

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

MEDIT CORPORATION,

Petitioner,

v.

3SHAPE A/S,

Patent Owner.

Case No.: IPR2023-01313

U.S. Patent No. 9,629,551

DECLARATION OF HASSAN FOROOSH, Ph.D.

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	QUALIFICATIONS	1
III.	MATERIALS CONSIDERED	9
IV.	TECHNOLOGY BACKGROUND.....	13
A.	Intraoral Scanning Technologies.....	13
1.	Passive Triangulation-Based Technologies	14
2.	Active Triangulation-Based Technologies	15
3.	Time-of-Flight-Based Technologies	17
4.	Confocal Scanning-Based Technologies	17
5.	Other Known Technologies	18
B.	From Raw Data to 3D Data.....	18
C.	From 3D Data to 3D Model	19
D.	Simultaneous Localization and Mapping (SLAM).....	21
E.	Background Modeling: Moving Object Exclusion	22
F.	Medical Applications of SLAM.....	23
V.	OVERVIEW OF THE '551 PATENT	24
A.	Specification of the '551 Patent	24
B.	Prosecution History	30
VI.	LEGAL PRINCIPLES USED IN ANALYSIS	32
A.	Prior Art.....	32
B.	Anticipation	32
C.	Obviousness.....	33
D.	Claim Construction	37
VII.	LEVEL OF ORDINARY SKILL IN THE ART	39
VIII.	CLAIM CONSTRUCTION	40
A.	The Preambles of the Independent Claims Are Limiting (Claims 1, 23, 25).....	40
B.	“excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation”	41

IX.	OVERVIEW OF THE PRIOR ART	41
A.	Newcombe	41
B.	Herbst	46
C.	Hähnel.....	51
D.	McQueston	53
X.	THE PRIOR ART RENDERS ALL OF THE CLAIMED ELEMENTS OF THE '551 PATENT OBVIOUS.....	57
A.	Claims 1, 6-12, 15-17, and 20-25 Are Rendered Obvious by Newcombe and Hähnel	57
1.	Claim 1 Is Rendered Obvious by Newcombe in view of Hähnel	57
2.	Claim 6 Is Rendered Obvious by Newcombe in view of Hähnel	75
3.	Claim 7 Is Rendered Obvious by Newcombe in view of Hähnel	76
4.	Claim 8 Is Rendered Obvious by Newcombe in view of Hähnel	77
5.	Claim 9 Is Rendered Obvious by Newcombe in view of Hähnel	78
6.	Claim 10 Is Rendered Obvious by Newcombe in view of Hähnel	80
7.	Claim 11 Is Rendered Obvious by Newcombe in view of Hähnel	81
8.	Claim 12 Is Rendered Obvious by Newcombe in view of Hähnel	83
9.	Claim 15 Is Rendered Obvious by Newcombe in view of Hähnel	84
10.	Claim 16 Is Rendered Obvious by Newcombe in view of Hähnel	85
11.	Claim 17 Is Rendered Obvious by Newcombe in view of Hähnel	86
12.	Claim 20 Is Rendered Obvious by Newcombe in view of Hähnel	87
13.	Claim 21 Is Rendered Obvious by Newcombe in view of Hähnel	88

14.	Claim 22 Is Rendered Obvious by Newcombe in view of Hähnel	90
15.	Claim 23 Is Rendered Obvious by Newcombe in view of Hähnel	91
16.	Claim 24 Is Rendered Obvious by Newcombe in view of Hähnel	109
17.	Claim 25 Is Rendered Obvious by Newcombe in view of Hähnel	111
B.	Claims 2-5, 18, and 19 Are Rendered Obvious By Newcombe in View of Hähnel, Further in View of McQueston	128
1.	Claim 2 Is Rendered Obvious By Newcombe in View of Hähnel, Further in View of McQueston	130
2.	Claims 3-5, 18, and 19 Are Rendered Obvious By Newcombe in View of Hähnel, Further in View of McQueston	130
C.	Claims 1, 6-12, 15-17, and 20-25 Are Rendered Obvious by Herbst	131
1.	Claim 1 Is Rendered Obvious by Herbst	131
2.	Claim 6 Is Rendered Obvious by Herbst	148
3.	Claim 7 Is Rendered Obvious by Herbst	150
4.	Claim 8 Is Rendered Obvious by Herbst	152
5.	Claim 9 Is Rendered Obvious by Herbst	155
6.	Claim 10 Is Rendered Obvious by Herbst	156
7.	Claim 11 Is Rendered Obvious by Herbst	157
8.	Claim 12 Is Rendered Obvious by Herbst	159
9.	Claim 15 Is Rendered Obvious by Herbst	160
10.	Claim 16 Is Rendered Obvious by Herbst	162
11.	Claim 17 Is Rendered Obvious by Herbst	163
12.	Claim 20 Is Rendered Obvious by Herbst	164
13.	Claim 21 Is Rendered Obvious by Herbst	165
14.	Claim 22 Is Rendered Obvious by Herbst	166
15.	Claim 23 Is Rendered Obvious by Herbst	167
16.	Claim 24 Is Rendered Obvious by Herbst	186
17.	Claim 25 Is Rendered Obvious by Herbst	187
D.	Claims 2-5, 18, and 19 Are Rendered Obvious By Herbst in View of McQueston	205

1.	Claim 2 Is Rendered Obvious By Herbst in View of McQueston	206
2.	Claims 3-5, 18, and 19 Are Rendered Obvious By Herbst in View of McQueston.....	207
XI.	DEMONSTRATION OF WRITTEN DESCRIPTION SUPPORT FOR MCQUESTON’S EFFECTIVE FILING DATE OF FEBRUARY 22, 2011	207
XII.	CONCLUSION.....	225

I, Hassan Foroosh, declare as follows:

I. INTRODUCTION

1. I have been retained by Medit Corporation (“Petitioner”) as an independent expert in this *inter partes* review before the United States Patent and Trademark Office (“PTO”) regarding U.S. Patent No. 9,629,551 (“the ’551 patent”), Ex-1001. Specifically, I have been asked to consider whether certain references disclose or suggest the features recited in claims 1-12 and 15-25 of the ’551 patent (“challenged claims”). As explained below, it is my opinion that the challenged claims are invalid as anticipated and/or obvious in view of the prior art references discussed in this Declaration.

2. I am being compensated at a consulting rate of \$550 per hour. No part of my compensation is dependent on the outcome of this proceeding or any other proceeding involving the ’551 patent or any of its related patents. I have no other interests in this proceeding.

II. QUALIFICATIONS

3. My qualifications are generally set forth in my Curriculum Vitae, which is attached as Ex-1004. My Curriculum Vitae also includes a list of the publications I have authored and a list of the other cases in which I have testified during the last four years.

4. I am a Professor of Computer Science at the University of Central Florida. My primary research interests include image/video processing, computer vision, image-based modeling, including 3D modeling, and machine learning (in particular deep learning), with applications in computer vision and natural language processing.

5. I received my Ph.D. in Computer Science from INRIA (French National Institute for Research on Informatics and Control), University of Nice Sophia Antipolis in 1996. I also earned a Master of Science in Computer Science from the University of Nice Sophia Antipolis in 1993, a Master of Science in Electrical Engineering from Kingston University, U.K., in 1990, and a Bachelor of Science in Electrical Engineering from Kingston University, U.K., in 1989.

6. Prior to research and teaching, I worked as an R&D Engineer at Innovative Engineering Solutions Ltd. in London from 1990 to 1992. Between 1993 and 1996, I served as a Research Assistant in the Scene Analysis and Symbolic Image Processing Group of INRIA, where I performed research in areas such as 3D super-resolution, imaging processes for registration, and shape perception. From 1997 to 2000, I served as an Assistant Research Scientist at the University of Maryland, College Park, where I investigated topics such as image registration (2D and 3D), super-resolution of video data for unmanned aerial vehicles, super-

resolution of Synthetic Aperture Radar (SAR) and Foliage Penetrating SAR (FOPEN-SAR), Automatic Target Detection and Tracking (ATD/R). During this time, I authored one of the most cited works on registration and alignment based on phase correlation: (Hassan Foroosh (Shekarforoush), Josiane Zerubia and Marc Berthod, Extension of Phase Correlation to Sub-pixel Registration, IEEE Transactions on Image Processing, vol. 11, Issue 3, pp. 188-200, 2002). Between 2000 to 2002, I served as Senior Research Scientist at the University of California, Berkeley where I also served as the Technical Coordinator and Principal and Research Scientist of a Multi-University Research Initiative (MURI) funded by the Office of Naval Research. As part of the project, I conducted research on multi-sensor acquisition of 3D data for autonomous 3D modeling and visualization.

7. I joined the Department of Computer Science at the University of Central Florida (UCF) as an Assistant Professor in 2002. Since then, I have served as an Associate Professor, Professor, and CAE Link Professor of Computer Science in the same department. I also served in various capacities in the department, such as the Graduate Program Coordinator, chair of various leadership committees, and most recently in the Faculty Senate Strategic Planning Council for UCF. I have taught a number of courses directed to computer modeling and 3D modeling, including 3D Computer Vision, Image Processing, Special Topics in 3D Computer

Vision, and Visual Simulation.

8. I have conducted extensive research in the field of 3D computer vision and computer modeling. Some example research projects that I have conducted include *Video Analysis Techniques for Terrain Modeling and Automated Texture Mapping* (2006-2007) and *Modeling and Visualization of Musculoskeletal Structures in Dental CT-Scan* (2006-2007). For the latter project, I developed algorithms for 3D registration, modeling and visualization of maxilla and mandible of patients for dental surgery using Computer Aided Tomography (CAT) scanning.

9. I have been the principal investigator or co-principal investigator of numerous research projects funded by the National Science Foundation (NSF), NASA, DoD, and industry, with total funding exceeding \$14M since joining UCF. As one example, I recently served as the principal investigator in two projects funded by the National Science Foundation (NSF) on 3D modeling and the role of 3D information in machine understanding of digitized world. One project funded by industry was directed to the registration/alignment of dental CT scans of patients over a period of time during the treatment of a patient, with the purpose of improving the quality of 3D models, e.g. corrections and noise removal in 3D CT scans.

10. I have authored or co-authored over 190 articles in peer reviewed journals, conference proceedings, and book chapters, many of which relate to 3D

reconstruction and metrology from uncalibrated camera images and depth cameras (such as RGB-D and LiDAR), as well as on combining shape-from-X and super-resolution (e.g., shape-from-stereo, shape-from-shading, and shape-from-photometric stereo). I also have several papers and a patent on optical system designs using dynamic focusing for optical coherence tomography (OCT) 3D scanning, as well as free-form optical surface design. Some examples include: Supraja Murali, Apurva Jain, Kye-sung Lee, Panomsak Meemon, Hassan Foroosh, Kevin Thompson, and Jannick Rolland, *Superresolution imaging combining the design of an optical coherence microscope objective with liquid-lens based dynamic focusing capability and computational methods*, Proc. of SPIE Conf. on Novel Optical Systems Design and Optimization XI, SPIE Vol. 7061 (2008); Ozan Cakmakci, Gregory E. Fasshauer, Hassan Foroosh, Kevin P. Thompson, and Jannick P. Rolland, *Meshfree Approximation Methods for Free-Form Surface Representation in Optical Design with Applications to Head-Worn Displays*, Proc. SPIE Conf. on Novel Optical Systems Design and Optimization XI, vol. 7061 (2008); O. Cakmakci, S. Vo, H. Foroosh, and J.P. Rolland, *Application of Radial Basis Functions to Shape Description in a Dual-Element Off-Axis Magnifier*, Optics Letters, Vol. 33, Issue 11, pp. 1237-1239 (2008); O. Cakmakci, B. Moore, H. Foroosh, and J.P. Rolland. *Optimal Local Shape Description for Rotationally Non-*

Symmetric Optical Surface Design and Analysis, Optics Express, Vol. 16, Issue 3, pp. 1583-1589 (2008); *Systems and Methods for Designing Optical Surfaces*, Patent pending, Ozan Cakmakci, Brendon Moore, Hassan Foroosh, and Jannick Rolland, US 2009/002825A1 (2009).

11. In addition, I have written several peer reviewed articles on camera self-calibration and distortion correction, including self-calibration and distortion correction for varying focal length cameras. Some example publications include: Xiaochun Cao and Hassan Foroosh, *Camera Calibration without Metric Information Using an Isosceles Trapezoid*, Proceedings of the IAPR International Conf. On Pattern Recognition (ICPR), Volume 1, pages 104-107 (2004); Xiaochun Cao, Jiangjian Xiao, Hassan Foroosh, and Mubarak Shah, *Self-Calibration from Turn Table Sequence in Presence of Zoom and Focus*, Computer Vision and Image Understanding (CVIU), 102(3):227-237 (June 2006); Nazim Ashraf, Imran Junejo, and Hassan Foroosh, *Near-Optimal Mosaic Selection for Rotating and Zooming Video Cameras*, Proc. Of Asian Conference on Computer Vision (ACCV), Vol. II, 63-72 (2007); Imran Junejo and Hassan Foroosh, *Optimizing PTZ Camera Calibration from Two Images*, Machine Vision and Applications (MVA), pp. 1-15 (Feb. 2011); Imran N. Junejo and Hassan Foroosh, *Refining PTZ Camera Calibration*, International Conference on Pattern Recognition (ICPR) (2008); Nazim

Ashraf and Hassan Foroosh, *Robust Auto-Calibration of a PTZ Camera with Non-overlapping FOV*, International Conference on Pattern Recognition (ICPR) (2008); and Imran Junejo and Hassan Foroosh, *Practical PTZ Camera Calibration Using Givens Rotations*, Proc. IEEE International Conference on Image Processing (ICIP) (2008).

12. I have also worked in the area of intraoral imaging and image processing, where I directed a Ph.D. student on this subject (described more below) and was Principal Investigator for the project *Modeling and Visualization of Musculoskeletal Structures in Dental CT-Scan* (2006-2007) funded by Imaging Sciences International, Inc. This project developed algorithms for 3D modeling and visualization of maxilla and mandible of patients of dental surgery using Computer Aided Tomography (CAT) scanning. My work for this project is described in M. Balci, M. Alnasser, and H. Foroosh, *Alignment of Maxillofacial CT Scans to Stone-Cast Models Using 3D Symmetry for Backscattering Artifact Reduction*, Proceedings of Medical Image Understanding and Analysis Conference (MIUA) (2006).

13. Since joining UCF, I have advised and chaired the dissertation of over 30 doctoral students, 25 of whom have already successfully defended their doctoral dissertation, and secured academic positions or industry jobs in major high-tech

corporations. As one example, I served as the advisor for a dissertation directed to “*Sub-pixel Registration in Computational Imaging and Applications to Enhancement of Maxillofacial CT Data*”. As I noted above, this research project focused on registration and alignment of 2D images and 3D digitized data obtained by dental CT scans and dental imprints of a patient (*e.g.*, casts), with the aim of improving the quality of CT scans, *e.g.* eliminating backscattering issues. Methods were developed to register CT models to the physical imprints of patients (*e.g.*, a stone cast model of the patient’s teeth).

14. I have been serving as the editor and/or committee member on numerous professional organizations. For instance, I have been involved with the following organizations: IEEE Transactions on Image Processing (Associate Editor, 2003-2008, 2011-2016); IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Member, Technical Committee, 2004-2012, 2015-2016, 2019-2020); IEEE Transactions on Pattern Analysis and Machine Intelligence (Reviewer, 1996-present).

15. For my work, I have been honored with numerous awards, including the *Piero Zamperoni Award*, International Association for Pattern Recognition (IAPR) (2004) for my work in Xiaochun Cao and Hassan Foroosh, *Camera Calibration without Metric Information Using an Isosceles Trapezoid*, Proceedings

of the IAPR International Conf. On Pattern Recognition (ICPR), Volume 1, pages 104-107 (2004). I have also received several Best Paper awards, including “*The Best Scientific Paper Award*” in the IAPR International Conference on Pattern Recognition in 2008, “*The Best Paper Award*” in the IEEE International Conference on Image Processing (ICIP) in 2018, and the *Distinguished Researcher*, UCF College of Engineering and Computer Science (2004-2005, 2005-2006)).

16. In sum, for decades, I have studied, taught, practiced, and conducted research in the area of 3D computer vision, covering pinhole camera model, camera self-calibration, camera distortion correction, pose estimation, 3D reconstruction, and many other related topics. I have also studied, taught, practiced, and conducted research in the area of image processing, covering various imaging modalities, electro-optical, sonar, radar, etc., algorithms for 3D reconstruction from projections, and many other related topics.

III. MATERIALS CONSIDERED

17. My opinions, as set forth in this Declaration, are based on the documents I have reviewed, my experience and background, and my knowledge and professional judgment. In forming these opinions, I have reviewed the '551 patent (Ex-1001), the prosecution history for the '551 patent (Ex-1002), the prior art discussed below, and any other material and information identified in this

Declaration. I have also relied on my experience and knowledge in the field of computer vision, in particular in 3D scanning, 3D modeling, registration and alignment methods (including intraoral), motion analysis (both in video and 3D scanners), and other related problems considered in computer vision and image processing, as well as computer memory and data storage devices.

