs within a 2D image (e.g., outlining the ROI(s) on a touchable display of 2D image using finger touches and/or moves, selecting the ROI(s) via head-mount display of 2D image based on gaze and/or gesture recognition, or the like). In these embodiments, the AR system 100 can determine fixed locations and/or shapes of one or more ROIs relative to each frame (e.g., a fixed-size square that is always located at the center of each frame). In some

embodiments, the AR system 100 can select ROI(s) based on automatic object detection within 2D images. For example, the AR system 100 can use suitable face detection methods to detect one or more regions that represent a human face within individual frames. In these embodiments, the AR system 100 can also estimate a measure of depth from the camera to the ROI(s) based on a comparison between the size of the detected object(s) in 2D image(s) and a known size of the object in the real world. Accordingly, 2D ROI determination can be performed regardless of movement of the AR device 102 or real world objects.

[0032] In some embodiments, the AR system 100 can enable a user to manually select one or more ROIs in a 3D reference system (3D coordinate system) associated with the AR device 102 (e.g., AR-HMD), for example, based on gaze and/or gesture recognition. In this regard, the AR system 100 can estimate camera motion once the AR system 100 has received and/or processed a sufficient amount of pose and map data. Given camera lens calibration data (e.g., provided by the AR device 102), the AR system 100 can estimate ROI(s) projected to any 2D frame using depth information of corresponding locations within the 3D reference system.

[0033] ROI determination does not have to be exact, because the presently disclosed technology includes functionalities designed to properly process "noise" and/or "outliers." In some embodiments, ROI determination is not performed, and the entire image, frame, and/or surrounding environment can be consider an ROI processed by the AR system 100 for superimposing AR content.

[0034] At block 510, the AR system 100 identifies and tracks 2D feature points from multiple frames. Illustratively, the AR system 100 can identify corners, edges, centroids, and/or other feature points within ROI(s) in each frame using suitable image processing methods known to those having skill in the art. As discussed above, 2D feature points can be tracked and correlated between or among frames using epipolar geometry based method(s), which is computationally more efficient than searching the entire image(s) for a feature point that can correlate to a corresponding feature point in another image. In some embodiments, the AR system 100 retrieves a pose estimation between two frames where feature points

are to be tracked, and determines matching epilines (with or without a threshold proximity included) based on that estimation, and then refine the pose estimation based on feature points that are tracked between the two frames. In some embodiments, various sensor data can be collected from the AR device 102 to generate more accurate pose estimation in a more efficient manner. For example, the AR system 100 can implement suitable sensor fusion methods known to those of skill in the art.

[0035] At block 515, the AR system 100 generates 2D point cloud corresponding to a target frame. As discussed above, the correlation of 2D feature points between images (e.g., based on epipolar geometry) can be used to determine relative rotations and/or translations between 2D reference systems (e.g., 2D coordinate system) of the respective images. Accordingly, the AR system 100 can project the 2D feature points extracted from multiple images captured during a sliding time window onto a 2D reference system associated with a target image (e.g., a frame captured at the beginning, mid-point, end, or other point within the sliding time window) to form a 2D point cloud corresponding to the target frame. This process can include de-duplication, weighting, or other smoothing actions known to those of skill in the art to avoid double counting 2D feature points that correspond to a same 3D feature point in real world environment.

[0036] At block 520, the AR system 100 generates 3D point cloud corresponding to the target frame based at least in on the 2D point cloud. As discussed above, the AR system 100 can convert the 2D point cloud into a 3D point cloud in a 3D reference system (e.g., 3D coordinate system) associated with the real-world environment or associated with the AR device 102 or its camera. As discussed above, locations of 3D feature points that correspond to extracted 2D feature points can be determined, for example, based on triangulation. In some embodiments, the conversion from 2D point cloud to 3D point cloud can be implemented based on any other suitable technique known to those of skill in the art. For example, Perspective- n -Point (PnP) based methods can be used. In

some embodiments, outliers of the feature points can be removed, for example, using bundle adjustment based methods.

[0037] At block 525, the AR system 100 matches one or more target 3D models with the 3D point cloud. Illustratively, the AR system 100 selects the target 3D model(s) from the model data service 106. The target 3D models can include, for example, a 3D mesh or point cloud model of a real-world object such as a human face, head, brain, or the like that has been pre-generated based on existing data, measurements, or design. In some embodiments, the AR system 100 converts the target 3D model into a simplified model point cloud with occlusion, that is, the AR system 100 determines which 3D points (e.g., corners, edges, centroids, or the like) included in or derived from the target 3D model would have been visible to the camera and/or may be detected as a feature point in a frame captured by the camera. The AR system 100 can then match the simplified model point cloud with the 3D point cloud generated at block 520 based, for example, on Iterative Closest Point (ICP) methods or other suitable methods known to those of skill in the art.

[0038] At block 530, the AR system 100 associates (1) the position and/or orientation of the matched target 3D model(s) with (2) the target frame. Illustratively, the matching process of block 525 determines the position and/or orientation of the target 3D model(s) so as to align with the 3D point cloud that corresponds to the target frame. The AR system 100 can record the position and/or orientation information and associate it (e.g., using pointers, address reference, and/or additional data fields of same data record) with the target frame (e.g., image data and/or 2D feature points extracted from the target frame). Such a record of association can be stored, indexed, and/or otherwise maintained by the association data service 108 for efficient search and retrieval.

[0039] At block 535, the AR system 100 renders AR content based on the position and/or orientation of the matched target 3D model. Illustratively, the target frame corresponds to a frame recently captured by a camera of the AR device 102. Using the position and/or orientation of the matched target 3D model as a basis, the AR system 100 can compute an estimated position and/or orientation of the

target 3D model to properly project or overlay AR content onto a current view of the user. In this regard, the estimation can be based on movement and/or rotation, in one or multiple dimensions, of the AR device 102 between the target frame and the current user view.

[0040] In some embodiments, the AR system 100 can use one or multiple types of data collected within a recent period of time (e.g., feature point tracking data, GPS data, IMU data, LiDAR data, or the like that can be provided by the AR device 102) to determine the movement and/or rotation. Given the estimated position and/or orientation of the target 3D model, the AR system 100 can render AR content based thereon and project or overlay AR content to align with a corresponding real world object within the user's current view.

