February 08, 2018

Spherical Light Field Capture with Rotating Camera Array

Paul Debevec

Jessica Busch

Xueming Yu

Daniel Erickson

Spherical Light Field Capture with Rotating Camera Array

Abstract

A light-field capture apparatus is provided that includes a camera housing having a plurality of cameras, each camera providing a different viewpoint of a scene. The apparatus may also include a drop-string mechanism coupled to a top portion of the camera housing and a rotary table pivotably coupled to a portion of the camera housing different than the top portion. The drop-string mechanism includes a string having a first end coupled to the top portion of the camera housing and a second end wound around a center column. The rotary table is rotatable around a center column to change horizontal positions of the camera housing. Additionally, rotating the rotary table causes the string of the drop-string mechanism to unwind around the center column in response to changes in horizontal positions of the camera housing, thereby causing a vertical drop in position of the camera housing.

Background

Conventional cameras render a three-dimensional (3D) scene onto a two-dimensional (2D) sensor. During operation, a conventional digital camera captures a 2D image representing a total amount of light that strikes each point on a photosensor within the camera. However, this 2D image contains no information about the direction of the light that strikes the photosensor. Thus, an image captured by a conventional camera integrates a radiance function over its angular portion, resulting in a two-dimensional intensity as a function of position. The angular information of the original radiance is lost. Thus, conventional cameras fail to capture a large amount of optical information.

Light field cameras capture information about the light field (i.e., vector function that describes the amount of light flowing in every direction through every point in space) emanating from a scene, and in doing so captures information about the intensity of light in a scene, and also the direction distribution of the light rays. Thus, the light field is a 4D record of all light rays in 3D. A light field camera captures radiance; therefore, light-field images originally taken out-of-focus may be refocused, noise may be reduced, viewpoints may be changed, and other light-field effects may be achieved.

Description

Cameras can be configured to have overlapping fields of view such that portions of scenes can be captured by multiple cameras, each from a different viewpoint of the scene. Spatial smoothness between pixels of rendered video frames can be improved by incorporating spatial information from the multiple cameras, such as by correlating pixels (each pixel representing a particular point in the scene) in an image to all other images from cameras that have also captured that particular point in the scene. Thus, light fields may be generated using camera imagery that captures rays of light and its direction from multiple different viewpoints of a scene—thus capturing a hologram of the scene as it would be viewed from within certain volume of 3D space.

In various embodiments, a light field capture apparatus includes a camera housing including a plurality of cameras, in which each camera provides a different viewpoint of a scene. The plurality of cameras is configured within the camera housing in an arc having a semicircular configuration. The plurality of cameras in the arc face outward from a rotation axis, thereby providing a configuration for capturing light field data for a sphere outlined by the cameras. A drop-string mechanism is coupled to a top portion of the camera housing and a rotary table is pivotably coupled to a portion of the camera housing different than the top portion. The drop-string mechanism includes a string having a first end coupled to the top portion of the camera housing and a second end wound around a center column, and the drop-string mechanism further including the center column which coincides with the rotation axis.

By panning the rotary table, the camera housing is rotatable around the rotation axis to change horizontal positions of the camera housing, the portion of the string between the camera housing and the center column increases in length as the second end of the string unwinds around the center column in response to changes in horizontal positions of the camera housing around the rotation axis. This increase in length of the portion of the string between the camera housing and the center column cause a drop in a vertical position of the camera housing. Rotating the arc of cameras about the vertical axis to sweep the cameras through a sampling of outward-looking viewpoints also results in a change in the inclination angles of the cameras. Thus, the drop-string mechanism of the described allows for rotation of the camera array in horizontal panning directions and also allows for vertical tilting to change camera inclination angles, using a single motor, for capturing a spherical light field. This provides the data necessary to render views of the scene from any position within the sphere outlined by the cameras.

FIG. 1 illustrates a side perspective view of a light field capture system 100 in accordance with some embodiments. The light field capture system 100 includes a plurality of cameras 102(1) through 102(N) mounted in an arc configuration facing outward from a rotation axis 104 and directed towards a surrounding 3D scene 106. The plurality of cameras 102 are mounted in a camera housing 108 such that each camera 102(1) through 102(N) captures a sequence of images (e.g., video frames) of the scene 106 and any objects (not shown) in the scene 106. Each camera has a different viewpoint or pose (i.e., location and orientation) with respect to the scene 106.