18. In particular, my opinions regarding the invalidity of the challenged claims are based on the following prior art references, as informed by my knowledge in the field.

Ex. No.	Description
1001	U.S. Patent No. 9,629,551 (“the ’551 patent”)
1002	Excerpts of the File History for U.S. Patent No. 9,629,551
1005	U.S. Patent No. 8,401,225 (“Newcombe”)
1006	Herbst, Evan, Peter Henry, Xiaofeng Ren, and Dieter Fox. “Toward Object Discovery and Modeling via 3-D Scene Comparison.” <i>2011 IEEE International Conference on Robotics and Automation</i> (vol. 3, pp. 2623-2629). 7 vols. Piscataway, NJ: IEEE, ©2011 (“Herbst”)
1007	Hähnel, Dirk, Rudolph Triebel, Wolfram Burgard, and Sebastian Thrun. “Map Building with Mobile Robots in Dynamic Environments.” <i>2003 IEEE International Conference on Robotics and Automation</i> (vol. 2, pp. 1557-1563). 3 vols. (“Hähnel”)
1008	U.S. Patent No. 9,245,374 (“McQueston”)
1009	U.S. Provisional Patent Application No. 61/445,499 (“McQueston Provisional”)

Ex. No.	Description
1011	Order on Claim Construction, <i>3Shape A/S v. Medit Corp.</i> , Case No. W-22-CV-00443-ADA-DTG, Dkt. No. 76 (W.D. Tex. May 2, 2023)
1012	F. Duret, B. Péliissier, Different Dental CAD/CAM Impression Methods, EMC (Elsevier Masson SAS, Paris), Médecine buccale, 28-740-R-10, 2010.
1013	R.I. Hartley and P. Sturm, Triangulation. Computer Vision and Image Understanding, vol. 68, Issue 2, pp. 146–157, 1997.
1014	N. Molton, and M. Brady. Practical Structure and Motion from Stereo When Motion is Unconstrained. International Journal of Computer Vision, vol. 39, Issue 1, pp: 5-23, 2000.
1015	S. Demey, A. Zisserman, and P. Beardsley, Affine and Projective Structure from Motion, Proc. of the British Machine Vision Conference, 1998.
1016	K. E. Ozden, K. Schindler and L. van Gool, Multibody Structure-from-Motion in Practice, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 6, pp. 1134-1141, June 2010.
1017	Jason Geng, Structured-light 3D surface imaging: a tutorial, Advances in Optics and Photonics, vol. 3, pp. 128-160, 2011.
1018	P. Henry, M. Krainin, E. Herbst, X. Ren, D. Fox, RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments, International Symposium on Experimental Robotics, 2010.
1019	K. Granström, Karl and T.B. Schön. Learning to close the loop from 3D point clouds, Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp: 2089-2095, 2010.
1020	U. Franke, C. Rabe, H. Badino, S. Gehrig, 6d-vision: Fusion of stereo and motion for robust environment perception, Pattern Recognition, pp. 216-223, 2005.

Ex. No.	Description
1021	M. Oshita and T. Matsunaga, Automatic learning of gesture recognition model using SOM and SVM, <i>Advances in Visual Computing</i> , pp. 751-759, 2010.
1022	N. Hautiere, R. Labayrade, M. Perrollaz and D. Aubert, Road Scene Analysis by Stereovision: a Robust and Quasi-Dense Approach, <i>International Conference on Control, Automation, Robotics and Vision</i> , pp. 1-6, 2006.
1023	R.A. Drebin, L. Carpenter, P. Hanrahan, Volume rendering, <i>Proceedings of the 15th annual conference on Computer graphics and interactive techniques</i> , 1988.
1024	H. Pfister, M. Zwicker, J. van Baar, M.Gross, Surfels: Surface Elements as Rendering Primitives, <i>SIGGRAPH 2000</i> .
1025	N. Karlsson, E. Di Bernardo, J. Ostrowski, L. Goncalves, P. Pirjanian, M. Munich, The vSLAM Algorithm for Robust Localization and Mapping. <i>Int. Conf. on Robotics and Automation (ICRA)</i> , 2005.
1026	Y. Jin, L. Tao, H. Di, N.I. Rao, G. Xu, Background modeling from a free-moving camera by Multi-Layer Homography Algorithm, <i>IEEE International Conference on Image Processing</i> , pp. 1572-1575, 2008.
1027	H. Kawachi, Y. Kawachi, C. Ikeda, R. Takagi, A. Katakura and T. Shibahara, Oral and Maxillofacial Surgery with Computer-assisted Navigation System, <i>Bulletin of Tokyo Dental College</i> , vol. 51, Issue 1, pp: 35-39, 2010.
1028	P. Mountney, D. Stoyanov, A. Davison, G.-Z. Yang, Simultaneous Stereoscope Localization and Soft-Tissue Mapping for Minimal Invasive Surgery, <i>International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)</i> , pp.: 347-54, 2006.

IV. TECHNOLOGY BACKGROUND

19. The '551 patent discloses methods and systems that come under the general umbrella of computer vision, with specific focus on intraoral scanning and 3D modeling of oral cavity in a dynamic setting, i.e., in the presence of moving objects in a static background. The background topics described briefly below provide a concise overview of some of the important elements of computer vision that are relevant to understanding what a person of ordinary skill in the art ("POSITA") would have known in these areas at the time of the invention of the '551 patent, which for purposes of this declaration I will assume to be the patent's claimed priority date of July 15, 2011.

A. Intraoral Scanning Technologies

20. This section gives a brief description of the state of the art for intraoral scanning technologies that were widely known to a POSITA at the time of filing of the '551 patent. As described below in Sections IV.B and IV.C, the same technologies were also widely known to a POSITA for 3D modeling and mapping in other application areas, such as robotics, where the scene being scanned and modeled was also dynamic, i.e. with both a static background and moving objects.

21. Intraoral scanning technologies that were known before the priority

date of '551 patent were based on a wide variety technologies. *See* Ex-1012. The list below highlights some of the important ones that are directly related to the prior art references used in this declaration.

1. Passive Triangulation-Based Technologies

22. These are scanning technologies that rely on standard RGB imaging technologies to capture a sequence of 2D images (or 2D image pairs), which are then used to create 3D models. The backbone of these scanning technologies is a classical method of computer vision known as stereovision, or more generally Structure from Motion (SfM). SfM refers to a class of 3D modeling methods in computer vision where a camera (or a pair of cameras) is freely moved in the scene to capture a sequence of images (or image pairs) from different positions and orientations. The goal in SfM is to determine from the images a 3D model of the scene and also determine the camera pose for each image (i.e. camera positions and orientations). A single pair of images captured by a stereovision system may be considered as the special case of SfM, where the number of views is limited to two, and the two views are taken simultaneously from distinct positions and orientations, where the physical distance between the two cameras is referred to as the stereoscopic baseline.

23. The main technique to determine the 3D structure of the scene from the images is known as the triangulation method, *see* Ex-1013, which for a calibrated

stereo system would provide a dense Euclidean depth map or equivalently a point cloud. This technique is referred to as being passive, since the imaging and triangulation rely completely on incoherent and unstructured lighting. As pointed out above, a generalization of passive stereovision is a freely moving stereo system used for 3D reconstruction. This corresponds to the use of stereovision-based intraoral scanners, where a stereo camera system is freely moved in the oral cavity for taking a sequence of stereo images (image pairs) from different positions and orientations. In computer vision, this is sometimes referred to as a binocular SfM, or structure and motion from stereo, *see* Ex-1014, as opposed to monocular SfM. *See* Ex-1015. The generalization to dynamic settings, including multiple moving objects is referred to as a multibody SfM. *See* Ex-1016.

2. Active Triangulation-Based Technologies

24. Unlike passive triangulation, these intraoral scanning technologies rely on actively illuminating the oral cavity with a light pattern, which is used by a stereovision system to determine the exact correspondences for solving the triangulation problem. *See* Ex-1017. Due to their nature, these technologies are often referred to as structured lighting methods. The structured light pattern projected in the oral cavity can be in the form of a collection of parallel lines or equidistant points, concentric patterns, or other patterns. But, regardless of the light

structure or the wavelength, the underlying concept remains the same, i.e. at each instant in time the projected pattern is used to streamline the triangulation process for a pair of calibrated cameras with a fixed baseline or across the sequence of stereo images as the scanner is moved around in the mouth. As the scanner is moved around inside the oral cavity different sections of mouth are scanned and therefore are combined to produce a full 3D model. The projected structured light in these scanners may be in an invisible wavelength (e.g., infrared), and the stereo cameras operating in that wavelength provide through triangulation a depth map or a point cloud that is used for 3D modeling. When such technologies are equipped with an additional RGB camera to provide also standard color images, the scanner is referred to as an RGB-D camera (where D stands for depth).

25. Outside intraoral scanning also well-known RGB-D scanning technologies were already developed at the time the '551 patent was filed, using the exact same principles for 3D modeling of indoor environments. The most popular of such technologies was probably the RGB-D scanner known as Microsoft Kinect developed in 2010 by Microsoft Corporation, which was based on an RGB-D camera developed in 2005 by PrimeSense. Although structured lighting as a scanning technology was known way before PrimeSense, *see* Ex-1017, PrimeSense popularized the technology by making it readily available off the shelf for gaming

and robotics applications. As described below these RGB-D cameras were widely studied in applications such as user interface design, robotics, and computer vision for 3D modeling of dynamic indoor scenes, i.e., scenes comprising of both static background and moving objects.

3. Time-of-Flight-Based Technologies

26. These technologies have the huge advantage that the 3D point cloud is directly provided by a transceiver that measures how long a wave (either electromagnetic, e.g. a light beam, or mechanical, e.g. ultrasound) takes to travel back and forth between the scanned objects and the transceiver. These types of scanning technologies typically rely on an active coherent light beam, but alternative versions may be based on ultrasound. These technologies are essentially working on the same principle as their counterparts used in applications such as mobile robotics and autonomous navigation, e.g. for modeling indoor or outdoor environments with both static and moving objects, e.g. on the road for autonomous driving.

4. Confocal Scanning-Based Technologies

27. A fourth class of technologies known at the time of '551 patent was based on the principle of confocal scanning. These technologies capture a stack of images as a lens (or possibly a group of lenses) is shifted along the optical path to

change the depth of field. Since the moving lens brings different points of the scanned surface into focus in a sensor array, by encoding the moving lens along the optical path, one can record the depth for different points on the scanned surface, and therefore produce a 3D point cloud for the scanned surface.

5. Other Known Technologies

28. Other intraoral scanning technologies that were known to a POSITA at the time when the '551 patent was filed are discussed in Exhibit 1012, including for instance the technologies based on interferometry or Moiré patterns. Most of these 3D scanning technologies were also known to be used at the time in applications such as robotics (stereovision, RGB-D, LiDAR), *see* Ex-1017, Ex-1019, Ex-1020, user interface design (stereovision, RGB-D), *see* Ex-1021, autonomous driving (stereovision, LiDAR), *see* Ex-1022, as well as other applications for 3D modeling of a dynamic scene, i.e. a scene with both static background and moving objects.

B. From Raw Data to 3D Data

29. Different 3D scanning technologies rely on different principles and modalities to capture raw data. Therefore, the backend processing of raw data in order to convert the raw data into 3D data varies from one technology to another. For instance, in 3D scanning technologies that rely on time of flight or confocal scanning, some aspects of producing 3D data are streamlined by the hardware,

reducing thus some of the backend processing of raw data. For time-of-flight technologies backend processing of data is mostly focused on filtering steps of the transmitted and reflected pulses, reduction of interference, and resolving multiple returns. Similarly, in confocal scanning the backend processing is mostly directed towards determining in-focus points in a stack of images and correction of distortions caused by lenses and a non-planar depth of field. For scanning technologies based on passive triangulation (e.g. stereovision and SfM), on the other hand, the backend processing is mainly focused on obtaining dense correspondences across images, in order to determine a depth map using the classical triangulation method, *see* Ex-1013, or a wide variety of SfM methods proposed in the literature. *See* Ex-1014, Ex-1015, Ex-1016.

C. From 3D Data to 3D Model

30. Except for time-of-flight technologies that directly produce a point cloud from sensor measurements, three-dimensional data produced by other scanning systems are typically in the form of a so-called depth map. A depth map is essentially an array of pixels with a pixel value representing a depth value. Since in the scanning technologies described above the scanning system capturing the depth map (i.e. the optics and the sensor array) would be calibrated, one can convert a depth map to a point cloud. Both depth maps and point clouds are considered as

3D representations of a scanned surface. Although, such 3D representations can be used for computer-based analysis and processing of 3D information, they are often considered as inconvenient for applications requiring visualization and navigation, e.g. in medical imaging, robotics, immersive environments (virtual/augmented reality), games, and simulations and training, to name a few. Therefore, conversion of 3D representations (e.g. point clouds) to 3D models (or 3D virtual models) has been a standard post-processing step in all these application areas, and widely studied, in particular by the computer graphics community.

31. The techniques for converting a point cloud 3D representation to a virtual 3D model have been driven by the tradeoff between hardware constraints (such as memory, processing, rendering time) and visual/physical fidelity (e.g. photo-realistic rendering, level of details, seamless navigation). The choice of 3D modeling may also be affected by the application area and even the choice of the scanning system. The 3D modeling techniques are broadly divided into two categories of volumetric models, *see* Ex-1023, and surface models. *See* Ex-1024. Regardless of the choice of the modeling technique, as mentioned above, the purpose in this step is to make 3D data usable for other applications. However, it is important to point out that there is no universally accepted description of what may or may not be considered as a 3D model. For instance, many authors may simply refer to a point

cloud, or a wire-frame representation as a 3D model. However, one may view a 3D model as a post-processed version of a point cloud so that it can be efficiently rendered and visualized (e.g. in restorative medicine, simulation and training, and gaming applications), or used as a 3D map (3D model) for high-level applications, such as robot navigation, and in particular seamless navigation. In fact, a group of techniques developed in computer vision and more specifically in the robotics community is referred to as Simultaneous Localization and Mapping (SLAM) with specific goal of generating 3D maps with both visual and physical fidelity that can be used for rendering and navigation. A second group of techniques developed in computer vision. These are discussed in the next section.

D. Simultaneous Localization and Mapping (SLAM)

32. SLAM refers to a class of algorithms that are used in mobile robotics and medical applications (e.g. minimally-invasive surgery) for simultaneous self-localization of a sensing unit (e.g. a robot or a laparoscope) and the mapping of its sensing environment. Although earlier SLAM methods were used mostly in robotics and produced 2D maps (i.e. a 2D top view of the environment) using for instance a 1D laser scanner and an inertial navigation unit, they were quickly replaced by 3D mapping as soon as 3D sensing technologies, such as stereovision, SfM, RGB-D sensors, and 3D time-of-flight sensors (e.g. 3D LiDAR) were deployed, with

applications extending beyond robotics, and most notably into medical interventions.

33. Also, earlier SLAM methods were limited to static navigation environments. *See* Ex-1025. However, they were soon replaced by SLAM methods that could handle both the static background and the dynamic foreground (i.e. excluding moving objects from the 3D map). *See* Ex-1006. Development of dynamic SLAM methods, *see id.*, were particularly important in the context of applications where the presence of moving objects could distort the 3D map and make it physically inaccurate for subsequent use. Therefore, dynamic SLAM methods were developed to exclude moving objects from the final 3D map (3D model). *See id.*

E. Background Modeling: Moving Object Exclusion

34. An old topic of interest in computer vision that is closely related to the '551 patent is a problem known as background modeling. The term background in this class of computer vision algorithms refers to the static parts of a scene. Earlier work in this area was focused on statistical modeling of the background with a stationary camera, since the goal was to detect and track moving objects. However, these methods quickly developed into the dual problem of modeling the background for excluding the moving object, for instance for generating panoramic images without including the disturbance caused by a moving object in the scene. To create

a panorama of course the camera was allowed to move (at least rotate). Soon the notion of background model was also extended to 3D maps (e.g. disparity maps), *see* Ex-1026, created by a freely moving camera while excluding moving objects in the scene. This is essentially a SfM algorithm that excludes moving objects from the 3D representation. Although independently developed by the computer vision community, these methods may also be viewed as equivalent or similar to SLAM using SfM or stereovision in a dynamic environment.

F. Medical Applications of SLAM

35. As described earlier, SLAM methods used in dynamic environments have been proven very successful to create 3D models and maps of the foreground by excluding regions of moving objects, which is a central concept in the '551 patent. The brief review of medical applications of SLAM in this section further highlights the fact that the use of SLAM in medical interventions has been known for almost two decades.

