

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 December 2009 (03.12.2009)

(10) International Publication Number  
**WO 2009/146044 A1**

(51) International Patent Classification:  
**G06T 17/00** (2006.01)

(21) International Application Number:  
PCT/US2009/039028

(22) International Filing Date:  
31 March 2009 (31.03.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
12/059,794 31 March 2008 (31.03.2008) US

(71) Applicant (for all designated States except US): **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, WA 98052-6399 (US).

(72) Inventors: **STEEDLY, Drew**; C/o Microsoft Corporation, Lca, International Patents, One Microsoft Way, Redmond, WA 98052-6399 (US). **SZELISKI, Richard**; C/o Microsoft Corporation, Lca, International Patents, One Microsoft Way, Redmond, WA 98052-6399 (US). **SIN-**

**HA, Sudipta**; C/o Microsoft Corporation, Lca, International Patents, One Microsoft Way, Redmond, WA 98052-6399 (US). **AGRAWALA, Manesh**; C/o Microsoft Corporation, Lca, International Patents, One Microsoft Way, Redmond, WA 98052-6399 (US).

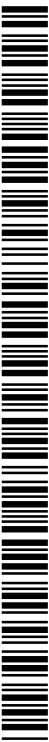
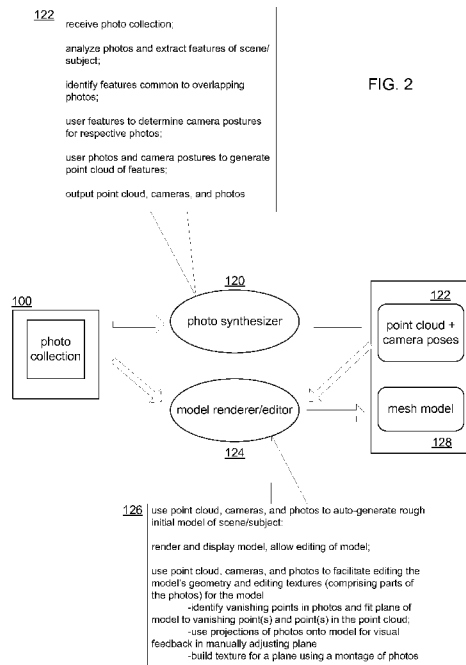
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: USING PHOTO COLLECTIONS FOR THREE DIMENSIONAL MODELING

(57) Abstract: A collection of photos and a three-dimensional reconstruction of the photos are used to construct and texture a mesh model. In one embodiment, a first digital image of a first view of a real world scene is analyzed to identify lines in the first view. Among the lines, parallel lines are identified. A three-dimensional vanishing direction in a three-dimensional space is determined based on the parallel lines and an orientation of the digital image in the three-dimensional space. A plane is automatically generated by fitting the plane to the vanishing direction. A rendering of a three-dimensional model with the plane is displayed. Three-dimensional points corresponding to features common to the photos may be used to constrain the plane. The photos may be projected onto the model to provide visual feedback when editing the plane. Furthermore, the photos may be used to texture the model.



WO 2009/146044 A1



ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,  
MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR),  
OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,  
MR, NE, SN, TD, TG).

**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))*

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

## USING PHOTO COLLECTIONS FOR THREE DIMENSIONAL MODELING

## BACKGROUND

[0001] Recently it has become possible to take collections of digital photos of a real world subject or scene and algorithmically reconstruct related information in a three-dimensional space. Such reconstructed information may relate to where a camera was situated relative to the subject or scene when a photo was taken, what the focal length of a photo is relative to its three-dimensional camera position, or where feature points of the subject or scene, as derived from the photos, are located in the three-dimensional space.

[0002] While this set of information is independently useful, such information has not been leveraged in the area of three-dimensional modeling. In particular, overlapping photos and three-dimensional information reconstructed from them have not been used to help rapidly generate and edit the geometry and texture of a three-dimensional mesh model of the subject or scene.

[0003] Techniques related to improved three-dimensional modeling and texturing are discussed below.

## SUMMARY

[0004] The following summary is included only to introduce some concepts discussed in the Detailed Description below. This summary is not comprehensive and is not intended to delineate the scope of the claimed subject matter, which is set forth by the claims presented at the end.

[0005] As discussed below, a collection of photos and a three-dimensional reconstruction of the photos may be used to construct and texture a mesh model. In one embodiment, a first digital image of a first view of a real world scene is analyzed to identify lines in the first view. Among the lines, parallel lines are identified. A three-dimensional vanishing direction in a three-dimensional space is determined

based on the parallel lines and an orientation of the digital image in the three-dimensional space. A plane is automatically generated by fitting the plane to the vanishing direction. A rendering of a three-dimensional model with the plane is displayed. Three-dimensional points corresponding to features common to the photos may be used to constrain the plane. The photos may be projected onto the model to provide visual feedback when editing the plane. Furthermore, the photos may be used to texture the model.

[0006] Many of the attendant features will be explained below with reference to the following detailed description considered in connection with the accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0007] The present description will be better understood from the following detailed description read in light of the accompanying drawings, wherein like reference numerals are used to designate like parts in the accompanying description.

[0008] Figure 1 shows a collection of photos.

[0009] Figure 2 shows a system for generating a model from a collection of photos.

[0010] Figure 3 shows photos and viewpoints of virtual cameras of a real world scene or subject.

[0011] Figure 4 shows photos sharing a common feature.

[0012] Figure 5 shows a plane fitting process.

[0013] Figure 6 shows parallel lines identified in a photo.

[0014] Figure 7 shows an overhead view of projections of photos onto a three-dimensional model.

[0015] Figure 8 shows a general process for using projections of photos onto a model to facilitate manual adjustment of a plane.

[0016] Figure 9 shows two renderings of a model with photos projected thereon.

[0017] Figure 10 shows a process for constructing a texture for a model based on multiple photos of the model.

[0018] Figure 11 shows photos with strokes identifying portions of the photos that are to be included in or excluded from a texture.

#### DETAILED DESCRIPTION

#### OVERVIEW

[0019] Embodiments discussed below relate to using digital photos of a subject, and three-dimensional information derived from the photos, to facilitate generating and editing the geometry and/or textures of a three-dimensional model of the real world subject in the photos. The following description will begin with discussion of the three-dimensional information, which may include postures (positions and orientations) of virtual cameras corresponding to postures of cameras that took the respective photos, as well as focal distances of the photos relative to their virtual cameras. The three-dimensional information may also include a cloud of three-dimensional points representing reconstructed locations of features of the photographic subject that are common to two or more of the photos. The description will then cover techniques for using this information to build the geometry, including fitting a plane of the model to one or more of the points in the point cloud and/or one or more directions of respective vanishing points derived from a photo. The description will then discuss techniques for constructing textures for the model using the photos and the three-dimensional information thereof.

## SYNTHESIS OF PHOTO COLLECTION

[0020] Figure 1 shows a collection of photos 100. The photos 100 are of a subject or scene, in this case, a shed. The photos 100 were taken from different positions relative to the shed, as seen in the differences between the photos 100. Some of the photos overlap. That is to say, parts of the shed in one photo are also in another photo. For example, overlapping photos 102 and 104 both show portions of a door. The collection of photos 100 may be taken from a same camera or from different cameras. The photos 100 might also be video frames captured from a video camera. Because of the overlap between photos, it is possible to reconstruct some basic three-dimensional information about the photos 100 and the shed.

[0021] Figure 2 shows a system for generating a model from a collection of photos. The system starts with the collection of photos 100. A photosynthesizer 120 (or synther) performs a process 122 for synthesizing the photos 100 to obtain three-dimensional information about the photos 100. More specifically, the photo synthesizer 120 receives the photos 100 and analyzes the photos to extract features of the real world subject in the photos 100. Features common to overlapping photos are identified. The features common to overlapping photos are used to determine camera postures for respective photos. These will be referred to as viewpoints or virtual cameras, which together are sometimes referred to as a "synth". A photo's virtual camera preferably lies on a line normal to the photo (preferably from the center of the photo) and at a distance known as the focal length. For an example, see Figure 3, showing a plurality of photos 100 of subject 150 and corresponding reconstructed three-dimensionally arranged virtual cameras 152, which in practice are stored as corresponding data structures or objects in computer memory or storage. More specifically, the photosynthesizer 120 reconstructs the positions and directions from which the photos were captured.