[0041] Figure 6 is a flowchart illustrating a process 600 that can be implemented by the AR system 100 for pre-computing and matching of 3D point clouds based on likely positions and/or orientations of a target 3D model in accordance with some embodiments of the presently disclosed technology. Process 600 can be implemented in combination, in parallel, or in sequence with the implementation of process 500. At block 605, the AR system 100 determines likely positions and/or orientations of a selected target 3D model for AR content rendering. Illustratively, for a target 3D model selected from the model data service 106, the AR system 100 can randomly generate a number of positions and/or orientations within a threshold proximity of a base position and/or orientation. In some embodiments, the base position and/or orientation can correspond to a position and/or orientation of the target 3D model that matches a 3D point cloud generated based on one or more recent frames. For example, the base position and/or orientation can be selected from the associations that have been added or updated with the association data service within a recent period of time. In some embodiments, the AR system 100 can determine likely positions and/or orientations of the selected target 3D model based on user input, historical statistics, positions and/or orientations of an associated 3D model for other AR content rendering, or the like.

[0042] At block 610, the AR system 100 pre-computes 3D point clouds based on the target 3D model in accordance with the determined likely positions and/or orientations. Illustratively, the AR system 100 converts the target 3D model into a simplified model point cloud with occlusion, that is, the AR system 100 determines which 3D points included in or derived from the target 3D model would have been visible to the camera and/or may be detected as a feature point in a frame captured by the camera. The AR system 100 can compute mathematical transformations (e.g., translation and/or rotation) between 3D reference systems of (1) the base position and/or orientation and (2) each likely positions and/or orientations. The AR system 100 can apply the mathematical transformations to the simplified model point cloud to obtain the pre-computed 3D point clouds that correspond to the target 3D model at various likely positions and/or orientations. In some embodiments, the AR system 100 can convert individual pre-computed 3D point clouds into one or more corresponding pre-computed 2D point clouds in accordance based on likely positions and/or orientation (e.g., a centered frontal orientation) of the AR device or camera.

[0043] At block 615, the AR system 100 matches the pre-computed 3D point clouds with a 3D point cloud that corresponds to a target frame. As discussed above with respect to blocks 515 and 520 of the process 500, the AR system 100 can generate a 3D point cloud corresponding to a target frame based at least in on a 2D point cloud. The AR system 100 can compare multiple pre-computed 3D point clouds with the 3D point cloud that corresponds to the target frame. The comparison can be achieved by calculating a comparative difference measure, such as an average pair-wise difference between nearest points in a pre-computed 3D point cloud and the 3D point cloud corresponding to the target frame. The AR system 100 can select one or more pre-computed 3D point clouds associated with the least comparative difference(s) as matching the 3D point cloud corresponding to the target frame after ordering or sorting the comparative differences. The comparison and matching can be computationally cheap and efficient because they do not need to include an optimization process (e.g., iterative gradient descent based methods). In the embodiments where the pre-computed 3D point clouds have been converted into corresponding pre-computed 2D point clouds, the

AR system 100 can alternatively match the pre-computed 2D point clouds with a 2D point cloud that corresponds to a target frame, which can be computationally more efficient.

[0044] At block 620, the AR system 100 associates the position and/or orientation of the target 3D model with the target frame based on the match of block 615. Illustratively, the AR system can record the position(s) and/or orientation(s) associated with the matching pre-computed 3D point cloud(s) as selected at block 615, and associate the recorded position(s) and/or orientation(s) with the target frame (or its 2D feature points). As discussed above, the AR system 100 can store the association with the association data service 108 for future search, retrieval, and/or other uses.

[0045] At block 625, the AR system 100 renders AR content based on the position and/or orientation of the target 3D model. As discussed above with respect to block 535 of the process 500, illustratively, the target frame can correspond to a frame recently captured by a camera of the AR device 102. Using the target 3D model position and/or orientation recorded at block 620, the AR system 100 can compute an estimated position and/or orientation of the target 3D model to properly project or overlay AR content onto a current view of the user. In this regard, the estimation can be based on movement and/or rotation, in one or multiple dimensions, of the AR device 102 between the target frame and the current user view.

[0046] In some embodiments, the AR system 100 can use one or multiple types of data collected within a recent period of time (e.g., feature point tracking data, GPS data, IMU data, LIDAR data, or the like that can be provided by the AR device 102) to determine the movement and/or rotation. Given the estimated position and/or orientation of the target 3D model, the AR system 100 can render AR content based thereon and project or overlay AR content to align with a corresponding real world object within the user's current view.

[0047] In other embodiments, the target frame can correspond to a frame reflecting the current view of the user. In these embodiments, the AR system 100 does not need to calculate additional estimated position and/or orientation of the

target 3D model, but can use the position and/or orientation recorded at block 620 for AR content rendering. These embodiments can be achieved, for example, due to the computational efficiency of the comparison and matching process of block 615.

[0048] Figure 7 is a flowchart illustrating a process 700 that can be implemented by the AR system 100 for identifying an existing association between a target 3D model and 2D feature points of a target frame in accordance with some embodiments of the presently disclosed technology. Process 700 can be implemented in combination, in parallel, or in sequence with the implementation of process 500 and/or process 600. At block 705, the AR system 100 processes a target frame to extract 2D feature points. In some embodiments, the AR system extracts a limited number of 2D feature points to ensure real-time execution of the process 700.

[0049] At block 710, the AR system 100 searches for an existing association between the position and/or orientation of a target 3D model and the target frame based at least in part on the extracted 2D feature points. Illustratively, the AR system 100 can search the association records maintained by the association data service 108 to identify a recorded set of 2D feature points that best matches (e.g. with the least matching error) the currently extracted 2D feature points. In some embodiments, multiple sets of matching 2D feature points can be identified. For example, the matching error for each set of the matching 2D feature points is below an acceptance threshold. In some embodiments, the search for existing association(s) can be based on a match between the target frame itself (e.g., 2D image data) and the image data maintained within association records of the association data service 108.

[0050] At block 715, the AR system 100 determines whether one or more exiting associations are identified. If so, the process 700 proceeds to block 720 where the AR system 100 renders AR content based at least in part on the identified position and/or orientation of the target 3D model. In embodiments where a single association is identified, the AR system 100 can use the position and/or orientation of the target 3D model associated with the matching set of 2D

feature points for AR content rendering. In embodiments where multiple associations are identified, the AR system 100 can use all their associated model positions and/or orientations by, for example, calculating a weighted average position and/or orientation of the target 3D model. In some embodiments, the AR content rendering at block 720 can be performed in a similar manner as block 535 of process 500.

[0051] In some embodiments, the target frame can correspond to a current user view and the AR system 100 does not need to compute another estimated position and/or orientation of the target 3D model for projecting AR content. In these embodiments, the AR system 100 can project AR content using the position and/or orientation obtained from the identified association(s). These embodiments can be enabled, at least in part, due to the relatively short processing time of the association identification process, which does not involve the computationally more expensive 2D and/or 3D point cloud generation and matching.