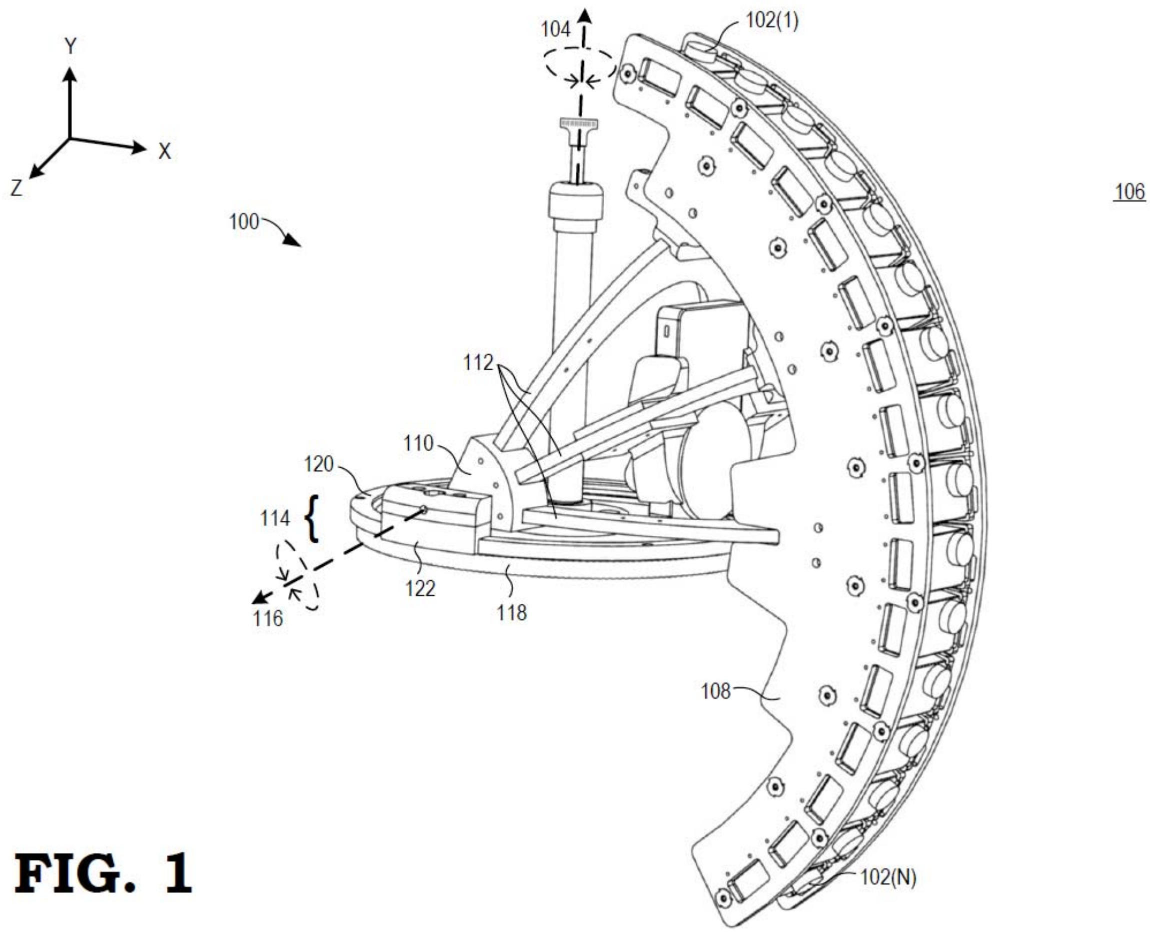

FIG. 1

The camera housing 108 is coupled to a hinge body 110 via one or more support arms 112. The hinge body 110 mounts the camera housing 108 to a rotary table 114, with the hinge body 110 pivotable about an axis 116 that is normal relative to the rotation axis 104. For example, if the rotation axis 104 corresponds to the y-axis of a Cartesian coordinate system as illustrated in this example, the axis 116 corresponds to the z-axis. Accordingly, the rotation axis 104 is interchangeable referred to as the “vertical rotation axis.” Although the embodiments are described here in the context of the rotation axis 104 corresponding to the y-axis and the camera housing 108 rotating about that vertical y-axis (i.e., camera housing sweeping horizontally), the rotation axis 108 can be oriented in other directions for other embodiments. For example, in other embodiments, the rotation axis 104 can be aligned with the z-axis of FIG. 1 such that the camera housing 108 sweeps vertically by rotating about the horizonal z-axis. In the other embodiments, sweeping the camera housing 108 vertically about the horizontal z-axis also causes vertical inclination of the cameras 102 in camera housing 108 to change, thereby resulting in changes to axial tilt.

The rotary table 114 includes a fixed plate 118 and a rotary plate 120 having one or more vertical slabs 122. In some embodiments, the hinge body 110 is coupled to the vertical slab 122 via a pin (not shown) such that the hinge body 110 pivots axially around the vertical slab 122 of the rotary plate 120. Although described in the context of a hinge body 110 to vertical slab 122 connection, any other coupling mechanism for forming a hinging relationship between the camera housing 108 and the rotary plate 120 may be used.

Although FIG. 1 illustrates an example implementation having sixteen cameras (that is, N=16), the number "N" of cameras in light field capture system 100 can include any positive number of cameras and which may account for parameters such as each camera's horizontal and/or vertical field of view, radius of the camera configuration R, etc. Further, a lightfield capture system is not limited to the vertical arc configurations described and various embodiments can include a different number and arrangements (e.g., cameras positioned on different planes relative to each other). For example, in an alternative embodiment, a light field capture system can include a plurality of cameras mounted around a spherical housing rather than in a single-plane, vertical arc configuration as illustrated in FIG. 1. In another alternative embodiment, the light field capture system 100 can include a second camera housing including a second array of cameras positioned opposite the camera housing 108 at a position 180 degrees around the rotary table 114. Such an embodiment provides for increased number of cameras for image capture and also provides for weight balancing in the light field capture system 100.