[0022] The photo synthesizer 120 also determines a sparse point cloud comprised of points in three-dimensional space, where the points are an estimation of the location in three-dimensional space of features of the subject found in the photos. A point may be found when computing the virtual cameras. More specifically, referring to Figure 4, this may involve identifying a feature 140A of the photographic subject in a photo 141, looking for the same feature 140B in another photo 142, and if it is found, simultaneously estimating the three-dimensional camera positions of the photos and a position of point 140C (an example of a point in the point cloud) such that when the point 140C is re-projected back into the corresponding photo's cameras it re-projects to points 140A, 140B where the feature is found in the respective photos 141, 142. Details on deriving a point cloud of features and postures of virtual cameras from a photo collection may be found elsewhere.

[0023] Referring back to Figure 2, the photosynthesizer 120 outputs the point cloud and camera poses/postures 122, which are made available to a model renderer/editor 124. The model renderer/editor 124 has one or more software modules for performing a process 126, which may include operations such as: maintaining and rendering a three-dimensional model; using points in the point cloud and/or vanishing points in the photos as constraints for fitting planes of the three-dimensional mesh model; interactively altering the geometry of the model and in particular orienting and translating planes of the model (perhaps as informed by the point clouds, photos, and virtual cameras); displaying photos for the purpose of selecting constraint points in the point cloud; marking the photos to define patches for generating corresponding textures for the model; etc. Details of these operations will be described in detail further below. Finally, the model renderer/editor 124 outputs a mesh model 128 which may also include textures for the model.

[0024] As mentioned above, Figure 3 shows photos 100 and viewpoints of virtual cameras 152 of a real world scene or subject 150. Though not shown, each photo 100 has image data of the subject 150, taken from varying perspectives and distances. As in Figure 1, the photos might be of a shed or some other building with planar surfaces. The virtual cameras 152 are three-dimensional positions and directions from which the subject 150 was photographed. For clarity, the pyramids in Figure 3 are shown distant from the subject 150, however, in practice a pyramid would be arranged such a point of its photo 100 would lie on or near a hypothetical reconstruction of the subject 150.

#### PLANE FITTING

[0025] Figure 5 shows a plane fitting process. The plane fitting process uses the photos 100 and point cloud 122 to constrain a fit of a plane of the three-dimensional model. The process starts by accessing 172 the photos, point cloud, and virtual cameras. One or more points from the point cloud are selected 172. A point may be selected by a variety of means. In one embodiment, one of the photos is displayed in two-dimensions, perhaps including re-projected points that happen to fall on or intersect the photo. A user selects a region of the photo containing re-projected points that the user deems to correspond to the plane that is being defined. The selection may involve a drag operation or interactively sketching a region of interest. A set of points may also be selected 174 automatically, for example, by a process that analyzes a photo for likely locations of a plane (e.g., regions between parallel lines, regions with a same estimated focal depth, etc.). Whether selected automatically or manually, one or more of the points may be selected 174 and used as an initial basis for constraining the plane.

[0026] Next, the photo is analyzed 176 to identify parallel lines and vanishing directions. Note that this analysis 176 may be done in advance during a



pre-processing stage during which each image is analyzed to find two-dimensional vanishing points, and from the two-dimensional vanishing points, three-dimensional vanishing directions are found (in a manner similar to that of finding three-dimensional feature points from two-dimensional feature points in the photos). As shown in Figure 6, known techniques can be used to find lines 190 of a photo's subject/scene that are parallel (as part of the photo's subject/scene) and converge on a vanishing point relative to the photo. Groups of parallel lines 190 may be found by choosing one line and repeatedly comparing it to other lines in the photo, thereby obtaining clusters of parallel lines that converge on respective vanishing points. In Figure 6, a vertical cluster is shown with solid lines, and a horizontal cluster is shown with dashed lines. The parallel lines in the photo can be used to derive a vanishing direction in three dimensions, a vanishing direction being a three-dimensional direction independent of any three-dimensional location. Note that the photo is three-dimensionally situated according to its virtual camera. Therefore, a photo's three-dimensional orientation (its normal projecting to its virtual camera) can be used to determine the three-dimensional direction of its vanishing point. Furthermore, by the same means, different vanishing directions of other photos may also be computed, thereby giving a set of potential vanishing directions in the locale of the plane that is to be fitted.

[0027] Given a vanishing direction and/or a point of the point cloud, a plane is fitted 176 as constrained by the vanishing direction and/or the point. If no other constraints are provided, the plane would ordinarily be constrained with at least one the vanishing direction and two three-dimensional points, or one point and an additional vanishing direction. Additionally or alternatively, two different vanishing directions can be used to completely constrain the orientation of the plane (without using a point in the point cloud). Depending on how many vanishing directions can be recovered for the plane of interest, there may be 1, 2 or 3 degrees of freedom in choosing the optimal plane. Thus the plane fit 176 may involve sampling different

orientations for the plane, using 1, 2, or 3 degree-of-freedom solutions. Three random points (of the relevant points in the point cloud) may be used to constrain the plane. One vanishing direction and two points may be used to compute the plane. Or, two vanishing directions and one point may be used to compute the plane. In other embodiments, only one or more points may be used, or only one or more vanishing directions may be used to constrain the fit 176 of the plane.

[0028] Regardless of what combination of points and/or vanishing directions are used to compute the plane, different candidate points and/or vanishing directions may be repeatedly tested to search for an optimal plane fit. Candidates can be measured by factors such as the complexity of the resulting model or how well the resulting model matches various of the photos. A photo based evaluation might involve, for example, projecting photos onto the resulting model, computing the average pixel color on the surface, and then evaluating how well that agrees with each image. A feature-based approximation could also be used to evaluate a potential plane fit. The photosynthesizer 120 might generate correspondences between two-dimensional interest points in the images as well as an estimated three-dimensional location for the point. For each candidate plane, there are a number of points associated with it and constrained to lie on the plane. The sum of the re-projection errors of these constrained 3D points can be used to evaluate the correctness of the model according to the candidate plane. This can be implemented by orthogonally projecting the original estimated three-dimensional point onto the plane to get the constrained three-dimensional location or by actually solving for the constrained position that minimizes the re-projection error.

[0029] As mentioned, one embodiment may involve fitting 176 a plane based only on one or more vanishing directions. Given some vanishing directions (three-dimensional directions), it can be assumed that the plane being fitted should be parallel to one direction and should be parallel to another direction (the directions possibly coming from different photos per their respective cameras), and then the

two vanishing directions will allow computation of the normal of the plane. From there, a depth (location of the plane along the normal) may be estimated using one or more points from the point cloud. The points, perhaps selected interactively from one of the photos, can then be used to snap the plane to a position that one or more points best agree with. Depths of different points can be tested (see the reconstruction tests discussed above) to find a preferred constraint point. In one embodiment, the plane depth that gives the most number of agreeing points is chosen; points that agree with that depth are used to perform a least squares optimization to find a final optimal depth.

#### PLANE ADJUSTMENT

[0030] Having computed a plane for the model under construction, there may be errors in the orientation or placement of the plane. An embodiment may be provided to allow a user to fine-tune a plane of the model. Figure 7 shows an overhead view of projections of photos 200 onto a three-dimensional model 201. The projections that can be used to facilitate manual manipulation of the geometry of the model, and in particular, the depth (location) and/or orientation of a plane of the model. The discussion of Figure 8 will cover how the projections of Figure 7 may be used.

[0031] In Figure 7, two or more photos 200 are oriented in a virtual three-dimensional space according to their respective viewpoints or virtual cameras 202. In a virtual three-dimensional space, the virtual cameras 202 have a position and orientation, and the respective photos 200 are shown according to their focal distance from their virtual cameras 202. The photos contain image data of a real world subject, such as a building 204. As discussed above, a vanishing point 206 (and corresponding vanishing direction) may be computed from a photo 200. In the example of Figure 7, each photo 200 contains a same feature 205 of the building

204. In this example, the feature 205 is a corner of building 204. A corresponding feature point 208 of a point cloud is also shown. The feature point 208 is a three-dimensional point where features 205 intersect when projected from their respective virtual cameras 202 and photos 200.