[0052] Referring back to block 715, if the AR system 100 fails to identify one or more existing associations, the process 700 proceeds to block 725 where the AR system 100 performs 3D point cloud based target model matching and AR rendering, for example, in accordance with process 600 and/or process 700 as discussed above.

[0053] Figure 8 is a block diagram illustrating an example of the architecture for a computer system (or computing device) 800 that can be utilized to implement various portions of the presently disclosed technology. In Figure 8, the computer system 800 includes one or more processors 805 and memory 810 connected via an interconnect 825. The interconnect 825 may represent any one or more separate physical buses, point to point connections, or both, connected by appropriate bridges, adapters, or controllers. The interconnect 825, therefore, may include, for example, a system bus, a Peripheral Component Interconnect (PCI) bus, a HyperTransport or industry standard architecture (ISA) bus, a small computer system interface (SCSI) bus, a universal serial bus (USB), IIC (I2C) bus, or an Institute of Electrical and Electronics Engineers (IEEE) standard 674 bus, sometimes referred to as "Firewire".

[0054] The processor(s) 805 may include central processing units (CPUs) to control the overall operation of, for example, the host computer. In certain embodiments, the processor(s) 805 accomplish this by executing software or firmware stored in memory 810. The processor(s) 805 may be, or may include, one or more programmable general-purpose or special-purpose microprocessors, digital signal processors (DSPs), graphics processing unit (GPU), mobile application processors, microcontrollers, application specific integrated circuits (ASICs), programmable gate arrays (PGAs), programmable controllers, programmable logic devices (PLDs), or the like, or a combination of such devices.

[0055] The memory 810 can be or include the main memory of the computer system. The memory 810 represents any suitable form of random access memory (RAM), read-only memory (ROM), flash memory, or the like, or a combination of such devices. In use, the memory 810 may contain, among other things, a set of machine instructions which, when executed by processor 805, causes the processor 805 to perform operations to implement various embodiments of the presently disclosed technology.

[0056] Also connected to the processor(s) 805 through the interconnect 825 is a (optional) network adapter 815. The network adapter 815 provides the computer system 800 with the ability to communicate with remote devices, such as the storage clients, and/or other storage servers, and may be, for example, an Ethernet adapter or Fiber Channel adapter. Additionally and optionally, transparent display device, depth camera or sensor, head tracking camera, video camera, other sensors, communication device, audio device, or the like can be connected to the processor(s) 805 (directly or indirectly) through the interconnect 825.

[0057] The machine-implemented operations described above can be implemented by programmable circuitry programmed/configured by software and/or firmware, or entirely by special-purpose circuitry, or by a combination of such forms. Such special-purpose circuitry (if any) can be in the form of, for example, one or more application-specific integrated circuits (ASICs),

programmable logic devices (PLDs), field-programmable gate arrays (FPGAs), system-on-a-chip systems (SOCs), etc.

[0058] Software or firmware for use in implementing the techniques introduced here may be stored on a machine-readable storage medium and may be executed by one or more general-purpose or special-purpose programmable microprocessors. A "machine-readable storage medium," as the term is used herein, includes any mechanism that can store information in a form accessible by a machine (a machine may be, for example, a computer, network device, cellular phone, personal digital assistant (PDA), manufacturing tool, any device with one or more processors, etc.). For example, a machine-accessible storage medium includes recordable/non-recordable media (e.g., read-only memory (ROM); random access memory (RAM); magnetic disk storage media; optical storage media; flash memory devices; etc.), etc.

[0059] The term "logic," as used herein, can include, for example, programmable circuitry programmed with specific software and/or firmware, special-purpose hardwired circuitry, or a combination thereof.

[0060] Some embodiments of the disclosure have other aspects, elements, features, and steps in addition to or in place of what is described above. These potential additions and replacements are described throughout the rest of the specification. Reference in this specification to "various embodiments," "certain embodiments," or "some embodiments" means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the disclosure. These embodiments, even alternative embodiments (e.g., referenced as "other embodiments") are not mutually exclusive of other embodiments. Moreover, various features are described which may be exhibited by some embodiments and not by others. Similarly, various requirements are described which may be requirements for some embodiments but not other embodiments.

[0061] Although the subject matter has been described in language specific to structural features and/or acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or

acts described above. Rather, the specific features and acts described above are disclosed as examples of implementing the claims and other equivalent features and acts are intended to be within the scope of the claims.

CLAIMS

I/We claim:

1. A computer-implemented method, comprising:
receiving a first image that captures at least a portion of a user's view at a first point in time;
detecting a set of feature points within the first image;
receiving one or more second images that capture at least a portion of the user's view at one or more second points in time;
tracking at least a subset of the set of feature points between the first image and the one or more second images based, at least in part, on epipolar geometry constructed between the first image and the one or more second images; and
rendering augmented reality (AR) content based, at least in part, on the tracked feature points.
2. The method of claim 1, further comprising determining a region of interest (ROI) for detecting the set of feature points.
3. The method of claim 2, wherein determining the ROI comprises automatically detecting at least one object of interest within the first image.
4. The method of claim 3, wherein the at least one object corresponds to a face of a subject.
5. The method of claim 2, wherein determining the ROI comprises receiving a user interaction that indicates the ROI via a user interface.
6. The method of claim 5, wherein the user interaction includes a gaze and/or gesture detected via a head-mount device.

7. The method of claim 1, wherein detecting the set of feature points comprises detecting at least one of corners, edges, or centroids of objects within the first image.

8. The method of claim 1, wherein tracking at least a subset of the set of feature points between the first image and the one or more second images comprises determining feature points along epilines on the one or more second images that correspond to the subset of feature points.

9. The method of claim 1, wherein the first and the one or more second images correspond to consecutive frames of a video.

10. The method of claim 1, wherein respective user's views captured by the first and the one or more second images differ, at least in some portions, due to rotational and/or translational motion of the user.

11. A non-transitory computer-readable medium storing computer-executable instructions that, when executed by one or more processors, cause the one or more processors to perform operations, the operations comprising:

for each image obtained by an augmented reality (AR) device during a period of time, detecting a set of two-dimensional feature points within the image;

correlating a plurality of two-dimensional feature points between at least two of the detected sets of two-dimensional feature points based, at least in part, on epipolar geometry constructed between the images obtained during the period of time; and

rendering augmented reality (AR) content based, at least in part, on the correlated two-dimensional feature points.