36. SLAM methods have been successfully applied to medical applications due to the fact that they are able to provide both localization and 3D mapping in parts of human body that are less accessible or not accessible without using invasive procedures. SLAM in medical applications grow out of a need in surgical procedures that were difficult or impossible to perform without resorting to invasive

methods. Earlier predecessors of SLAM in medical intervention includes computer-assisted surgery systems that used computer-assisted navigation (i.e. a localization method similar to SLAM) along with a 3D model obtained prior to surgical procedures. Some examples include intraoral surgeries, *see* Ex-1027, and other computer-assisted surgeries. Later, full SLAM methods were successfully deployed in minimally-invasive endoscopic and laparoscopic surgeries, *see* Ex-1028, and ever since their use has been growing in the medical community.

V. OVERVIEW OF THE '551 PATENT

A. Specification of the '551 Patent

37. The '551 patent generally describes “a method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object.” Ex-1001 at 1:6-9. In particular, the patent describes an application of such a method that involves “scanning of a patient’s set of teeth in the mouth of the patient by means of a handheld scanner.” *Id.* at 1:9-11.

38. The '551 patent purports to address the problem of “distinguish[ing] between movable objects and rigid objects, when both movable objects and rigid objects are present in a location, when scanning in the location for obtaining a virtual 3D model of the rigid object.” *Id.* at 2:26-30. For example, the rigid objects may

be teeth, and the movable object may be soft tissue or a dentist's instrument or finger.

Id. at 3:50-60.

39. Accordingly, the '551 patent describes a method, as shown in Figure 1 below, including the following steps: 1) providing a first 3D representation; 2) providing a second 3D representation; 3) determining a first excluded volume; 4) determining a second excluded volume; and 5) testing if a portion of the surface in either 3D representation should be disregarded. *Id.* at Figure 1.

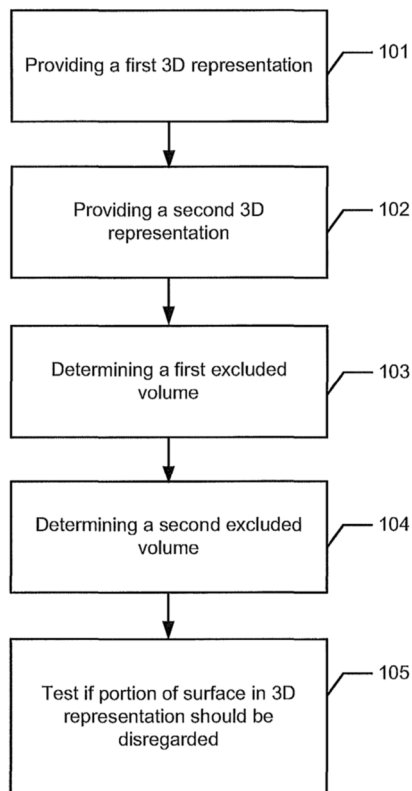
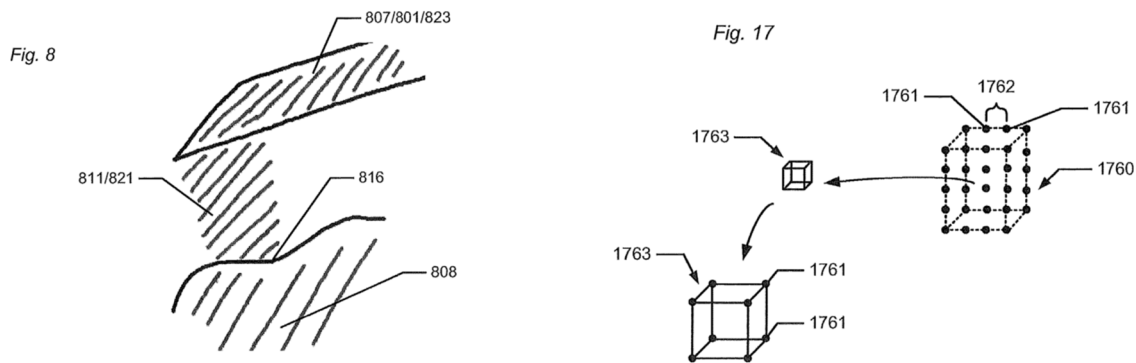


Fig. 1

40. The '551 patent states that obtaining a 3D representation of an object can generally be accomplished by using a 3D scanner to collect information about surfaces within the scanner's field of view and then developing a mathematical representation of the three-dimensional surface of the object via specialized software. *Id.* at 15:21-35. The patent states that "[t]he purpose of a 3D scanner is usually to create a point cloud of geometric samples on the surface of the object." *Id.* at 15:29-31. A variety of scanners may be used, including, for example, stationary desktop scanners and handheld scanners, and the scanner may rely on various technologies, including, for example, triangulation, structured-light, or confocal or focus scanning. *Id.* at 15:44-16:20. A "conceptually simple" and "commonly used" algorithm, such as iterative closest point (ICP), which "minimize[s] the difference between two clouds of points," may then be used to "reconstruct 2D or 3D surfaces from different scans or sub-scans." *Id.* at 16:21-58. The ICP algorithm "or similar methods" may further include an alignment process as part of generating a virtual 3D model from the different scans or sub-scans. *Id.*

41. During the scan, the scanner obtains a "scan volume," which the '551 patent defines as "the physical volume which the scanner is adapted to scan relative to the view position and the orientation of the scanner, such as relative to the scan head of the scanner." *Id.* at 4:32-35. Similarly, an "excluded volume" "is the volume

in space where no surface can be present.” *Id.* at 22:65-66. As shown in Figure 8 below, “[a]t least a part of the excluded volume 821 may correspond to the scan volume 811 of a 3D representation, since the space between the scan head 807 or the focusing optics of the 3D scanner and the captured surface 816 must be an empty space” *Id.* at 22:66-23:3. The ’551 patent further discloses that two or more excluded volumes may be determined for two or more 3D representations. *See id.* 23:14-61; Figs. 9, 10 (shown and discussed below).



42. Figure 17, as shown above, illustrates how an excluded volume is determined:

The space may be quantized in a 3D volume grid 1760. The distance 1762 between the corners 1761 in the 3D grid 1760 may be equidistant. The single cells 1763 in the grid each comprises eight corners 1761, and when each of the eight corners 1761 has been covered by a representation, then this cell 1763 is marked as seen. Thus if all eight corners 1761 of a cell 1763 is in the scan volume of a representation, then this cell 1763 may be marked as excluded volume.

Id. at 27:50-60.

43. The '551 patent then describes the process for determining whether any portion of either 3D representation should be disregarded. The patent defines “disregarding” as “not tak[ing] into account, such as deleting or not adding, when generating the 3D virtual model.” *Id.* at 3:16-18. Thus, the patent states that:

[A] portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, if the portion of the surface in the first 3D representation is located in space in the second excluded volume, and/or a portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model, if the portion of the surface in the second 3D representation is located in space in the first excluded volume.

Id. at 21:4-12.

44. The '551 patent further explains:

Thus it is an advantage that the method provides a determination of whether a detected surface portion is a point in space where there should be no surface, by detecting the space of the surface portion in both a first representation and in the second representation. If the surface portion is only present in one of the representations and the representations cover the same space of the surface portion, then the surface portion must represent an object which was only present when one of the representations were acquired, and therefore the surface

portion must originate from a movable object, which has moved during the acquisitions of the two representations.

Id. at 2:62-3:6.

45. The '551 patent illustrates this concept in Figures 9 and 10. Figures 9a and 9b below show, respectively, a first excluded volume 921 determined for a first 3D representation 916, and a second excluded volume 922 determined for a second 3D representation 917. *See id.* at 23:14-30. In this case, each of the representations was acquired at a different angle between the scanner and the tooth. *Id.* Thus, because “[n]o surface portion of the first representation 916 lies in the second excluded volume 922, and no surface portion of the second representation 917 lies in the first excluded volume 921, . . . no surface portion(s) are disregarded in the generation of the virtual 3D model in this case.” *Id.* at 23:31-35.

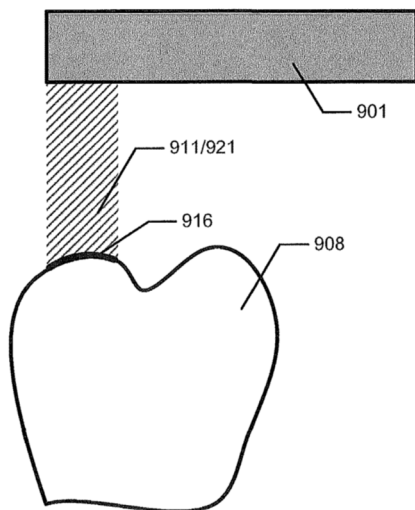


Fig. 9a)

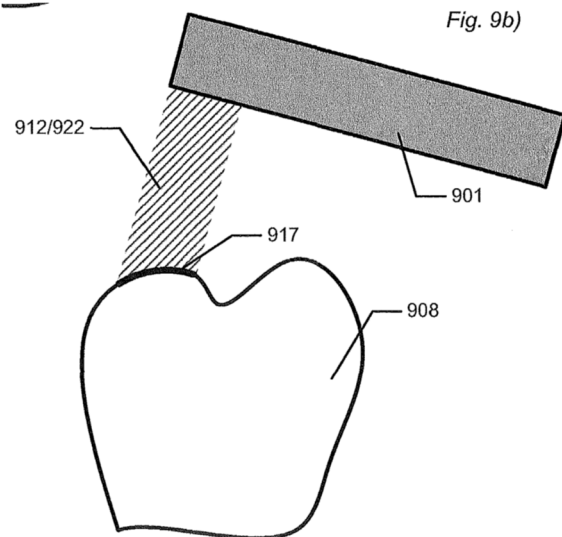
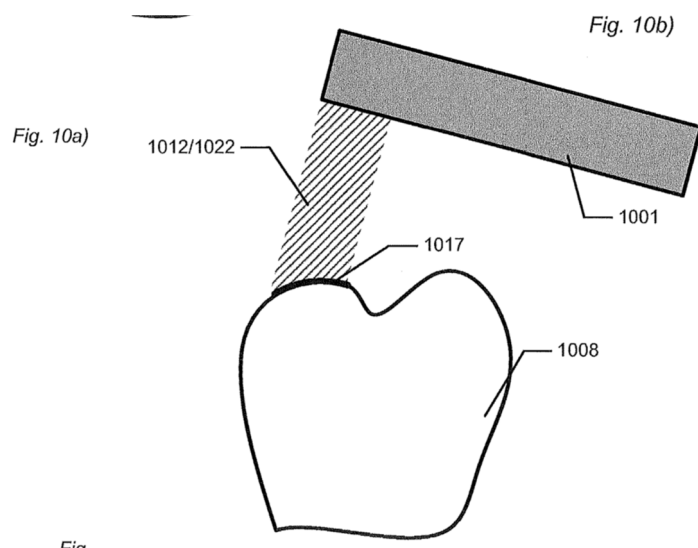
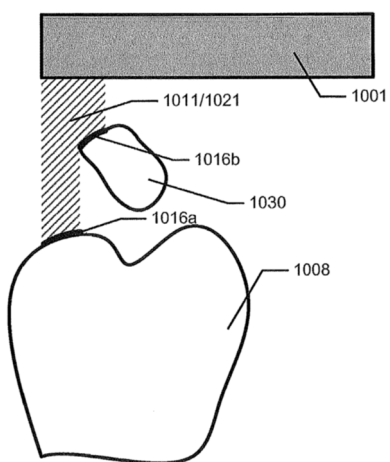


Fig. 9b)

46. Similarly, Figures 10a and 10b below show, respectively, a first excluded volume 1021 determined for a first 3D representation 1016, and a second excluded volume 1022 determined for a second 3D representation 1017. *See id.* at 23:36-57. Again, each of the representations was acquired at a different angle between the scanner and the tooth. *Id.* However, in this case “a movable object is captured in part of the first representation.” *Id.* at 23:38-39. The first 3D representation is therefore divided into a part 1016a comprising a surface of the scanned tooth and a part 1016b comprising a surface of the moveable object. *Id.* at 23:42-46. The '551 patent explains that “[s]ince the surface portion 1016b of the first representation 1016 lies in the second excluded volume 1022, this surface portion 1016b is disregarded in the generation of the virtual 3D model.” *Id.* at 23:58-61.



B. Prosecution History

47. The application underlying the '551 patent was filed on January 21, 2014. Ex. 1001. During examination, Patent Owner made the following amendments to claim 1 on November 21, 2016, to overcome a final rejection, in which the Examiner asserted that the applied-for claims were anticipated by or obvious in view of the prior art:

A method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the method comprises:

...

- determining for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;

- determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;

....

Ex-1002 at 1-26. Patent Owner amended then-pending claims 51 and 57, which correspond to claims 23 and 25 in the issued patent, with identical language. Ex-1002 at 20-26.

48. In response, the Examiner issued a Notice of Allowance on December 16, 2016. Ex-1002 at 32-36. Although the Examiner did not provide

reasons for allowing the claims in the Notice of Allowance, an interview summary dated December 5, 2016, states, “Applicant and Examiner discussed claimed invention in the amendment submitted 11/21/2016 vs. teaching of Marshall. Examiner agrees the amendment overcomes the current rejection.” Ex-1002 at 30.

VI. LEGAL PRINCIPLES USED IN ANALYSIS

49. I am not an attorney and I offer no opinions on the law. I have been advised by Petitioner’s counsel of certain legal principles relevant to an invalidity analysis, as summarized below. I have conducted my analysis in accordance with these principles.

A. Prior Art

50. I understand that a patent or other publication must qualify as prior art before it can be used to invalidate a patent claim.

51. I understand that a U.S. or foreign patent qualifies as prior art to a challenged claim if the patent was issued prior to the invention date of the challenged claim. I also understand that a U.S. or foreign patent or published patent application qualifies as prior art to a challenged claim if it was published before the invention date of the challenged claim.

B. Anticipation

52. I understand that a prior art reference “anticipates” an asserted claim, thereby rendering the claim invalid, if all elements of the claim are disclosed in the

reference, either expressly or inherently. I understand that an element is disclosed inherently in a prior art reference if it is necessarily present in the prior art reference or is the natural result flowing from the disclosure of the prior art reference. I have been informed that a prior art reference does not need to use the same words as the challenged claim in order to anticipate the claim.

C. Obviousness

53. I also understand that a patent is invalid if the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious to a person of ordinary skill in the pertinent art at the time the invention was made. I understand that a POSITA provides a reference point from which the prior art and claimed invention should be viewed. This reference point prevents one from using his or her own personal insight or hindsight in deciding whether a claim is obvious.

54. I understand that an obviousness determination includes the consideration of various factors, such as (1) the scope and content of the prior art, (2) the differences between the prior art and the asserted claims, (3) the level of ordinary skill in the pertinent art, and (4) the existence of secondary considerations such as commercial success, long-felt but unresolved needs, failure of others, etc.

55. I understand that an obviousness evaluation can be based on a

combination of multiple prior art references. I understand that the prior art references themselves may provide a suggestion, motivation, or reason to combine, which does not need to be explicit. I further understand that any need or problem known in the field and addressed by the patent can provide a reason for combining the elements in the manner claimed. To determine whether there was an apparent reason to combine the known elements in the way a patent claims, I have been informed that I may look to interrelated teachings of multiple patents, to the effects of demands known to the relevant community or present in the marketplace, and to the background knowledge possessed by a POSITA. I also understand that practical and common-sense considerations should guide a proper obviousness analysis because familiar items may have obvious uses beyond their primary purposes. I further understand that a POSITA looking to overcome a problem will often be able to fit together the teachings of multiple publications. I understand that an obviousness analysis therefore accounts for the inferences and creative steps that a POSITA would employ under the circumstances.

56. I understand that the nexus linking two or more prior art references can also be simple common sense. I further understand that the obviousness analysis recognizes that market demand, rather than scientific literature, often drives innovation, and that a motivation to combine references may be supplied by the

direction of the marketplace.

57. I understand that if a technique has been used to improve one device, and a POSITA would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill.

58. I understand that a particular combination may be proven obvious merely by showing that it was obvious to try the combination. For example, when there is a design need or market pressure to solve a problem and there are a finite number of identified, predictable solutions, a POSITA has good reason to pursue the known options within his or her technical grasp because the result is likely not the product of innovation, but of ordinary skill and common sense.

59. I understand that the combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results. When a work is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or a different field. If a POSITA can implement a predictable variation, the patent claim is likely obvious.

60. As noted above, I understand that a proper obviousness analysis focuses on what was known or obvious to a POSITA, not just the patentee. Accordingly, I

understand that any need or problem known in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining the elements in the manner claimed.

61. I understand that a claim can be obvious in light of a single reference, without the need to combine references, if the elements of the claim that are not found explicitly or inherently in the reference can be supplied by the common sense or knowledge of one of skill in the art.

62. I also understand that secondary indicia of non-obviousness may include: (1) a long-felt but unmet need in the prior art that was satisfied by the invention of the patent; (2) commercial success of processes covered by the patent; (3) unexpected results achieved by the invention; (4) praise of the invention by others skilled in the art; (5) taking of licenses under the patent by others; (6) deliberate copying of the invention; (7) failure of others to find a solution to the long-felt need; and (8) skepticism by experts.