[0032] Figure 7 also shows a top-down two-dimensional view of the model 201, which is a simple three-dimensional mesh model for which a plane 210 is being fitted. The model 201 is a model corresponding to the building 204. Note that the vanishing direction 206 is also shown near model 201. Vanishing direction 206 is a three-dimensional direction derived from one of the photo's 200 parallel lines (vanishing point) and oriented according to the photo's virtual camera 202, or equivalently, according to the orientation and placement of the photo in the three-dimensional space. As will be discussed with reference to Figure 8, the photos 200 can be projected onto the model 201 by computing where pixels of the photos 200, as projected from their virtual cameras 202, fall on the model. Overlap of the projected pixels can be used to help a user align and orient the plane 210. The content of Figure 7 is only an example, and arbitrary photos 100/200, models 128/201, camera poses 122/cameras 202, and points 122/208 may be used in arbitrary three-dimensional arrangements.

[0033] Figure 8 shows a general process for using projections of photos 200 onto model 201 to facilitate manual adjustment of plane 208. Initially, photos 200 that relate to the plane are selected 250. This may be done automatically by identifying photos whose projected pyramids intersect plane 210. Selecting 250 may also be done manually, possibly by choosing from among photos that have some minimal projection onto the plane. The selected 250 photos 200 are then projected 252 onto the model 201 and plane 210. Projecting 252 can be conceptually thought of as similar to an optical slide projector projecting a photographic slide (i.e., image 200), with the light of the projector (i.e., camera 202) projecting out as a pyramid and illuminating whatever it hits (model 201) with the projected image of the slide

(photo 200). As seen in image 280 of Figure 9, ghosting occurs when multiple photos 200 are projected onto the model 201 while the plane 210 is not positioned or oriented in close correspondence with its "true" orientation/position (a position/orientation mirroring the corresponding real world plane of real world subject 204). This ghosting effect is a result of parallax; the plane-induced shift of a projected point away from where it would be if the plane were at its true orientation/position.

[0034] To explain further, consider a single photo 200 which contains an image of a particular view of the building 204 and a wall that plane 210 represents. The virtual orientation and location of the photo's 200 virtual camera 202 is comparable to the real location and orientation of the actual camera that took the photo 200. If the plane 210 is arranged with a location and orientation that corresponds to the wall it represents, then the projection of the photo 200 will match the plane 210. In other words, the lines and features (e.g. corners, edges, borders) of the wall in photo 200, when projected from its virtual camera 202, will align with fall upon any corresponding lines and features of the plane 210. When the plane 210 is not arranged to match the arrangement of the wall it represents, images of the wall projected onto the plane 210 will create shadows, ghosts, and other artifacts. See Figure 9, which will be discussed later.

[0035] Returning to Figure 8, the model 201 and plane 210, with the photos 200 projected 252 thereon, are rendered 254 and displayed. The user manually adjusts 256 the plane 210, for example, by using a mouse to adjust the orientation and/or the location of the plane. In particular, the depth of the plane 210 may be adjusted. This involves the user translating the plane 210 along its normal. Furthermore, when adjusting 256 the plane 210, the plane 210 can be automatically snapped to various positions and/or orientations, for example, based on the point cloud or constraints derived from the point cloud. Plane adjustment 256 may also involve revising the border or perimeter 284 of the surface defined by the plane 210.

[0036] The projecting 252, rendering 254, and adjusting 256 may be repeated until the user, according to the visual feedback of the rendering 254 and displaying, has finally arranged the plane 210 to minimize the visual artifacts and accurately model the wall that the plane 210 represents. At this point, the user might also select one or more photos for texturing the plane 210.

[0037] Figure 9 shows two renderings 280, 282 of the model 201 with photos 200 projected 252 thereon. The various photos 200 are shown individually, in two-dimensions, on the right hand side of the renderings 280, 282. In the rendering 280, plane 210 of the model 201 is out of alignment; its depth (or orientation) is incorrect. Consequently, the photos 200 simultaneously projected 252 thereon appear out of focus as a group. As the user adjusts 256 the depth of the plane 210 (or orientation), the parallax of each projected 252 photo 200 is reduced until the displayed rendering 282 shows few collective artifacts on the plane 210. Naturally, due to parallax, projected portions of the photos 200 that are not part of the wall/plane may continue to exhibit artifacts and parallax.

[0038] While rendering of photo projections onto a model can reflect the accuracy of the model, other techniques can also be used to provide visual feedback to the user. For example, the plane, textured by the projected photos, can be analyzed for artifacts. Parallax of feature points, such as point 208, as projected onto the model/plane can be measured and used as a basis to provide visual feedback. A graph might be displayed indicating how closely the plane agrees with the relevant points of the point cloud. Furthermore, a user may be provided with means to control which photo(s) are currently projected, thereby enabling fine-tuning the plane to a particular photo.

## TEXTURE GENERATION

[0039] Figure 10 shows a process for constructing a texture for a model based on multiple photos of the model. While it may be possible for some planes of a model to be textured by only one photo (from among many possible photos), it will often be desirable to create a texture for a plane using different patches or portions from different photos. For example, one photo, from one perspective, may show a left half of a wall unoccluded, while another photo of the wall, due to its different perspective, shows the left half of the wall occluded by a bush or tree, for example. The process of Figure 10 allows portions of photos to be stitched together to create a texture image for the model.

[0040] The process may begin by displaying 300 a photo. The photo may be one of several that correspond to the plane being textured. User input or other information is received that identifies 302 one or more regions to exclude or include in the texture. Referring to Figure 11, such input might include strokes 320 to identify portions of a photo that are to be included in the texture. The input might also include strokes 322 that mark underlying patches or regions as portions to be omitted from the texture. While some parts of photos may be identified 302 manually, a patch or portion may also be automatically identified 302 for inclusion or exclusion automatically, based, for example, on its image qualities or how well it matches nearby textures of other parts of the model. Some image portions may be automatically identified as occlusions (e.g., trees) using corresponding points of the point cloud. The points may indicate such image portions are not near the relevant plane. Some occlusions might be identified using an image recognition algorithm to recognize types of objects not typically part of a building, for example.

[0041] Once various regions of photos have been identified 302 for inclusion/inclusion in the texture, the regions are extracted 304 and merged 306. Any number of known algorithms can be used for generating an image from

different regions or patches of photos of a same subject. Such algorithms typically perform color and tone balancing, and may identify seams or other areas where patches may be joined with minimal visual artifacts.

## CONCLUSION

[0042] Embodiments and features discussed above can be realized in the form of information stored in volatile or non-volatile computer or device readable media. This is deemed to include at least media such as optical storage (e.g., CD-ROM), magnetic media, flash ROM, or any current or future means of storing digital information. The stored information can be in the form of machine executable instructions (e.g., compiled executable binary code), source code, bytecode, or any other information that can be used to enable or configure computing devices to perform the various embodiments discussed above. This is also deemed to include at least volatile memory such as RAM and/or virtual memory storing information such as CPU instructions during execution of a program carrying out an embodiment, as well as non-volatile media storing information that allows a program or executable to be loaded and executed. The embodiments and featured can be performed on any type of computing device, including portable devices, workstations, servers, mobile wireless devices, and so on.



## CLAIMS

1. A computer implemented method of fitting a plane for a three-dimensional model, the model comprising vertices interconnected by line segments, the plane and model being represented by data in computer memory, the method comprising:
  - accessing: a collection of digital photos and a cloud of points, where the photos are of a real world scene, where the photos are three-dimensionally arranged in a three-dimensional space containing the three-dimensional model, where the points are also arranged in the three-dimensional space, and where each point in the cloud corresponds to a feature of the scene that exists in at least two of the photos;
  - receiving information identifying a plurality of points corresponding to a first photo in the collection of photos;
  - obtaining a vanishing direction in the three-dimensional space by analyzing the first photo to identify a corresponding vanishing point of the scene in the first photo, and using the vanishing point and the orientation of the photo in the three-dimensional space to compute the vanishing direction;
  - arranging the plane in the three-dimensional space by fitting the plane to at least the vanishing direction and/or at least one of the points in the plurality of points; and
  - displaying a rendering of the model according to the fitted plane.
2. A computer implemented method according to claim 1, further comprising:
  - displaying the first photo and superimposing onto the first photo points that correspond to points in the three-dimensional model, including the plurality of points; and
  - generating the information identifying the plurality of points based on user interaction with the displayed first photo and superimposed points.