In various embodiments, the light field capture system 100 includes a motor (not shown) or other rotary mechanism that rotates the rotary plate 120 (and therefore the cameras 102 in camera housing 108) in a circular path 360 degrees about the vertical rotation axis 104. Rotating the cameras 102 through a number of positions results in a series of images of the scene 106. Each of the cameras 102 has a particular field of view that define the outer edges of their respective fields of view. The field of view for each camera overlaps with the field of view of at least one other camera to form a stereoscopic field of view. Accordingly, images from the two cameras can be provided to the viewpoints of a viewer's eyes as a stereoscopic pair for providing a stereoscopic view of objects in the overlapping field of view. However, the overlapping field is not restricted to being shared between only two cameras.

Each pixel in a camera image corresponds to a ray in space and captures light that travels along that ray to the camera. Light rays from different portions of the three-dimensional scene 106 are directed to different pixel portions of 2D images captured by the cameras 102, with each of the cameras 102 capturing the 3D scene 106 visible with their respective fields of view from a different viewpoint. With circular projection models, if rays of all directions from each viewpoint can be captured, a stereoscopic image pair can be provided for any viewing direction to provide for full view coverage that is both stereographic and covers 360-degree coverage of the scene 106. In some embodiments, the camera housing 108 includes a sufficient number of cameras 102 to encompass a 180-degree vertical field-of-view. Accordingly, sweeping the camera housing 108 horizontally 360-degrees enables capture of a full spherical light field image.

The horizontal resolution for a spherical light field captured by light field capture system 100 may be varied based on the distance the camera housing 108 is rotated about the vertical rotation axis 104 between each image capture. For example, the light field capture system 100 may be configured to capture an image for each one degree turn of the camera housing 108. By completing a full rotation, each of the cameras 102 captures 360 outward-looking viewpoints of a spherical, outward-looking light field. Additionally, in some embodiments, the motor of the rotary table 114 rotates the camera housing 108 (e.g., by starting and stopping) at M-number of evenly-spaced divisions around the circular 360-degree coverage of the scene 106, thereby enabling the capture of imagery with no motion blur. Alternatively, in other embodiments, the rotary table 114 rotates the camera housing 108 in a continuously panning motion. By increasing or decreasing the number of images captured for each one degree turn of the camera housing 108 about the vertical rotation axis 104, the horizontal resolution for the spherical light field may be adjusted.