63. I also understand that there must be a relationship between any such secondary considerations and the invention. I further understand that contemporaneous and independent invention by others is a secondary consideration supporting an obviousness determination.

64. I further understand that an obviousness determination does not require

an actual, physical substitution of elements from multiple references. Instead, the primary consideration is whether the claimed inventions are rendered obvious by the teaching of the prior art of the whole, not whether the references could be physically combined.

65. In sum, it is my understanding that prior art teachings are properly combined where a POSITA—having the understanding and knowledge reflected in the prior art and being motivated by the general problem facing the inventor—would have been led to make the combination of elements recited in the claims. Under this analysis, the prior art references themselves, or any need or problem known in the field of endeavor at the time of the invention, can provide a reason for combining the elements of multiple prior art references in the claimed manner.

D. Claim Construction

66. I understand that, before an invalidity analysis can be performed, the scope and meaning of the challenged claims must be determined. It is my understanding that this is accomplished through the process of claim construction.

67. I have been informed and understand that claims are generally construed from the perspective of a POSITA at the time of the invention, which I have been asked to assume is the '551 patent's priority date of July 15, 2011. More specifically, I understand that the claims should be afforded a meaning in accordance

with the ordinary and customary meaning of the claim, as understood by a POSITA at the time of the invention.

68. I have been advised that, in construing a claim term, one must first look to the “intrinsic” patent evidence, which includes the claims themselves, the patent specification, and the prosecution history of the patent. I have also been informed that “extrinsic” evidence, which is outside the patent and prosecution history, may be useful in construing patent claims. It is my understanding that such extrinsic evidence may include dictionaries, textbooks, treatises, and similar sources of information.

69. I have been advised that a claim may recite a function to be performed rather than a structure for performing the function. I have been informed that such claim terms are known as “means-plus-function” claim terms. I understand that a term introduced by the phrase “means for” is presumed to be a means-plus-function claim term. I have also been informed that a claim term may be construed as a means-plus-function term where it fails to recite sufficiently definite structure, or else recites function without reciting sufficient structure for performing that function.

70. I have been advised that the construction of a means-plus-function claim term is a two-step process. I understand that the first step requires an identification of the function recited in the claim. I understand that the second step

requires determining what structure is disclosed in the specification for performing the recited function. I have been advised that a structure disclosed in the patent specification qualifies as corresponding structure if the intrinsic evidence clearly links or associates that structure to the function recited in the claim. I have been advised that the proper construction for a means-plus-function term is restricted to the corresponding structure identified in the specification and equivalents thereof.

71. I understand that claims must be construed in an *inter partes* review using the same claim construction standard that would be used to construe the claim in a civil action before a district court.

VII. LEVEL OF ORDINARY SKILL IN THE ART

72. My opinions have been guided by my appreciation of how a POSITA (described below) would have understood the claims of the '551 patent as of the invention date, which I have been asked to assume is the patent's priority date, July 15, 2011.

73. It is my understanding that a POSITA is one who is presumed to be aware of all pertinent art and is a person of ordinary creativity. I further understand that the following factors may be considered in establishing the level of ordinary skill in the art in the relevant field: the type of problems encountered in the art; the prior art solutions to those problems; the rapidity with which innovations are made;

sophistication of the technology; and the educational level of active workers in the field.

74. It is my opinion that the level of ordinary skill in the art in the July 2011 timeframe of the invention would have had a bachelor's degree in computer engineering, computer science, computer vision, or an equivalent field as well as at least one or two years of industry experience. A POSITA at this time would have experience with three-dimensional imaging systems and three-dimensional modeling techniques.

VIII. CLAIM CONSTRUCTION

75. I have set forth my understanding of relevant claim construction principles above, which I have followed in my analysis. Below I discuss two claim terms that were construed by the parties and the district court in the related litigation. In my view, the remaining claim terms should be given their plain and ordinary meaning.

A. The Preambles of the Independent Claims Are Limiting (Claims 1, 23, 25)

76. Counsel informs me that the parties in the district court litigation have agreed that the preambles of claims 1, 23, and 25 are limiting. Counsel has also informed me that the district court has adopted this construction in its order. Ex-1011 at 2. For the purposes of this declaration, I have been asked to adopt this

construction in my analysis.

B. “excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation”

77. Claims 1, 23, and 25 recite a “first” and “second” “excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation.” I am informed by counsel that the district court has construed these terms to mean an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present.” Ex-1011 at 2. For the purposes of this declaration, I have adopted the district court’s construction of this term.

IX. OVERVIEW OF THE PRIOR ART

A. Newcombe

78. Newcombe is directed to systems and methods for segmenting moving objects from the background of a depth image of a scene. Ex-1005 at Abstract. Newcombe discloses a depth camera that is capable of capturing depth images of a room or scene and objects within the room. *Id.* at 2:46-53. A real-time camera tracking system is used to monitor the location and orientation of the camera within the room. *Id.* at 2:54-67. The depth information utilized by Newcombe’s system and methods can be obtained from any “suitable technique including, but not limited to, time of flight, structured light, and stereo images.” *Id.* at 4:11-50.

79. Newcombe also provides a moving-object segmentation system, which “receives the depth images captured by the mobile depth camera 102, and performs a segmentation operation to identify objects in the depth images that are moving or being manipulated and which are independent of the relatively unchanging background.” *Id.* at 3:1-6. Newcombe discloses that the “moving-object segmentation system 114 detects foreground objects and labels them in the depth map, enabling these objects to be separately processed as desired.” *Id.* at 3:51-54. Newcombe further discloses that the depth camera, real-time tracking system and moving-object segmentation system can be used to “allow individual depth images to be built up into an overall 3D model” utilizing the “simultaneous and localization and mapping (SLAM)” technique. *Id.* at 3:28-37.

80. According to the process disclosed by Newcombe, “a depth image of the scene to be segmented (as captured by the mobile depth camera) is received.” *Id.* at 6:49-50. A previous depth image, which may be the preceding frame captured by the depth camera and which overlaps with at least part of the same scene, is also retrieved, along with a capture pose for the previous image, which is typically the pose of the depth camera at the time the previous frame was captured. *Id.* at 6:54-7:9. A frame alignment engine is then used to align the current and previous frames and to determine the current pose of the depth camera. *Id.* at 7:10-18. This process

involves the use of an iterative closest point (ICP) algorithm, which “is an iterative process that determines a set of corresponding points between the two depth images, determines a transformation that minimizes an error metric over the corresponding points, and repeats the process using the determined transformation until convergence is reached.” *Id.* at 7:18-26. Any divergent points are deemed to be “outliers,” which can be determined by two compatibility tests. *Id.* at 7:27-30; 8:27-9:16.

81. Newcombe’s compatibility tests are illustrated by Figures 5 and 6 of the reference, shown below. Specifically, Newcombe explains that “[i]n the first compatibility test, a Euclidean distance is determined 414 between the source and destination points,” which “is illustrated in Fig. 5 as distance 510 and in Fig. 6 as distance 604.” *Id.* at 8:29-32. The source point is indicated by points 502 and the destination points are indicated by points 508 and 602 in Figures 5 and 6 below. Newcombe further explains:

The Euclidean distance is then compared 416 to a predefined distance threshold. If the destination point is too far away from the source point, then it is considered to be an outlier (i.e. non-corresponding) for the purposes of the ICP algorithm, and is labeled 418 as such.

The example of Fig. 5 shows the case where the distance 510 between the source and destination points is small, and is less than the predefined

distance threshold (and hence is not labeled as an outlier at this stage). Conversely, the example of Fig. 6 shows the case where differences between the previous depth image 500 and the current depth image 600 (e.g. due to new or moving objects) result in the distance 604 between the source and destination points being greater than the predefined threshold, and hence the destination point is labeled as an outlier.

Id. at 8:32-49.

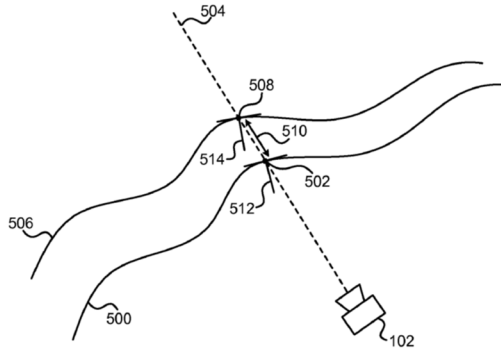


FIG. 5

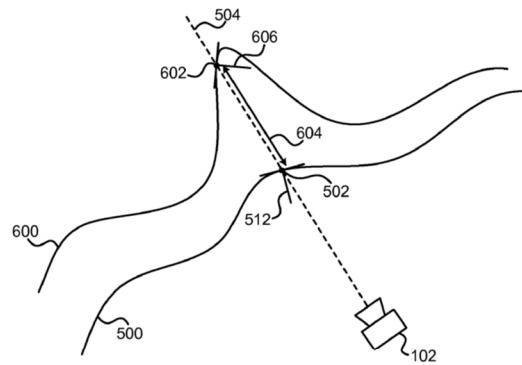


FIG. 6

82. In the case of Fig. 5 above, where the Euclidean distance is within the predefined distance threshold, Newcombe discloses that a second compatibility test is then performed. *Id.* at 8:49-51. Newcombe states that, “[t]o perform this test, the surface normals are calculated 419 at the source and destination points. For example, this is achieved for a given point by finding the four (or more) nearest neighbor

points in the depth image and computing a surface patch which incorporates those neighbors and the point itself. A normal to that surface patch is then calculated at the location of the point.” *Id.* at 8:51-57. According to Newcombe, “[t]he angle between the surface normal of the source point and the destination point is then determined 420. . . . The calculated angle (or other measure) is then compared 422 to a predefined angle threshold. If the angle is greater than the predefined angle threshold, then the destination point is considered to be an outlier, and labeled 418 as such. Conversely, if the angle is not greater than the predefined angle threshold, then the destination point is selected 424 as a corresponding point to the source point.” *Id.* at 8:61-9:4. Thus, for example, in Figure 5 above, “the angle between the source surface normal 512 and destination surface normal 514 is small, and hence these are considered to be corresponding points.” *Id.* at 9:7-9. On the other hand, “in Fig. 6, the angle between the source surface normal 512 and the destination surface normal 606 is larger, and taken to be greater than the predefined angle threshold . . . , and hence the destination point is labeled as an outlier (ignoring that this destination point was already labeled as an outlier due to exceeding the distance threshold).” *Id.* at 9:10-16.

83. The outlying point “indicate points in the depth image where something has changed since the previous depth image” and are provided as input to the object

segmentation system. *Id.* at 9:26-28. For example, the object segmentation system can segment all outlying points or a subset of outlying points and label them as foreground objects in the depth image. *Id.* at 9:29-41. These segmented portions of depth images may correspond to foreground objects, including moving objects, and can be tracked separately and independently of the background. *Id.* at 10:28-38. Independent 3D models of the moving objects may also be generated. *Id.* at 10:57-59. Newcombe further discloses that moving objects may be “instantly segmented upon removal from the scene.” *Id.* at 11:21-23.

B. Herbst

84. Herbst is directed to methods that improve the performance of robots that scan indoor locations to create 3D maps and images. Ex-1006 at 1. Herbst discloses a method of detecting moving objects within a scene by comparing color and depth information for surfaces scanned in two separate 3D scans of a scene. *Id.* Herbst further discloses that such moving objects within a scene, once scanned, detected, and segmented, can be extracted from a 3D model of the scene. *Id.* at 5.

85. Herbst discloses the use of an RGB-D camera, which can provide dense color and depth measurements, to obtain 3D scans of a location. *Id.* at 1. Thus, “[t]he inputs to [Herbst’s] system are two or more RGB-D videos representing separate visits to the same location.” *Id.* at 2. Herbst further discloses the use of

RGB-D Mapping, which is a Simultaneous Localization and Mapping (SLAM) technique for generating dense 3D maps. *Id.* Each frame obtained by the RGB-D cameras consists of multiple points in a point cloud, each of which is converted to a surfel according to Herbst's method. *Id.* Herbst then "probabilistically models the position, color and orientation of each surfel." *Id.*

86. Herbst's probabilistic model relies on the RGB-D camera's ability to "provide a dense matrix of pixels, each measuring the color and distance of the closest surface in the pixel direction." *Id.* Thus, Herbst's model "make[s] use of dense depth information to locally estimate surface normals" *Id.* at 2-3. Herbst's probabilistic model considers two different cases: one in which Herbst models an RGB-D pixel measurement when no prior information about surfaces in a scene are known and one in which information about scene surfaces is known. *Id.* at 3-4.

87. Herbst further discloses that the estimated pixel measurements from two 3-D maps can be used to detect differences between the maps after the scenes in each map have been aligned, and specifically to calculate a probability that any object surface in a scene moved from one map to the next. *Id.* at 4. Thus, objects within a scene that have moved locations "are detected as high-probability areas, using [a] combination of depth, color, and orientation cues." *Id.* at 5.

88. Herbst also provides details of experiments designed to evaluate the model and to demonstrate the ability to discover and model objects, including moveable objects, within a tabletop scene. Herbst discloses that its method can find all moveable objects within a scene with high accuracy. *Id.* at 5. For example, Herbst provides the following two figures, Figure 1(a) and 1(b), which show two different 3D representations of the same tabletop scene.

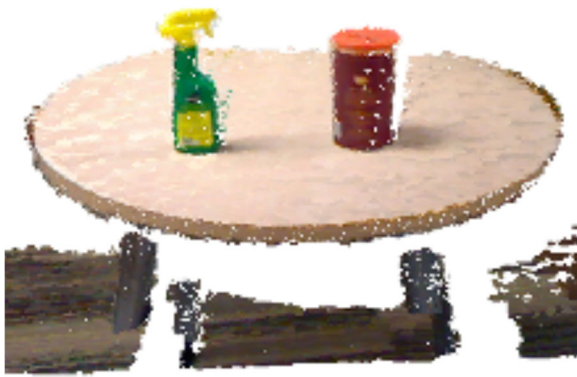


Fig. 1(a)



Fig. 1(b)

89. As can be seen in the images above, each representation shows different objects resting on the same table. As shown in Figure 7(a) of Herbst below, when the two representations are merged into a single 3D model, Herbst's disclosed method is able to accurately determine that each of the objects on the table is a moveable object because each object is moved between the two scenes. These objects are then extracted and shown separately around the top edge of the merged 3D model. *See id.* at 5-7; Figs. 1(a)-(b) & 7(a).



Fig. 7(a)

The table itself, however, is correctly identified as a stationary, or non-moveable object, and therefore remains in place in the merged 3D model. Further, some of the “holes” – i.e., portions of the table occluded by the objects in the first two images – are “filled in” in the merged 3D model. *Id.*

90. Herbst demonstrates the model’s ability to detect moveable objects again in Figures 6(a), (d) and 7(b). Figures 6(a) and 6(d) show the same scene with various objects resting on a table. Although the table and some of the objects remain stationary between the two scans, some of the objects change locations or move, as shown below.



Fig. 6(a)

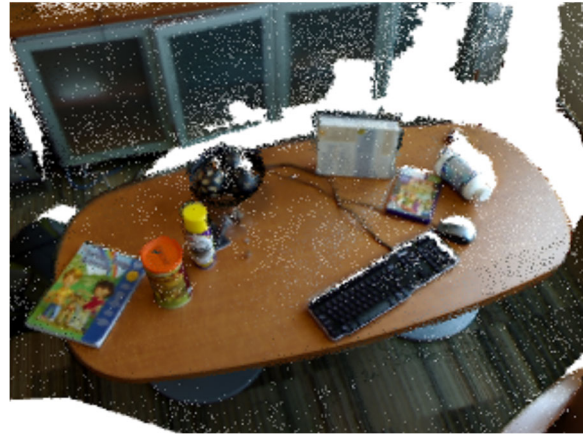


Fig. 6(d)

91. In the merged 3D model shown in Figure 7(b) below, the table and stationary objects remain in place in the merged 3D model. The movable objects, however, are accurately detected and extracted and shown in separately around the top edge of the 3D model. Also, “holes” on the table created by moveable objects are again “filled in” in the merged 3D model. *See id.* at 5-7; Figs. 6(a), (d) & 7(b).



Fig. 7(b)

C. Hähnel

92. Hähnel is directed to methods for creating maps with mobile robots in dynamic environments, including, for example, environments with moving objects or people. Ex-1007 at 1. The approach disclosed by Hähnel can generate accurate 2d and 3d models in a variety of dynamic indoor and outdoor environments and builds on the simultaneous localization and mapping (SLAM) technique, which “includes both estimating the position of the robot relative to the map and generating a map using the sensory input and the estimates about the robot’s pose.” *Id.* Hähnel further discloses that, by filtering and extracting the dynamic objects from the environment, accurate textured 3d models of the dynamic objects may also be generated. *Id.*

93. Hähnel discloses the use of laser-range scanners for obtaining measurements of the dynamic environment to be mapped. *Id.* An expectation-maximization algorithm is used to detect dynamic objects in the environment. *Id.* at 1-2. To determine whether or not a beam generated by the scanner is reflected by a dynamic object, Hähnel’s algorithm considers factors including the likelihood of a measurement in the current environment, the pose of the robot, and whether the measured length is reflected by a maximum range reading. *Id.* at 2. An iterative SLAM process is used that includes obtaining an initial map through an incremental

mapping approach, determining expectation values based on prior probabilities that a measurement is caused by a static object, computing new expectation values, and then using the new expectations values to generate a new map. *Id.* at 2-4. Hähnel discloses that its system is robust in a variety of highly dynamic environments, including, for example, environments with moving people or with a fast-moving car. *Id.* at 5.