3. A computer implemented method according to claim 1, further comprising analyzing a second of the photos, identifying lines therein, computing directions of the lines based on the orientation of the photo in the three-dimensional space, and if the direction of a line corresponds to the vanishing direction computed from the first photo, then using such line to perform the fitting of the plane or to fit another plane.
4. A computer implemented method according to claim 1, wherein the vanishing point of the first photo is found by analyzing near-parallel lines in the first photo.
5. A computer implemented method according to claim 1, the process further comprising displaying one of the digital photos, receiving strokes directed to and displayed on the one of the digital photos, and using the strokes to identify parts of the one of the digital photos to use in a texture for the plane.
6. A computer implemented method according to claim 1, the process further comprising generating a texture image for the plane by using points in the cloud of points to automatically identify a portion of the first photo to include or omit in the texture image.
7. One or more volatile and/or non-volatile computer readable media storing information to enable a computer to perform a process, the process comprising:
  - analyzing a first digital image of a first view of a real world scene to identify lines in the first view of the scene, and identifying among the lines a set of lines that are determined to be parallel in the first view of the scene;
  - determining a three-dimensional vanishing direction in a three-dimensional space based on the parallel lines and an orientation of the digital image in the three-dimensional space;
  - automatically fitting a plane of a model in the three-dimensional space by constraining the plane to the vanishing direction; and

displaying a rendering of a three-dimensional model in the three-dimensional space, where the model includes the fitted plane.

8. One or more computer readable media according to claim 7, the process further comprising:

accessing a plurality of points in the three-dimensional space, where each point comprises a respective feature in the scene that is common to a group of corresponding digital images in the three-dimensional space;

selecting one of the plurality of points, where the selected point comprises an intersection, in the three-dimensional space, of a feature in the first view of the real world scene and the same feature in a second digital image comprising a second view of the real world scene; and

using the point to further constrain the plane.

9. One or more computer readable media according to claim 8, further comprising displaying the first digital image, allowing a user to interactively define a region of interest in the first digital image, and the selecting the one of a plurality of points comprises selecting a point that maps to the region of interest defined by the user.

10. One or more computer readable media according to claim 9, wherein the region of interest comprises a polygon drawn by the user on the display of the first digital image, and where a side of the polygon is automatically aligned with the set of lines.

11. One or more computer readable media according to claim 7, wherein the set of lines is computed by testing different pairs of lines in the first view, which groups lines into different sets of lines that are parallel in the real world scene.

12. One or more computer readable media according to claim 7, the process further comprising:

repeatedly re-projecting a plurality of three-dimensionally arranged digital images of the real world scene onto the model while the orientation and/or position of the plane is repeatedly modified.

13. One or more computer readable media according to claim 7, the process further comprising translating the plane along the normal of the plane in accordance with user input.

14. A computer implemented method for interactively editing a three-dimensional model, the method comprising:

providing a plurality of virtual cameras oriented and positioned in three dimensions relative to the three-dimensional model, each virtual camera having a corresponding digital photo of a real world scene corresponding the three-dimensional model;

displaying a rendering of the three-dimensional model and a plane of the three-dimensional model while a user interactively makes changes to the orientation and/or position of the plane relative to the three-dimensional model;

as the orientation and/or position of the plane is changed relative to three-dimensional model, re-computing projections of the digital photos from their respective virtual cameras onto the three-dimensional model and the plane thereof; and

displaying visual indicia based on the computed projections of the digital photos relative to the plane.

15. A computer implemented method according to claim 14, wherein the displaying visual indicia comprises rendering and simultaneously displaying the digital images as projected onto the plane.

16. A computer implemented method according to claim 14, wherein ghosting due to multiple photos being projected on the plane increases or decreases as the orientation and/or position of the plane changes.

17. A computer implemented method according to claim 14, further comprising creating a texture image for the plane by merging different portions of different digital photos into the texture image.

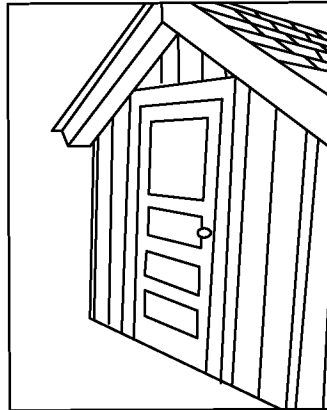
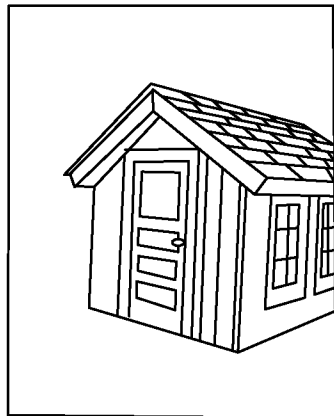
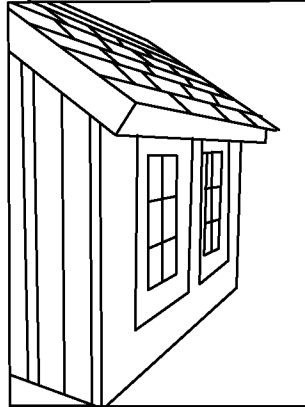
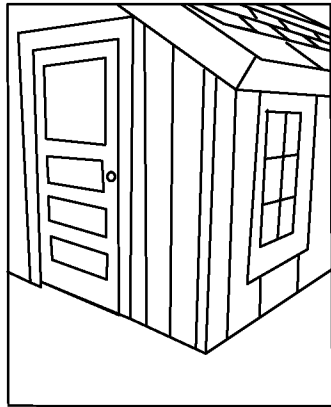
18. A computer implemented method according to claim 17, further comprising identifying the portions based on user input directed to a display of one of the digital photos.

19. A computer implemented method according to claim 14, wherein two digital photos contain image data corresponding to a same feature of the real world scene that is represented by the plane, the process further comprising creating a texture image for the plane by combining different portions of the two digital photos.

20. A computer implemented method according to claim 19, further comprising:  
providing a cloud of three-dimensional points that represent features common to groups of digital photos; and  
using points of the cloud to identify portions of the digital photos that are to be omitted from the texture image.

