

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
30 July 2009 (30.07.2009)

PCT

(10) International Publication Number  
**WO 2009/093136 A2**

- (51) International Patent Classification:  
*G06T 15/20* (2006.01)      *G06T 17/40* (2006.01)
  - (21) International Application Number:  
PCT/IB2009/000119
  - (22) International Filing Date: 23 January 2009 (23.01.2009)
  - (25) Filing Language: English
  - (26) Publication Language: English
  - (30) Priority Data:  
0801297.3      24 January 2008 (24.01.2008)      GB
  - (71) Applicant (for all designated States except US): **ARE-OGRAPH LTD** [NZ/NZ]; 6 Fea Street, Dunedin, New Zealand (NZ).
  - (71) Applicant (for MG only): **EIP LIMITED** [GB/GB]; Fairfax House, 15 Fulwood Place, London WC1V 6HU (GB).
  - (72) Inventor; and
  - (75) Inventor/Applicant (for US only): **REID, Luke** [NZ/NZ]; 6 Fea Street, Dunedin, New Zealand (NZ).
  - (74) Agent: **SPAARGAREN, Jerome**; EIP, FairFax House, 15 Fulwood Place, London WC1V 6HU (GB).
  - (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, 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:**  
— without international search report and to be republished upon receipt of that report

(54) Title: IMAGE CAPTURE AND MOTION PICTURE GENERATION

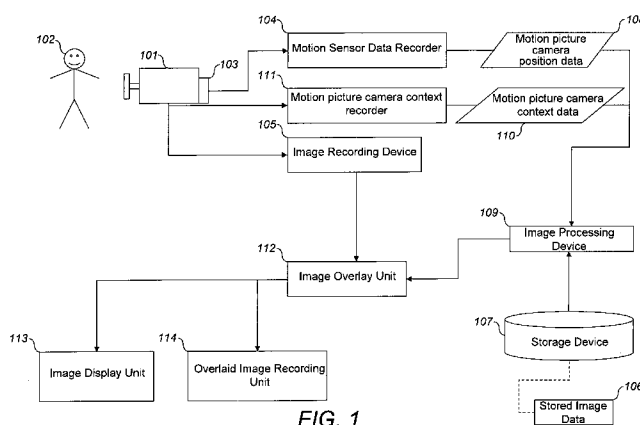


FIG. 1

(57) Abstract: A method of generating a moving image in the form of a series of playback frames, wherein said moving image is generated from: a first set of image data captured using a motion picture camera system having an associated movement capture system; movement capture data generated by said movement capture system; and a second set of scenic image data captured using a scene-scanning imaging system which provides image data from which images, representing viewpoints distributed across a scene through which a virtual viewer is capable of navigating, are derivable, the method comprising: on the basis of said movement capture data, selecting a first viewpoint position; deriving first scenic viewpoint image data from said second set of image data, based on the selection of said first viewpoint position; combining image data from said first image data with said first scenic viewpoint image data to generate a first playback frame; on the basis of said movement capture data, selecting a next viewpoint position from a plurality of potential next viewpoint positions distributed relative to the first viewpoint position across said scene; deriving second scenic viewpoint image data from said second set of image data, based on the selection of said next viewpoint position; and combining image data from said first image data with said second scenic viewpoint image data to generate a second playback frame.



WO 2009/093136 A2

## Image Capture and Motion Picture Generation

### Field of the Invention

5 The present invention relates to capturing image data and subsequently generating a moving picture in the form of a series of playback frames.

### Background of the Invention

10 Traditional motion picture image capture and playback uses a motion picture camera which captures images in the form of a series of image frames, commonly referred to as footage, which is then stored as playback frames and played back in the same sequence in which they are captured. A motion picture camera may be either a film camera or a video camera (including digital video cameras). Furthermore the sequence of image frames may be stored as a video signal, and the resulting motion pictures may be edited or unedited motion picture sequences which are used for motion picture film, TV, or other playback channels. Whilst developments in recording and playback technology allow the frames to be accessed separately, and in a non-sequential order, the main mode of playback is sequential, in the order in which they are recorded and/or edited. In terms of accessing frames in non-sequential order, interactive video techniques have been developed, and in optical recording technology, it is possible to view selected frames distributed through the body of the content, in a preview function. This is, however, a subsidiary function which supports the main function of playing back the frames in the order in which they are captured and/or edited.

25 Computer generation is an alternative technique for generating video signals. Computer generation is used in simulators and motion picture films. In computer generation the video signals are computer-generated from a three dimensional (3D) representation of the scene, typically in the form of an object model, and by then applying geometry, viewpoint, texture and lighting information. Rendering may be conducted non-real time, in which case it is referred to as pre-rendering, or in real time. Pre-rendering is a computationally

30

intensive process that is typically used for motion picture film creation, while real-time rendering is used for simulators. For simulators, the playback equipment typically uses graphics cards with 3D hardware accelerators to perform the real-time rendering.

5           The process of capturing the object model for a computer-generated scene has always been relatively intensive, particularly when it is desired to generate photorealistic scenes, or complex stylized scenes. It typically involves a very large number of man hours of work by highly experienced programmers. This applies not only to the models for the moving characters and other moving  
10           objects within the scene, but also to the background environment. As computers and motion picture film generation techniques become more capable of generating complex scenes, and capable of generating scenes which are increasingly photorealistic, the cost of capturing the object model has correspondingly increased, and the initial development cost of a simulator or  
15           computer generated motion picture film, is constantly increasing. Also, the development time has increased, which is particularly disadvantageous when time-to-market is important.

          Image-based-rendering (IBR) is an alternative technique to 3D geometric object modeling, for generating different viewpoint image data of an object  
20           and/or scene. In IBR geometric data of an object and/or scene is derived from previously captured 2D images of the object and/or scene. Provided these captured images are taken from different angles with respect to the object and/or scene, geometrical data regarding the object and/or scene can be derived from them. The computer-generated 3D object model of the scene allows one to  
25           deduce different viewpoint images of the scene that have not been previously captured.

          Apple Computer Inc.'s proprietary QuickTime VR™ software system generates panoramic images of a scene or object from pre-captured images of different viewpoints of the scene or object. This is done by stitching together  
30           the different viewpoint images whose totality represent a 360° viewpoint image of the scene or object. In addition QuickTime VR™ can be used to generate a

virtual walkthrough of the captured scene. One can imagine the observer's viewing position as being the centre of a cylinder or sphere. By projecting portions of the stitched 360° scenic image on the interior surface of a cylinder or sphere, the viewer has the impression he/she is within this virtually constructed scene. Only selected portions of the stitched image are projected at any one time, corresponding to the viewer's field of vision. As the viewer rotates about his/her viewing position different portions of the 360° scenic image are displayed accordingly. This method has certain limitations: it cannot be used to generate new perspectives of a captured scene – it can only display information captured in the pre-captured images of the scene and/or object. Accordingly one can zoom in or out, however a new perspective that was not captured in the pre-captured images cannot be generated without the use of a 3D object model, since the optical data defining such a new perspective is not known.

IBR techniques include methods based on the principle of light fields. A light field of an object and/or scene can be described as a field of light reflected from its surface, containing optical information characterising the scene. This light field may be represented by a set of light rays reflected from the scene and/or object. The light field is represented by a mathematical function called the plenoptic function, describing the radiance of all required light rays, in all required directions, at any required point in space. This relates to a technique of IBR called Light Field Rendering (LFR). By manipulating information contained within the light field (quantified by the plenoptic function), it is possible to generate a desired perspective of a scene and/or object. This can be achieved by sampling the light field at an appropriate rate such that the plenoptic function for a particular light field can be suitably defined in a region of space. In practice this is achieved by capturing many images of a scene, with suitable apparatus, from different perspectives. Each captured light ray's characteristics are stored as a pixel, including: colour, brightness, directional and positional data of the incident ray. When a suitable number of light ray characteristics have been captured, new perspectives can be generated by selecting and combining the pixels corresponding to the characteristics of individual light rays

passing through a desired viewpoint position, to generate the image of the scene and/or object as it appears from the chosen viewpoint position, without a physical presence of the capture device. Levoy and Hanrahan describe in their paper entitled "Light Field Rendering" (Proc. ACM SIGGRAPH '96) a theoretical principle behind light field sampling and rendering, and illustrate how a method of light field sampling can be used to generate new viewpoint images of a static object. Light field rendering is discussed in Levoy's article entitled "Light Fields and Computational Imaging" (published in the August 2006 issue of the IEEE Computer Society journal).

US2005/0285875A1 relates to a process for generating and rendering an interactive viewpoint video wherein a user can watch a video sequence and change the viewpoint at will during playback. This is achieved by recording a dynamic scene from a plurality of video cameras capturing a plurality of video streams representing different viewpoints of the scene. The different viewpoints are used to generate a 3D model of the scene and to generate disparity maps. Disparity mapping is a technique for recovering crude 3D information of a scene, however it is inherently limited and the quality of images derived in this way will not generally suffice for motion picture films or TV production.

When making motion picture films and TV shows (and other motion picture sequences) a large part of the development cost is in post-production. Once sets have been destroyed or are no longer available no new footage can be recorded. In such cases sets are either rebuilt, or a director may use poor footage and special effects to overcome the issues with the footage. Both of these solutions are expensive.

Chroma key techniques are used in both the motion picture and television industries. Complex computer-generated background scenes can be compiled with footage shot in front of a green screen. The time-consuming editing work is still mostly undertaken manually by specialist motion picture film editors. A significant amount of time is invested in ensuring the background image is consistent with the foreground image. Only in the simplest of applications, such as the weather forecast does the keying occur in real-time.

It is one objective of the invention to improve computer generation techniques for motion pictures.

### Summary of the Invention

5           The present invention is set out in the appended claims.

The present invention provides a method of combining image data generated by scene scanning a set, with image data captured with a motion picture camera.

10           An advantage of the invention is that highly photorealistic scenes can be computer-generated in correspondence with a motion picture capture sequence, according to movement of the motion picture camera through an environment corresponding to the scenes in any direction in at least a two dimensional space. The computer-generation can be based on real photographic images, for example of a motion picture film or TV set. The ability to virtually reproduce a set would greatly reduce the cost, and allow a director more artistic freedom  
15           during post-production.

In one embodiment of the present invention at each viewpoint position, a stored image is used to generate the associated viewpoint image of the scene.

20           In another embodiment of the present invention a light field of a particular scene is sampled by capturing image data of the scene, where said captured image data represent images of the scene taken from different viewpoints, this sampled light field data is then used to generate any desired new viewpoint image of the scene.

25           Using the present invention, scenes can be captured with a fraction of the cost and time required using known techniques. Also, the scenes can be played back at highly photorealistic levels, without requiring as much rendering as computer generation techniques relying purely on object models.

30           The techniques of the present invention may also be used in conjunction with object modelling techniques. For example, stored images may be used to generate the background scene whilst moving objects may be overlaid on the background scene using object models, in addition to the objects which are

captured using the motion picture camera. In this regard, object model data is preferably stored with the stored images, and used for overlaying moving object images correctly on the computer-generated scenes generated from the stored images.

5           In certain embodiments the second set of image data comprise captured images with a horizontal field of view which is more than 100°. The method of the present invention preferably involves selecting a suitable part of the captured image for playback, once the captured image has been selected on the basis of the current location of view. In this way, the viewing direction can be altered, at  
10 each viewpoint position.

In preferred embodiments of the present invention, a chroma key technique is used to combine the image data, and provides a means of performing the image overlaying in real-time.

15           Further features and advantages of the invention will become apparent from the following description of preferred embodiments of the invention, given by way of example only, which is made with reference to the accompanying drawings.

#### Brief Description of the Drawings

20           Figure 1 shows a schematic block diagram of apparatus according to a first embodiment of the invention.

Figure 2 shows a plan view of a grid pattern used for image capture and playback according to an embodiment of the invention.

25           Figure 3 shows a plan view of a triangular grid pattern used for image capture and playback according to an alternative embodiment of the present invention.

Figure 4 shows a perspective view of a set and/or scene and a planar 2D grid image data capture pattern according to a first embodiment of the invention.

30           Figure 5 shows a perspective view of a set and/or scene and a 3D volumetric grid image data capture pattern in accordance with an alternative embodiment of the present invention.

Figure 6 shows a flow diagram of a method of capturing image data at each node position contained within an image data capture grid, according to an embodiment of the invention.

5 Figure 7 shows a schematic block diagram of apparatus, comprising amongst other elements a camera mounted on a robotic arm, used for image data capture within a grid pattern according to an embodiment of the invention.

Figure 8 shows a perspective view of apparatus used to capture image data of a set and/or scene at each node position contained within the grid, according to an embodiment of the invention.

10 Figure 9 shows a panoramic lens arrangement for use in the image capture apparatus shown in Figure 8.

Figure 10 is a schematic block diagram of elements of the image capture apparatus shown in Figure 8.

15 Figure 11 is a schematic block diagram depicting components of video playback apparatus in accordance with an embodiment of the present invention.

Figure 12a shows a schematic representation of image data as captured and stored by a panoramic camera, in an embodiment of the invention.

Figure 12b shows a schematic representation of an image frame as played back in an embodiment of the invention.

20 Figure 13 shows a flow diagram of a method of generating image sequences from stored images according to an embodiment of the invention.

Figure 14 shows a flow diagram of a method of processing stored image data to generate a first scenic viewpoint image according to an embodiment of the invention.

25 Figure 15 shows a flow diagram of a method of generating a first scenic viewpoint image by extrapolation using ray tracing techniques, according to an alternative embodiment of the present invention.

30 Figure 16 is a schematic perspective view of the principle behind the method of extrapolation of a first scenic viewpoint image using ray tracing techniques, from stored image data, in accordance with an alternative embodiment of the invention.

Figure 17 shows a schematic block diagram of apparatus according to an embodiment of the present invention used in conjunction with green screen techniques.