12. The computer-readable medium of claim 11, wherein the operations further comprise determining three-dimensional feature points that correspond to the plurality of two-dimensional feature points.

13. The computer-readable medium of claim 11, wherein the operations further comprise generating a two-dimensional point cloud based, at least in part, on the sets of two-dimensional feature points detected during the period of time.

14. The computer-readable medium of claim 11, wherein the AR device includes a head-mount device.

15. The computer-readable method of claim 11, wherein rendering AR content comprises projecting AR content via the AR device to a user's current view.

16. A system, comprising:

one or more processors;

a memory configured to store a set of instructions, which when executed by the one or more processors cause the system to:

for each image obtained during a period of time, detect a set of two-dimensional feature points within the image;

correlate a plurality of two-dimensional feature points between at least two of the detected sets of two-dimensional feature points based, at least in part, on epipolar geometry constructed between the images obtained during the period of time; and

render augmented reality (AR) content based, at least in part, on the correlated two-dimensional feature points.

17. The system of claim 16, wherein the set of instructions, which when executed by the one or more processors further cause the system to generate a three-dimensional point cloud based, at least in part, on the sets of two-dimensional feature points obtained during the period of time.

18. The system of claim 17, wherein the three-dimensional point cloud corresponds to a particular point of time within the period of time.

19. The system of claim 18, wherein the set of instructions, which when executed by the one or more processors further cause the system to determine an estimated positional change between the particular point of time and a present time.

20. The system of claim 19, wherein rendering AR content is further based, at least in part, on the estimated positional change.

21. A computer-implemented method, comprising:
generating a plurality of pre-computed three-dimensional point clouds based, at least in part, on a plurality of likely positions and/or orientations of a three-dimensional model for rendering augmented reality (AR) content, wherein each pre-computed three-dimensional point cloud corresponds to a distinct likely position and/or orientation of the three-dimensional model;
selecting one of the pre-computed three-dimensional point clouds that matches a target three-dimensional point cloud that corresponds to a first image capturing a user's view at a first point in time; and
rendering augmented reality (AR) content based, at least in part, on the likely position and/or orientation of the three-dimensional model that corresponds to the selected pre-computed three-dimensional point cloud.

22. The method of claim 21, wherein generating the plurality of pre-computed three-dimensional point clouds comprises determining the plurality of likely positions and/or orientations of the three-dimensional model based, at least in part, on a base position and/or orientation.

23. The method of claim 22, wherein the base position and/or orientation is determined based, at least in part, on one or more second images capturing the user's view at one or more second points in time.

24. The method of claim 23, wherein the one or more second points in time precede the first point in time.

25. The method of claim 22, wherein determining the plurality of pre-computed three-dimensional point clouds comprises randomly generating the plurality of pre-computed three dimensional point clouds within a proximity of the base position and/or orientation.

26. The method of claim 21, further comprising generating and maintaining an association between an indication of the first image and the likely position and/or orientation of the three-dimensional model that corresponds to the selected pre-computed three-dimensional point cloud.

27. The method of claim 26, wherein the indication of the first image includes a plurality of two-dimensional feature points detected from the first image.

28. The method of claim 27, wherein the two-dimensional feature points include at least one of detected corners, edges, or centroids of objects within the first image.

29. The method of claim 21, wherein the first and the one or more second images correspond to frames of a video.

30. The method of claim 21, wherein respective user's views captured by the first and the one or more second images differ, at least in some portions, due to rotational and/or translational motion of the user.

31. A non-transitory computer-readable medium storing computer-executable instructions that, when executed by one or more processors, cause the one or more processors to perform operations, the operations comprising:

generating a plurality of pre-computed point clouds based, at least in part, on a plurality of likely positions and/or orientations of a three-dimensional model for rendering augmented reality (AR) content, wherein each pre-computed point cloud corresponds to a likely position and/or orientation of the three-dimensional model;

selecting a pre-computed point cloud that matches a target point cloud that corresponds to a first image of a user's view captured by an augmented reality (AR) device at a first point in time; and
rendering augmented reality (AR) content based, at least in part, on the likely position and/or orientation of the three-dimensional model that corresponds to the selected pre-computed point cloud.

32. The computer-readable medium of claim 31, wherein the pre-computed point clouds are three-dimensional or two-dimensional.

33. The computer-readable medium of claim 31, wherein the selected pre-computed point cloud has a least comparative difference with the target point cloud among the plurality of pre-computed point clouds.

34. The computer-readable medium of claim 31, wherein the AR device includes a head-mount device.

35. The computer-readable method of claim 31, wherein rendering AR content comprises projecting AR content via the AR device to a user's view at a second point in time, wherein the first point in time precedes the second point in time.

36. A system, comprising:
one or more processors;
a memory configured to store a set of instructions, which when executed by the one or more processors cause the system to:
generate a plurality of pre-computed point clouds based, at least in part, on a plurality of likely positions and/or orientations of a three-dimensional model for rendering augmented reality (AR) content, wherein each pre-computed point cloud corresponds to a likely position and/or orientation of the three-dimensional model;
select one or more pre-computed point clouds that match a target point cloud that corresponds to a user's view captured by an augmented reality (AR) device at a target point in time; and

render augmented reality (AR) content based, at least in part, on one or more likely positions and/or orientations of the three-dimensional model that each correspond to the selected one or more pre-computed point clouds.

37. The system of claim 36, wherein generating the plurality of pre-computed point clouds is further based on at least one of user input, historical statistics, positions and/or orientations of an associated three-dimensional model.

38. The system of claim 36, wherein the set of instructions, which when executed by the one or more processors further cause the system to determine an estimated positional change between the target point in time and a present time.

39. The system of claim 38, wherein rendering AR content is further based, at least in part, on the estimated positional change.

40. The system of claim 38, wherein determining the estimated positional change is based, at least in part, on movement and/or rotation, in one or multiple dimensions, of the AR device between the target point in time and the present time.

