

CLAIMS

I/We claim:

1. A method of determining the depth of a scene, the method comprising:
capturing depth data of the scene with a depth sensor;
capturing image data of the scene with a plurality of cameras;
generating a point cloud representative of the scene based on the depth data;
identifying a region of the point cloud;
generating depth data for the region based on the image data; and
merging the depth data for the region with the depth data from the depth sensor to
generate a merged point cloud representative of the scene.
2. The method of claim 1 wherein the region of the point cloud is a missing region
of the point cloud in which the point cloud includes no data or sparse data.
3. The method of claim 2 wherein identifying the missing region of the point cloud
includes determining that the missing region of the point cloud has fewer than a predetermined
threshold number of data points.
4. The method of claim 2 wherein identifying the missing region of the point cloud
includes identifying a hole in the point cloud that is larger than a user-defined threshold.
5. The method of claim 2 wherein generating the depth data for the missing region
is further based on a portion of the depth data captured by the depth sensor that surrounds the
missing region.
6. The method of claim 1 wherein the depth data for the region has a greater
resolution than the depth data captured with the depth sensor.
7. The method of claim 1 wherein the method further comprises generating a three-
dimensional mesh representative of the scene based on the merged point cloud.

8. The method of claim 1 wherein the scene is a surgical scene.
9. The method of claim 1 wherein the plurality of cameras each have a different position and orientation relative to the scene, and wherein the image data is light field image data.
10. The method of claim 1 wherein the method further comprises:
processing the image data and the merged point cloud to synthesize an output image of the scene corresponding to a virtual camera perspective; and
transmitting the output image to the display for display to a user.
11. The method of claim 10 wherein the display is a head-mounted display worn by the user, and wherein identifying the region of the scene is based on at least one of a position and an orientation of the head-mounted display.
12. A system for imaging a scene, comprising:
multiple cameras arranged at different positions and orientations relative to the scene and configured to capture image data of the scene;
a depth sensor configured to capture depth data about a depth of the scene; and
a computing device communicatively coupled to the cameras and the depth sensor, wherein the computing device has a memory containing computer-executable instructions and a processor for executing the computer-executable instructions contained in the memory, and wherein the computer-executable instructions include instructions for—
receiving the image data from the cameras;
receiving the depth data from the depth sensor;
generating a point cloud representative of the scene based on the depth data;
identifying a region of the point cloud;
generating depth data for the region based on the image data; and
merging the depth data for the region with the depth data from the depth sensor to generate a merged point cloud representative of the scene.

13. The system of claim 12 wherein the region of the point cloud is a missing region of the point cloud in which the point cloud includes no data or sparse data.
14. The system of claim 12 wherein the region of the point cloud is user-selected.
15. The system of claim 12, further comprising a display, wherein the computing device is communicatively coupled to the display, and wherein the computer-executable instructions further include instructions for—
processing the image data and the merged point cloud to synthesize an output image of the scene corresponding to a virtual camera perspective; and
transmitting the output image to the display for display to a user.
16. The system of claim 15 wherein identifying the region of the scene is based on at least one of a position and an orientation of the display.
17. A method of determining the depth of a scene, the method comprising:
capturing depth data of the scene with a depth sensor;
generating a point cloud representative of the scene based on the depth data;
identifying a region of the point cloud;
registering the point cloud with three-dimensional (3D) medical scan data; and
merging at least a portion of the 3D medical scan data with the depth data from the depth sensor to generate a merged point cloud representative of the scene.
18. The method of claim 17 wherein the region of the point cloud is a missing region of the point cloud in which the point cloud includes no data or sparse data.
19. The method of claim 18 wherein the scene is a medical scene including a portion of a patient, wherein the missing region of the point cloud corresponds to the portion of the patient, and wherein the portion of the 3D medical scan data portion corresponds to the same portion of the patient.
20. The method of claim 17 wherein the 3D medical scan data is a computed tomography (CT) data.