However, with the fixed nature of the cameras 102 within the camera housing 108, the vertical resolution of the captured spherical light field is more limited due to distances between lenses of adjacent cameras. For example, as illustrated in FIG. 1, the different vertical camera orientations come from the semicircular array of sixteen cameras in the camera housing 108. Accordingly, as described below in more detail relative to FIGS. 2-3, the light field capture system 100 further includes a drop-string mechanism that changes the inclination angles of the cameras 102 as the camera housing 108 rotates about vertical rotation axis 104, which enables the light field capture system 100 to capture images with a greater resolution of inclination angles.

In some embodiments, such as described here and further in detail with respect to FIG. 4, the light field capture system 100 further includes an electronic processing device (e.g., electronic processing device 400 of FIG. 4) communicably coupled to the cameras 102. The electronic processing device generates viewpoints using, for example, multi-view synthesis (i.e., more than two images used to generate a viewpoint) by corresponding pixels (each pixel representing a particular point in the scene 106) in an image to all other images from cameras that have also captured that particular point in the scene 106. For any given view (i.e., image captured by one of the cameras 102(N)), the electronic processing device determines the 3D position of that point in the scene 106. Further, the electronic processing device 118 generates a depth map that maps depth distance to each pixel for any given view. In some embodiments, the electronic processing device determines the 3D position of a point in space and its depth information, thereby extending images and/or video captured by the light field capture system 100 to six degrees of freedom (i.e., both head rotation and translation).

FIGS. 2-3 illustrate side views of the light field capture system 100 of FIG. 1 in accordance with some embodiments.

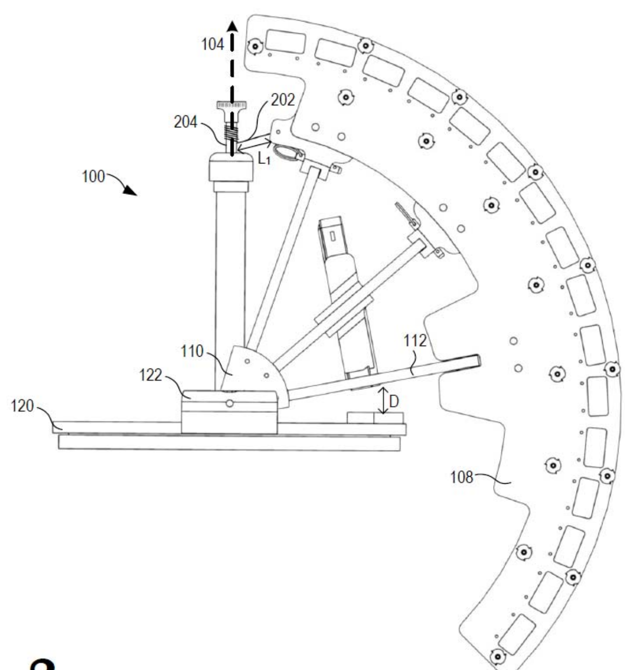

FIG. 2

The light field capture system 100 includes a drop-string mechanism that controls the vertical height of the camera housing 108. The drop-string mechanism includes a string 202 (or any other long flexible structure, such as wires, cords, and the like) that couples the camera housing 108 to a center column 204. As illustrated in FIG. 2, a first end of the string 202 is coupled to a top portion of the camera housing 108 and a second end of the string 202 is wound around the center column 204, leaving a portion of string having length L_1 between the camera housing 108 and the center column 204. Due to the camera housing 108 hanging from the center column 204 while also being hingeably connected to vertical slab 122 (as previously discussed in more detail relative to FIG. 1), the support arm 112 is elevated above the rotary plate 120 by a distance D.