94. Hähnel further discloses that its method can be used to generate textured 3d maps. *Id.* at 6. Using a laser scanner and camera, textured 3d maps of an environment may be obtained. *Id.* Hähnel further discloses that a horizontally scanning laser range finder may be used to determine dynamic objects in the environment. *Id.* Hähnel discloses that dynamic objects can be filtered from the scene to generate a 3d model without the dynamic objects. *Id.*

95. For example, the upper image below, from Figure 11 of Hähnel, includes an arrow pointing to “a polygon whose texture contains fractions of an image of a person which walked through the scene while the robot was scanning it.” *Id.* at 6-7. The lower image was generated after applying the approach disclosed by Hähnel and shows the same scene with the walking person (i.e., the dynamic object) filtered out of the scene. *Id.*



Fig. 11

96. Hähnel further discloses that, in addition to “filtering dynamic objects and learning static aspects of environments[, its] algorithm can also be used to separate dynamic objects from the environment.” *Id.* at 6.

D. McQueston

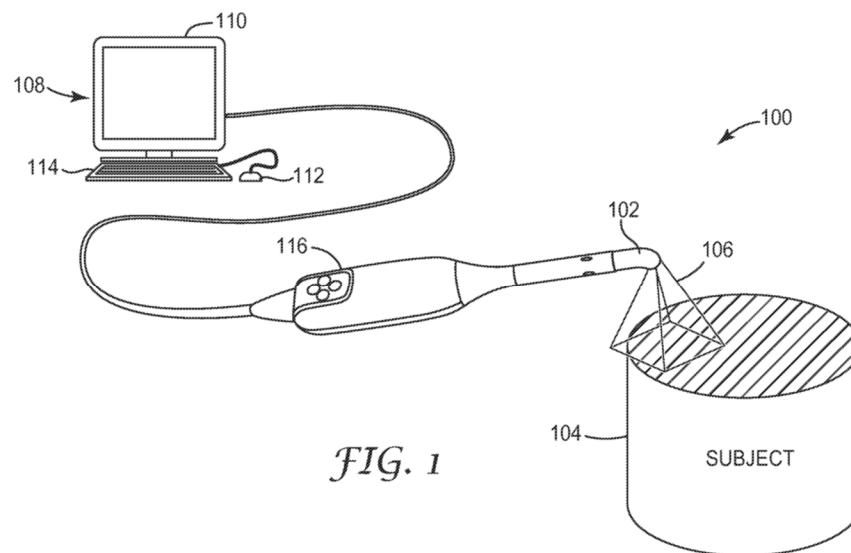
97. McQueston discloses the use of “space carving” in three-dimensional scanning to “exclude certain scan results from processing and display.” Ex-1008 at

1:27-29. The space carving techniques, which McQueston describes as “generally known in the art,” are used to maintain a spatial matrix for storing data on volumetric regions, known as voxels, that are known to be empty. *Id.* at 1:29-13, 15:9-13. “By excluding or modifying processing of subsequent outlier data from within these unoccupied voxels, a three-dimensional reconstruction process can achieve concurrent improvements in accuracy and speed.” *Id.* at 1:31-34.

98. McQueston describes the possibility that its disclosures can be used with any “scanning technology,” in which “the calculation and display of three-dimensional data can become encumbered by transient obstructions within a scanning volume.” *Id.* at 1:14-16. A POSITA could readily use McQueston’s teachings “to improve any three-dimensional acquisition or processing in which the capture of three-dimensional data occurs over a period of time.” *Id.* at 3:14-30. In particular, the reference specifically identifies the need for the disclosed technology in dental scans stating that, “[f]or example, in a dental scan, a patient’s tongue or other soft tissue may temporarily obstruct or move adjacent to a scanned surface. A technician may similarly obstruct a scanner’s view of the scanned surface with a finger, tongue depressor, or other instrument.” *Id.* at 1:16-21.

99. As Figure 1 of McQueston illustrates, the reference discloses an “image capture system 100,” including “a scanner 102 that captures images of a subject 104

within an image plane 106, and forwards the images to a computer 108, which may include a display 110 and one or more user input devices such as a mouse 112 or a keyboard 114.” *Id.* at 4:59-64. “The scanner 102 may include any camera or camera system suitable for capturing images from which a three-dimensional point cloud may be recovered,” including “a stereoscopic, triscopic or other multi-camera or other configuration. . . .” *Id.* at 4:65-5:13. For dental applications, the scanner may “be shaped and sized for intraoral scanning and data capture.” *Id.* at 5:33-34.



100. McQueston’s three-dimensional scanning system employs a spatial matrix, which “may define a mesh of voxels that together form the scan volume or scan space in which data is processed” *Id.* at 14:26-28. The scan volume may be any suitable size or shape according to the specific object being scanned. *Id.* at 14:30-32. The spatial matrix may further “include a binary flag for each one of the

plurality of voxels that indicates whether processing is permitted (for example, unmodified) or forbidden (for example, modified) for a volume of the scan space.” “While the ‘forbidden’ voxels may be generally understood as voxels that are unoccupied, it will be appreciated that this is not universally true and that voxels in which points are detected may be forbidden (such as where a surface temporarily moves into a scan space) and voxels that are empty may be permitted” *Id.* at 15:25-31. McQueston further teaches that “any suitable techniques for characterizing a voxel as containing or not containing the surface of interest may be suitably employed to identify prohibited space as contemplated herein.” *Id.* at 15:67-16:3.

101. McQueston discloses that “any three-dimensional process that does not use space carving or the like” may be used for the permitted, or unmodified, processing, while the modified processing “may include any processing suitable for data in the ‘prohibited volume.’” *Id.* at 18:13-25. “For example, a reconstruction process may treat data within the prohibited volume as outliers that are excluded from processing.” *Id.* at 18:29-31. Alternatively, data in the prohibited volume may be “de-weighted,” “de-emphasized,” “excluded from display,” or otherwise controlled with respect to “how data in the prohibited volume is used in a surface mesh of the three-dimensional data.” *Id.* at 18:31-47.

X. THE PRIOR ART RENDERS ALL OF THE CLAIMED ELEMENTS OF THE '551 PATENT OBVIOUS

A. Claims 1, 6-12, 15-17, and 20-25 Are Rendered Obvious by Newcombe and Hähnel

102. It is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, every element of claims 1, 6-12, 15-17, and 20-25 of the '551 patent.

1. Claim 1 Is Rendered Obvious by Newcombe in view of Hähnel

103. As explained below, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, every element of claim 1.

a. [Claim 1 pre]: “A method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the method comprises:”

104. As discussed above, in section VIII.A, for the purposes of this declaration, I have treated the preamble of claim 1 as limiting. It is my opinion that Newcombe teaches this claim element.

105. Newcombe teaches detecting and identifying movable objects from a series of depth images of a scene using segmentation. Ex-1005 at Abstract. Multiple depth images are obtained using a depth camera that gathers both image and depth data. *Id.* at 2:46-53; 3:25-27. It is my opinion that this depth camera corresponds to

the claimed “3D scanner.” Newcombe teaches detecting moving objects and segmenting them from a scene comprised of fixed objects that is being scanned by depth images. *Id.* at 1:49-65. It is my opinion that this corresponds to the claimed “method for detecting a movable object in a location when scanning a rigid object in the location.” Newcombe also teaches generating a dense 3D environment model. *Id.* at 3:21-37. It is my opinion that this corresponds to “generating a virtual 3D model of the rigid object,” as claimed.

106. Accordingly, it is my opinion that Newcombe teaches “detecting a movable object in a location when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object,” as recited by the preamble of claim 1.

b. [Claim 1a]: “providing a first 3D representation of at least part of a surface by scanning at least part of the location;”

107. Newcombe teaches this claim element. Newcombe teaches using a mobile depth camera to obtain depth images of a location. *Id.* at 2:46-53. These depth images include both image data and depth data. *See id.* at 3:64-4:10. The depth images are used to form the dense 3D model of the environment. *Id.* at 3:25-27. It is my opinion that a depth image of the location, as described by Newcombe, corresponds to a “first 3D representation of at least part of a surface scanning at least

part of the location,” as claimed.

108. Accordingly, it is my opinion that Newcombe teaches “providing a first 3D representation of at least part of a surface of the location,” as recited in this element of claim 1.

c. [Claim 1b]: “providing a second 3D representation of at least part of the surface by scanning at least part of the location;”

109. Newcombe teaches this claim element. As discussed above, Newcombe describes obtaining multiple depth images of the location to build the dense 3D model. It is my opinion that one of the additional depth images of the location, as described by Newcombe, corresponds to a “second 3D representation of at least part of the surface by scanning at least part of the location,” as claimed.

110. Furthermore, Newcombe discloses two separate scanning processes, through which multiple depth images may be provided. Specifically, Newcombe discloses the following:

Firstly, a depth image of the scene to be segmented (as captured by the mobile depth camera) is received 400. If the segmentation operation is making use of the camera-tracking results, then this current depth image is received at the frame alignment engine 318. The current depth image is also known as the destination depth image. A previous depth image of at least a portion of the scene is then retrieved 402. The previous depth image is also known as the source depth image. In some

examples, the previous depth image may be the preceding frame captured by the depth camera and stored at a storage device. In alternative examples, the previous depth image may be generated from the dense 3D model stored in graphics memory by requesting the previous depth image from the GPU. The GPU can generate the previous depth image from the dense 3D model stored on the graphics memory by determining the view from a virtual camera at the last known location and orientation of the real depth camera looking in the dense 3D model. The use of the dense 3D model provides a higher resolution previous depth image.

Id. at 6:49-67. It is my opinion that each of the previous and current depth images disclosed by Newcombe correspond to the first and second 3D representations recited in the claim.

111. Accordingly, it is my opinion that Newcombe teaches “providing a second 3D representation of at least part of the surface by scanning at least part of the location,” as recited by this element of claim 1.

- d. **[Claim 1c]: “determining for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

112. Newcombe teaches this claim element. Newcombe describes identifying, through segmentation, moving objects in the foreground of a location. It is my opinion that a POSITA at the time the '551 patent was filed would

understand Newcombe to disclose this claim element in connection with the disclosed segmentation process. As described above, Newcombe discloses a previous 3D depth image and a current 3D depth image, which a POSITA would understand correspond to first and second 3D representations. *See supra*, §§ X.A.1.b-c. It is further my opinion that a POSITA at the time the '551 patent was filed would understand that all spaces within the obtained 3D depth images for the location that *do not* correspond to a surface within the background of the location would collectively constitute an excluded volume, as recited by the claims, and more particularly would constitute an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present, according to the court's claim construction. For example, it is my opinion that a POSITA would have understood that the empty space shown in Figure 1, shown below, between camera 102 and the stationary objects in the room would constitute an excluded volume for the first 3D representation, according to the court's construction.

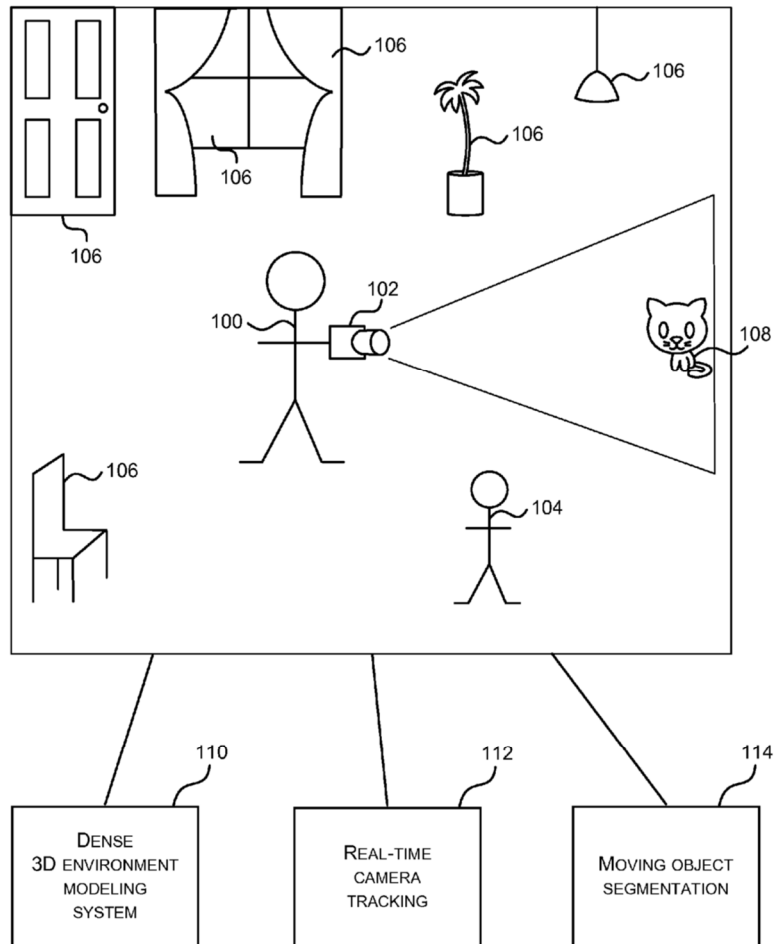


FIG. 1

113. It is my opinion that this understanding of Newcombe is supported by the reference's disclosure of the performance of compatibility tests to determine whether points in two depth images are corresponding points, which can be attributed to rigid objects, or are outliers, which can be attributed to moving objects. *See* Ex-1005 at 8:27-29. Specifically, it is my opinion that Newcombe discloses that at least the first compatibility test can be used to determine an excluded volume

where no rigid object should be present in any depth image of the location, and consequently, where any object detected is therefore a moving or “non-rigid” object at the location. *See, e.g., id.* at 9:26-29.

114. As described by Newcombe, the first compatibility test relies on defining a threshold distance for a measurement between points in two 3D depth images – i.e., a “source point” in a previous depth image and a “destination point” in a current depth image. *Id.* at 8:3-33. As Newcombe explains, “[i]f the destination point is too far away from the source point, then it is considered to be an outlier (i.e. non-corresponding)” *Id.* at 8:33-35. Newcombe further explains that “outliers” or “outlying points” refer to “points in the depth image where something has changed since the previous depth image” and “can be caused by dynamic or non-rigid objects.” *Id.* at 9:26-29.

115. This scenario is demonstrated in Figure 6, shown below, of Newcombe: “[T]he example of Fig. 6 shows the case where differences between the previous depth image 500 and the current depth image 600 (e.g. due to new or moving objects) result in the distance 604 between the source and destination points being greater than the predefined distance threshold, and hence the destination point is labeled as an outlier.” *Id.* at 8:42-48. Newcombe further explains the process as follows:

Moving object segmentation using depth images is described. In an

example, a moving object is segmented from the background of a depth image of a scene received from a mobile depth camera. A previous depth image of the scene is retrieved, and compared to the current depth image using an iterative closest point algorithm. The iterative closest point algorithm includes a determination of a set of points that correspond between the current depth image and the previous depth image. During the determination of the set of points, one or more outlying points are detected that do not correspond between the two depth images, and the image elements at these outlying points are labeled as belonging to the moving object. In examples, the iterative closest point algorithm is executed as part of an algorithm for tracking the mobile depth camera, and hence the segmentation does not add substantial additional computational complexity.

Id. at 1:49-65.

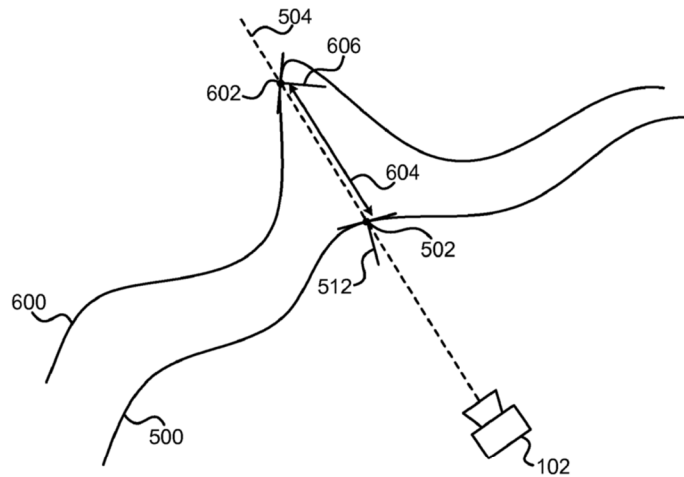


FIG. 6

116. Thus, it is my opinion that a POSITA would understand that by defining a distance threshold for the two three-dimensional depth images beyond which any point will be deemed an outlier, Newcombe teaches the determination of an excluded volume, and particularly, an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present,” as construed by the district court.