5 Figure 18 shows a perspective view of a wall mounted camera on rails used to capture image data of a set and/or scene, according to an alternative embodiment of the invention.

### Detailed Description of the Invention

10 The invention provides for a method of generating a moving image in the form of a series of playback frames. The moving image represents movement of a camera through a generated virtual scene, in certain preferred embodiments a computer is used to generate the virtual scene. The moving image is composed of a sequence of discretely captured images, captured in a sequential order.

15 A first set of image data is captured using a motion picture camera, which can be a video camera in certain preferred embodiments. The motion picture camera has an associated motion recorder, such that the motion data of the motion picture camera can be recorded.

20 A second set of image data is captured using a scene-scanning device. The scene-scanning device provides image data of a scene, which in this embodiment is a motion picture filming set, for example a motion picture film set or TV set, taken at different positions along the scene, therefore providing different perspectives or viewpoints of the scene. Using the motion data and context data from the motion recorder and camera context recorder respectively,  
25 of the motion picture camera, the appropriate scenic viewpoint image can be selected for each image frame captured by the motion picture camera. The first image data set and selected parts of the scenic image data set are then combined to form playback frames.

30 Figure 1 depicts a preferred embodiment of the present invention. A motion picture camera 101 captures images of an actor 102 or other object. A motion sensor unit 103 is attached to the motion picture camera 101. In this

embodiment the motion sensor includes an accelerometer. The motion data of the motion picture camera 101 is recorded by a motion sensor data recorder 104. The motion sensor data recorder 104 takes motion picture camera motion data, from the motion sensor unit 103, as an input and processes this to output motion picture camera position data 108 associated with each captured image frame. An image recording device 105 records the images captured by the motion picture camera 101. In certain embodiments the image recording device 105 may be part of the motion picture camera 101. Stored image data 106 of a set or scene, having been captured previously, are stored on a storage device 107.

10 An image processing device 109 receives motion picture camera position data for a particular captured frame, and uses this data to recover a corresponding stored image data 106 of a motion picture filming set from the storage device 107. The motion picture camera context data 110 relates to data such as zoom state and focal state. The motion picture camera context recorder 15 111 derives the motion picture camera context data 110 by analysing the command signals of the motion picture camera to derive zoom state and focus state of the motion picture camera. The motion picture camera context data 110 is sent to the image processing device 109, where it is used to process the selected scenic image data 106 to generate the corresponding scenic viewpoint image. In this way the zoom state and focal state of the generated scenic viewpoint image is consistent with the image captured by the motion picture camera 101. For example if the captured image is a close-up of an actor 102, the background scenic image data may be processed to generate a background scenic viewpoint image which is out-of-focus. An in-focus stored image data 20 106 is processed to depict the correct zoom and focal states.

Having generated the background viewpoint image, the image captured by the motion picture camera 101 and the processed background viewpoint image are overlaid in the image overlay unit 112. The image overlay unit 112 correctly overlays the captured image in the foreground of the generated background scenic viewpoint image. The resulting overlaid image can be 30 displayed on a display unit 113 and/or recorded by the overlaid image recording

unit 114. In certain embodiments this process may occur in real-time, such that an overlaid image of actor and scenic background image can be viewed on a display 113 in real-time as the actor 102 is filmed.

5 In another embodiment of the present invention the computer-generated virtual scene is generated using captured images by taking the captured images to have different viewpoints within the virtual scene, the viewpoints corresponding to different perspectives of the scene captured from different points of capture. An image may be stored for each of the viewpoints, by capturing a plurality of images based on the selection of a plurality of points of  
10 capture.

In another embodiment the computer-generated virtual scene is generated by extrapolating a particular viewpoint image of the scene from the sampled light field data of the scene. Where said sampled light field data corresponds to image data, positional and orientational data of incident light  
15 rays, captured at different positions along the scene. In this embodiment it is understood that both colour and intensity are part of said image data. The different positions corresponding to capture nodes forming an array of points of view or perspectives of said scene or set.

At least some of said points of capture (capture nodes) are distributed  
20 with a substantially constant or substantially smoothly varying average density across a first two-dimensional area. The capture nodes are distributed in at least two dimensions, and may be distributed in three dimensions.

At least some of said points of capture are distributed in a regular pattern including a two-dimensional array in at least one two-dimensional area, for  
25 example in a grid pattern, if possible depending on the scene scanning imaging device. One suitable grid formation is illustrated in Figure 2, which in this example is a two dimensional square grid. The viewpoints are located at each of the nodes of the grid.

The intended application of the generated background scene may  
30 condition the choice of scene scanning imaging device, as the captured field of view is dependent on the characteristics of said device. In many embodiments it

is desirable to capture images with a horizontal field of view greater than 140°. In other preferred embodiments the captured images comprise images with a 360° horizontal field of view. For example, this might be the desired choice if one intended to generate a virtual scene, where it is desirable that a director has the freedom to view in all directions about any viewpoint position. Each stored image may be composed from more than one captured image. More than one photograph may be taken at each capture node, taken in different directions, with the captured images being stitched together into a single stored image for each viewpoint or node position. It is preferable to use a single shot image capture process to reduce geometrical errors in the stored image data which will be amplified on playback as many images are played back per second. Where the captured images are photographic images, these will have been captured at a plurality of points of capture in a real scene using camera equipment. In one embodiment the captured images will preferably have been captured using panoramic camera equipment.

During playback, the video frames are preferably generated at a rate of at least 30 frames per second. The spacing of the points of capture in the virtual scene, and also the real scene from which the virtual scene is initially captured, is determined not by the frame rate but the rate at which the human brain is capable of detecting changes in the video image. Preferably, the image changes at a rate less than the frame rate, and preferably less than 20Hz. The viewpoint position spacing is determined by the fact that the brain only really notices up to 14 changes in images per second. While we can see 'flicker' at rates up to 70-80Hz. Thus the display needs to be updated regularly, at the frame rate, but the image only needs to really change at about 14Hz. The viewpoint position spacing is determined by the speed in meters per second, divided by the selected rate of change of the image. For instance at a walking speed of 1.6m/s images are captured around every 50mm to create a fluid playback. For a driving game this might be something like one every meter (note that the calculation must be done for the slowest speed one moves in the simulation). In any case, the points of capture, at least in some regions of said real scene, are preferably spaced less

than 5m apart, at least on average. In some contexts, requiring slower movement through the scene during playback, the points of capture, at least in some regions of said real scene, are spaced less than 1m apart, at least on average. In other contexts, requiring even slower movement, the points of capture, at least in some regions of said real scene, are spaced less than 10cm apart, at least on average. In other contexts, requiring yet slower movement, the points of capture, at least in some regions of said real scene, are spaced less than 1cm apart, at least on average.

The capturing comprises recording data defining the locations of viewpoint positions in the virtual scene. For example in most embodiments, the viewpoints' locations may correspond to the locations of points of capture in said real scene. A position of each point of capture may thus be recorded as location data associated with each captured scenic image data, for subsequent use in selecting the viewpoint. In a preferred embodiment, when the position of the motion picture camera is close to a particular node position, then the scenic image data associated with the node position is selected and the appropriate viewpoint is selected therefrom when moving through the virtual scene.

Returning to Figure 2, it can be seen that the nodes of the grid, representing a plurality of points of capture and stored image data, are distributed relative to a first point of capture, let us take for example point n1, in at least two spatial dimensions. The points of capture are distributed around point n1, across four quadrants around the first point of capture.

Whilst Figure 2 illustrates a square grid, at least some of the points of capture may be distributed in a non-square grid across the first two-dimensional area. In an alternative embodiment, at least some of the points of capture are distributed in a triangular grid across the first two-dimensional area, as shown in Figure 3.

Alternatively, or in addition, at least some of the points of capture may be distributed in an irregular pattern across the first two-dimensional area – this may simplify the capture process. In this case, images are captured which irregularly, but with a constant or smoothly varying average density, cover the

area. This still allows the playback apparatus to select the nearest image at any one time for playback - or blend multiple adjacent images, as will be described in further detail below

5 Different areas may be covered at different densities. For example, an area in a virtual environment which is not often visited may have a lower density of coverage than a more regularly visited part of the environment. Thus, the points of capture may be distributed with a substantially constant or smoothly varying average density across a second two-dimensional area, the second two-dimensional area being delineated with respect to the first two-dimensional area and the average density in the second two-dimensional area being different to the average density in the first two-dimensional area. If one is sampling the light field then a relatively dense capture node distribution will allow one to generate more new viewpoint images without having to use interpolation or other approximations.

15 The viewpoint positions may be distributed across a planar surface, for example in a virtual scene representing an in-building environment. Alternatively, or in addition, the viewpoint positions may be distributed across a non-planar surface, for example in a virtual scene representing rough terrain in a driving game for example. If the surface is non-planar, the two dimensional array will be parallel to the terrain i.e. it will move with the ground. The terrain may be covered using an overlay mesh - each part of the mesh may be divided into a triangle which includes a grid pattern inside the triangle similar to that shown in Figures 2 or 3, and the surface inside each triangle will be flat (and the triangles will in some, and perhaps all cases, not be level). All triangles will be at different angles and heights with respect to each other, to cover the entire terrain. During the capture process, it is possible to survey the area prior to scanning it, and create a 2D mesh of triangles, where all neighbouring triangle edges and vertices line up. The capture apparatus can be repositioned, sequentially collecting data in each of the triangles.

30 In another embodiment, the array of viewpoint positions or nodes may be a two-dimensional planar surface, substantially perpendicular to the ground.

This is particularly useful when a scene only needs to be captured from a particular range of directions. For example, during conventional filming of a scene in a motion picture, the motion picture film camera captures a set from a particular viewpoint angle. The objective of capturing many images of the scene is to permit one to reproduce viewpoint images, hence it is unnecessary to capture more viewpoints of the set than will potentially be used. Rather, restricting the array to cover the area the motion picture film camera may move in will suffice. This planar, perpendicular orientation of the capture grid is also used in a preferred embodiment when sampling the light field of a scene (including a motion picture filming set) and/or object. Figure 4 illustrates one such embodiment. A capture plane 401 is oriented perpendicular to the ground and parallel to a set 403. The capture plane 401 consisting of a number of nodes 402 designating the positions at which image data of the set will be captured by a camera.

The previous argument is also valid for 3D arrays of viewpoint positions or nodes. If a director is aware of the general range of viewpoint images that are required, then the array of capture node positions can be restricted to cover the viewpoint positions of this selected range of viewpoint images.

In a further embodiment, the viewpoint positions or nodes are distributed across a three-dimensional volume, for example for use in a flight simulator or otherwise. The node positions may be arranged in a regular 3D array. Figure 5 depicts an example of a volumetric array for capturing a scene. The capture volume 501 consists of a number of nodes 502, representing the positions where image data of a set 503 is captured by a camera.

Figure 6 is a flow chart illustrating a method 600 of sequentially capturing scenic image data at each node inside a chosen capture space. The capture space can be a 2D area, or a 3D volume, as mentioned above. The 2D area may be planar or non-planar. The 3D volume may be divided into a 3D grid composed of many parallel planes, or a more complex grid pattern, such as a pyramidal 3D grid. Both types of capture spaces are composed of nodes, designating the positions where image data is captured. The method involves

defining the capture space 601, whether it is a 2D area or a 3D volume and the shape of said chosen space. The next step is to define the spacing of the nodes 602 within the chosen capture space. A scene scanning imaging device is placed at a designated starting node 603 where a first image of the scene is captured 604. The captured image is then stored 605. It is determined 606 if there remain any nodes wherefrom image data has not been captured. If nodes remain, then the scene scanning imaging device is repositioned to the next node 607 where another image of the scene is captured and stored. This process is repeated until image data at all nodes in the designated capture space have been captured, at which point the process is ended 608.

In another embodiment of the invention, the images are captured using an automated mechanically repositionable camera. The automated mechanically repositionable camera is moved in a regular stepwise fashion across the real scene.

Figure 7 depicts an image capture device 700 used in an embodiment of the invention. A wide-angle camera 701 mounted on a servo 702, which in turn is mounted on a dolly 703 having a robotic arm 704, is used to capture different image data at different viewpoint or node positions. The servo 702 controls three angular degrees of freedom (e.g. pan, tilt and yaw) of the camera 701, whilst movement along three locational degrees of freedom (e.g. x, y and z-axis) are controlled by the dolly unit 703. In particular movement in the vertical direction (along the z-axis) is controlled by the robotic arm 704 which is itself part of the dolly unit 703. A motion sensor 705, in this embodiment including a plurality of sensors capable of detecting changes in any one of the six degrees of freedom, is attached to the wide-angle camera 701. The motion sensor 705 continuously captures and sends motion data to the motion controller 706. The motion data is used to calculate the current position and viewing direction of the camera 701. The motion controller is connected (either physically or wirelessly) to both the servo control unit 707 and dolly control unit 708. Using the motion data from the motion sensor 705, the motion controller 706 can control all movement of the servo 702 and dolly unit 703 to a high degree of precision. In

certain embodiments the motion controller 706 can be configured to receive positional instructions remotely. For example, one may generate a number of positional instructions on a computer or other processing device, and send these positional instructions to the motion controller 706. The motion controller 706  
5 will translate these positional instructions into independent instructions destined for the servo 707 and dolly control units 708. In this manner the capture process is fully automated and does not require manual control. Furthermore, when an image is captured and recorded on the image recorder 709, the image recorder 709 will record the position data from the motion controller 706 associated with  
10 the particular recorded image. The recorded image with associated position data is stored on a storage medium 710.

Figure 8 shows an image capture device 802 in another embodiment, comprising a base 804, a moveable platform 806, a turret 808, and a camera 809. The base 804 is mounted on wheels 812 whereby the device is moved  
15 from one image capture position to another. The moveable platform 806 is mounted on rails 814 running along the base 804 to provide scanning movement in a first direction X. The turret 808 is mounted on a rail 816 which provides scanning movement in a second direction Y, which is perpendicular to the first direction X. Note that the rails 814 may be replaced by high-tension wires, and  
20 in any case the moveable platform 806 and the turret 808 are mounted on the rails or wires using high precision bearings which provide sub-millimetre accuracy in positioning both the first and second directions X, Y.