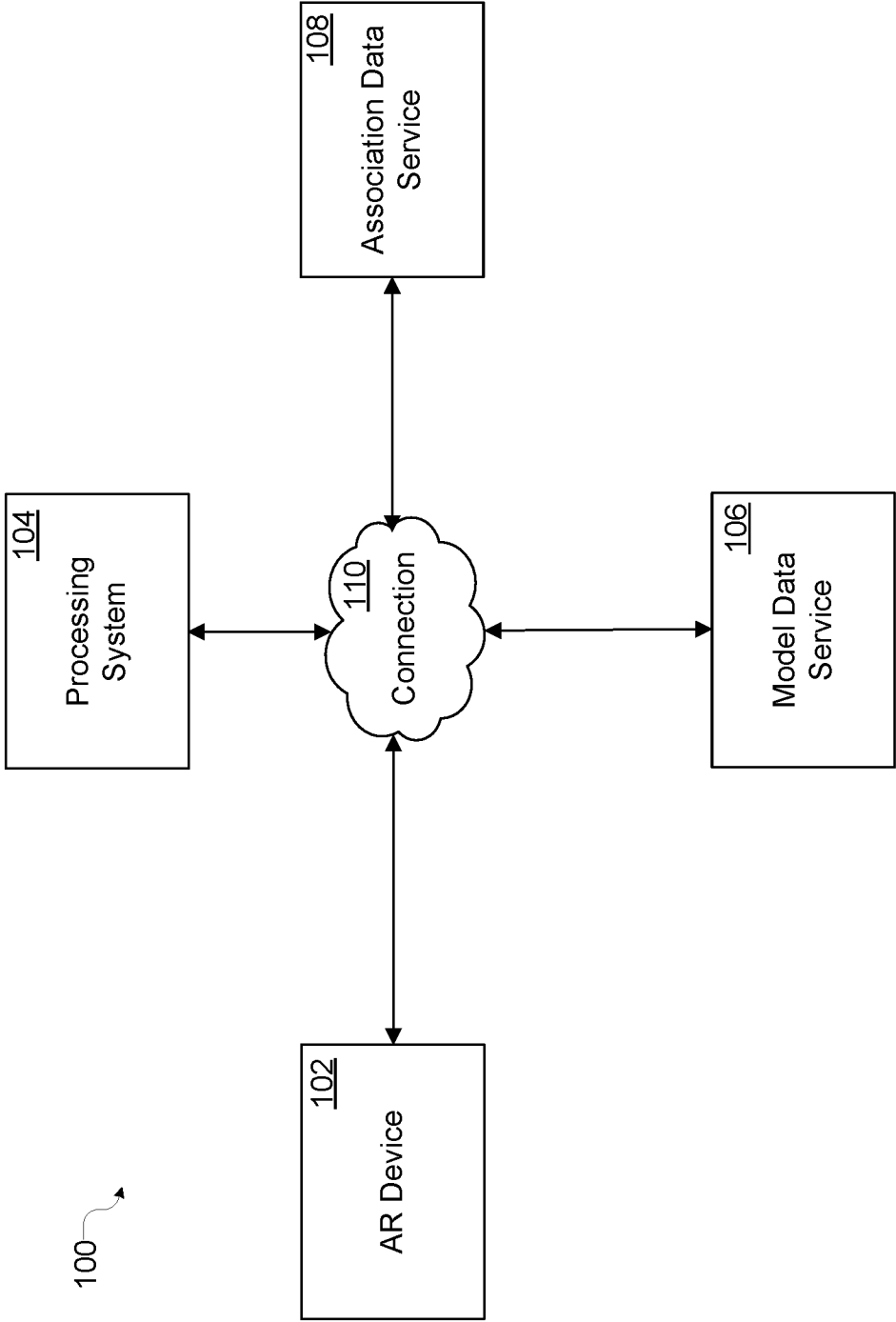


FIG. 1

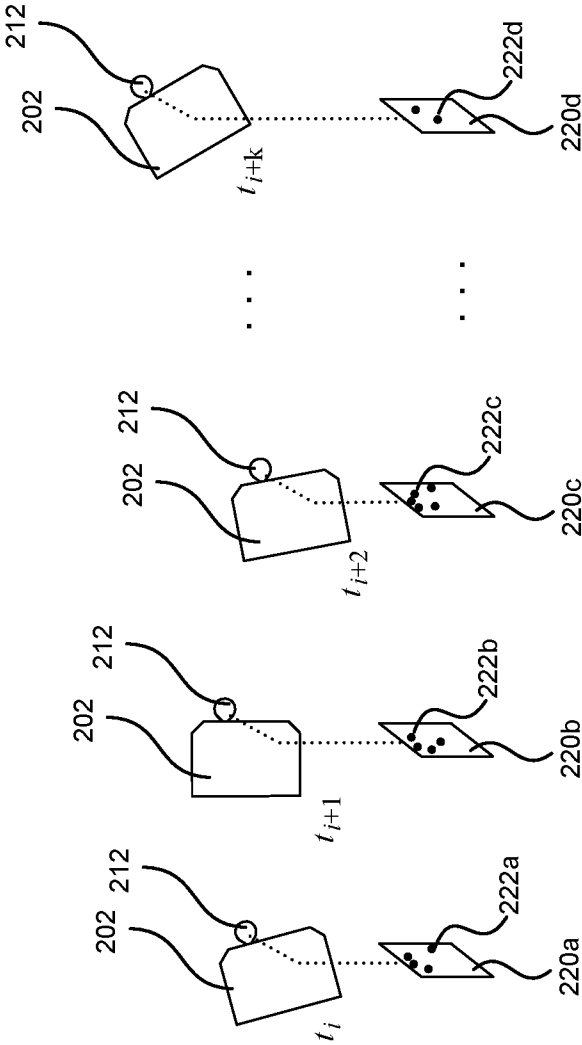


FIG. 2

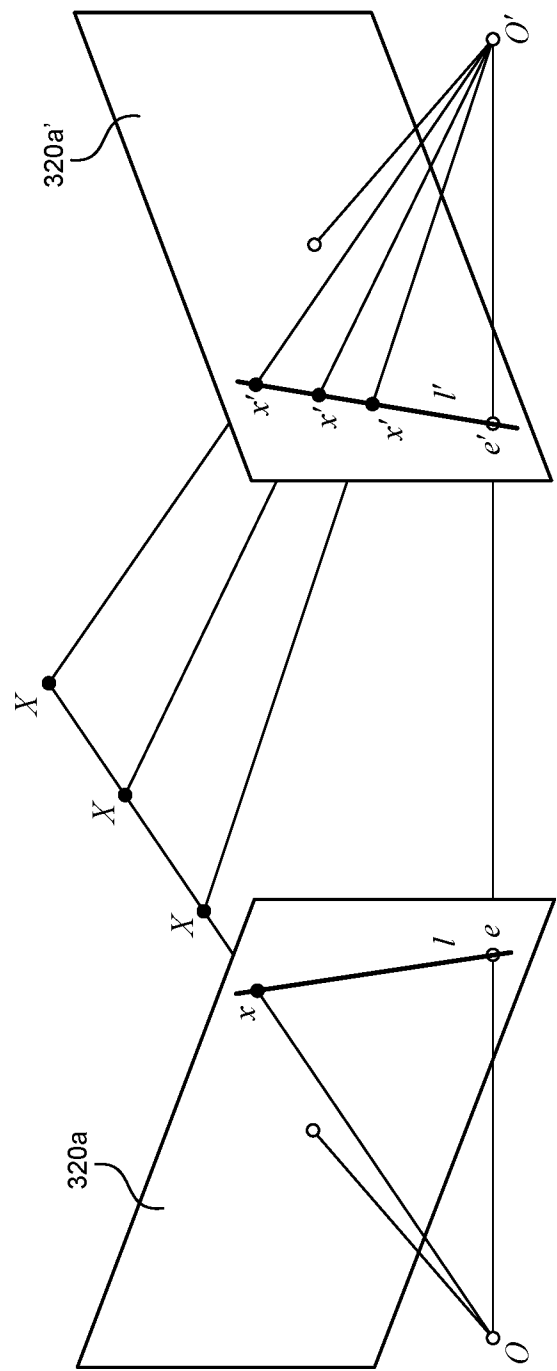