117. Accordingly, it is my opinion that Newcombe teaches “determining for the first 3D representation a first excluded volume in space where no surface can be

present in both the first 3D representation and the second 3D representation,” as recited in claim 1.

- e. **[Claim 1d]: “determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

118. Newcombe teaches this claim element. My analysis and discussion in the preceding section of Newcombe’s segmentation process and the compatibility tests for determining corresponding points and outliers applies equally to this claim element. *See supra*, § X.A.1.d. As described above, Newcombe discloses a previous 3D depth image and a current 3D depth image. *Id.* In my opinion, A POSITA would understand these depth images to correspond to the first and second 3D representations of this claim element. As I also discuss above, it is my opinion that a POSITA would understand Newcombe’s disclosure of the space between the camera and the stationary objects in the background of a location in the preceding section as teaching an excluded volume necessarily applies equally to both a first depth image and a second depth image. *See supra*, § X.A.1.d. Thus, it is my opinion that a POSITA would have understood that the empty space shown in the figure – i.e., the space in the scene that is not occupied by the fixed objects in the background would constitute an excluded volume for the second 3D representation, according to the court’s construction.

119. Again, as described in the preceding section, this understanding of Newcombe by a POSITA is supported by Newcombe's disclosure of a compatibility test to determine an excluded volume where no rigid object should be present in any depth image of the location, and consequently, where any object detected is therefore a moving or "non-rigid" object at the location. *See, e.g., id.* at 9:26-29. As described above, the compatibility test necessarily incorporates a first 3D representation and a second 3D representation (i.e., a previous 3D depth image and a current 3D depth image). *Supra*, § X.A.1.d. Furthermore, it is my opinion that Newcombe teaches that the excluded volume determined by the compatibility test – i.e., the space beyond the distance threshold where any observed points will be deemed outliers corresponding to non-rigid objects – is also necessarily common to both depth images, because points in both images are required to define the threshold distance and compare the corresponding Euclidean distance between the source and destination points, as is further evidenced visually by Newcombe's Figure 6 below. *See id.* at 8:3-48.

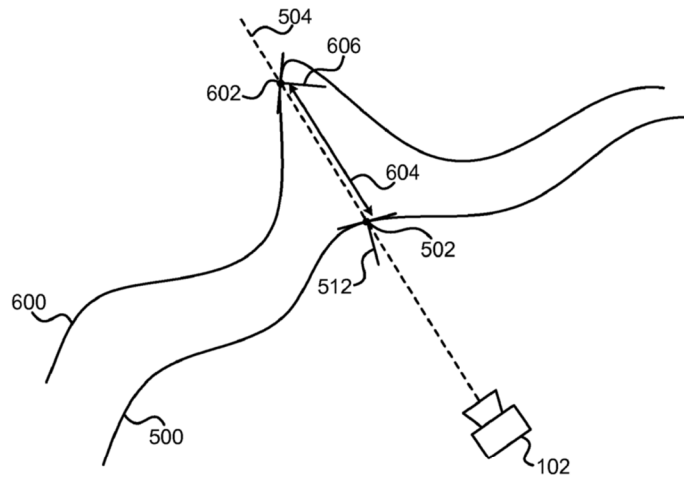


FIG. 6

120. Accordingly, it is my opinion that Newcombe teaches “determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 1.

- f. **[Claim 1e]: “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, and/or”**

121. Newcombe in view of Hähnel teaches this claim element. As discussed

above, a POSITA would understand Newcombe to teach a method for determining excluded volumes corresponding to previous and current 3D depth images. *See supra*, §§ X.A.1.d-e. As is also discussed above, Newcombe discloses that points corresponding to surfaces are identified as outliers if they are in an excluded volume (i.e., in a volume outside a predefined threshold distance), or as corresponding points if they are not. *See id.* Thus, Newcombe teaches that a surface in the first 3D representation that is located in space in the second excluded volume will be identified as an outlier. *See id.*

122. Newcombe further teaches that the outliers, which correspond to moving objects, can be “stored and provided as input to the object segmentation system.” Ex-1005 at 9:25-26. Newcombe discloses that the segmentation of the image can include “segment[ing] all the outlying points and label[ing] these as foreground objects in the depth image.” *Id.* at 9:30-32. The segmentation may be completed using “segmentation algorithms.” *See, e.g., id.* at 10:1-4.

123. Newcombe discloses that moving objects that have been segmented “can be separately processed as desired.” *Id.* at 3:51-54. For example, Newcombe teaches that “the segmentation of moving objects enables these moving objects to be tracked separately and independently of the background. This can be achieved executing a further ICP algorithm on the portion of the depth image that was

identified as a moving object by the segmentation system.” *Id.* at 10:31-36.

124. Newcombe further discloses that, in addition to “independently tracking a moving object in the depth image, an independent dense 3D model of the object can also be generated.” *Id.* at 10:57-59. Newcombe teaches that the creation of a “dense 3D model of a segmented object can be performed in a number of ways,” including a technique that “enables objects to be instantly segmented upon removal from the scene.” *Id.* at 11:3-23.

125. Additional examples of the processing of segmented objects disclosed by Newcombe include “input[ting] to an object recognition system,” classifying according to the “behavior in the scene,” and “control[ling] further aspect of the operation of the system.” *Id.* at 11:28-51. With respect to the last example, Newcombe discloses that a segmented, moving object may be either “arranged to interact” with other objects or may be “deliberately ignored.” *Id.* at 11:51-53.

126. In my opinion, Newcombe’s disclosure that segmented objects may be “separately processed as desired,” as well as the disclosure of the various ways that segmented moving objects may be processed, including with respect to the “removal from the scene” and or being “deliberately ignored,” would be understood by a POSITA to teach that these objects, which necessarily includes the surfaces of the objects, may be “disregarded” or otherwise removed or omitted from a virtual 3D

model generated according to the methods disclosed by Newcombe.

127. However, to the extent that Newcombe may be deemed not to disclose the “disregarding” of a portion of a surface in a 3D representation, Hähnel supplies this teaching. It is my opinion that a POSITA would have been motivated to combine Newcome and Hähnel. Both references concern the use of mobile imaging devices to generate 3D maps of environments utilizing the simultaneous localization and mapping (SLAM) technique. Ex-1005 at 3:28-37; Ex-1007 at 1. Both references also recognize the utility of identifying or segmenting dynamic, moving objects in a surrounding static scene or environment when creating such maps. *See* Ex-1005 at 3:1-6 (“A moving-object segmentation system 114 receives the depth images captured by the mobile depth camera 102, and performs a segmentation operation to identify objects in the depth images that are moving or being manipulated and which are independent of the relatively unchanging background.”); Ex-1007 at 5 (“As can be seen, our approach can reliably identify dynamic aspects and is able to learn maps that include the static aspects only.”). Newcombe teaches that once a dynamic object is identified and segmented from the surrounding environment, it can be “separately processed as desired.” Ex-1005 at 3:51-54. Similarly, Hähnel teaches that once dynamic objects that have been identified, they can be filtered from a 3D map of the surrounding environment. Ex-1007 at

6. Hähnel further teaches that such filtering is advantageous for creating accurate 3D maps and accurate 3D models of dynamic objects. *Id.* at 1. Thus, in a context of mapping a location in a dynamic environment using 3D representation analogous to that disclosed by Newcombe, Hähnel teaches that moving objects (i.e., dynamic objects), can lead to serious errors in mapping a location. *See, e.g.*, Ex-1007 at 1. Hähnel further teaches that these moving objects can be reliably filtered out to obtain accurate models of a location. *Id.* at 1, 6; Fig. 11. It is my opinion that a POSITA would have understood Hähnel’s description of filtering out of dynamic objects to correspond to “disregard[ing]” a “portion of a surface,” as claimed.

128. As discussed above, Newcombe teaches that moving objects can be detected as outliers and segmented and integrated into separate 3D volumes. *Supra*, §§ X.A.1.d-e. Newcombe also discloses that segmented objects can be separately processed as desired. Ex-1005, at 3:51-54. Thus, it is my opinion that it would have been obvious to a POSITA to process the moving objects identified by segmentation in the depth image-based system of Newcombe with the filtering of Hähnel in order to reduce errors in the dense 3D environment model.

129. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D

representation is disregarded in the generation of the virtual 3D model,” as recited by this element of claim 1.

- g. [Claim 1f]: “if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model.”**

130. Newcombe in view of Hähnel teaches this claim element. My analysis of Newcombe in view of Hähnel in the preceding section equally applies to this claim element. As discussed above, Newcombe discloses that points corresponding to surfaces are identified as outliers if they are in an excluded volume (i.e., in a volume outside a predefined threshold distance), or as corresponding points if they are not. *See supra*, §§ X.A.1.d-f. Thus, Newcombe teaches that a surface in the second 3D representation that is located in space in the first excluded volume will be identified as an outlier. *See id.*

131. Again, it is my opinion that a POSITA would understand Newcombe to teach that segmented, moving objects, identified as outliers in an excluded volume of a depth image, can be processed in a variety of ways. *See supra*, § X.A.1.f. Such processing can include removal from a scene or deliberately ignoring the segmented objects. *Id.* Newcombe further teaches that the identification of moving objects and subsequent processing can be performed repeatedly with respect to any pair of depth

images. *See* Ex-1005 at 8:3-9:23. It is my opinion that a POSITA would therefore understand Newcombe's teachings with respect to the determination of excluded volumes and subsequent disregarding of points within those excluded volumes to be equally applicable to both first and second 3D representations. In other words, it is my opinion that a POSITA would understand Newcombe to teach that, if a portion of a surface in a second 3D representation is located in space in the corresponding first excluded volume, the portion of the surface in the second 3D representation should be disregarded, just as a POSITA would understand Newcombe to teach, as described above, that if a portion of a surface in a first 3D representation is located in space in the corresponding second excluded volume, the portion of the surface in the first 3D representation should be disregarded. *See supra*, § X.A.1.f. Thus, it is my opinion that Newcombe teaches this element.

132. As described in the preceding section, it is also my opinion that a POSITA would have found it obvious to combine Hähnel's filtering process with Newcombe's process of identifying and segmenting moving objects in a scene. *Supra*, § X.A.1.f. As further described above, Newcombe discloses that its segmenting process can be applied repeatedly to multiple depth images to obtain multiple separate 3D volumes with respect to the multiple depth images. Ex-1005 at 8:3-9:23. In my opinion, a POSITA therefore would have been motivated and

found it obvious to apply Hähnel's filtering process repeatedly to multiple segmented objects in the separate 3D volumes disclosed by Newcombe's iterative process, in order to reduce any errors in location mapping caused by the segmented moving objects. Accordingly, Newcombe in view of Hähnel teaches this element.

133. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches "if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model," as recited by this element of claim 1.

2. Claim 6 Is Rendered Obvious by Newcombe in view of Hähnel

134. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 6 obvious.

- a. [Claim 6]: "The method according to claim 1, wherein at least part of the surface captured in the first representation and at least part of the surface captured in the second representation are overlapping the same surface part on the rigid object."**

135. Newcombe in view of Hähnel teaches this claim element. As discussed above, Newcombe teaches providing first and second 3D representations of at least part of a surface by scanning at least part of the location with depth images in order to generate the model of the dense 3D environment. For example, Newcombe

teaches that the acquired depth images of the location at least partially overlap while the user scans the environment with the depth camera. Ex-1005 at 5:28-42.

136. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches “wherein at least part of the surface captured in the first representation and at least part of the surface captured in the second representation are overlapping the same surface part on the rigid object,” as recited in claim 6.

3. Claim 7 Is Rendered Obvious by Newcombe in view of Hähnel

137. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 7 obvious.

- a. **[Claim 7]: “The method according to claim 1, wherein the first representation of at least part of the surface is defined as the first representation of at least a first part of the surface, and the second representation of at least part of the surface is defined as the second representation of at least a second part of the surface.”**

138. Newcombe in view of Hähnel teaches this claim element. As discussed above, Newcombe teaches providing first and second 3D representations of a location by scanning at least part of the location using depth images, which are at least partially overlapping. *See supra*, §§ X.A.1.b-c. A POSITA would understand that these overlapping depth images would comprise two separate representations of a first and a second part of the scanned surface.

139. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches “wherein the first representation of at least part of the surface is defined as the first representation of at least a first part of the surface, and the second representation of at least part of the surface is defined as the second representation of at least a second part of the surface,” as recited in claim 7.

4. Claim 8 Is Rendered Obvious by Newcombe in view of Hähnel

140. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 8 obvious.

- a. [Claim 8] “The method according to claim 1, wherein the method comprises determining a first scan volume in space related to the first representation of at least part of the surface, and determining a second scan volume in space related to the second representation of at least part of the surface.”**

141. Newcombe teaches this claim element. As discussed above, Newcombe teaches providing first and second 3D representations of at least part of a surface by scanning at least part of the location. *See supra*, §§ X.A.1.b-c. Newcombe teaches that each depth image taken by the depth camera contains image data with corresponding distance data. *See, e.g.*, Ex-1005 at 4:17-35. In this example, because the distance data is determined by a time-of-flight reflected infrared light, a POSITA would understand that each depth image defines corresponding

scan volume between the depth camera and the surface being scanned.

142. Further, in describing the disclosed segmentation system, Newcombe refers to the “camera’s frustum (i.e. not only the front faces which the camera sees).” Ex-1005 at 11:15-16. In my opinion, a POSITA at the time the ’551 patent was filed would understand Newcombe’s reference to a frustum to refer to the entire field of view of the camera used to obtain the 3D depth images (i.e., the first and second 3D representations). It is further my opinion that a POSITA would understand that this field of view of the camera corresponds to a scan volume in space that can be determined for each of the first and second representations of a surface.

143. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches “wherein the method comprises determining a first scan volume in space related to the first representation of at least part of the surface, and determining a second scan volume in space related to the second representation of at least part of the surface,” as recited in claim 8.

5. Claim 9 Is Rendered Obvious by Newcombe in view of Hähnel

144. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 9 obvious.

- a. **[Claim 9]: “The method according to claim 1, wherein the scan volume is defined by focusing optics in the 3D scanner and the distance to the surface**

which is captured.”

145. Newcombe in view of Hähnel teaches this claim element. As discussed above, Newcombe teaches a scan volume with respect to each of the first and second 3D representations. *See supra*, § X.A.4.a. Furthermore, as discussed above, Newcombe discloses the use of a depth camera for 3D scanning. *See supra*, §§ X.A.1.b-c. In my opinion, it would have been obvious to a POSITA that the scan volume created when scanning a surface to create a 3D representation would have been defined by the focusing optics of the depth camera and the distance to the surface which is captured.

146. Newcombe describes that “[t]he depth information may be obtained using any suitable technique including, but not limited to, time of flight, structured light, and stereo images.” Ex-1005 at 4:11-13. In my opinion, a POSITA at the time the ’551 patent was filed would know that these technologies use one or more electro-optical sensors with associated focusing optics. Therefore, it is my opinion that a POSITA would know that the scan volume created by the depth camera when scanning a surface to create 3D representations would depend on the field of view and the depth of field of the electro-optical sensor, both of which depend on the focusing optics and the distance to the surface captured by the depth camera.

147. Accordingly, it is my opinion that Newcombe in view of Hähnel

teaches “wherein the scan volume is defined by focusing optics in the 3D scanner and the distance to the surface which is captured,” as recited in claim 9.

6. Claim 10 Is Rendered Obvious by Newcombe in view of Hähnel

148. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 10 obvious.

- a. [Claim 10] “The method according to claim 1, wherein the first scan volume related to the first representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the first representation; and the second scan volume related to the second representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the second representation.”**

149. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to claim 8, Newcombe describes determining a first and second scan volume related to the first and second representations of at least a first and second part of the surface. *See supra*, § X.A.4.a. As discussed with respect to claim 9, it is my opinion that the depth camera taught by Newcombe would necessarily have focusing optics that a POSITA would understand to define a scan volume. *See supra*, § X.A.5.a.

150. Further, as described above, it is my opinion that a POSITA at the time

the '551 patent was filed would have understood Newcombe's disclosure of a camera frustum to correspond to a scan volume defined by the focusing optics of the camera and the distance from the camera to the surface captured by the camera. *See supra*, §§ X.A.4.a-5.a. As is also discussed above, Newcombe discloses obtaining first and second 3D representations, or depth images, with the disclosed camera. *See supra*, §§ X.A.1.b-c.

151. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches "wherein the first scan volume related to the first representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the first representation; and the second scan volume related to the second representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the second representation," as recited by claim 10.

7. Claim 11 Is Rendered Obvious by Newcombe in view of Hähnel

152. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 11 obvious.