100

102



104

FIG. 1

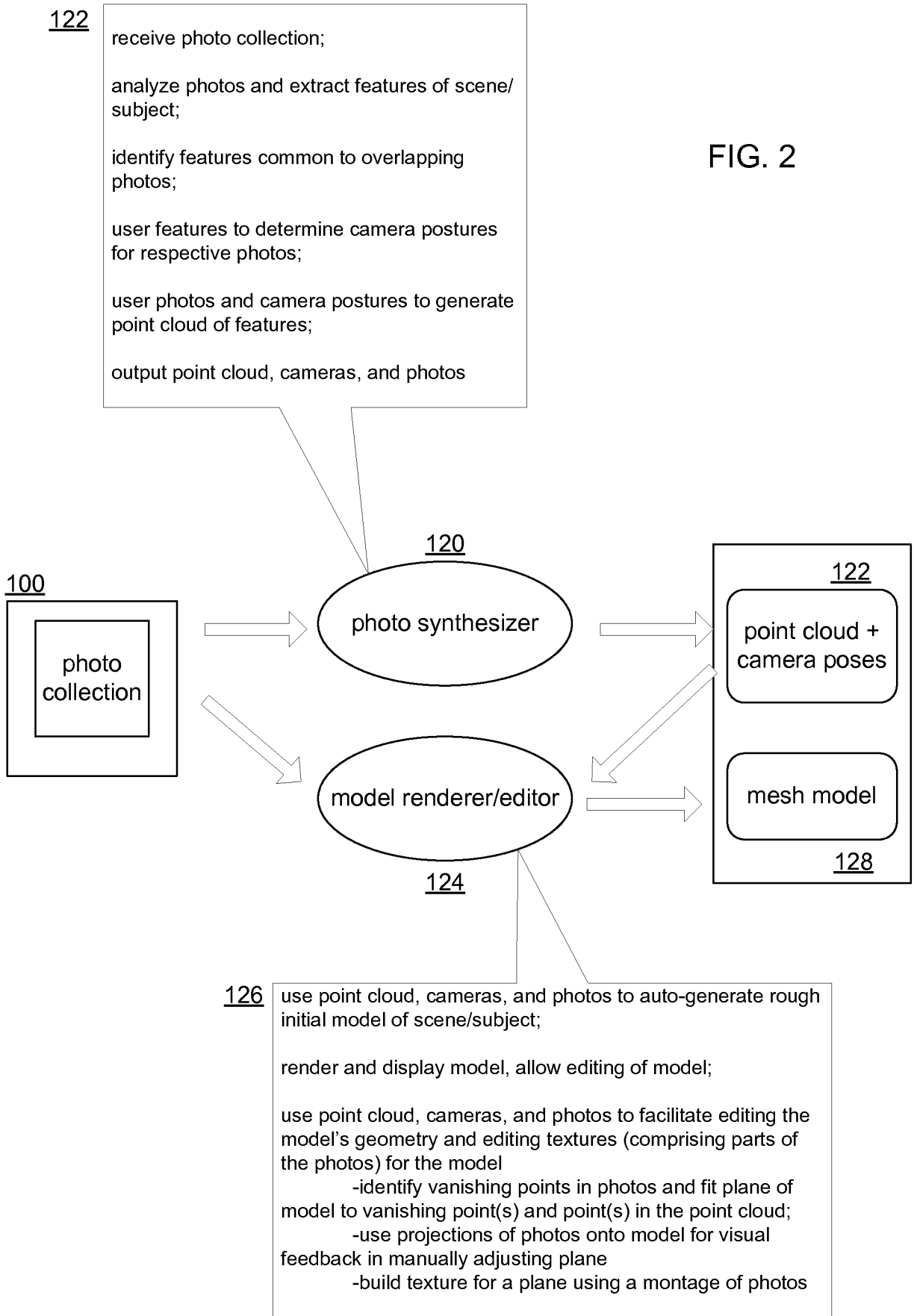


FIG. 2

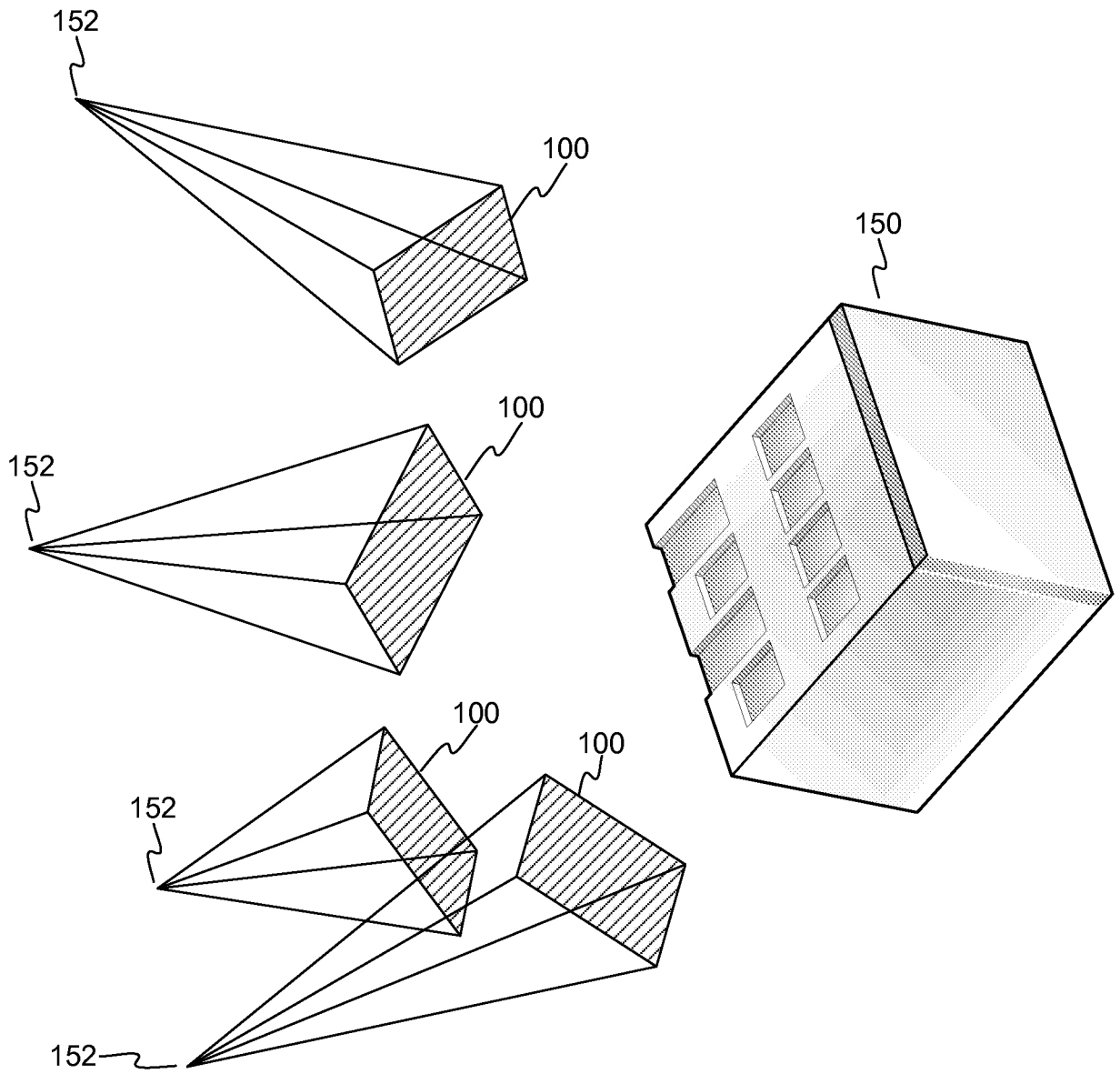


FIG. 3



4/11

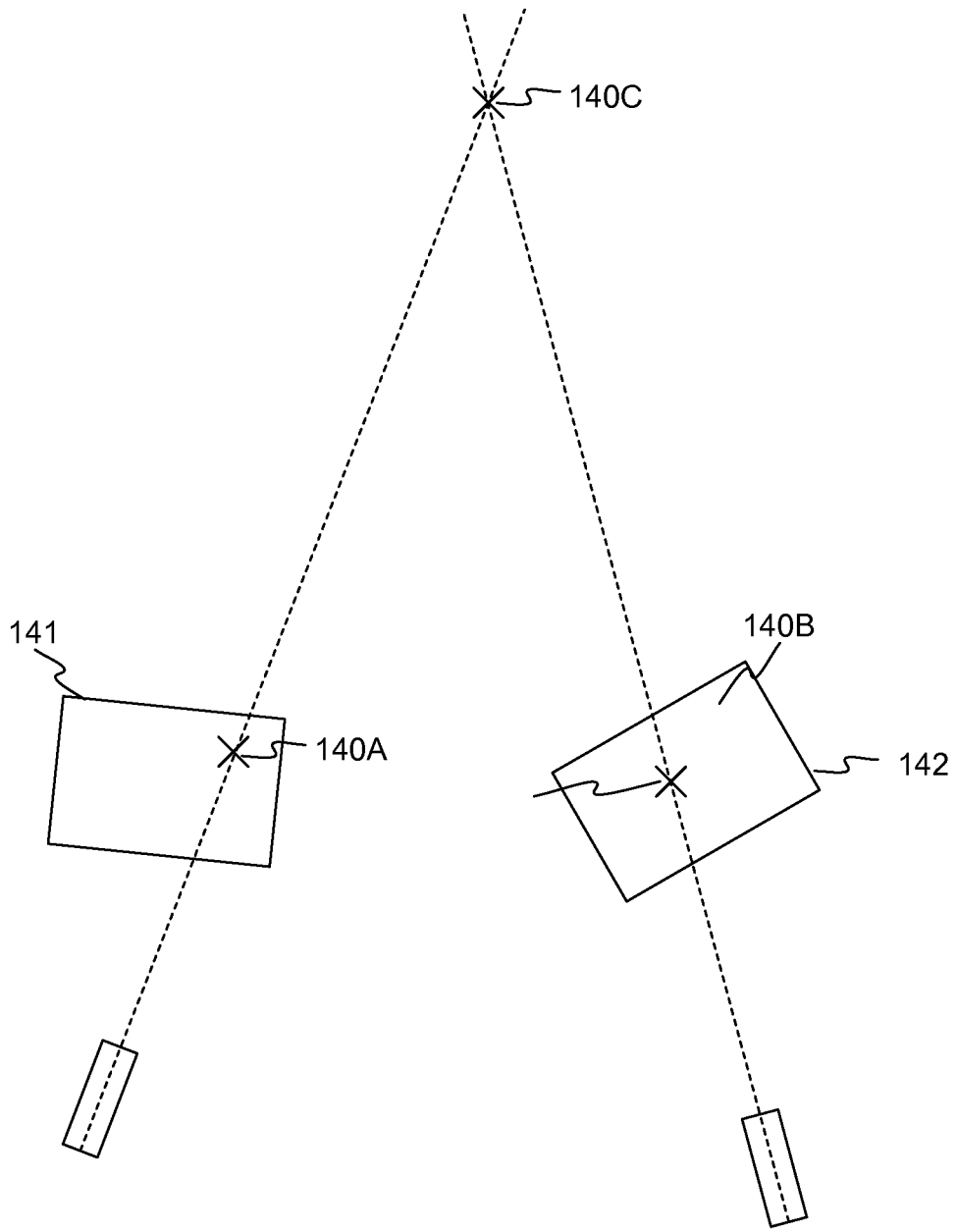