Mounted above the camera 9 is a panoramic imaging mirror 810, for example the optical device called "The 0-360 One-Click Panoramic Optic"<sup>TM</sup>  
25 shown on the website [www.0-360.com](http://www.0-360.com). This is illustrated in further detail in Figure 9. The optical arrangement 810 is in the form of a rotationally symmetric curved mirror, which in this embodiment is concave, but may be convex. The mirror 810 converts a 360 degree panoramic image captured across a vertical field of view 926 of at least 90 degrees into a disc-shaped image captured by the  
30 camera 9. The disc-shaped image is shown in Figure 12a and described in more detail below.

In the image capture device shown in Figure 8, the base may have linear actuators in each corner to lift the wheels at least partly off the ground, in order to substantially level the image capture apparatus on uneven terrain. Lifting the wheels at least partly off the ground also helps to transfer vibration through to the ground – to reduce lower frequency resonance of the whole machine during image capture. A leveling system may also be provided on the turret. This allows fine calibration to make sure the images are substantially level.

Figure 10 shows a control arrangement 1000 for the device illustrated in Figure 8. The arrangement includes image capture apparatus 1002 including the panoramic camera 9, x- and y-axis control arrangement including stepper motors 1020, 1030, and corresponding position sensors 1022, 1032, tilt control arrangement 1006 including x-axis and y-axis tilt actuators 1040, and corresponding position sensors 1042, and drive arrangement 1008, including drive wheels 812 and corresponding position sensors 1052. The control arrangement is controlled by capture and control computer 1012, which controls the position of the device using drive wheels 812. When in position, the turret 808 (Figure 8) scans in a linear fashion, row by row, to capture photographic images, which are stored in media storage device 1014, in a regular two-dimensional array across the entire area of the base 804. The device is then moved, using the drive wheels 812, to an adjacent position, and the process is repeated, until the entire real area to be scanned has been covered.

Returning to Figure 2, during playback, a video signal comprising a moving image in the form of a series of playback frames is generated using stored images by taking the stored images, which are stored for each of the nodes  $n$  of the grid, according to the current position  $P$  (defined by two spatial coordinates  $x,y$ ) of the viewer. Take for example an initial position of the viewer  $P_1(x,y)$ , as defined by a control program which is running on the playback apparatus, which tracks the location of the motion picture camera as the camera moves through the computer-generated scene. The position of the viewer is shown using the symbol  $x$  in Figure 2. A first stored image based on the selection of a first viewpoint position  $n_1$  which is closest to the initial

position  $P1(x,y)$  is selected. The playback apparatus then generates a first playback frame using the first stored image. More than one playback frame may be generated using the same first stored image. The position of the viewer may change. The viewer, in a preferred embodiment, may move in any direction in at least two dimensions. A plurality of potential next viewpoints  $n_p$ , shown using the symbol "o" in Figure 2, are distributed around the initial viewpoint position  $n_1$ . These are distributed in all four quadrants around the initial viewpoint position  $n_1$  across the virtual scene. The viewer is moved to position  $P2(x,y)$ . The playback apparatus selects a next viewpoint position  $n_2$  from the plurality of potential next viewpoint positions distributed relative to the first viewpoint position across the computer-generated scene, on the basis of proximity to the current position of the viewer  $P2(x,y)$ . The playback apparatus then selects a second stored image on the basis of the selected next viewpoint position; and generates a subsequent playback frame using the second stored image.

In certain embodiments the viewer may be the motion picture camera. In such an embodiment the motion of the motion picture camera is recorded and the positional coordinates of the camera's motion trace out a movement path. This movement path can be used within the computer-generated scene to select and display the correct scenic viewpoint images. This corresponds to selecting and displaying the viewpoint images captured at the capture nodes, within the array of capture nodes, that intersect (or are nearest to) the motion picture camera's motion path, when said motion path is embedded on the array of capture nodes.

In another embodiment, interpolation can be used to generate the motion picture camera's associated scenic viewpoint image. Stored images nearest to the motion picture camera's position are selected, and from these stored images the viewpoint at the motion picture camera's current position is interpolated. Metadata can be calculated and stored in advance to aid the interpolation between multiple images.

The generation of playback frames may comprise generating playback frames based on selected portions of the stored images. The playback frame's field of view is less than the captured scenic image data's field of view. The captured images may have a field of view of more than 140° whilst the selected portions may have a field of view of less than 140°, and the playback equipment in this example also monitors the current viewing direction in order to select the correct portion of the image for playback. In an embodiment involving the overlaying of a moving image with a computer-generated background image, the position, and context data of the motion picture camera is used to select the correct portion of the stored image, to generate a scenic viewpoint image for playback. In particular the positional data provides information used in selecting the correct viewpoint and viewing direction. The zoom state helps in establishing the correct field of view of the scenic image intended for playback. The focal data is used to achieve the required focal state of the scenic image. In one embodiment, the selected portions have a field of view of approximately 100°.

As described above, in one embodiment the playback method comprises receiving data indicating a position of the motion picture camera, and selecting a next viewpoint on the basis of the position. The selecting comprises taking into account a distance between the position and the plurality of potential next viewpoint positions in the virtual scene. The method preferably comprises taking into account the nearest potential next viewpoint position to the current position and comprises taking into account a direction of travel of the camera, in addition to the position. The playback apparatus may receive a directional indication representing movement of the camera, and calculating the position on the basis of at least the directional indication.

Figure 11 illustrates playback equipment 1100, according to an embodiment of the invention. The playback equipment 1100 includes a control unit 1110, a display 1120 and a man-machine interface 1130. The control unit 1110 may be a computer, such as a PC, or a game console. The control unit 1110 is comprised of the following components: conventional I/O interface

1152, processor 1150, memory 1154, storage 1160, and operating system 1162. In certain embodiments the control unit 1110 additionally comprises control software 1164 and stored photographic images 1172, along with other graphics data 1174. The control software 1164 operates to monitor the position of the viewer in a virtual scene, as controlled by the user using man-machine interface 1130. As described above, the control software generates video frames using the stored images 1172, along with the other graphics data 1174, which may for example define an object model associated with the stored images 1172, using the process described above. In another embodiment, motion picture camera position data 1156, zoom state data 1158 and focal state data 1157, are fed to the playback equipment 1100 through the I/O interface 1152. The image processing/editing software 1176 uses the positional data 1156 and the zoom state data 1158 to retrieve the correct scenic image from the stored images 1172, and to select the correct viewpoint image therefrom. The focal state data 1157 is used to generate the correct focal resolution. The resulting scenic viewpoint image can then be displayed on a display unit 1120.

Figure 12a illustrates an image 1200 as stored in one embodiment, when a 360° viewpoint image is used. The image 1200 includes image data covering an annular area, corresponding to the view in all directions from a particular viewpoint position. When the viewpoint position is selected by the playback apparatus 1100 (Figure 11), the playback apparatus selects a portion 1220 of the stored image 1172 (Figure 11) corresponding to the current direction of view of the viewer. The playback apparatus 1100 then transforms the stored image portion 1220 into a playback image viewpoint 1220', by performing a dewarp coordinate transform on it and placing the data as regularly spaced pixels within a rectangular image frame 1270, shown in Figure 12b. When conducting the transformation, a good way to do it is to map the stored image portion onto a shape which recreates the original environment. For some camera setups, this will mean projecting the stored image portion on the inside of a sphere or cylinder. On others it might mean just copying the stored image to the display surface.

Figure 13 shows a flow diagram of a method 1300 of generating image sequences from stored image data 1172 (Figure 11). Motion picture camera position data 1156 is received 1310, for an individual frame, by a control unit 1110, and used to select the image data 1320, from stored images 1172, corresponding to the position of the motion picture camera. Motion picture camera context data 1350 is used to process 1330 the selected stored image data. This may include manipulating the zoom and focus state of the selected stored image. The processed image is then stored 1340 either in a work memory 1154 or on a storage medium 1160 for future use, and represents one image frame within an image sequence, consisting of a plurality of image frames. It is determined 1360 whether any more images remain to be processed in the sequence. If there are remaining images to be processed then steps 1310-1360 are repeated for all remaining images. Once all the images within a sequence have been processed, the individual processed images may be merged to form an image sequence, this merged image sequence is then stored 1370 and may be displayed 1380 on a display unit such as 1120. A dynamic image sequence is created by individually processing a plurality of pre-captured images stored on a storage media.

Figure 14 is a detailed flow chart of a method 1400 of processing stored image data, in accordance with the method described in Figure 13, to generate a first scenic viewpoint image for use in creating an image sequence. It is to be noted that the method applies to each individual image used to create a sequence. The correct stored image is received 1320 by using the motion picture camera position data. The motion picture camera position data 1430 is used to select the portion of the stored image with the correct field of view 1420. For example this may ensure that the orientation and the aspect ratio of the stored image correspond with the motion picture camera data. The correct magnification of the stored image is determined 1440, by comparison with the motion picture camera zoom state data 1450. If required the stored image magnification is increased or decreased appropriately 1460. This ensures there is no optical discontinuity between the foreground image (as captured by the

motion picture camera) and the background image when overlaid. The correct focal clarity of the stored image being processed is determined 1470 by using the motion picture camera focal state data 1490. If required the focal clarity is manipulated 1480 to be consistent with the motion picture camera settings. The process is ended 1411 and is repeated for each individual stored image within the sequence.

In another embodiment of the present invention, the generated scenic image viewpoint position is not restricted to being contained within that of the corresponding positions of the array of capture nodes. In this embodiment a desired viewpoint position is selected which lies outside a defined grid of image data capture nodes. The chosen viewpoint position may lie between the grid of capture nodes and the scene, or further away from the scene. The range of viewpoint positions available and the quality of the resulting generated image is dependent on the resolution of the captured scenic image data, and the density of the capture nodes. Special scenic image data capturing devices are used when capturing scenic image data of the scene. The capture device is in this embodiment constructed such that each pixel on the image plane of the capture device is associated to one light ray within the light field, such as in a plenoptic camera arrangement or equivalent.

For each light ray incident on the front element of the capture device, the position and angle of incidence is recorded along with the corresponding pixel.

For each captured ray the following data is recorded: colour (wavelength), intensity (brightness), position and angular direction.

The scenic viewpoint image associated with the desired viewpoint position is then extrapolated from the captured scenic image data captured at each capture node, using ray tracing techniques, to identify individual light rays that pass through the chosen position from the set of pre-captured scenic images, and to combine the pixels associated to the identified light rays to form a coherent scenic viewpoint image of the scene as would be observed at the chosen viewpoint position.

The method 1500, shown in Figure 15, involves defining a virtual or imaginary camera position 1501, virtual is used in this context to indicate the absence of a real physical camera. This position can be arbitrarily selected and corresponds to the position where the desired scenic viewpoint image of the scene and/or set is to be taken from. The optical properties of this virtual camera are defined, step 1502. This may include defining the optical arrangement of lenses inside the virtual camera, and defining quantities such as focal length, aperture size, etc. With the optical imaging properties of the camera defined one can model the light ray capturing process, or imaging process of the virtual camera. Imaginary rays of light are backwards-traced 1503 from the virtual camera lens's front element to the grid of nodes of captured image data, where the point 1504 and angle 1505 of intersection between grid and imaginary ray are recorded. The point and angle of intersection of the imaginary light ray with the capture grid are used in identifying and relating pixels within captured scenic image data to specific light rays.

The backwards ray-tracing technique is conceptually illustrated in Figure 16. Light rays 1601 are back traced from a virtual camera 1602, through a capture plane 1603, to a set or scene 1604. It is determined 1506 (Figure 15) if the point of intersection between the imaginary ray and grid corresponds to a node position 1605. If the point of intersection between the imaginary ray and grid corresponds to a node position, then the associated image data captured at the node is recovered 1507. If the point of intersection does not correspond to an existing node position, then the image data at the node nearest 1508 to the point of intersection between imaginary ray and grid is selected and recovered 1509. The scenic image data at the point of intersection can be interpolated 1510 from the scenic image data captured at the nodes nearest to the point of intersection.

The angle of intersection 1505 between imaginary ray and grid can then be used to identify the pixel 1511 on the selected image data, caused by a captured light ray having intersected the current node in substantially the same

direction and with substantially the same angle as the imaginary ray. The process is repeated for all imaginary rays entering the virtual camera, such that a pixel can be associated with each imaginary ray 1512.

5 The set of pixels are then compiled 1513 together to form one scenic viewpoint image which can be stored 1514. Furthermore the colour and intensity of each pixel may be averaged proportionately to the deviation of the virtual ray from the real ray.

10 The image quality of the generated scenic image viewpoint using this backwards ray-tracing technique is dependent on the angular resolution of the scenic image data captured at each capture node. When the angular resolution is high, the method is more accurate as less approximation is made when relating pixels on captured scenic image data to light rays.

15 In a preferred embodiment the capture node spacing is relatively small and the pixel density of the scene-scanning imaging device is relatively high. Since optical characteristics of the virtual camera are arbitrarily defined, one can define an optical system having any number of optical characteristics including aberrations such as: coma, spherical aberration, astigmatism, chromatic aberration etc. This may be desirable for achieving desired visual effects.

20 Another embodiment of the present invention involves an improvement to chroma key, also known as colour keying, techniques. Figure 17 illustrates a method of using the current invention with a chroma key technique. An actor 1701 or other moving object, placed in front of a green screen 1702 (or other monochromatic colour screen) is captured by a first camera system 1703, which is a motion picture camera system.