- a. **[Claim 11]: "The method according to claim 1, wherein the first excluded volume and second excluded volume in space where no surface can be present corresponds to the first scan volume and the second scan volume, respectively."**

153. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to the excluded volume and scan volume elements of claims 1 and 8-10, a POSITA would understand the scanned depths that do not correspond to a stationary surface in the background to be excluded volumes where the surfaces of the background being modeled in the dense 3D environment cannot be present.

154. Further, as discussed above, a POSITA would understand the camera frustum disclosed by Newcombe to correspond to first and second scan volumes. *See supra*, §§ X.A.4.a.-6.a. Newcombe's disclosure of the camera frustum is made with respect to a technique for creating a dense 3D model of a segmented object, which as described above, Newcombe teaches correspond to outlier points in a first and/or second excluded volume. *See supra*, §§ X.A.1.a-g. Specifically, Newcombe states that "after the outliers have been detected (as described above), every voxel in the dense 3D model within the camera's frustum (i.e. not only the front faces which the camera sees) is projected into the current depth map." Ex-1005 at 11:13-17. It is my opinion that a POSITA would understand this disclosure in Newcombe to teach that the first and second excluded volumes would correspond to the first and second scan volumes, respectively.

155. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches "wherein the first excluded volume and second excluded volume in space

where no surface can be present corresponds to the first scan volume and the second scan volume, respectively,” as recited by claim 11.

8. Claim 12 Is Rendered Obvious by Newcombe in view of Hähnel

156. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 12 obvious.

a. [Claim 12]: “The method according to claim 1, wherein the volume of the 3D scanner itself is defined as an excluded volume.”

157. To the extent that Newcombe in view of Hähnel does not explicitly describe that the 3D scanner is defined as an excluded volume, it is my opinion that this claim element would be obvious to a POSITA in light of the teachings of Newcombe. A POSITA would recognize that the depth camera of Newcombe would necessarily be a part of “an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present,” per the district court’s construction, because it is impossible for the depth camera to be included in any scan volume. Indeed, as discussed above with respect to claims 8-11, a POSITA would understand Newcombe’s disclosure of a camera frustum to correspond to a scan volume and would further understand that a camera itself is not included in its own viewing frustum, which automatically makes the 3D camera part of an excluded volume.

Therefore, it is my opinion that Newcombe in view of Hähnel renders claim 12 obvious.

158. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches “wherein the volume of the 3D scanner itself is defined as an excluded volume,” as recited by claim 12.

9. Claim 15 Is Rendered Obvious by Newcombe in view of Hähnel

159. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 15 obvious.

- a. [Claim 15]: “The method according to claim 1, wherein the first representation of at least part of a surface is a first subscan of at least part of the location, and the second representation of at least part of the surface is a second subscan of at least part of the location.”**

160. It is my opinion that a POSITA would understand that Newcombe in view of Hähnel describes the elements of this claim. Although Newcombe does not explicitly use the term “subscan” with respect to the first and second 3D representations comprising depth images that it teaches, a POSITA would understand that these multiple depth images are subscans of at least a part of the location depicted in the scene and that together these subscans will be used to generate a virtual 3D model that is the dense 3D model. *See, e.g.,* Ex-1001 at 3:44-

49 (“Typically when scanning an object, such as a set of teeth, more representations or subscans may be provided . . . during a complete scanning process.”).

161. Newcombe further discloses that a scene in a current depth frame may “at least partially overlap” with a scene in a previous depth frame. Ex-1005 at 5:36-37. It is my opinion that a POSITA would understand that “partially overlap,” as used by Newcombe, implies that each scan provides a subset of 3D points representing the scanned surface (some overlapping and some not overlapping), and therefore each scan is a subscan of the overall surface being scanned.

162. Accordingly, it is my opinion that Newcombe in view of Hähnel renders obvious “wherein the first representation of at least part of a surface is a first subscan of at least part of the location, and the second representation of at least part of the surface is a second subscan of at least part of the location,” as recited by claim 15.

10. Claim 16 Is Rendered Obvious by Newcombe in view of Hähnel

163. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 16 obvious.

- a. **[Claim 16]: “The method according to claim 1, wherein the first representation of at least part of a surface is a provisional virtual 3D model comprising the subscans of the location acquired already, and the second representation of at least part of the surface is**

a second subscan of at least part of the location.”

164. It is my opinion that a POSITA would understand that Newcombe in view of Hähnel renders obvious the elements of claim 16. As discussed with respect to claims 1 and 15 above, a POSITA would understand that the first and second 3D representations of Newcombe can be considered subscans of the final dense 3D model of the location because the model is not complete until all the depth images are used to finalize the three-dimensional model. *See, e.g.*, Ex-1005 at 3:28-30.

165. Accordingly, it is my opinion that Newcombe in view of Hähnel renders obvious “wherein the first representation of at least part of a surface is a provisional virtual 3D model comprising the subscans of the location acquired already, and the second representation of at least part of the surface is a second subscan of at least part of the location,” as recited by claim 16.

11. Claim 17 Is Rendered Obvious by Newcombe in view of Hähnel

166. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 17 obvious.

- a. [Claim 17]: “The method according to claim 1, wherein acquired subscans of the location are adapted to be added to the provisional virtual 3D model concurrently with the acquisition of the subscans.”**

167. It is my opinion that a POSITA would find the elements of this claim obvious in light of Newcombe’s teachings. It would have been obvious to a POSITA

to update a real-time provisional model of a dense 3D environment in order to obtain a preview of the final model.

168. Accordingly, it is my opinion that Newcombe in view of Hähnel, in combination with the knowledge and creativity of a POSITA renders obvious, “wherein acquired subscans of the location are adapted to be added to the provisional virtual 3D model concurrently with the acquisition of the subscans,” as recited by claim 17.

12. Claim 20 Is Rendered Obvious by Newcombe in view of Hähnel

169. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 20 obvious.

a. [Claim 20 pre]: “The method according to claim 1, where in the method includes:”

[Claim 20a]: “disregarding the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model, and”

170. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to second excluded volume element of claim 1, applying the filtering of dynamic objects described by Hähnel to the moving objects identified by Newcombe’s segmentation would remove, and therefore disregard any moving objects in the excluded volumes and reduce errors in the dense 3D environment

model. *See supra*, § X.A.1.f.

- b. [Claim 20b]: “disregarding the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model.”**

171. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to first excluded volume element of claim 1, the filtering method described by Hähnel removes, and therefore disregards, any surface measurements corresponding to moving objects in the first excluded volume of the first representation. *See supra*, § X.A.1.g.

172. Accordingly, it is my opinion that Newcombe in view of Hähnel renders obvious “disregarding the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model, and disregarding the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model,” as recited by claim 20.

13. Claim 21 Is Rendered Obvious by Newcombe in view of Hähnel

173. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 21 obvious.

- a. [Claim 21 pre]: “The method according to claim 1, wherein:”**

[Claim 21a]: “the first excluded volume is determined to be a space where no surface of the rigid object is present, and”

174. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to first excluded volume element of claim 1, Newcombe discloses a compatibility test to determine a first excluded volume that does not include any points corresponding to surfaces of rigid objects. *See supra*, § X.A.1.d. Further, the filtering method described by Hähnel removes, and therefore disregards, any surface measurements corresponding to moving objects in the first excluded volume of the first representation. *See supra*, § X.A.1.g.

175. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches “wherein . . . the first excluded volume is determined to be a space where no surface of the rigid object is present,” as recited by claim 21.

b. [Claim 21b]: “the second excluded volume is determined to be a space where no surface of the rigid object is present.”

176. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to second excluded volume element of claim 1, Newcombe discloses a compatibility test to determine a second excluded volume that does not include any points corresponding to surfaces of rigid objects. *See supra*, § X.A.1.e. Further, the filtering method described by Hähnel removes, and therefore disregards,

any surface measurements corresponding to moving objects in the second excluded volume of the second representation. *See supra*, § X.A.1.f.

177. Accordingly, it is my opinion that Newcombe in view of Hähnel renders obvious, “wherein the first excluded volume is determined to be a space where no surface of the rigid object is present, and the second excluded volume is determined to be a space where no surface of the rigid object is present,” as recited by claim 21.

14. Claim 22 Is Rendered Obvious by Newcombe in view of Hähnel

178. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 22 obvious.

- a. **[Claim 22]: “A nontransitory computer readable medium encoded with a computer program product comprising program code for causing a data processing system to perform the method of claim 1, when said program code is executed on the data processing system.”**

179. Newcombe in view of Hähnel teaches this claim element. Newcombe teaches that the steps of its algorithm and method are all performed on a computer or computer gaming system. *See, e.g.*, Ex-1005 at 13:53-59. In my opinion, a POSITA would therefore understand Newcombe in view of Hähnel to teach or render obvious a nontransitory computer readable medium encoded with a computer

program product comprising program code for causing a data processing system to perform the method disclosed by Newcombe in view of Hähnel.

180. Accordingly, Newcombe in view of Hähnel renders obvious, “a nontransitory computer readable medium encoded with a computer program product comprising program code for causing a data processing system to perform the method of claim 1, when said program code is executed on the data processing system,” as recited by claim 22.

15. Claim 23 Is Rendered Obvious by Newcombe in view of Hähnel

181. As explained below, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, every element of claim 23.

- a. **[Claim 23 pre]: “A system for detecting a movable object in a location when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the system comprises a hardware processor configured to:”**

182. As discussed above, in Section VIII.A, for the purposes of this declaration, I have treated the preamble of claim 23 as limiting. It is my opinion that Newcombe teaches this claim element.

183. Newcombe teaches detecting and identifying movable objects from a series of depth images of a scene using segmentation. Ex-1005 at Abstract. Multiple

depth images are obtained using a depth camera that gathers both image and depth data. *Id.* at 2:46-53; 3:25-27. It is my opinion that this depth camera corresponds to the claimed “3D scanner.” Newcombe teaches detecting moving objects and segmenting them from a scene comprised of fixed objects that is being scanned by depth images. *Id.* at 1:49-65. It is my opinion that this corresponds to the claimed “system for detecting a movable object in a location when scanning a rigid object in the location.” Newcombe also teaches generating a dense 3D environment model. *Id.* at 3:21-37. It is my opinion that this corresponds to “generating a virtual 3D model of the rigid object,” as claimed.

184. Newcombe further discloses a “computing-based device,” which “comprises one or more processors” and which “may be implemented as any form of a computing and/or electronic device,” in which the image segmentation techniques and methods taught by the reference may be implemented. Ex-1005 at 12:29-13:64; Fig. 7. It is my opinion that a POSITA would understand Newcombe to therefore disclose that its methods may be implemented on a computer system comprising a hardware processor.

185. Accordingly, it is my opinion that Newcombe teaches “a system for detecting a movable object in a location when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object,

wherein the system comprises a hardware processor” as recited by the preamble of claim 23.

b. [Claim 23a]: “provide a first 3D representation of at least part of a surface by scanning at least part of the location;”

186. Newcombe teaches this claim element. Newcombe teaches using a mobile depth camera to obtain depth images of a location. *Id.* at 2:46-53. These depth images include both image data and depth data. *See id.* at 3:64-4:10. The depth images are used to form the dense 3D model of the environment. *Id.* at 3:25-27. It is my opinion that a depth image of the location, as described by Newcombe, corresponds to a “first 3D representation of at least part of a surface scanning at least part of the location,” as claimed.

187. Accordingly, it is my opinion that Newcombe teaches “a hardware processor configured to provide a first 3D representation of at least part of a surface of the location,” as recited in this element of claim 23.

c. [Claim 23b]: “provide a second 3D representation of at least part of the surface by scanning at least part of the location;”

188. Newcombe teaches this claim element. As discussed above, Newcombe describes obtaining multiple depth images of the location to build the dense 3D model. It is my opinion that one of the additional depth images of the

location, as described by Newcombe, corresponds to a “second 3D representation of at least part of the surface by scanning at least part of the location,” as claimed.

189. Furthermore, Newcombe discloses two separate scanning processes, through which multiple depth images may be provided. Specifically, Newcombe discloses the following:

Firstly, a depth image of the scene to be segmented (as captured by the mobile depth camera) is received 400. If the segmentation operation is making use of the camera-tracking results, then this current depth image is received at the frame alignment engine 318. The current depth image is also known as the destination depth image. A previous depth image of at least a portion of the scene is then retrieved 402. The previous depth image is also known as the source depth image. In some examples, the previous depth image may be the preceding frame captured by the depth camera and stored at a storage device. In alternative examples, the previous depth image may be generated from the dense 3D model stored in graphics memory by requesting the previous depth image from the GPU. The GPU can generate the previous depth image from the dense 3D model stored on the graphics memory by determining the view from a virtual camera at the last known location and orientation of the real depth camera looking in the dense 3D model. The use of the dense 3D model provides a higher resolution previous depth image.

Id. at 6:49-67. It is my opinion that each of the previous and current depth images

disclosed by Newcombe correspond to the first and second 3D representations recited in the claim.

190. Accordingly, it is my opinion that Newcombe teaches “a hardware processor configured to ... provide a second 3D representation of at least part of the surface by scanning at least part of the location,” as recited by this element of claim 23.

- d. **[Claim 23c]: “determine for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

191. Newcombe teaches this claim element. Newcombe describes identifying, through segmentation, moving objects in the foreground of a location. It is my opinion that a POSITA at the time the '551 patent was filed would understand Newcombe to disclose this claim element in connection with the disclosed segmentation process. As described above, Newcombe discloses a previous 3D depth image and a current 3D depth image, which A POSITA would understand correspond to first and second 3D representations. *See supra*, §§ X.A.15.b-c. It is further my opinion that a POSITA at the time the '551 patent was filed would understand that all spaces within the obtained 3D depth images for the location that *do not* correspond to a surface within the background of the location would collectively constitute an excluded volume, as recited by the claims, and more

particularly would constitute an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present, according to the court's claim construction. For example, it is my opinion that a POSITA would have understood that the empty space shown in Figure 1, shown below, between camera 102 and the stationary objects in the room would constitute an excluded volume for the first 3D representation, according to the court's construction.

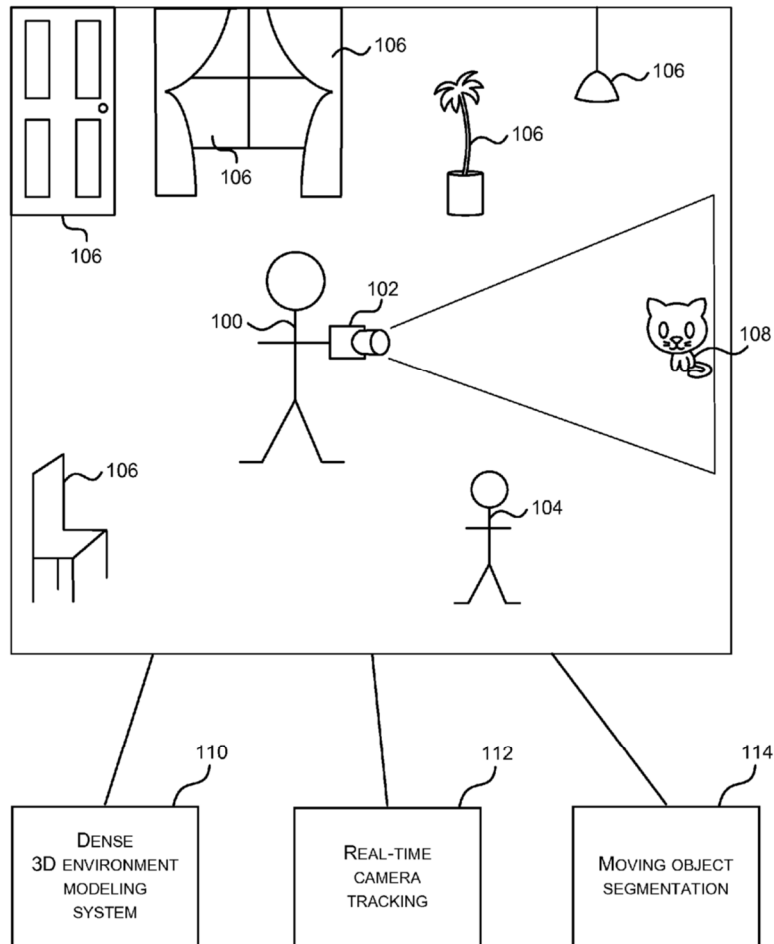


FIG. 1

192. It is my opinion that this understanding of Newcombe is supported by the reference's disclosure of the performance of compatibility tests to determine whether points in two depth images are corresponding points, which can be attributed to rigid objects, or are outliers, which can be attributed to moving objects. *See* Ex-1005 at 8:27-29. Specifically, it is my opinion that Newcombe discloses that at least the first compatibility test can be used to determine an excluded volume

where no rigid object should be present in any depth image of the location, and consequently, where any object detected is therefore a moving or “non-rigid” object at the location. *See, e.g., id.* at 9:26-29.