In operation, as the camera housing 108 is rotated about the vertical rotation axis 104, the string 202 also unwinds around the center column 204. This unwinding increases the length of the portion of string between the camera housing 108 and the center column 204, thereby lowering the height of the camera housing 108. Referring now to FIG. 3, illustrated is a side view of the light field capture system 100 of FIG. 2 after the camera housing 108 has been rotated about the vertical rotation axis 104 for a full rotation of 360 degrees.

As shown, the rotation of the camera housing 108 results in unwinding of string 202 such that the portion of string between the camera housing 108 and the center column 204 increases from length L_1 (as shown in FIG. 2, representing length prior to rotation) to a length of $L_1 + L_2$ (as shown in FIG. 3, representing length subsequent to rotation). Due to the camera housing 108 hanging from the center column 204 while also being hingeably connected to vertical slab 122 (as previously discussed in more detail relative to FIG. 1), the increased length of the portion of string between the camera housing 108 and the center column 204 causes the support arm 112 to drop into contact with the rotary plate 120, thereby eliminating the gap having distance D illustrated in

FIG. 2. This dropping of the support arm 112 (and the camera housing 108 coupled thereto) therefore changes the inclination angles of the cameras 102 as the camera housing 108 rotates about vertical rotation axis 104.

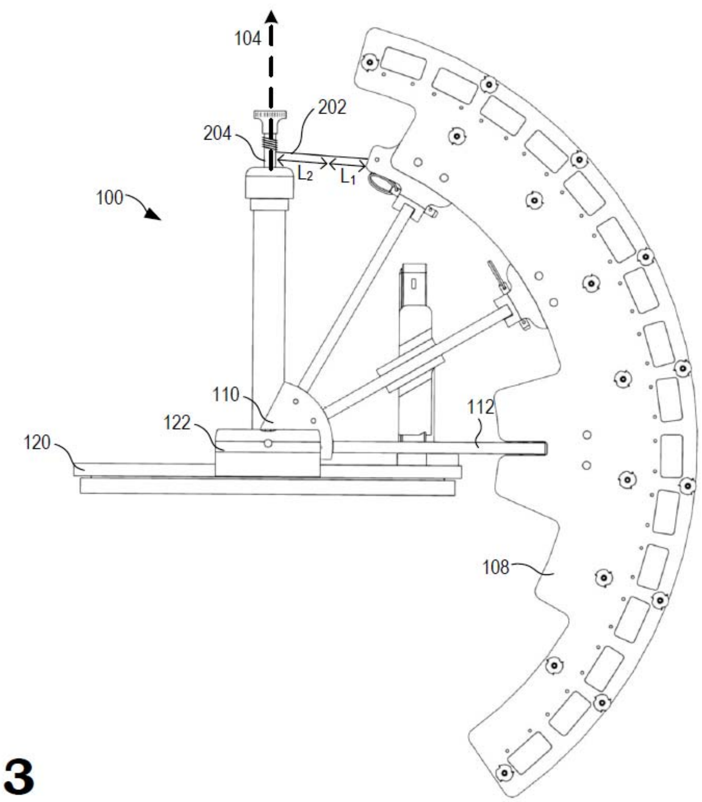

FIG. 3

The change in length L_2 is proportional to the outer circumference of the center column 204. Accordingly, increasing the outer circumference of the center column (i.e., a larger center column) results in larger changes to the length of the portion of string between the camera housing 108 and the center column 204 per rotation as the string 202 unwinds. Similarly, decreasing the outer circumference of the center column (i.e., a smaller center column) results in smaller changes to the length of the portion of string between the camera housing 108 and the center column 204 per rotation as the string 202 unwinds. In various embodiments, the vertical resolution of a captured spherical light field may be increased by rotating the camera housing 108 about the vertical rotation axis 108 multiple times to gradually change the inclination of the cameras 102 as the string 202 unwinds about the center column 204. Thus, as the outer circumference of the center column 204 decreases, the vertical resolution of a spherical light field captured by the light field capture system 100 increases.