25 The motion picture camera system 1703 is composed of a motion picture camera 1404 mounted on a servo device 1707, which is mounted on a dolly unit 1705. A motion sensor 1709 sends motion picture camera motion data to a motion sensor data recorder 1710, whilst the motion picture camera 1704 films the actor 1701. The motion picture camera's captured images are recorded on  
30 an image recording device 1711. Motion picture camera zoom state data, focal

state data and other defining state data of the motion picture camera's optical system are recorded by a motion picture camera context recorder 1712.

5 A virtual camera system 1713 which navigates within a virtual motion picture filming set 1714, is comprised of a camera 1715, a servo device 1716 mounted on a dolly unit 1717 (also comprising a robotic arm), an associated dolly control unit 1718, servo control unit 1719, and a camera control unit 1720. The camera control unit 1720 is able to control all functions of the virtual camera 1715, such as zoom state, focal state, aperture size etc. The virtual motion picture filming set 1714 is generated by sampling the light field of the  
10 corresponding real motion picture filming set from a plurality of different viewpoint positions. Light field sampling of the real motion picture filming set can either be achieved by a single repositionable camera or in an alternate embodiment, by an array of stationary cameras.

Motion data from the motion sensor data recorder 1710 of the first  
15 camera system 1703 is sent to the motion controller 1721 controlling the virtual camera system 1713. The motion controller 1721 translates the motion sensor data into separate instructions destined for the servo control unit 1719 and/or dolly control unit 1718 of the virtual camera system 1713. Zoom state data and focal state data for each captured image frame of the first camera system 1703,  
20 is sent to the camera control unit 1720 of the virtual camera system 1713, from the motion picture camera context recorder 1712. This data allows the virtual camera system 1713 to reproduce the first camera system's 1703 movement and contextual behaviour, such as zoom and focal resolution, for each captured image frame, within the virtual set 1714. This is achieved by selecting scenic  
25 image data, from the sampled light field data, on the basis of the first camera system's 1703 motion data; and using the context data to generate the associated scenic viewpoint image data. The virtual camera system uses the motion and context data of the first camera system, to navigate and generate new scenic viewpoint image data, from the set of scenic image data captured during  
30 sampling of the motion picture filming set's light field. The result of this

process is that the virtual camera has reproduced exactly, the first camera system's motion and context data within the virtual motion picture filming set.

The generated scenic viewpoint images of the motion picture filming set are recorded on an image recording device 1722. Both camera systems' recorded images are then overlaid in the image editing device/overlay system 1723. This involves keying out the green screen 1702 from the image recorded by the image recording device 1711, and replacing it with the generated background image of the virtual set 1714 recorded by the image recording device 1722. The resulting overlaid image can then either be displayed on a display unit 1724, or alternately recorded on an appropriate storage medium (not pictured). This process is repeated for every image frame captured by the first camera system 1703.

The method can be performed in real-time, wherein the first camera system 1703 and the virtual camera system 1713 perform their respective tasks simultaneously, and the resulting overlaid image sequence is displayed on the display unit 1724 as the camera systems capture/generate their respective images. Note that the first camera system 1703 and the virtual camera system 1713 need not be in geographic proximity. It is envisioned that the two systems may be remote with a suitable communication channel being provided between the two systems.

In an alternative embodiment LFR (light field rendering) is not used to generate new scenic viewpoint images, rather scenic viewpoint images are generated by processing stored scenic image data using motion and context data from the first camera system.

In embodiments of the invention where the motion picture camera 1704 does not have an electronic control interface, such as older video cameras used for TV, the motion picture camera's context data can be derived from image analysis of the image data captured with the motion picture camera. For example, neighbouring captured image frames may be compared to derive motion picture camera zoom state and focal state data.

In an alternate embodiment of the present invention used in motion picture production, specific scenes can be re-filmed once the motion picture filming sets have been destroyed or are no longer available. For example, a director could select the motion picture footage representing a particular scene which he/she wishes to re-film. This consists in identifying the individual motion picture image frames of the scene in question. The selected plurality of image frames are removed from the original motion picture footage, where said footage is comprised of an actor filmed on a motion picture filming set, so-called "on-scene", and replaced with new footage filmed "off-scene." This new footage can be generated by re-filming the actor or plurality thereof in front of a monochromatic background screen. New scenic viewpoint image data of the motion picture filming set is generated by rendering the desired viewpoint images from within the virtual scene, using the sampled light field data of the real motion picture filming set. These computer-generated scenic viewpoint images are then overlaid with the footage of the actor captured in front of the monochromatic background, using chroma keying. The overlaid footage can then be re-inserted into the original motion picture footage in the appropriate location.

#### 20 Further embodiments of capture apparatus

In a further embodiment of the invention, the image capture apparatus may be ceiling-mounted within a building, as depicted in Figure 18. A camera 1801 is mounted on a rail 1803, itself suspended on rails 1802 lining a wall of the closed set 1806. A capture plane 1804 is defined consisting of a number of capture nodes 1805 distributed throughout the capture plane. The camera 1801 is positioned at an initial node 1805 where image data of the set 1806 is captured. The camera 1801 is sequentially repositioned to neighbouring nodes by moving along the rail 1803 in the horizontal direction, sequentially capturing image data of the set 1806 at each node. Camera movement in the vertical direction is controlled by moving rail 1803 vertically along rails 1802. The horizontal image capture sequence is repeated for each vertical position of the

rail 1803. This allows a plurality of image data of the set 1806 to be captured from different viewpoint positions. This arrangement may be used for capturing an artificial scene constructed from miniatures (used for flight simulators for instance).

5           In a further embodiment, the image capture apparatus is wire-mounted or otherwise suspended or mounted on a linear element, such as a pole or a track. The capture device obtains a row of images when the linear element is moved. This can be used for complex environments like rock faces or over areas a ground-mounted image capture apparatus is unable to be placed. The wire or  
10 other linear element may be removed from the images digitally.

          A two-step photographing process may be used – at each point two photographs are captured rather than one. This may be done by using a wide angle lens (8mm or 180 degrees). The image capture apparatus takes all photographs in its grid area, then rotates the camera a half turn, then takes  
15 photographs at each node again.

          The number of points of capture is preferably at least 400 per square meter, and in a preferred embodiment the number per square meter is 900, and where two photographs are taken per point, there are 1800 raw photographs per  
20 square meter.

20

#### Recovering physics data from the images

          An object model accompanying the stored images may be generated from the stored images themselves. 3D point/mesh data may be recovered from the images for use in physics, collision, occlusion and lighting calculations.  
25 Thus, a 3D representation of the scene can be calculated using the images which have been captured for display. A process such as disparity mapping can be used on the images to create a 'point cloud' which is in turn processed into a polygon model. Using this polygon model which is an approximation of the real scene, we can add 3D objects just like we would in any 3D simulation. All  
30 objects, or part objects, that are occluded by the static captured environment are (partially) overwritten by the static image.

Alternatively, or in addition, the 3D representation of the scene may be captured by laser scanning of the real scene using laser-range-finding equipment.

5           The above embodiments are to be understood as illustrative examples of the invention. Further embodiments of the invention are envisaged. For example, in the above embodiments, the image data is stored locally on the playback apparatus. In an alternative embodiment, the image data is stored on a server and the playback apparatus requests it on the fly. It is to be understood  
10 that any feature described in relation to any one embodiment may be used alone, or in combination with other features described, and may also be used in combination with one or more features of any other of the embodiments, or any combination of any other of the embodiments. Furthermore, equivalents and modifications not described above may also be employed without departing  
15 from the scope of the invention, which is defined in the accompanying claims.

Claims

1. A method of generating a moving image in the form of a series of playback frames, wherein said moving image is generated from:

5 a first set of image data captured using a motion picture camera system having an associated movement capture system;

movement capture data generated by said movement capture system; and

a second set of scenic image data captured using a scene-scanning imaging system which provides image data from which images, representing viewpoints distributed across a scene through which a virtual viewer is capable  
10 of navigating, are derivable, the method comprising:

on the basis of said movement capture data, selecting a first viewpoint position;

15 deriving first scenic viewpoint image data from said second set of image data, based on the selection of said first viewpoint position;

combining image data from said first image data with said first scenic viewpoint image data to generate a first playback frame;

20 on the basis of said movement capture data, selecting a next viewpoint position from a plurality of potential next viewpoint positions distributed relative to the first viewpoint position across said scene;

deriving second scenic viewpoint image data from said second set of image data, based on the selection of said next viewpoint position; and

25 combining image data from said first image data with said second scenic viewpoint image data to generate a second playback frame.

2. A method according to claim 1, wherein said motion picture camera system has a motion picture camera context capture system, and wherein said moving image is further generated from motion picture camera context capture data generated by said motion picture camera context capture system.

3. A method according to claim 2, wherein said motion picture camera context capture data is used with said motion picture camera movement capture data to derive a first scenic viewpoint image data from said second set of scenic image data, the method comprising:

5 on the basis of said motion picture camera movement capture data, selecting a first viewpoint position; and

on the basis of said selected first viewpoint position and said motion picture camera context capture data, deriving a first scenic viewpoint image data associated with said first viewpoint position.

10

4. A method according to claim 2 or 3, wherein said motion picture camera context capture data is used with said motion picture camera movement capture data to derive a second scenic viewpoint image data, the method comprising:

15 on the basis of said motion picture camera movement capture data, selecting a next viewpoint position from a plurality of potential next viewpoint positions distributed relative to the first viewpoint position across said scene; and

20 on the basis of said motion picture camera context capture data, deriving second scenic viewpoint image data associated with said selected next viewpoint position.

5. A method according to any preceding claim, wherein said generating of playback frames comprises generating playback frames based on  
25 selected portions of images in said second set of scenic image data.

6. A method according to claim 5, wherein said selected portions have a field of view of less than 140°.

30 7. A method according to claim 5 or 6, wherein said selected portions have a field of view of more than 100°.

8. A method according to any preceding claim, wherein said second set of scenic image data comprises photographic images which have been captured at a plurality of points of capture in a real scene using camera equipment.

9. A method according to claim 8, wherein said photographic images have been captured using panoramic camera equipment.

10. A method according to claim 8 or 9, wherein said points of capture, at least in some regions of said real scene, are spaced less than 5m apart.

11. A method according to claim 10, wherein said points of capture, at least in some regions of said real scene, are spaced less than 1m apart.

12. A method according to claim 11, wherein said points of capture, at least in some regions of said real scene, are spaced less than 10cm apart.

13. A method according to claim 12, wherein said points of capture, at least in some regions of said real scene, are spaced less than 1cm apart.

14. A method according to any preceding claim, comprising receiving data indicating a position of said viewer in said virtual scene, and selecting a next viewpoint position on the basis of said position.

15. A method according to claim 14, wherein said selecting comprises taking into account a distance between said indicated position and said plurality of potential next viewpoint positions in said virtual scene.

16. A method according to claim 15, comprising taking into account the nearest potential next viewpoint position to said indicated position.

5 17. A method according to claim 14, 15 or 16, comprising taking into account a direction of travel of said viewer, in addition to said indicated position.

10 18. A method according to any of claims 14 to 17, comprising receiving a directional indication representing movement of the viewer, and calculating said indicated position on the basis of at least said directional indication.

15 19. A method according to any preceding claim, wherein said plurality of potential next viewpoint positions are distributed relative to the first viewpoint position across said virtual scene in at least two spatial dimensions.

20 20. A method according to claim 19, wherein said plurality of potential next viewpoint positions are distributed across at least two adjacent quadrants around said first viewpoint position, in said virtual scene.

21. A method according to claim 20, wherein said plurality of potential next viewpoint positions are distributed across four quadrants around said first viewpoint position, in said virtual scene.

25 22. A method according to any preceding claim, wherein at least some of said viewpoint positions are distributed with a substantially constant or substantially smoothly varying average density across a first two-dimensional area in said virtual scene.

23. A method according to claim 22, wherein said at least some of said viewpoint positions are distributed in a regular pattern including a two-dimensional array in said first two-dimensional area.

5 24. A method according to claim 23, wherein said at least some of said viewpoint positions are distributed in a square grid across said first two-dimensional area.

10 25. A method according to claim 23, wherein said at least some of said viewpoint positions are distributed in a non-square grid across said first two-dimensional area.

15 26. A method according to claim 25, wherein said at least some of said viewpoint positions are distributed in a triangular grid across said first two-dimensional area.

20 27. A method according to claim 22, wherein said at least some of said viewpoint positions are distributed in an irregular pattern across said first two-dimensional area.

25 28. A method according to any of claims 22 to 27, wherein said viewpoint positions are distributed with a substantially constant or smoothly varying average density across a second two-dimensional area, said second two-dimensional area being delineated with respect to said first two-dimensional area and the average density in said second two-dimensional area being different to the average density in said first two-dimensional area.

30 29. A method according to any preceding claim, wherein said viewpoint positions are distributed across a planar surface.

30. A method according to any of claims 1 to 28, wherein said  
viewpoint positions are distributed across a non-planar surface.

31. A method according to any preceding claim, wherein said  
5 viewpoint positions are distributed across a three-dimensional volume.

32. A method according to any preceding claim, wherein said first  
image data includes images of one or more objects captured against a  
substantially monochrome background, and said combining comprises laying  
10 said objects over said first scenic viewpoint image data on the basis of detection  
of said substantially monochrome background.

33. A method according to claim 2, 3 or 4, wherein said motion  
picture camera context capture system comprises an image zoom state recording  
15 system for generating image zoom state data, wherein said method comprises:  
on the basis of said image zoom state data, selecting portions of said  
second scenic image data.

34. A method according to claim 2, 3, 4 or 33, wherein said motion  
20 picture camera context capture system comprises an associated image focus  
state recording system for generating image focus state data, wherein said  
method comprises:  
on the basis of said image focus state data, manipulating said first scenic  
viewpoint image data, until said manipulated first scenic viewpoint image data's  
25 focal state is consistent with said first set of image data's focal state.

35. A method according to claims 2, 3 or 4, wherein said motion  
picture camera context capture data includes data related to variables affecting  
the optical imaging properties of the motion picture camera.