FIG. 4

FIG. 4

5/11

plane fitting

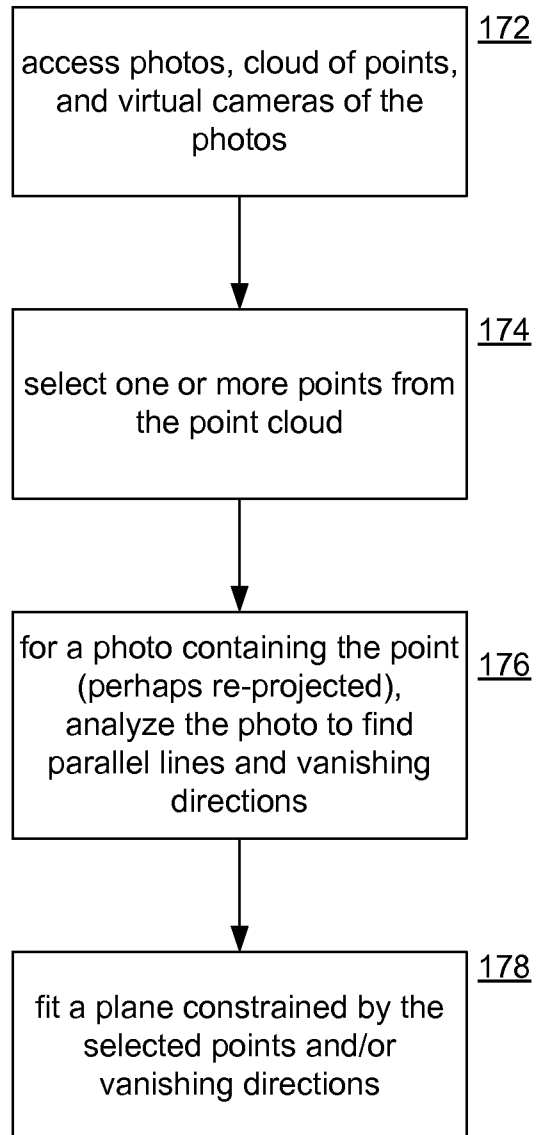


FIG. 5

6/11

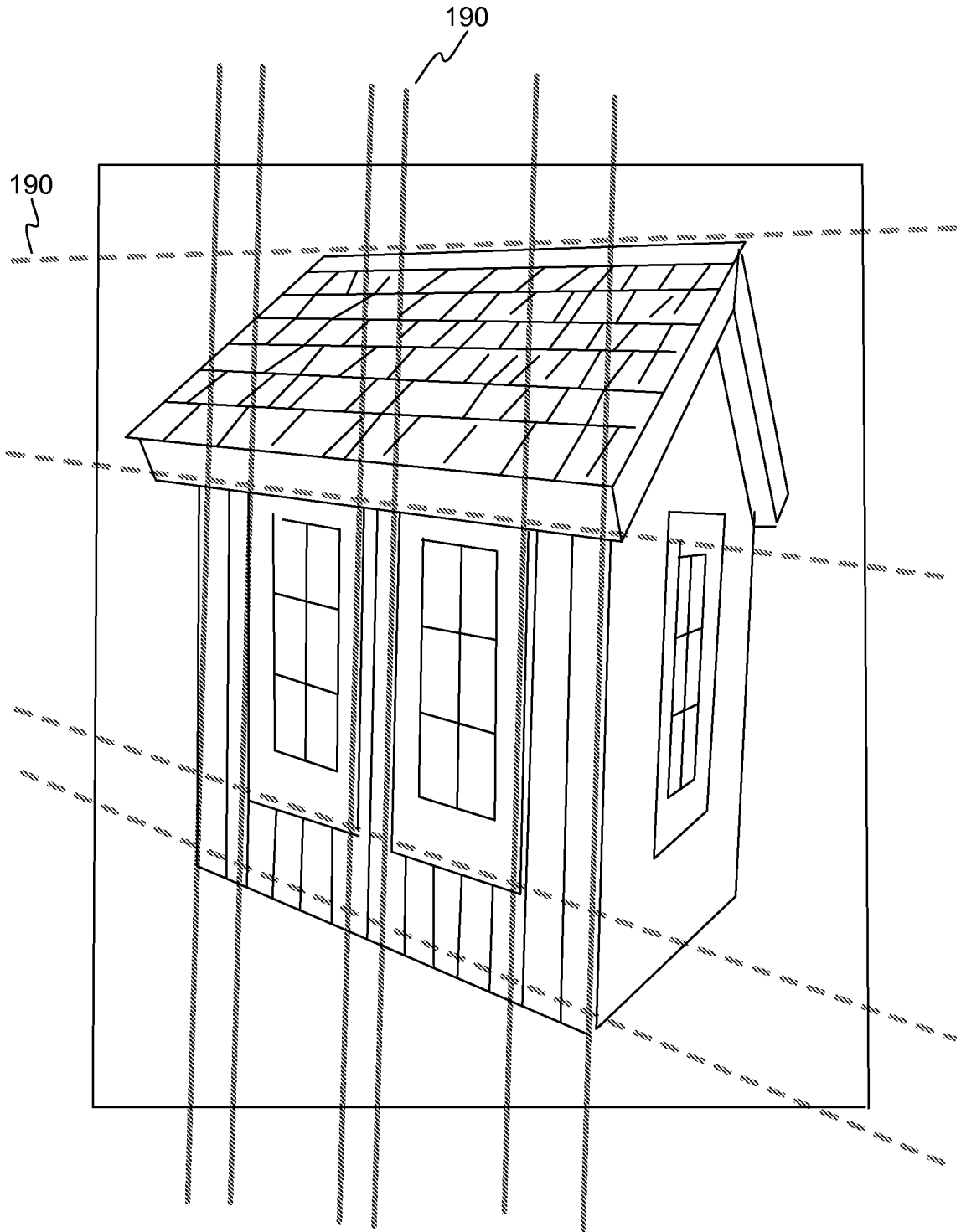


FIG. 6

7/11