Although the embodiments are described relative to an example implementation having sixteen cameras in a single vertical arc array, in other embodiments, the light field capture system 100 can include multiple arrays placed at different azimuths around the rotary plate 120. It should be appreciated that because both the vertical and horizontal positions of each camera 102 changes as the camera housing 108 rotates around the vertical rotation axis 104, changes in the number of cameras and/or camera housings holding camera arrays will change the rate of light field capture and the vertical resolution of the captured light field. For example, in an alternative embodiment, the light field capture system 100 can include a second camera housing including a second array of cameras positioned opposite the camera housing 108 at a position 180 degrees around the rotary table 114. Such an embodiment provides for increased number of cameras for image capture, thereby speeding up image capture and also increasing the vertical resolution.

Thus, the light field capture system 100 described enables spherical light field capture using only a single motor that not only rotates the camera array within the camera housing 108 in a horizontal panning direction (i.e., rotatably around the vertical rotation axis 104), but also changes the vertical orientation (i.e., inclination angles) of the cameras 102 via the drop-string mechanism. In other words, the rotation plane of the motor is perpendicular to the vertical rotation axis (i.e., direction of gravity in this embodiment) and the change in vertical orientation of the cameras 102 (e.g., dropping motion caused by the drop-string mechanism) yields an opposite direction of motion to the spinning. In this matter, the drop-string mechanism enables the camera array within camera housing 108 to achieve a greater resolution of inclination angles without introducing a second motor for tilt, thereby reducing system complexity and weight.

FIG. 4 is a diagram illustrating an example hardware implementation of an electronic processing device 400 configured to operate the light field capture system 100 in accordance with at least some embodiments.

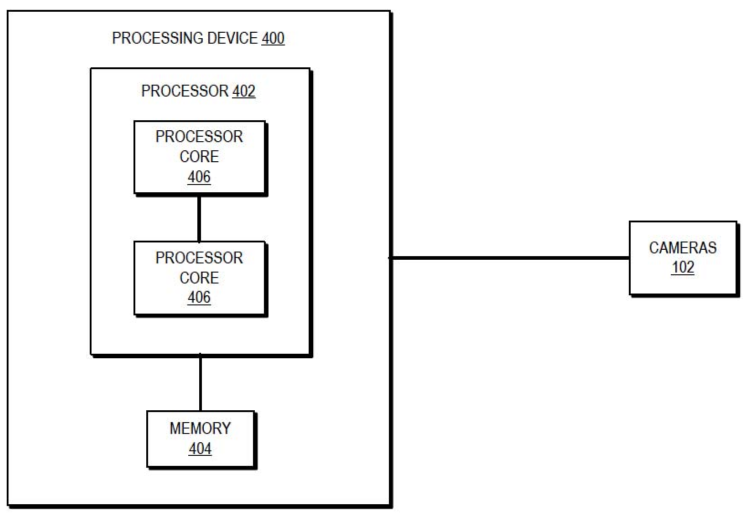

FIG. 4

In the depicted example, the electronic processing device 400 includes a processor 402 and a non-transitory computer readable storage medium 404 (i.e., memory 404). The processor 402 includes one or more processor cores 406. The electronic processing device 400 can be incorporated in any of a variety of electronic devices, such as a server, personal computer, tablet, set top box, gaming system, and the like. The processor 402 is generally configured to execute software that manipulate the circuitry of the processor 402 to carry out defined tasks. The memory 404 facilitates the execution of these tasks by storing data used by the processor 402.

In some embodiments, the software comprises one or more sets of executable instructions stored or otherwise tangibly embodied on the non-transitory computer readable storage medium 404. The software can include the instructions and certain data that, when executed by the one or more processor cores 406, manipulate the one or more processor cores 406 to perform one or more aspects of the operations described. The non-transitory computer readable storage medium 404 can include, for example, a magnetic or optical disk storage device, solid state storage devices such as Flash memory, a cache, random access memory (RAM) or other non-volatile memory device or devices, and the like. The executable instructions stored on the non-transitory computer readable storage medium 404 may be in source code, assembly language code, object code, or other instruction format that is interpreted or otherwise executable by one or more processor cores 406.

FIG. 5 is a flow diagram illustrating a method 500 of capturing a light field in accordance with some embodiments.