36. Computer software arranged to conduct the method of any of claims 1 to 35.

5 37. Apparatus arranged to conduct the method of any of claims 1 to 35.

10 38. A method of capturing image data for subsequently generating a video signal comprising a moving image in the form of a series of playback frames, the moving image representing movement of a viewer through a computer-generated virtual scene, the scene including a representation of a fixed background set and a representation of a moving object within the scene, wherein said computer-generated virtual scene is capable of being generated using captured images of said fixed background by taking said captured images to have different viewpoints relative to said fixed background set, said  
15 viewpoints corresponding to different points of capture, and by combining said captured images with images of said moving object which are filmed at a location separate from the location of said fixed background, the method comprising:

20 capturing a plurality of images of said fixed background set based on the selection of a plurality of points of capture,

wherein at least some of said points of capture are distributed with a substantially constant or substantially smoothly varying average density across a first two-dimensional area.

25 39. A method according to claim 38, wherein said captured images comprise images with a greater than 180° horizontal field of view.

30 40. A method according to claim 38 or 39, wherein said captured images comprise images with a 360° horizontal field of view.

41. A method according to any of claims 38 to 40, wherein said captured images comprise photographic images which have been captured at a plurality of points of capture in a real scene using camera equipment.

5 42. A method according to claim 41, wherein said captured images have been captured using panoramic camera equipment.

43. A method according to claim 41 or 42, wherein said points of capture, at least in some regions of said real scene, are spaced less than 5m  
10 apart.

44. A method according to claim 43, wherein said points of capture, at least in some regions of said real scene, are spaced less than 1m apart.

15 45. A method according to claim 44, wherein said points of capture, at least in some regions of said real scene, are spaced less than 10cm apart.

46. A method according to claim 45, wherein said points of capture, at least in some regions of said real scene, are spaced less than 1cm apart.  
20

47. A method according to any of claims 38 to 46, wherein said capturing comprises recording data defining viewpoint positions in said virtual scene, the viewpoint positions corresponding to the locations of said points of capture in said real scene.  
25

48. A method according to any of claims 38 to 47, wherein said images are captured using an automated mechanically repositionable camera.

49. A method according to claim 48, wherein said automated  
30 mechanically repositionable camera is moved in a regular stepwise fashion across said real scene.

50. A method according to any of claims 38 to 49, wherein said captured images comprise images which are rendered from a computer-generated scene.

5

51. A method according to any of claims 38 to 50, wherein a plurality of points of capture are distributed relative to said first point of capture in at least two spatial dimensions.

10

52. A method according to claim 51, wherein said plurality of points of capture are distributed around said first point of capture.

15

53. A method according to claim 52, wherein said plurality of points of capture are distributed across four quadrants around said first point of capture.

20

54. A method according to any of claims 38 to 53, wherein said at least some of said points of capture are distributed in a regular pattern including a two-dimensional array in said first two-dimensional area.

25

55. A method according to claim 54, wherein said at least some of said points of capture are distributed in a square grid across said first two-dimensional area.

30

56. A method according to claim 55, wherein said at least some of said points of capture are distributed in a non-square grid across said first two-dimensional area.

57. A method according to claim 56, wherein said at least some of said points of capture are distributed in a triangular grid across said first two-dimensional area.

58. A method according to any of claims 38 to 53, wherein said at least some of said points of capture are distributed in an irregular pattern across said first two-dimensional area.

5

59. A method according to any of claims 38 to 58, wherein said points of capture are distributed with a substantially constant or smoothly varying average density across a second two-dimensional area, said second two-dimensional area being delineated with respect to said first two-dimensional area and the average density in said second two-dimensional area being different to the average density in said first two-dimensional area.

10

60. A method according to any of claims 38 to 59, wherein said viewpoint positions are distributed across a planar surface.

15

61. A method according to any of claims 38 to 59, wherein said viewpoint positions are distributed across a non-planar surface.

62. A method according to any of claims 38 to 61, wherein said viewpoint positions are distributed across a three-dimensional volume.

20

63. Computer software arranged to conduct the method of any of claims 38 to 62.

25 64. Apparatus arranged to conduct the method of any of claims 38 to 62.

65. A method of generating a moving image in the form of a series of playback frames for use in a motion picture production environment, wherein said moving image is generated from:

30

a first set of image data captured using a motion picture camera system having associated movement and context capture systems;

5 movement capture data and motion picture camera context capture data generated by said movement capture system and said context capture system respectively; and

a set of sampled light field data of a scene captured using a scene-scanning imaging system which samples the light field of said scene at a plurality of different capture positions, the method comprising:

10 on the basis of said movement capture data selecting a first viewpoint position; and

generating a first scenic viewpoint image data light field rendering, on the basis of said selected first viewpoint position, said motion picture camera context capture data and said captured sampled light field data;

15 combining image data from said first set of image data with said generated first scenic viewpoint image data to generate a first playback frame;

on the basis of said movement capture data, selecting a next viewpoint position relative to said first viewpoint position; and

20 on the basis of said motion picture camera context capture data and said captured sampled light field data, using light field rendering to generate a second scenic viewpoint image data of said scene; and

combining image data from said first image data with said second scenic viewpoint image data to generate a second playback frame.

25 66. A method according to claim 65, wherein said sampled light field data includes scenic image data of said scene; wherein each pixel in said captured light field data is associated with a light ray in said scene's sampled light field.

30 67. A method according to claim 66, wherein said scenic image data comprise photographic images which have been captured at a plurality of different capture positions in a real scene using plenoptic camera equipment.

68. A method according to any of claims 65 to 67, wherein said generating of said first scenic viewpoint image data using ray tracing, includes tracing rays from said selected scenic viewpoint position to said viewpoint  
5 positions of said captured sampled light field data; and

identifying pixels caused by said traced rays from said selected first scenic viewpoint position.

69. A method according to any of claims 65 to 68, wherein a  
10 plurality of potential next sampled light field's viewpoint positions are distributed relative to the first sampled light field's viewpoint position across said light field in at least two spatial dimensions.

70. A method according to claim 69, wherein said plurality of  
15 potential next sampled light field's viewpoint positions are distributed across at least two adjacent quadrants around said first viewpoint position, in said virtual scene.

71. A method according to claim 70, wherein said plurality of  
20 potential next sampled light field's viewpoint positions are distributed across four quadrants around a first sampled viewpoint position, in said light field.

72. A method according to any of claims 65 to 71, wherein said  
25 viewpoint positions are distributed across a three-dimensional volume.

73. A method according to any of claims 65 to 72, wherein said first  
image data includes images of one or more objects captured against a  
substantially monochrome background, and said combining comprises laying  
said objects over said generated first scenic viewpoint image data on the basis of  
30 detection of said substantially monochrome background.

74. A method according to any of claims 65 to 73, wherein said motion picture camera's context capture data includes data related to variables affecting the optical imaging properties of said motion picture camera.

5 75. A method according to claim 74, wherein said motion picture camera's context capture data includes image zoom state data.

76. A method according to claim 74 or 75, wherein said motion picture camera's context capture data includes image focus state data.

10

77. A method according to claim 67, wherein said sampled light field data is captured using an automated mechanically repositionable plenoptic camera.

15 78. A method according to claim 77, wherein said automated mechanically repositionable plenoptic camera is moved in a regular stepwise fashion across said light field of said real scene.

20 79. Computer software arranged to conduct the method of any of claims 65 to 78.