FIG. 3

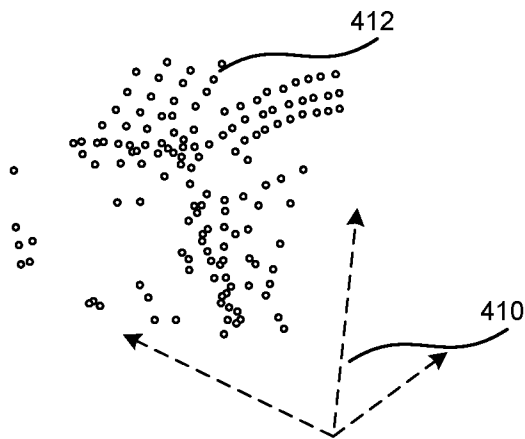
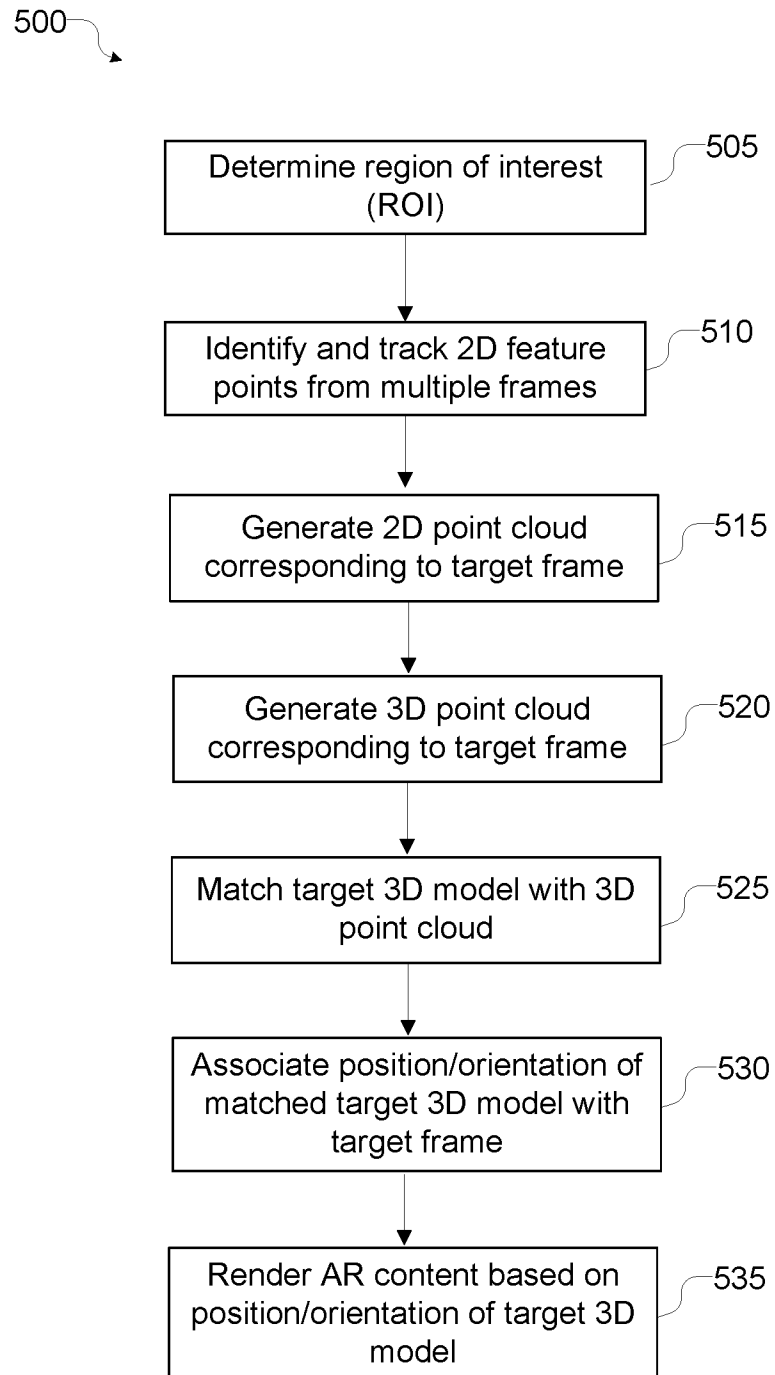
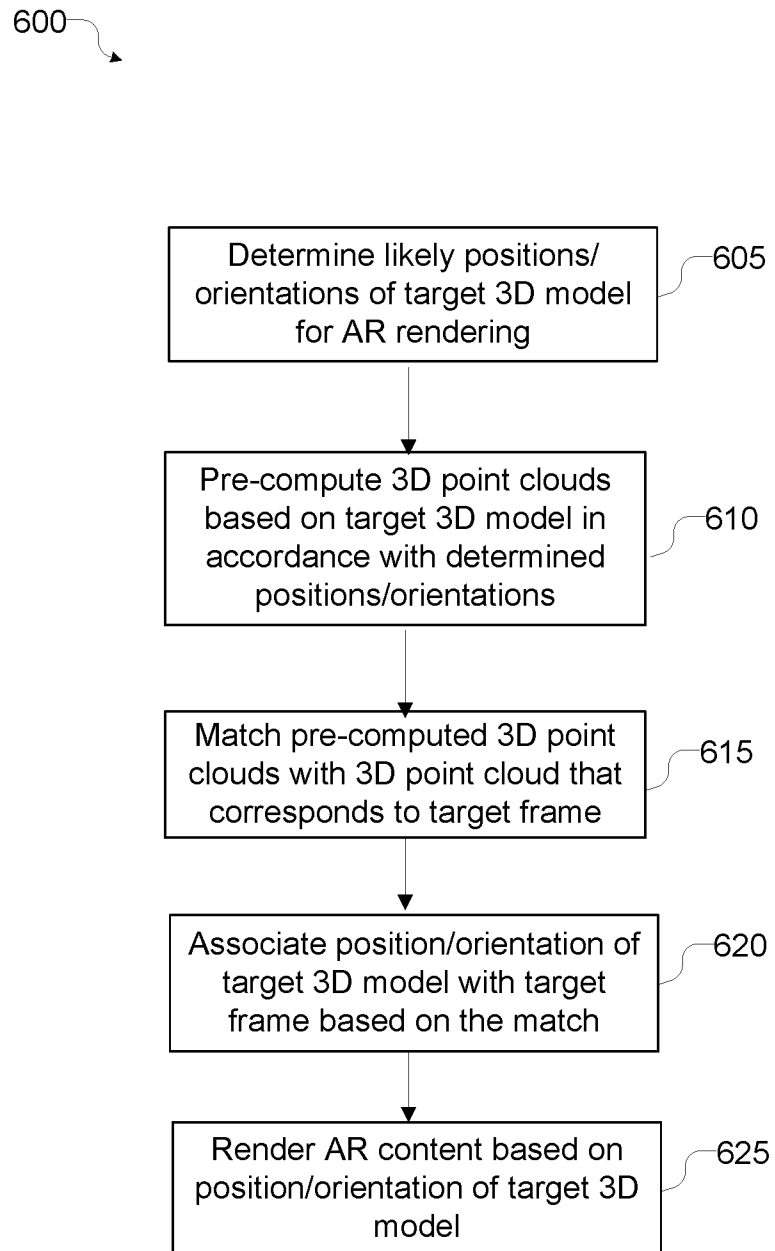