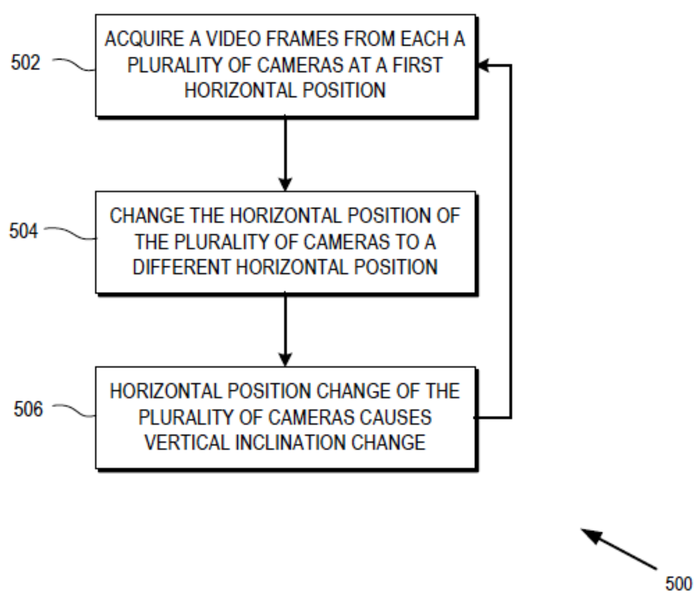

FIG. 5

The method 500 begins at block 502 by acquiring, with a plurality of cameras, a plurality of video frames. Each camera captures a video frame that provide a different viewpoint of a scene, such as described above with respect to cameras 102 of FIG. 1. In some embodiments, the plurality of cameras are mounted in a vertical arc configuration within a camera housing and directed towards a surrounding 3D scene. Each camera captures one or more images (e.g., video frames) of the scene and any objects in the scene at a first horizontal position for the camera housing.

At block 504, the electronic processing device 400 instructs a motor to change the horizontal position of the plurality of cameras mounted within camera housing (e.g., camera housing 108 of FIG. 1). For example, in one embodiment, the electronic processing device 400 instructs the motor to rotate the camera housing 108 about the vertical rotation axis 104 to pan the camera housing 108 by one degree. In other embodiments, the electronic processing device 400 instructs the motor to rotate the camera housing 108 by a different amount. Accordingly, the horizontal resolution for a spherical light field captured by light field capture system 100 may be varied based on the distance the camera housing 108 is rotated about the vertical rotation axis 104 between each image capture. By increasing or decreasing the number of images captured for each one degree turn of the camera housing about the vertical rotation axis 104, the horizontal resolution for the spherical light field may be adjusted.

At block 506, the horizontal change in position of the camera housing (e.g., camera housing 108 of FIG. 1) also results in a vertical change of position for the camera housing. As described in more detail relative to FIGS. 2-3, as the camera housing 108 changes horizontal position around the vertical rotation axis 104, the string 202 of the drop-string mechanism unwinds around the center column 204. This unwinding increases the length of the portion of string between the camera housing 108 and the center column 204, thereby lowering the height of the camera housing 108. Due to the camera housing 108 hanging from the center column 204 while also being hingeably connected to vertical slab 122 (as previously discussed in more detail relative to FIG. 1), the increased length of the portion of string between the camera housing 108 and the center column 204 causes the support arm 112 to drop in position relative to the rotary plate 120. This dropping of the support arm 112 (and the camera housing 108 coupled thereto) therefore changes the inclination angles of the cameras 102 as the camera housing 108 rotates about vertical rotation axis 104. Subsequently, the method 500 returns to block 502 and the electronic processing device 400 repeats this single motor drive process to capture light field images with a plurality of horizontal positions and camera inclinations.

Based on the captured spherical light field, the electronic processing device 400 can render the scene 106 from any view, including omni-directional stereo (ODS) views used for virtual reality video. In one embodiment, the electronic processing device 400 renders a global shutter image of a viewpoint of the scene at one point in time. In another embodiment, the electronic processing device 400 renders a stereoscopic pair of images (e.g., each one having a slightly different viewpoint of the scene) to provide stereoscopic video. The electronic processing device 400 can further stitch the images rendered together to generate ODS video.