80. Apparatus arranged to conduct the method of any of claims 65 to 78.

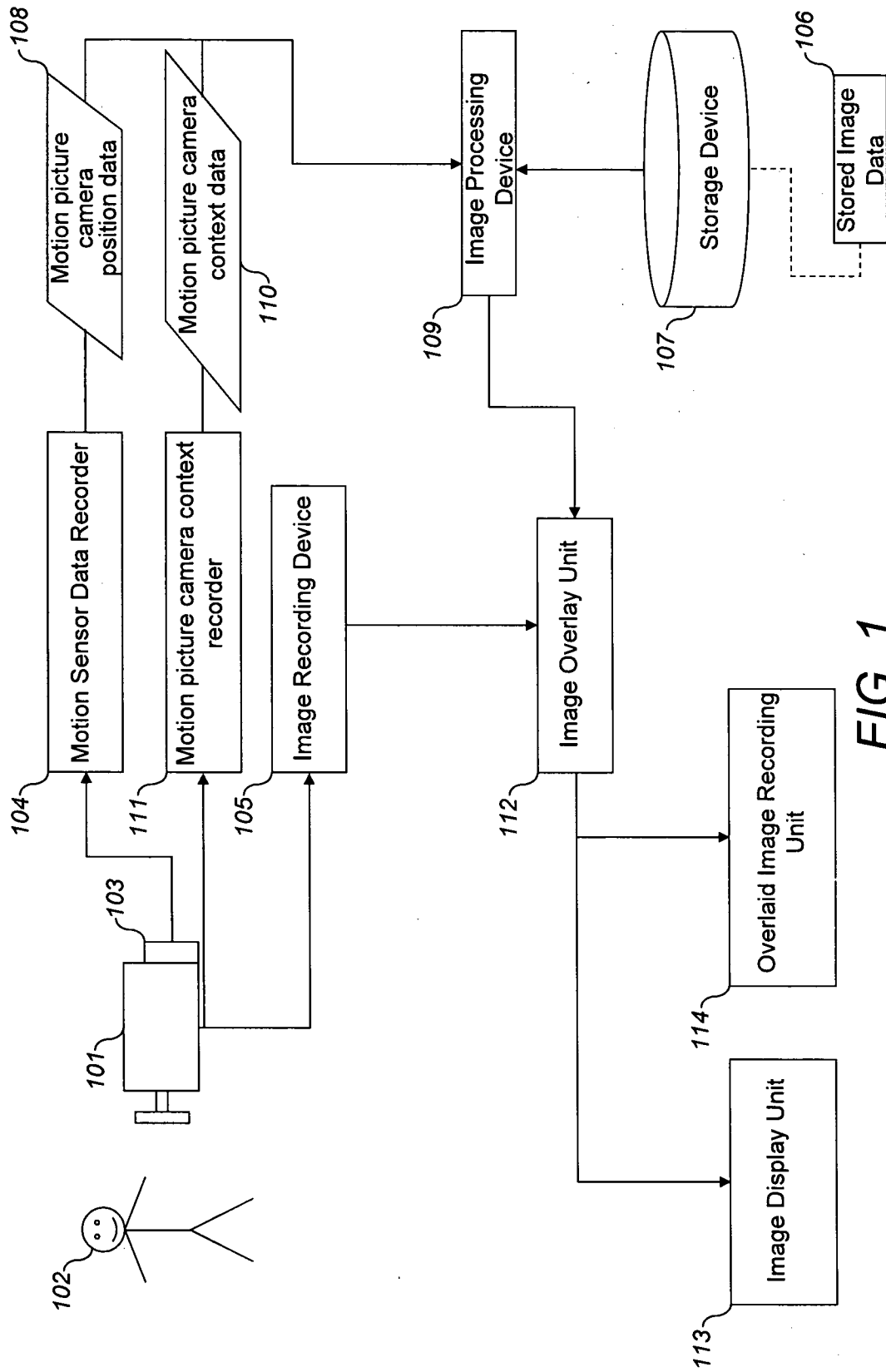


FIG. 1

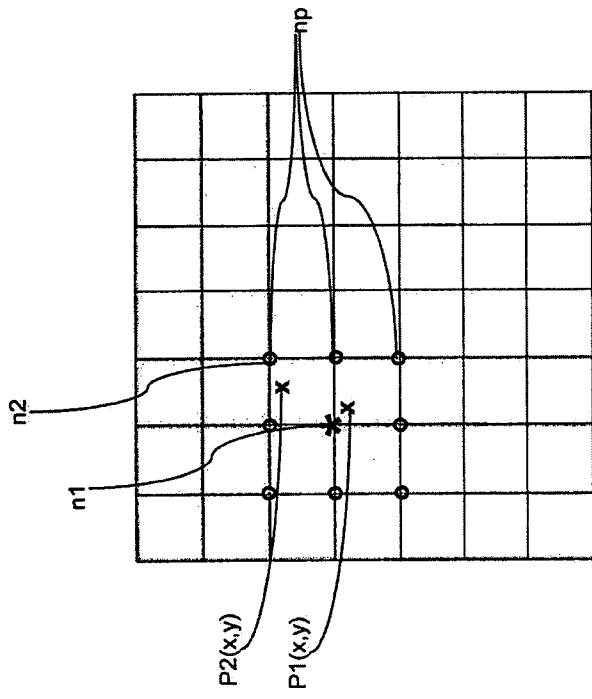


Figure 2

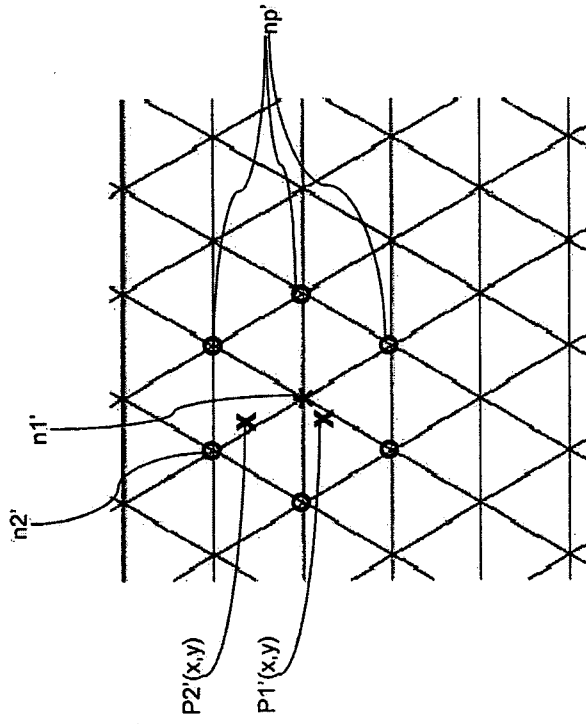


Figure 3

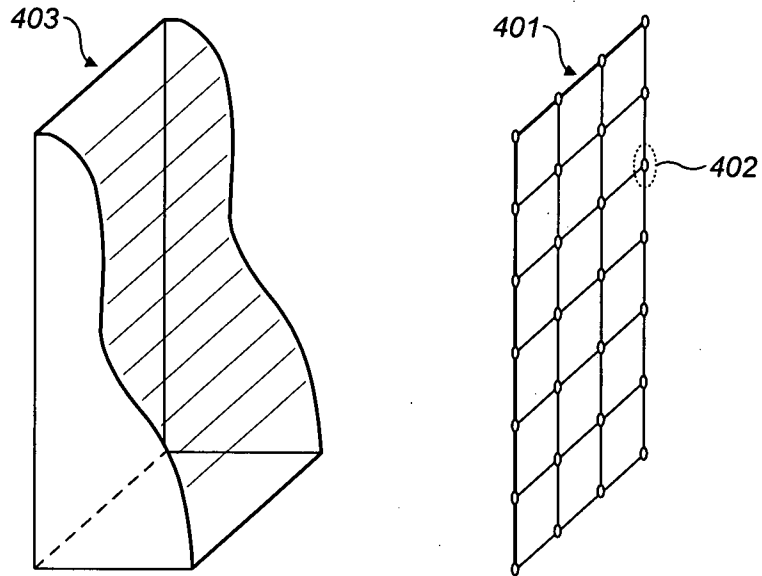


FIG. 4

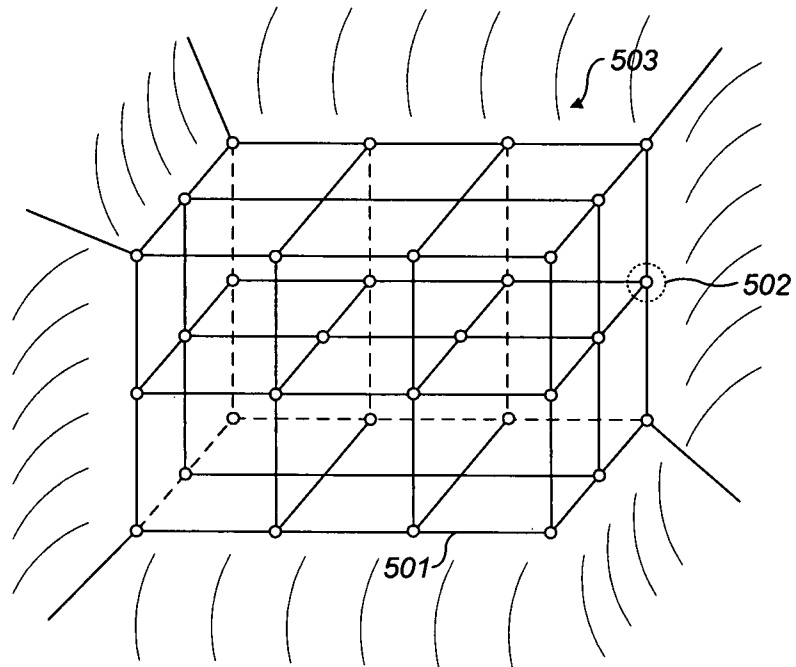


FIG. 5