FIG. 4

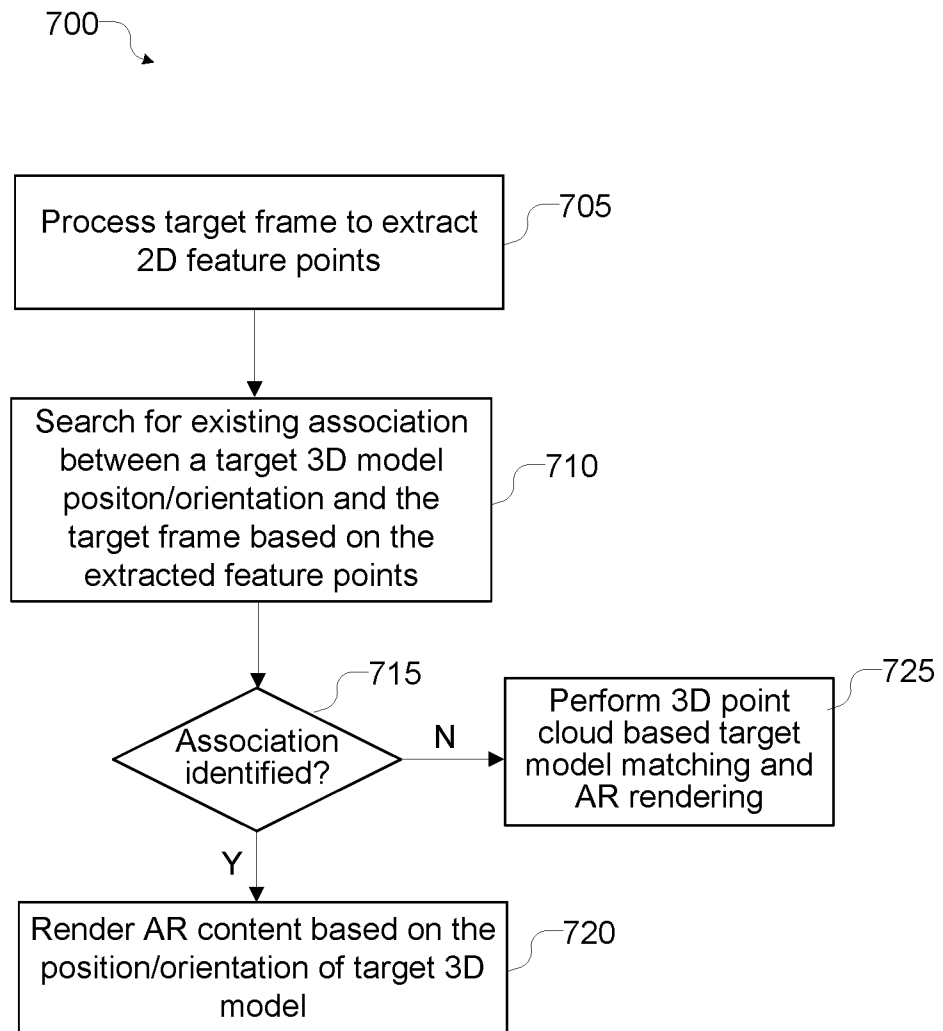
5/8

**FIG. 5**

6/8

**FIG. 6**

7/8

**FIG. 7**

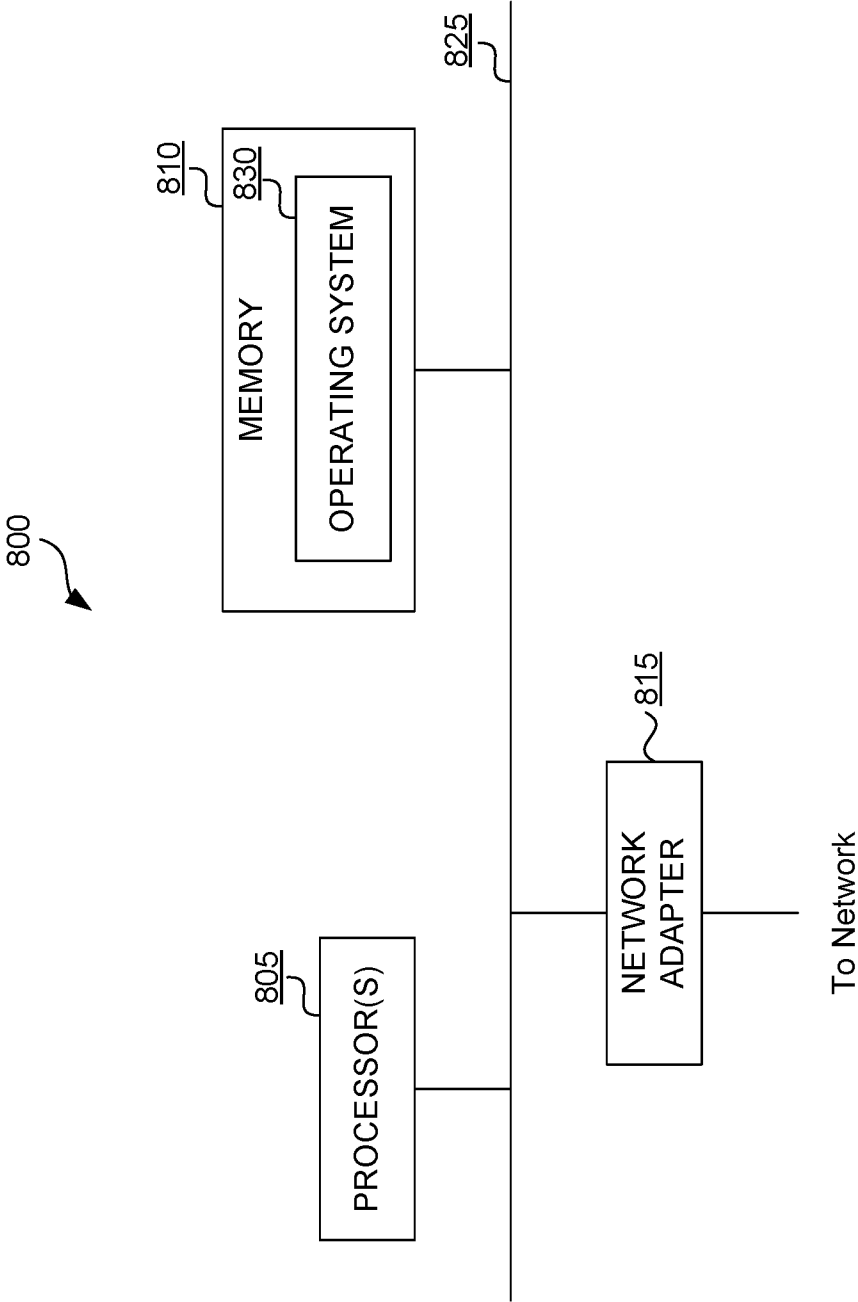


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 18/43164

A. CLASSIFICATION OF SUBJECT MATTER
 IPC(8) - G09G 5/00 (2018.01)
 CPC - G06T 19/006, G06T 7/246, G06T 2207/10016

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)

See Search History Document

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

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Sanjuan et al. "ONLINE REGISTRATION TOOL AND MARKERLESS TRACKING FOR AUGMENTED REALITY." In: Proceeding of WIAMIS 2005, Montreux, Switzerland, April, 2005, [online] [retrieved on 31 October 2018 (31.10.2018)] Retrieved from the Internet <URL: https://www.researchgate.net/publication/37436032 >, entire document, especially Abstract; page 2-4	1-20
Y	US 2016/0253805 A1 (Google Inc), 01 September 2016 (01.09.2016), entire document, especially Abstract; para [0004], [0032]-[0037]	1-20
Y	US 2014/0055342 A1 (FUJITSU LIMITED), 27 February 2014 (27.02.2014), entire document, especially Abstract; para [0008]	4
Y	US 2015/0130740 A1 (TOBII TECHNOLOGY AB), 14 May 2015 (14.05.2015), entire document, especially Abstract; para [0136]-[0140], [0208]-[0209]	6
Y	US 2005/0200624 A1 (Lachner et al.), 15 September 2005 (15.09.2005), entire document, especially Abstract; para [0027], [0033]-[0036]	8
Y	US 2017/0193331 A1 (Autodesk, Inc), 06 July 2017 (06.07.2017), entire document, especially Abstract; para [0006], [0054]	13, 17-20
A	US 2013/0335529 A1 (Metaio GmbH), 19 December 2013 (19.12.2013), entire document	1-20
A	US 2011/0090252 A1 (YOON et al.), 21 April 2011 (21.04.2011), entire document	1-20

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

01 November 2018

Date of mailing of the international search report

26 NOV 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
 P.O. Box 1450, Alexandria, Virginia 22313-1450
 Facsimile No. 571-273-8300

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 18/43164

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

--- (See Supplemental Box) ---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-20

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 18/43164

Continuation of:

Box III. Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I - Claims 1-20 are directed to a method, program product and system for rendering augmented reality content based on tracked feature points.

Group II - Claims 21-40 method, program product and system for rendering augmented reality content based on the likely position and orientation of a three-dimensional model.

