

CLAIMS

I/We claim:

1. A camera array, comprising:
a support structure having a center;
a depth sensor mounted to the support structure proximate to the center;
a plurality of cameras mounted to the support structure radially outward from the depth sensor, wherein the cameras are configured to capture image data of a scene; and
a plurality of trackers mounted to the support structure radially outward from the cameras, wherein the trackers are configured to capture positional data of a tool within the scene.
2. The camera array of claim 1 wherein the plurality of cameras includes four cameras, wherein the plurality of trackers includes four trackers, and wherein the cameras and the trackers are positioned symmetrically about the support structure.
3. The camera array of claim 1 wherein the support structure includes a central axis and a longitudinal axis extending orthogonal to the central axis, wherein the depth sensor is aligned along the central axis, and wherein each of the cameras and each of the trackers are equally spaced apart from the central axis and the longitudinal axis.
4. The camera array of claim 1 wherein the trackers are infrared imagers.
5. The camera array of claim 1 wherein the trackers are imagers having a different spectral sensitivity than the cameras.
6. The camera array of claim 1 wherein the cameras and the trackers each have a field of view, wherein the fields of view of the cameras at least partially overlap to define an imaging volume, and wherein the fields of view of the trackers at least partially overlap to define a tracking volume that is larger than the imaging volume.

7. The camera array of claim 1 wherein the cameras are each angled radially inward toward the center of the support structure.

8. The camera array of claim 7 wherein the depth sensor has a focal plane, wherein the cameras each have a focal axis, and wherein the focal axes of the cameras converge at a point below the focal plane of the depth sensor.

9. A mediated-reality system, comprising:

a camera array including—

a support structure having a center;

a depth sensor mounted to the support structure proximate to the center;

a plurality of cameras mounted to the support structure radially outward from the depth sensor, wherein the cameras are configured to capture image data of a scene; and

a plurality of trackers mounted to the support structure radially outward from the cameras, wherein the trackers are configured to capture positional data of a tool within the scene;

an input controller configured to control a position and orientation of a virtual perspective of the scene;

a processing device communicatively coupled to the camera array and the input controller, wherein the processing device is configured to—

synthesize a virtual image corresponding to the virtual perspective based on the image data from at least two of the cameras; and

determine a position of the tool based on the positional data from at least two of the trackers; and

a display device communicatively coupled to the processing device, wherein the display device is configured to display a graphical representation of the tool at the determined position in the virtual image.

10. The mediated-reality system of claim 9 wherein the processing device is further configured to determine the position of the tool based on the image data from at least one of the cameras.

11. The mediated-reality system of claim 10 wherein the processing device is configured to determine the position of the tool by—
estimating an initial three-dimensional (3D) position of the tool based on the positional data from the at least two of the trackers;
defining a region of interest in the image data from the at least one of the cameras based on the initial 3D position of the tool;
processing the image data in the region of interest to determine a position of the tool in the region of interest; and
determining an updated 3D position of the tool based on the determined position of the tool in the region of interest.
12. A method of imaging a subject within a scene, the method comprising:
aligning a depth sensor of a camera array with a portion of interest of the subject such that the portion of interest is positioned proximate a focal depth of the depth sensor;
capturing, with a plurality of cameras of the camera array, image data of the scene including the portion of interest of the subject;
capturing, with a plurality of trackers of the camera array, positional data of a tool within the scene;
receiving input regarding a selected position and orientation of a virtual perspective of the scene;
synthesizing a virtual image corresponding to the virtual perspective of the scene based on the image data;
determining a position of the tool based on the positional data; and
displaying, at a display device, a graphical representation of the tool at the determined position in the virtual image.
13. The method of claim 12 wherein the scene is a surgical scene, and wherein the portion of interest of the subject is a spine of the subject.
14. The method of claim 12 wherein the cameras each have a focal axis, and wherein the focal axes of the cameras converge at a point below the focal depth of the depth sensor.

15. The method of claim 12 wherein determining the position and orientation of the tool is further based on the image data.

16. The method of claim 15 wherein determining the position of the tool includes—
estimating an initial three-dimensional (3D) position of the tool based on the positional data;
defining one or more regions of interest in the image data based on the initial 3D position of the tool;
processing the image data in the one or more regions of interest to determine a position of the tool in the one or more regions of interest; and
determining an updated 3D position of the tool based on the determined position of the tool in the one or more regions of interest.

17. A method for determining the position of a tip of a tool within a scene, the method comprising:
receiving positional data of the tool from at least two trackers;
estimating a three-dimensional (3D) position of the tip of the tool based on the positional data;
receiving an image of the scene from each of one or more cameras;
for the image from each of the one or more cameras—
defining a region of interest in the image based on the estimated 3D position of the tip of the tool; and
processing the image in the region of interest to determine a position of the tip of the tool in the region of interest; and
determining an updated 3D position of the tip of the tool based on the determined position of the tool in the region of interest of the one or more images.

18. The method of claim 17 wherein receiving the image from each of the one or more cameras includes receiving images from corresponding ones of a plurality of cameras, and wherein determining the updated 3D position of the tip of the tool includes triangulating the updated 3D position based on the determined positions of the tip of the tool in the regions of interest in the images from the plurality of cameras.

19. The method of claim 17 wherein estimating the 3D position of the tip of the tool includes—

determining a primary axis of the tool; and
estimating the 3D position based on a known offset along the primary axis.

20. The method of claim 19 wherein receiving the image from each of the one or more cameras includes receiving an image from one camera, and wherein determining the updated 3D position of the tip of the tool includes—

projecting the position of the tip of the tool in the region of interest into a 3D line based on a calibration of the camera; and
determining the updated 3D position of the tip of the tool based on an intersection of the 3D line and the primary axis of the tool.