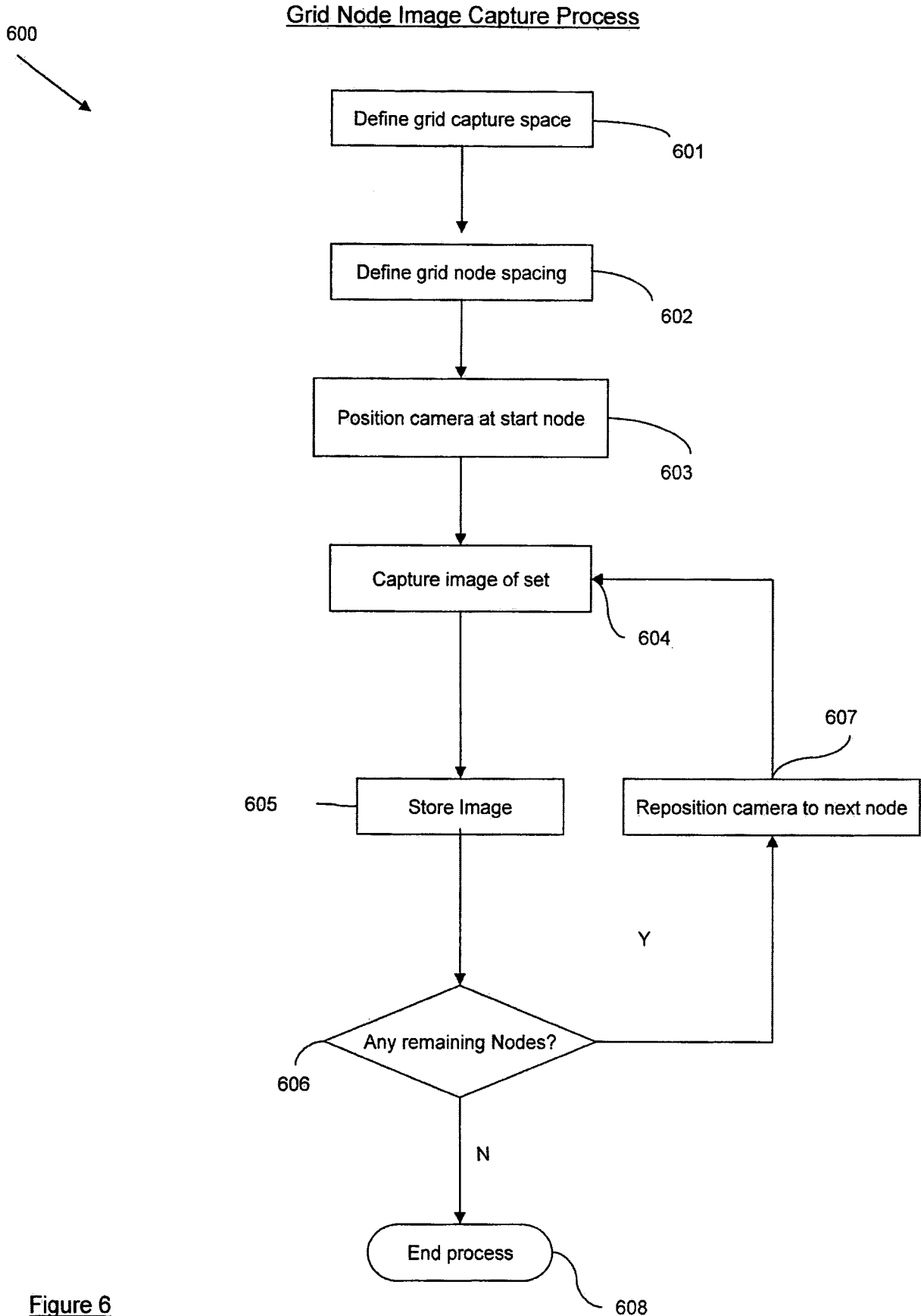


Figure 6



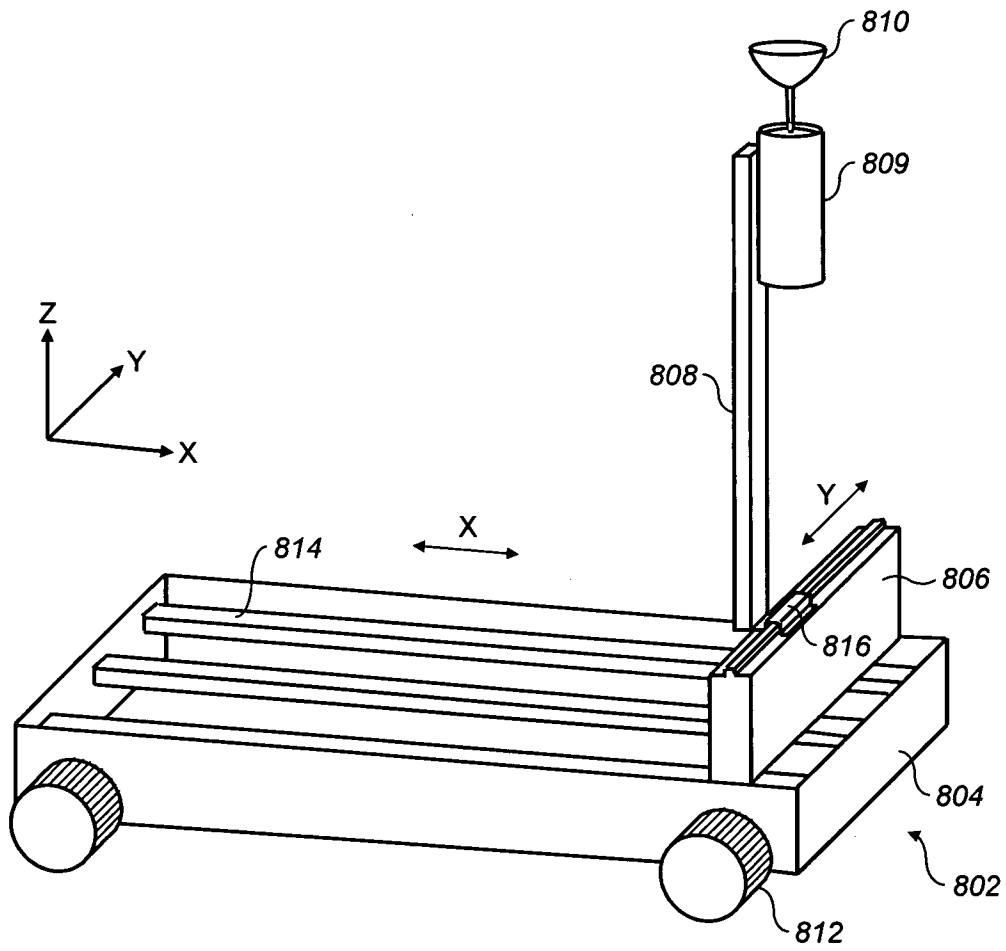


FIG. 8

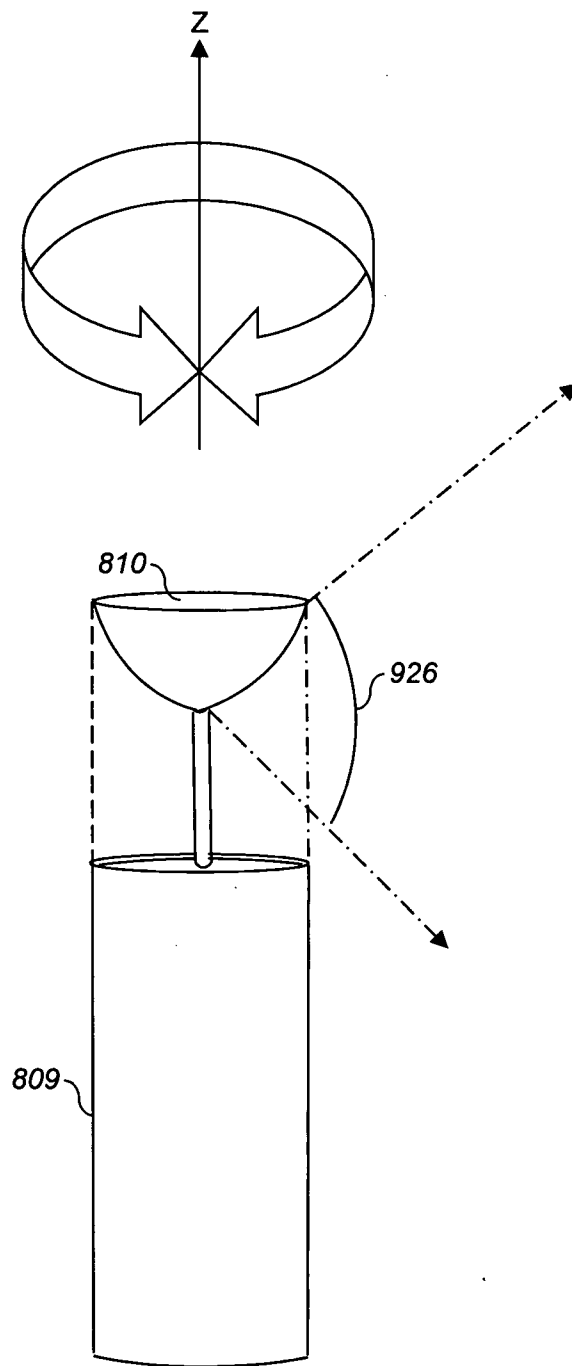


FIG. 9

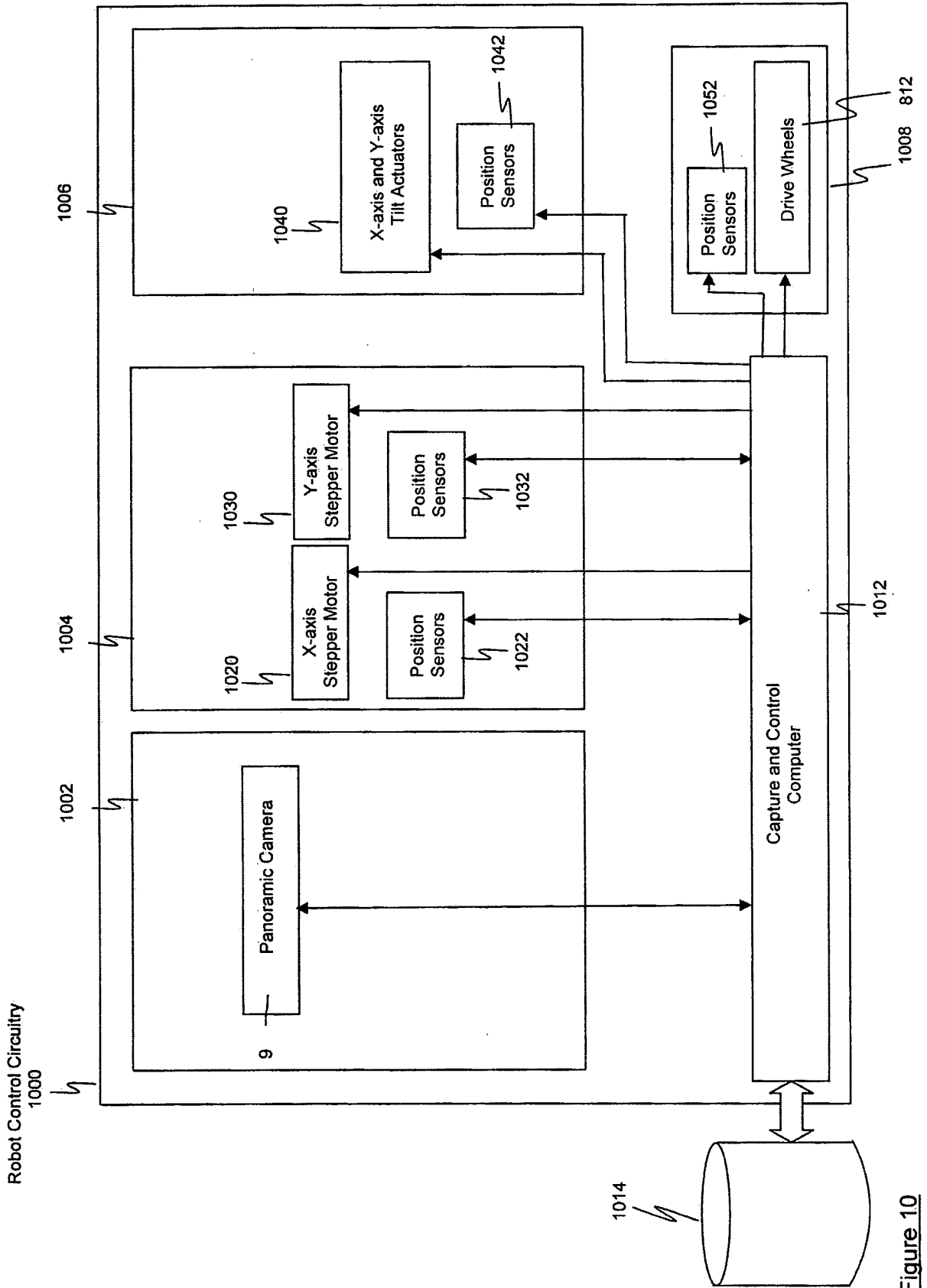


Figure 10

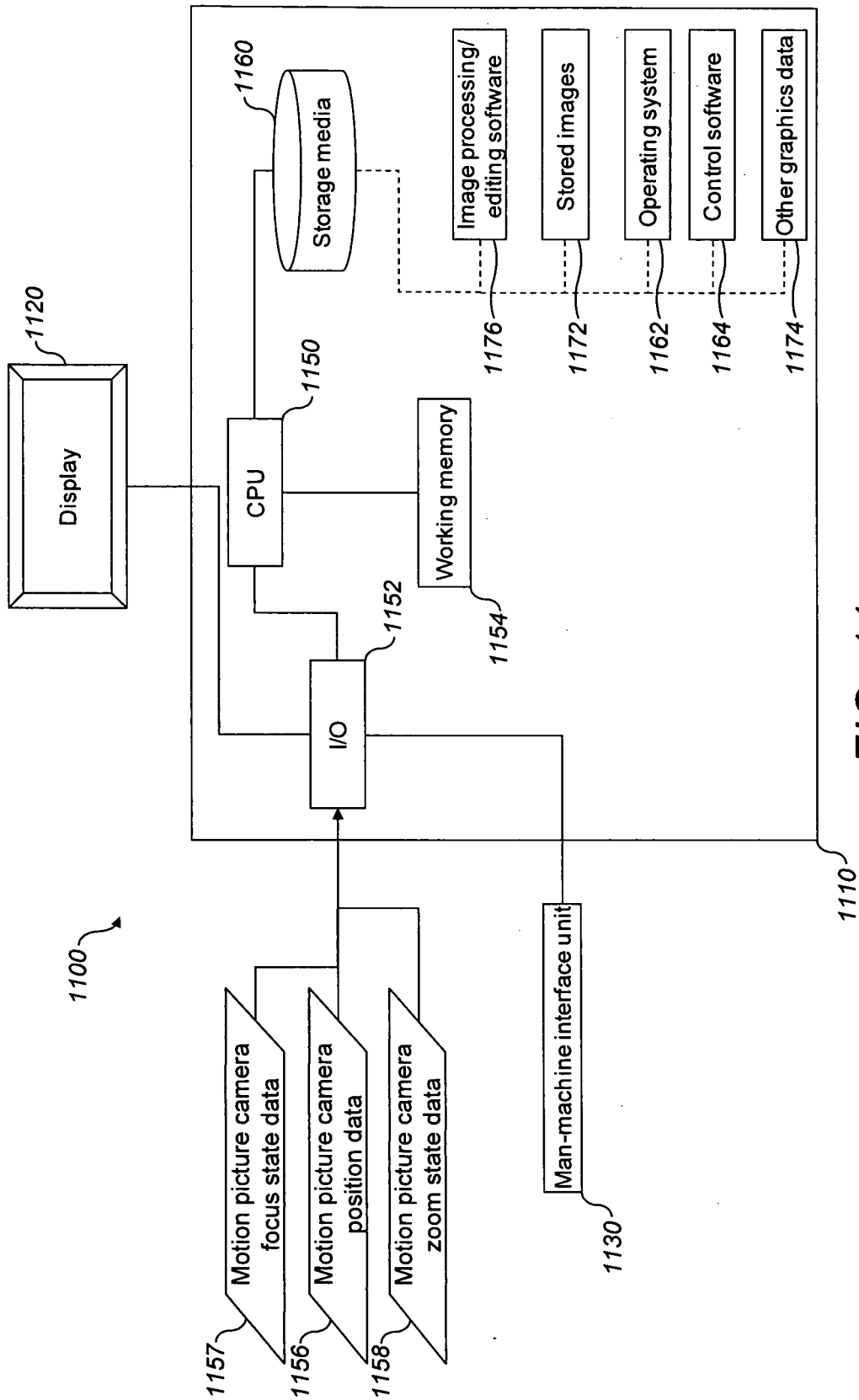
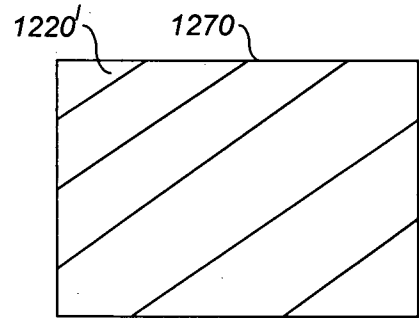
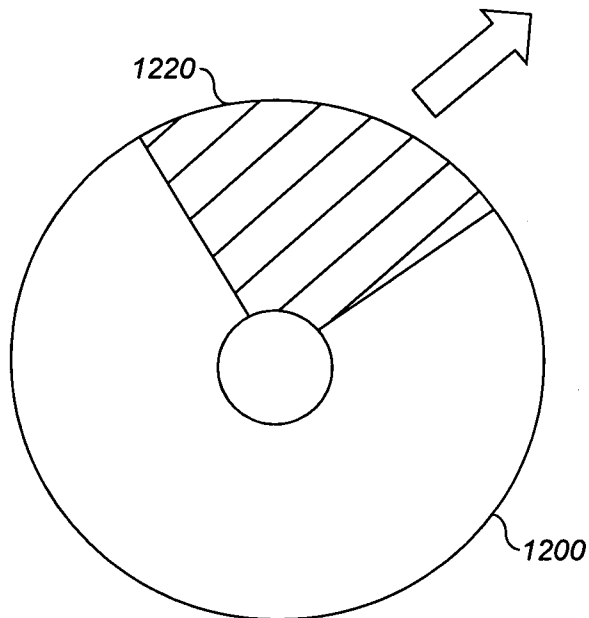