The inventions listed as Groups I-II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Features:

The invention of Group I included the features of receiving a first image that captures at least a portion of a user's view at a first point in time; detecting a set of feature points within the first image; receiving one or more second images that capture at least a portion of the user's view at one or more second points in time; images obtained by an augmented reality device during a period of time; detecting a set of two-dimensional feature points within the image; tracking/correlating at least a subset of the set of feature points between the first image and the one or more second images based, at least in part, on epipolar geometry constructed between the first image and the one or more second images; and rendering augmented reality content based, at least in part, on the tracked/correlated feature points, not required by group II.

The invention of Group II included the features of generating a plurality of pre-computed three-dimensional point clouds based, at least in part, on a plurality of likely positions and/or orientations of a three-dimensional model for rendering augmented reality (AR) content, wherein each pre-computed three-dimensional point cloud corresponds to a distinct likely position and/or orientation of the three-dimensional model; selecting one of the pre-computed three-dimensional point clouds that matches a target three-dimensional point cloud that corresponds to a first image capturing a user's view at a first point in time; and rendering augmented reality (AR) content based, at least in part, on the likely position and/or orientation of the three-dimensional model that corresponds to the selected pre-computed three-dimensional point cloud, not required by group I.

Common Technical Features

Groups I-II share the features of one or more processors and a memory configured to store a set of instructions; rendering augmented reality (AR) content.

However, the shared technical features do not represent a contribution over prior art as being anticipated by US 2015/0332512 A1 (NANT VISION, INC.), 19 November 2015 (19.11.2015).

NANT VISION, INC teaches one or more processors and a memory configured to store a set of instructions (para [0032]- processor and memory storing software instructions); rendering augmented reality content (para [0032]- the device to render AR content as disclosed.).

As the common features were known in the art at the time of the invention, this cannot be considered a common technical feature that would otherwise unify the groups. Therefore, Groups I-II lack unity under PCT Rule 13.