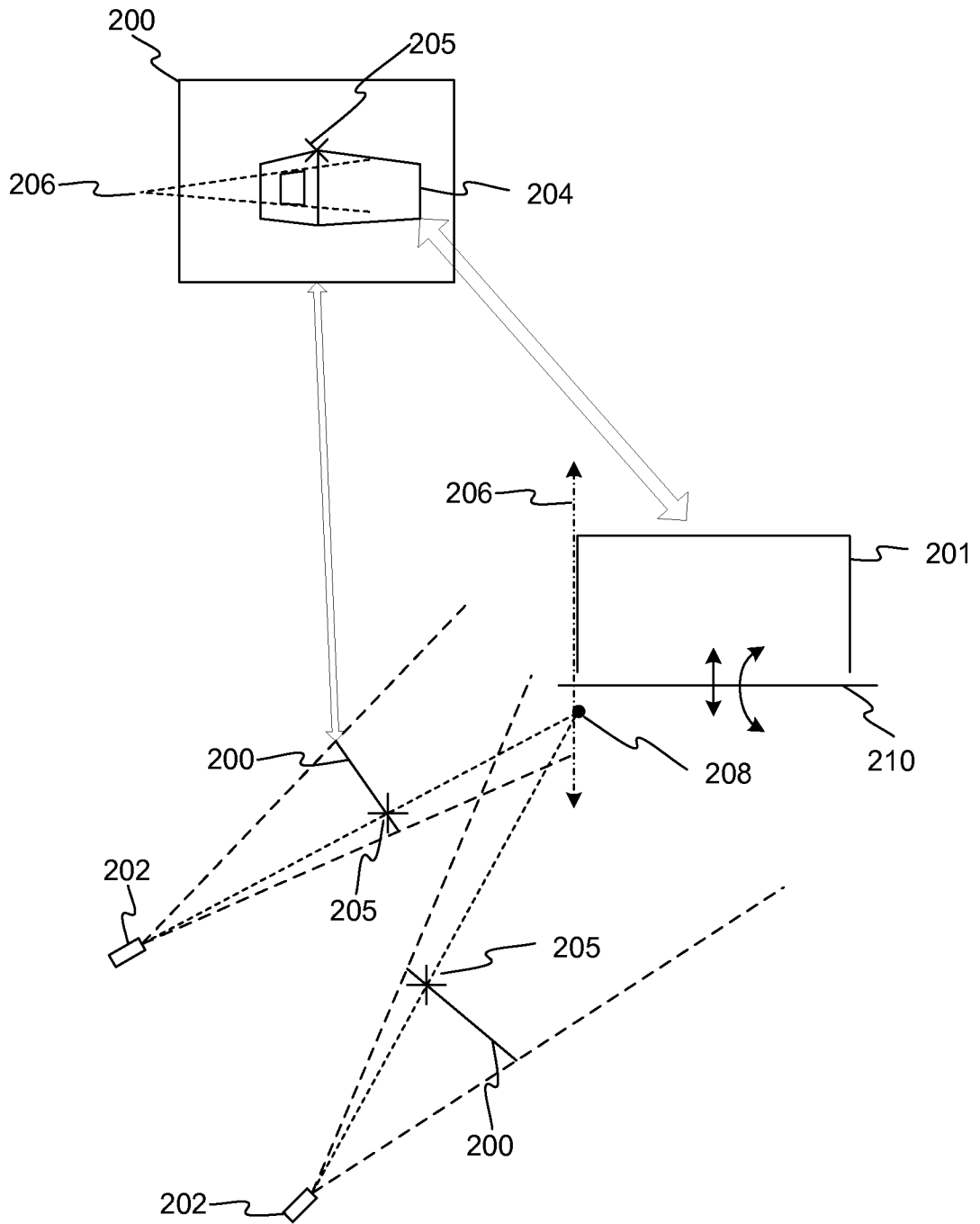


FIG. 7

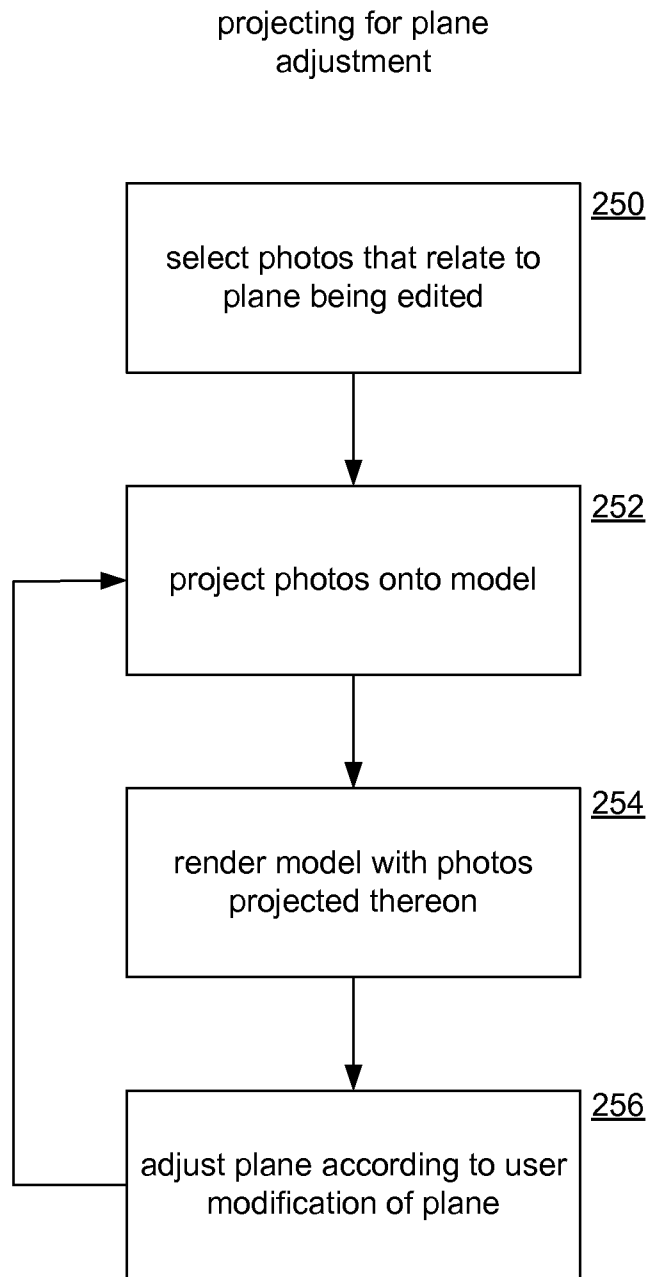


FIG. 8

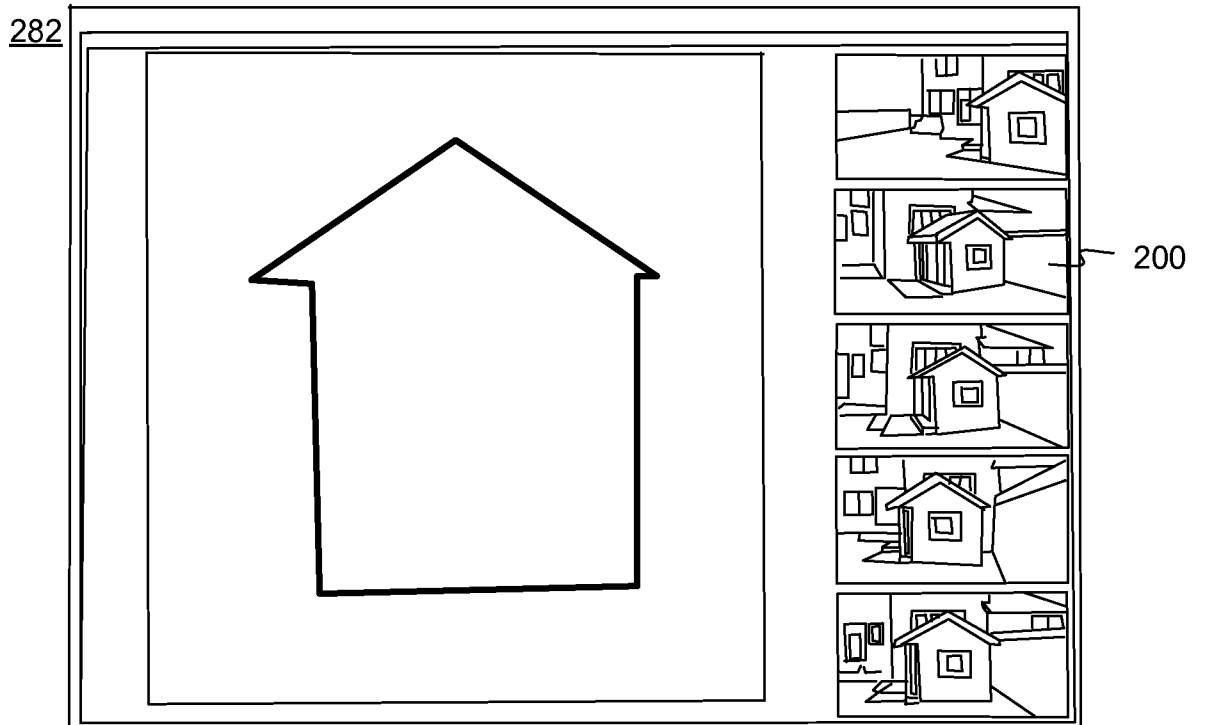
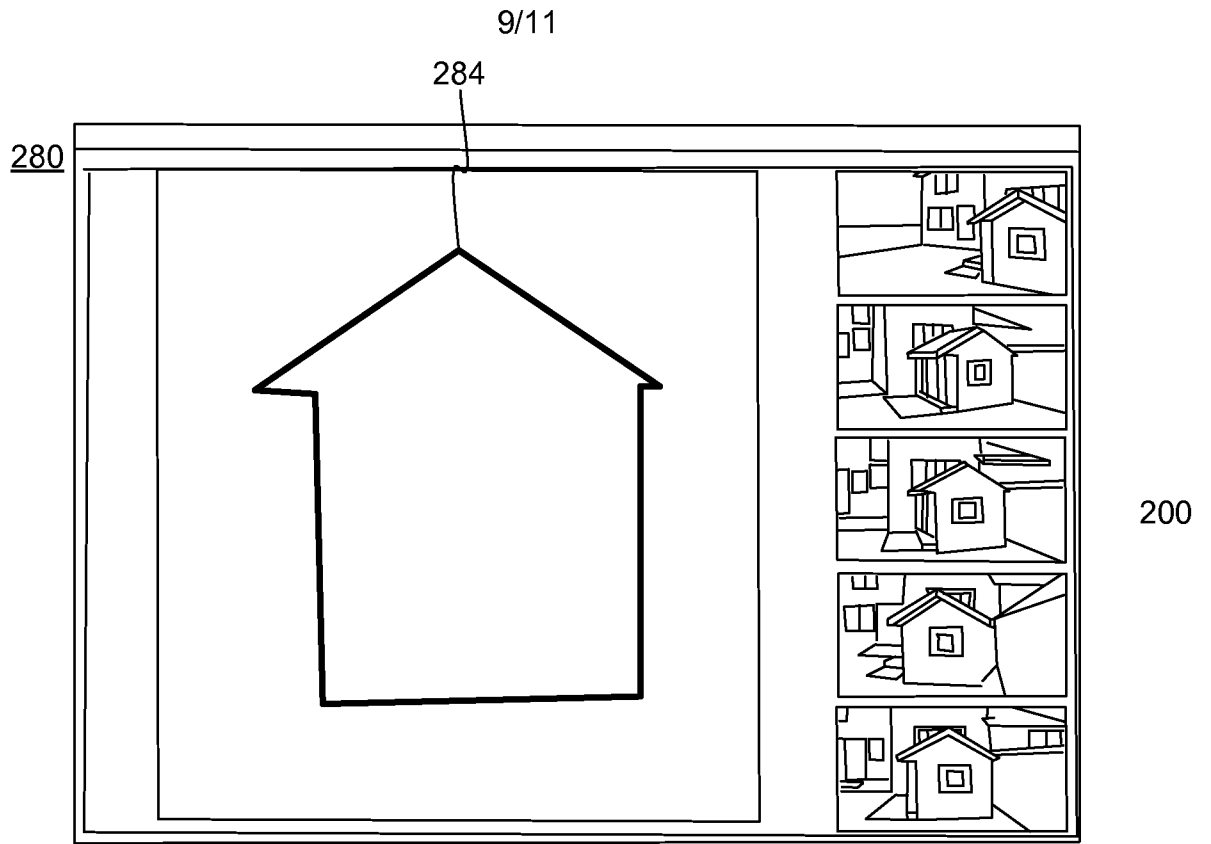