FIG. 11



**FIG. 12b**



**FIG. 12a**

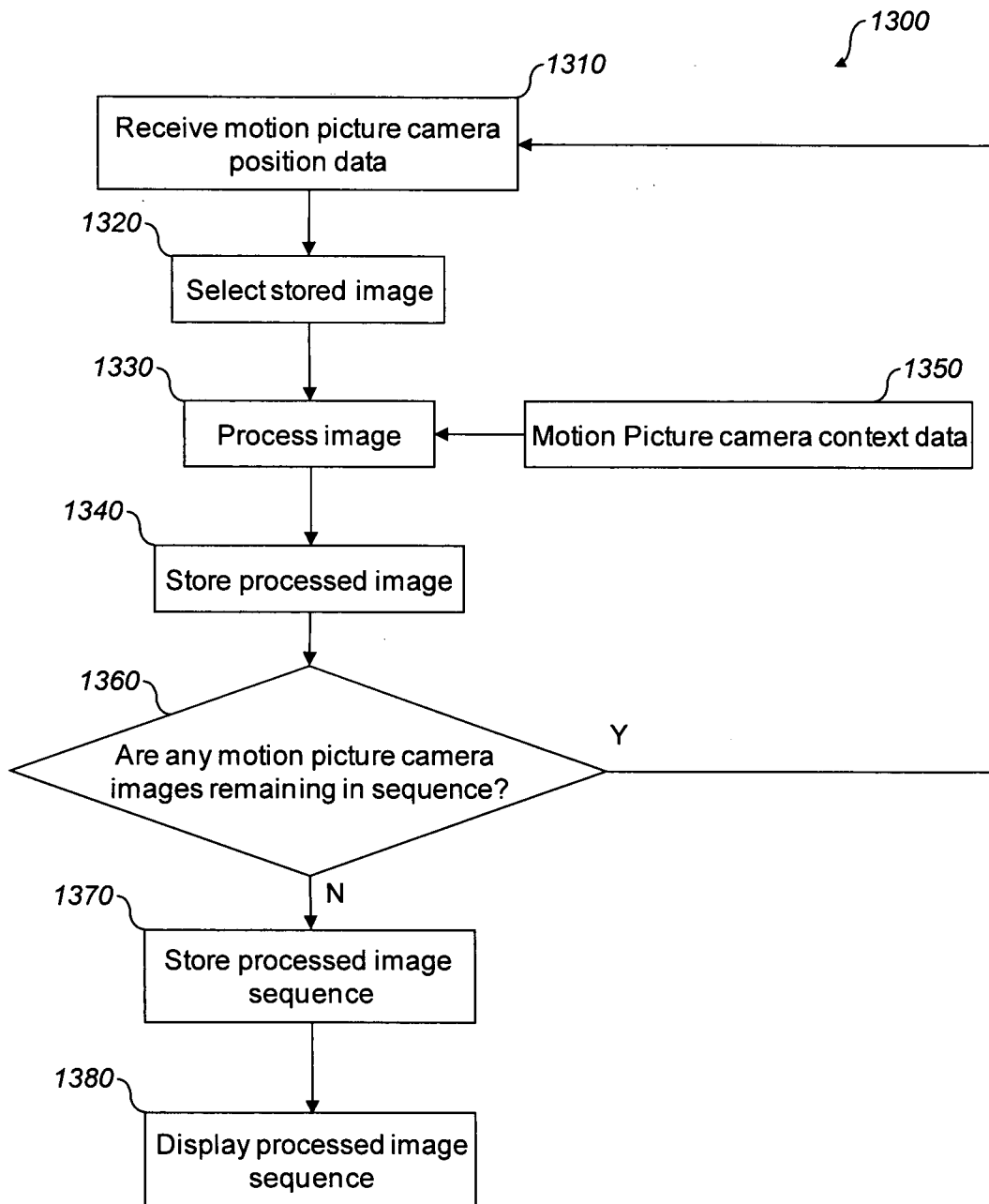


FIG. 13

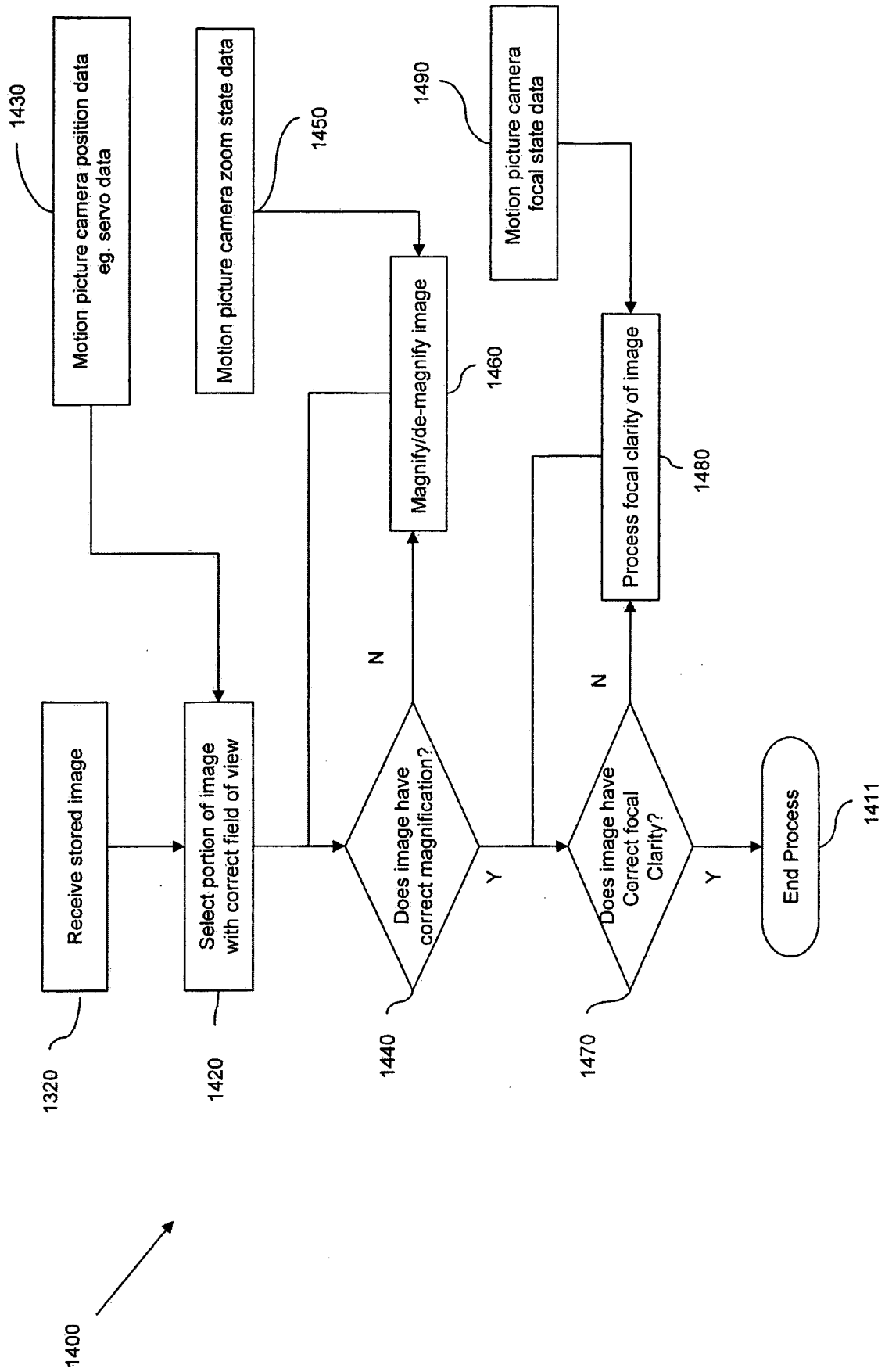


Figure 14

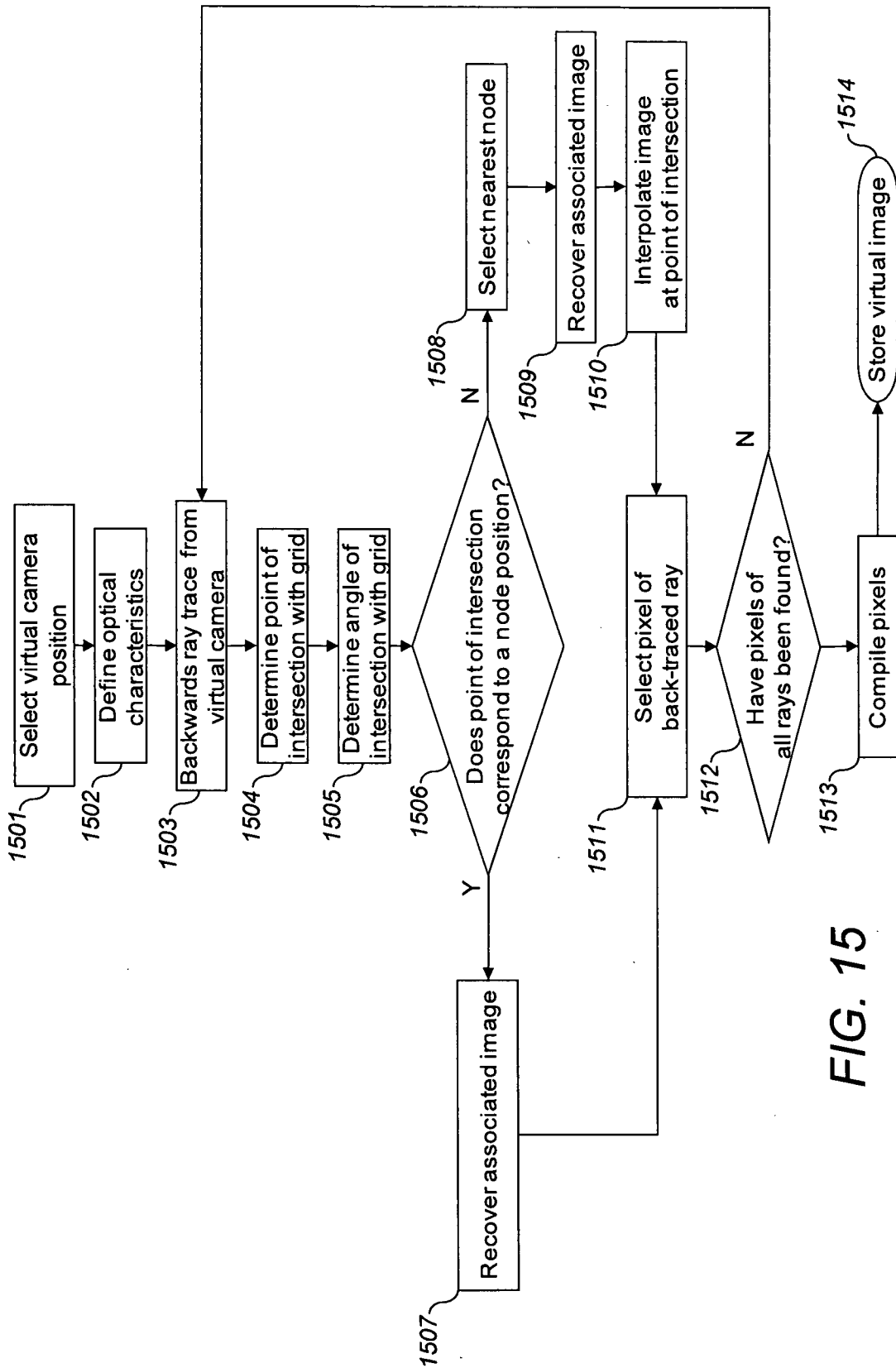


FIG. 15

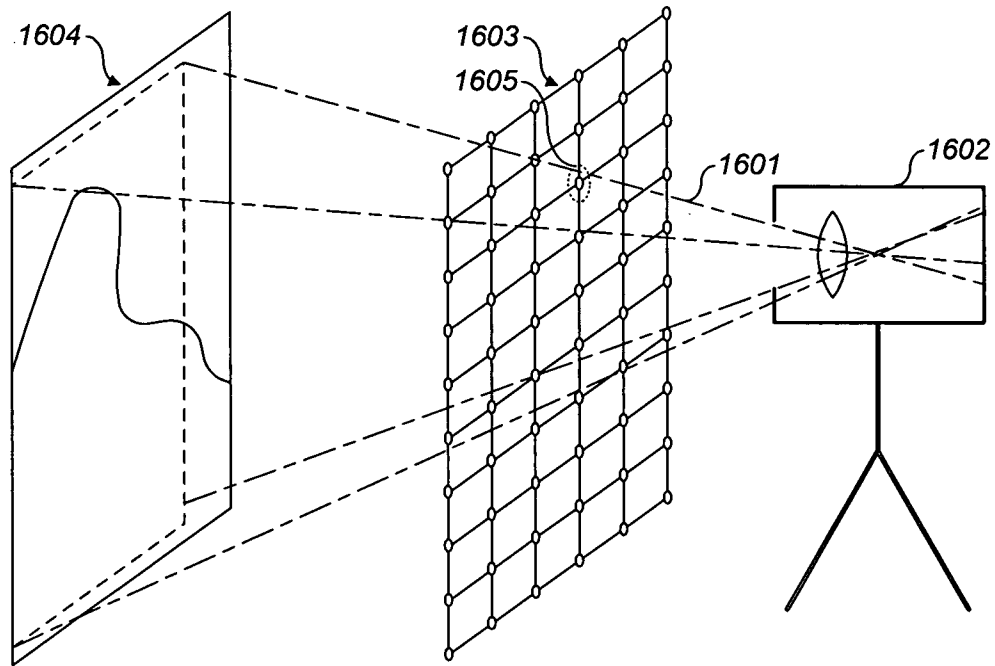


FIG. 16

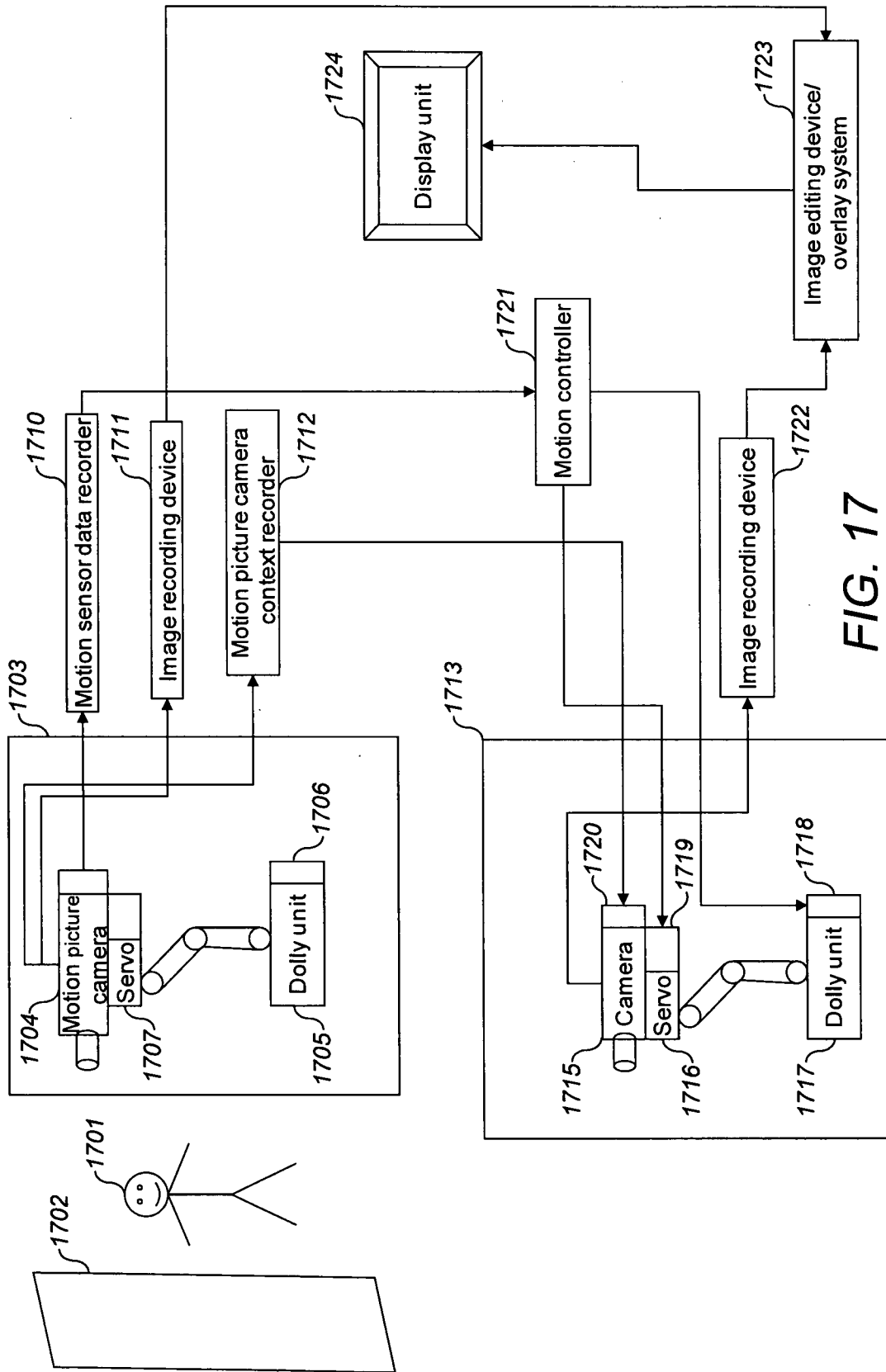


FIG. 17

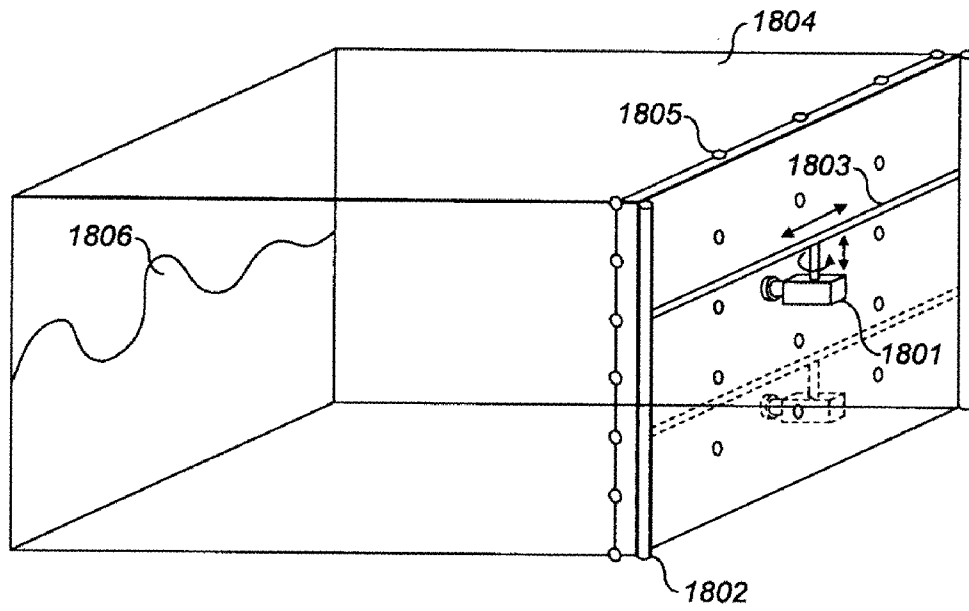


FIG. 18