FIG. 9

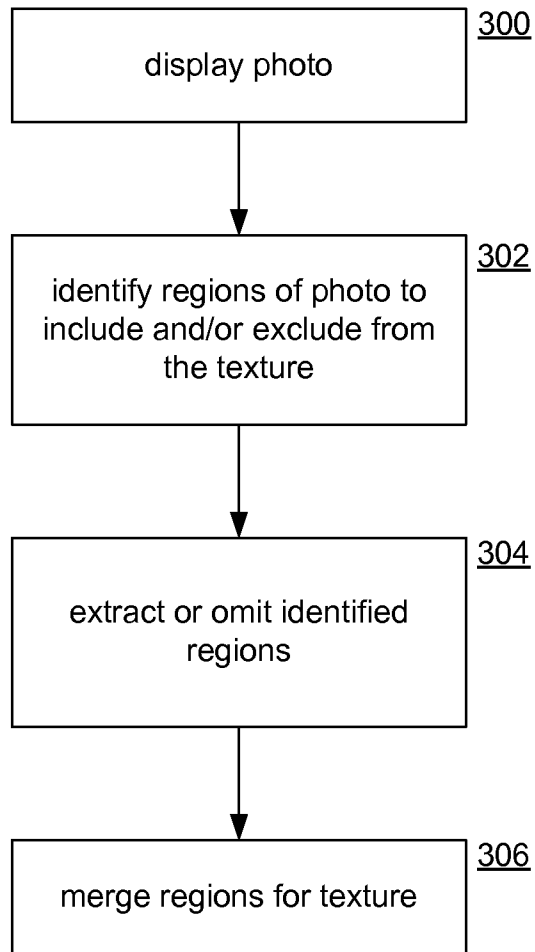


FIG. 10

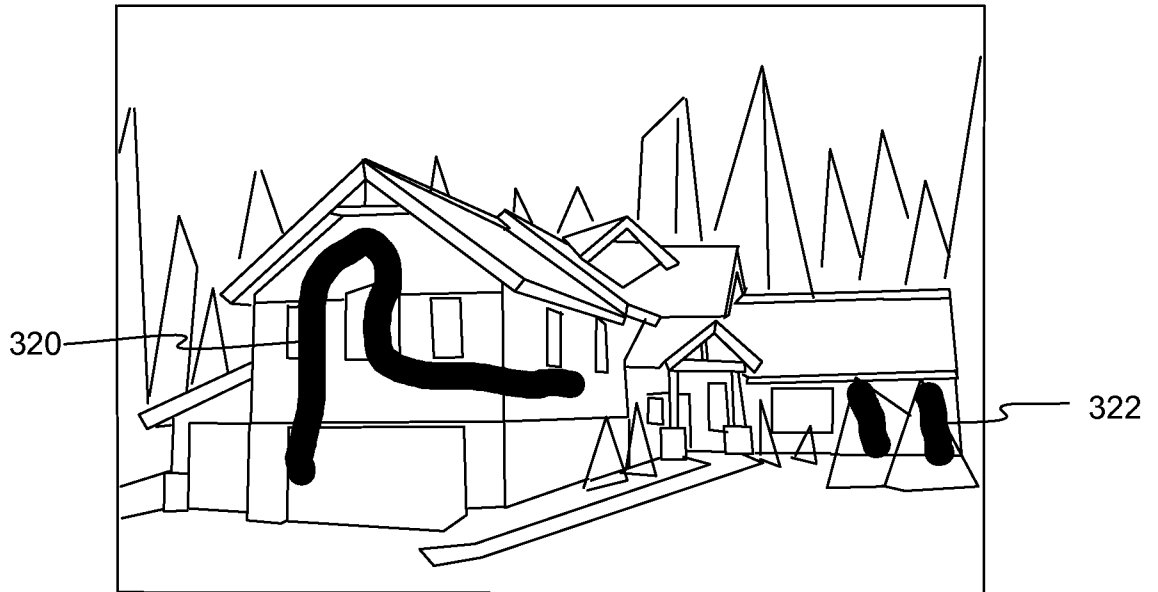


FIG. 11



**A. CLASSIFICATION OF SUBJECT MATTER****G06T 17/00(2006.01)i**

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)

IPC 8: G06T

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

Korean Utility models and applications for Utility Models : IPC as above

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

eKOMPASS (KIPO internal) &amp; keywords: vanishing direction, 3D, plane

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5990900 A1 (SEAN W. SEAGO) 23 November 1999 See abstract; claims 1 and 7; figure 2.	1-20
A	US 6124859 A1 (YOUICHI HORII et al.) 26 September 2000 See abstract; claims 1-8; figures 1A and 1B.	1-20
A	KR 10-2004-0043280 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 24 May 2004 See abstract; claim 1; figures 4-7.	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 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

"&amp;" document member of the same patent family

Date of the actual completion of the international search

09 NOVEMBER 2009 (09.11.2009)

Date of mailing of the international search report

**11 NOVEMBER 2009 (11.11.2009)**

Name and mailing address of the ISA/KR

Korean Intellectual Property Office  
Government Complex-Daejeon, 139 Seonsa-ro, Seo-  
gu, Daejeon 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Moon, Nam Doo

Telephone No. 82-42-481-5751



**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2009/039028**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5990900 A1	23.11.1999	None	
US 06124859 A1	26.09.2000	JP 10-049704 A	20.02.1998
KR 10-2004-0043280 A	24.05.2004	KR 10-0443552 B1	09.08.2004
		US 2004-0095385 A1	20.05.2004
		US 07174039 B2	06.02.2007