193. As described by Newcombe, the first compatibility test relies on defining a threshold distance for a measurement between points in two 3D depth images – i.e., a “source point” in a previous depth image and a “destination point” in a current depth image. *Id.* at 8:3-33. As Newcombe explains, “[i]f the destination point is too far away from the source point, then it is considered to be an outlier (i.e. non-corresponding)” *Id.* at 8:33-35. Newcombe further explains that “outliers” or “outlying points” refer to “points in the depth image where something has changed since the previous depth image” and “can be caused by dynamic or non-rigid objects.” *Id.* at 9:26-29.

194. This scenario is demonstrated in Figure 6, shown below, of Newcombe: “[T]he example of Fig. 6 shows the case where differences between the previous depth image 500 and the current depth image 600 (e.g. due to new or moving objects) result in the distance 604 between the source and destination points being greater than the predefined distance threshold, and hence the destination point is labeled as an outlier.” *Id.* at 8:42-48. Newcombe further explains the process as follows:

Moving object segmentation using depth images is described. In an

example, a moving object is segmented from the background of a depth image of a scene received from a mobile depth camera. A previous depth image of the scene is retrieved, and compared to the current depth image using an iterative closest point algorithm. The iterative closest point algorithm includes a determination of a set of points that correspond between the current depth image and the previous depth image. During the determination of the set of points, one or more outlying points are detected that do not correspond between the two depth images, and the image elements at these outlying points are labeled as belonging to the moving object. In examples, the iterative closest point algorithm is executed as part of an algorithm for tracking the mobile depth camera, and hence the segmentation does not add substantial additional computational complexity.

Id. at 1:49-65.

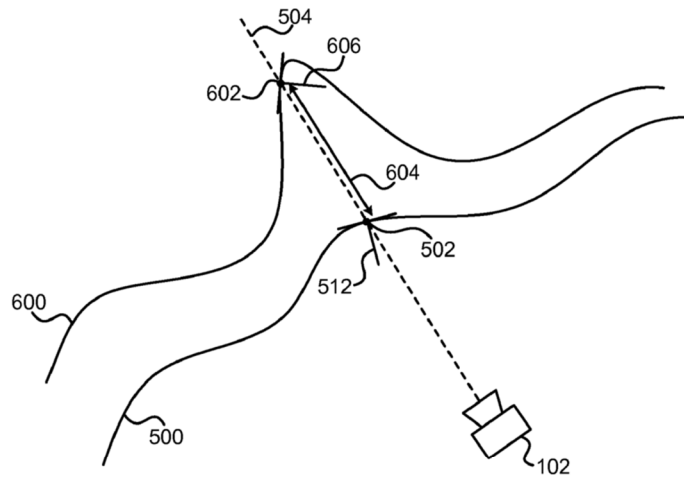


FIG. 6

195. Thus, it is my opinion that a POSITA would understand that by defining a distance threshold for the two three-dimensional depth images beyond which any point will be deemed an outlier, Newcombe teaches the determination of an excluded volume, and particularly, an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present,” as construed by the district court.

196. Accordingly, it is my opinion that Newcombe teaches, and therefore renders obvious, “a hardware processor configured to ... determine for the first 3D

representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 23.

- e. **[Claim 23d]: “determine for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

197. Newcombe teaches this claim element. My analysis and discussion in the preceding section of Newcombe’s segmentation process and the compatibility tests for determining corresponding points and outliers applies equally to this claim element. *See supra*, § X.A.15.d. As described above, Newcombe discloses a previous 3D depth image and a current 3D depth image. *Id.* In my opinion, a POSITA would understand these depth images to correspond to the first and second 3D representations of this claim element. As I also discuss above, it is my opinion that a POSITA would understand Newcombe’s disclosure of the space between the camera and the stationary objects in the background of a location in the preceding section as teaching an excluded volume necessarily applies equally to both a first depth image and a second depth image. *See supra*, § X.A.15.d. Thus, it is my opinion that a POSITA would have understood that the empty space shown in the figure – i.e., the space in the scene that is not occupied by the fixed objects in the background would constitute an excluded volume for the second 3D representation,

according to the court's construction.

198. Again, as described in the preceding section, this understanding of Newcombe by a POSITA is supported by Newcombe's disclosure of a compatibility test to determine an excluded volume where no rigid object should be present in any depth image of the location, and consequently, where any object detected is therefore a moving or "non-rigid" object at the location. *See, e.g., id.* at 9:26-29. As described above, the compatibility test necessarily incorporates a first 3D representation and a second 3D representation (i.e., a previous 3D depth image and a current 3D depth image). *Supra*, § X.A.15.d. Furthermore, it is my opinion that Newcombe teaches that the excluded volume determined by the compatibility test – i.e., the space beyond the distance threshold where any observed points will be deemed outliers corresponding to non-rigid objects – is also necessarily common to both depth images, because points in both images are required to define the threshold distance and compare the corresponding Euclidean distance between the source and destination points, as is further evidenced visually by Newcombe's Figure 6 below. *See id.* at 8:3-48.

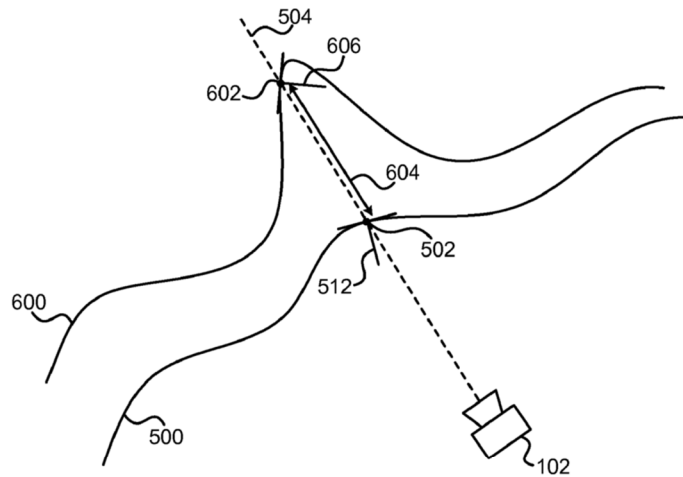


FIG. 6

199. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, “a hardware processor configured to ... determine for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 23.

- f. **[Claim 23e]: “disregard the portion of the surface in the first 3D representation in the generation of the virtual 3D model, if a portion of the surface in the first 3D representation is located in space in the second excluded volume, and/or”**

200. Newcombe in view of Hähnel teaches this claim element. As discussed above, a POSITA would understand Newcombe to teach a method for determining excluded volumes corresponding to previous and current 3D depth images. *See supra*, §§ X.A.15.d-e. As is also discussed above, Newcombe discloses that points corresponding to surfaces are identified as outliers if they are in an excluded volume (i.e., in a volume outside a predefined threshold distance), or as corresponding points if they are not. *See id.* Thus, Newcombe teaches that a surface in the first 3D representation that is located in space in the second excluded volume will be identified as an outlier. *See id.*

201. Newcombe further teaches that the outliers, which correspond to moving objects, can be “stored and provided as input to the object segmentation system.” Ex-1005 at 9:25-26. Newcombe discloses that the segmentation of the image can include “segment[ing] all the outlying points and label[ing] these as foreground objects in the depth image.” *Id.* at 9:30-32. The segmentation may be completed using “segmentation algorithms.” *See, e.g., id.* at 10:1-4.

202. Newcombe discloses that moving objects that have been segmented “can be separately processed as desired.” *Id.* at 3:51-54. For example, Newcombe teaches that “the segmentation of moving objects enables these moving objects to be tracked separately and independently of the background. This can be achieved

executing a further ICP algorithm on the portion of the depth image that was identified as a moving object by the segmentation system.” *Id.* at 10:31-36.

203. Newcombe further discloses that, in addition to “independently tracking a moving object in the depth image, an independent dense 3D model of the object can also be generated.” *Id.* at 10:57-59. Newcombe teaches that the creation of a “dense 3D model of a segmented object can be performed in a number of ways,” including a technique that “enables objects to be instantly segmented upon removal from the scene.” *Id.* at 11:3-23.

204. Additional examples of the processing of segmented objects disclosed by Newcombe include “input[ting] to an object recognition system, classification according the “behavior in the scene,” and “control[ling] further aspect of the operation of the system.” *Id.* at 11:28-51. With respect to the last example, Newcombe discloses that a segmented, moving object may be either “arranged to interact” with other objects or may be “deliberately ignored.” *Id.* at 11:51-53.

205. In my opinion, Newcombe’s disclosure that segmented objects may be “separately processed as desired,” as well as the disclosure of the various ways that segmented moving objects may be processed, including with respect to the “removal from the scene” and or being “deliberately ignored,” would be understood by a POSITA to teach that these objects, which necessarily includes the surfaces of the

objects, may be “disregarded” or otherwise removed or omitted from a virtual 3D model generated according to the methods disclosed by Newcombe.

206. However, to the extent that Newcombe may be deemed not to disclose the “disregarding” of a portion of a surface in a 3D representation, Hähnel supplies this teaching. Like Newcombe, Hähnel discloses 3D modeling processes based on simultaneous localization and mapping (SLAM). *See* Ex-1005 at 3:35-37; Ex-1007 at 1. Thus, in a context of mapping a location in a dynamic environment using 3D representation analogous to that disclosed by Newcombe, Hähnel teaches that moving objects (i.e., dynamic objects), can lead to serious errors in mapping a location. *See, e.g.*, Ex-1007 at 1. Hähnel further teaches that these moving objects can be reliably filtered out to obtain accurate models of a location. *Id.* at 1, 6; Fig. 11. It is my opinion that a POSITA would have understood Hähnel’s description of filtering out of dynamic objects to correspond to “disregard[ing]” a “portion of a surface,” as claimed.

207. As discussed above, Newcombe teaches that moving objects can be detected as outliers and segmented and integrated into separate 3D volumes. *See supra*, §§ X.A.15.d-e. Newcombe also discloses that segmented objects can be separately processed as desired. Ex-1005, at 3:51-54. Thus, it is my opinion that it would have been obvious to a POSITA to process the moving objects identified by

segmentation in the depth image-based system of Newcombe with the filtering of Hähnel in order to reduce errors in the dense 3D environment model.

208. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious “a hardware processor configured to ... disregard the portion of the surface in the first 3D representation in the generation of the virtual 3D model, if a portion of the surface in the first 3D representation is located in space in the second excluded volume,” as recited by this element of claim 23.

- g. [Claim 23f]: “disregard the portion of the surface in the second 3D representation in the generation of the virtual 3D model, if a portion of the surface in the second 3D representation is located in space in the first excluded volume.”**

209. Newcombe in view of Hähnel teaches this claim element. My analysis of Newcombe in view of Hähnel in the preceding section equally applies to this claim element. As discussed above, Newcombe discloses that points corresponding to surfaces are identified as outliers if they are in an excluded volume (i.e., in a volume outside a predefined threshold distance), or as corresponding points if they are not. *See supra*, §§ X.A.15.d-f. Thus, Newcombe teaches that a surface in the second 3D representation that is located in space in the first excluded volume will be identified as an outlier. *See id.*

210. Again, it is my opinion that a POSITA would understand Newcombe to teach that segmented, moving objects, identified as outliers in an excluded volume of a depth image, can be processed in a variety of ways. *See supra*, § X.A.15.f. Such processing can include removal from a scene or deliberately ignoring the segmented objects. *Id.* Newcombe further teaches that the identification of moving objects and subsequent processing can be performed repeatedly with respect to any pair of depth images. *See* Ex-1005 at 8:3-9:23. It is my opinion that a POSITA would therefore understand Newcombe's teachings with respect to the determination of excluded volumes and subsequent disregarding of points within those excluded volumes to be equally applicable to both first and second 3D representations. In other words, it is my opinion that a POSITA would understand Newcombe to teach that, if a portion of a surface in a second 3D representation is located in space in the corresponding first excluded volume, the portion of the surface in the second 3D representation should be disregarded, just as a POSITA would understand Newcombe to teach, as described above, that if a portion of a surface in a first 3D representation is located in space in the corresponding second excluded volume, the portion of the surface in the first 3D representation should be disregarded. *See supra*, § X.A.15.f. Thus, it is my opinion that Newcombe teaches this element.

211. As described in the preceding section, it is also my opinion that a

POSITA would have found it obvious to combine Hähnel's filtering process with Newcombe's process of identifying and segmenting moving objects in a scene. *Supra*, § X.A.15.f. As further described above, Newcombe discloses that its segmenting process can be applied repeatedly to multiple depth images to obtain multiple separate 3D volumes with respect to the multiple depth images. Ex-1005 at 8:3-9:23. In my opinion, a POSITA therefore would have been motivated and found it obvious to apply Hähnel's filtering process repeatedly to multiple segmented objects in the separate 3D volumes disclosed by Newcombe's iterative process, in order to reduce any errors in location mapping caused by the segmented moving objects. Accordingly, Newcombe in view of Hähnel teaches this element.

212. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, "a hardware processor configured to ... disregard the portion of the surface in the second 3D representation in the generation of the virtual 3D model, if a portion of the surface in the second 3D representation is located in space in the first excluded volume," as recited by this element of claim 23.

16. Claim 24 Is Rendered Obvious by Newcombe in view of Hähnel

213. As explained below, it is my opinion that Newcombe in view of Hähnel renders every element of claim 24 obvious.

- a. **[Claim 24 pre]: “The system according to claim 23, where in the hardware processor is configured to:”**

[Claim 24a]: “disregard the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model, and”

214. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to claim elements [1e]-[1f], it is my opinion that a POSITA would understand Newcombe to disclose disregarding a portion of a surface in a first 3D representation that is located in space in a second exclude volume. *See supra*, §§ X.A.1.f-g. Further, as is also discussed with respect to claim elements [1e]-[1f], applying the filtering of dynamic objects described by Hähnel to the moving objects identified by Newcombe’s segmentation would remove, and therefore disregard any moving objects in the excluded volumes and reduce errors in the dense 3D environment model. *Id.*

215. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious “a hardware processor configured to ... disregard the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model,” as recited by this element of claim 24.

- b. **[Claim 24b]: “disregard the portion of the surface in the second 3D representation that is located in space**

in the first excluded volume in the generation of the virtual 3D model.”

216. Newcombe in view of Hähnel teaches this claim element. As discussed above with respect to claim elements [1e]-[1f], it is my opinion that a POSITA would understand Newcombe to disclose disregarding a portion of a surface in a first 3D representation that is located in space in a second exclude volume. *See supra*, §§ X.A.1.f-g. Further, as is also discussed with respect to claim elements [1e]-[1f], applying the filtering of dynamic objects described by Hähnel to the moving objects identified by Newcombe’s segmentation would remove, and therefore disregard any moving objects in the excluded volumes and reduce errors in the dense 3D environment model. *See supra*, §§ X.A.1.f-g.

217. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, “a hardware processor configured to ... disregard the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model,” as recited by this element of claim 24.

17. Claim 25 Is Rendered Obvious by Newcombe in view of Hähnel

218. As explained below, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, every element of claim 25.

- a. **[Claim 25 pre]: “A method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the method comprises:”**

219. As discussed above, in section VIII.A, for the purposes of this declaration, I have treated the preamble of claim 25 as limiting. It is my opinion that Newcombe teaches this claim element.

220. Newcombe teaches detecting and identifying movable objects from a series of depth images of a scene using segmentation. Ex-1005 at Abstract. Multiple depth images are obtained using a depth camera that gathers both image and depth data. *Id.* at 2:46-53; 3:25-27. It is my opinion that this depth camera corresponds to the claimed “3D scanner.” Newcombe teaches detecting moving objects and segmenting them from a scene comprised of fixed objects that is being scanned by depth images. *Id.* at 1:49-65. It is my opinion that this corresponds to the claimed “method for detecting a movable object in a location when scanning a rigid object in the location.” Newcombe also teaches generating a dense 3D environment model. *Id.* at 3:21-37. It is my opinion that this corresponds to “generating a virtual 3D model of the rigid object,” as claimed.

221. Accordingly, it is my opinion that Newcombe teaches a “method for detecting a movable object in a location, when scanning a rigid object in the location

by means of a 3D scanner for generating a virtual 3D model of the rigid object,” as recited by the preamble of claim 25.

b. [Claim 25a]: “providing a first 3D representation of at least part of a surface by scanning at least part of the location;”

222. Newcombe teaches this claim element. Newcombe in view of Hähnel teaches this claim element. Newcombe teaches using a mobile depth camera to obtain depth images of a location. *Id.* at 2:46-53. These depth images include both image data and depth data. *See id.* at 3:64-4:10. The depth images are used to form the dense 3D model of the environment. *Id.* at 3:25-27. It is my opinion that a depth image of the location, as described by Newcombe, corresponds to a “first 3D representation of at least part of a surface scanning at least part of the location,” as claimed.

223. Accordingly, it is my opinion that Newcombe teaches “providing a first 3D representation of at least part of a surface of the location,” as recited in this element of claim 25.

c. [Claim 25b]: “providing a second 3D representation of at least part of the surface by scanning at least part of the location;”

224. Newcombe teaches this claim element. As discussed above, Newcombe describes obtaining multiple depth images of the location to build the

dense 3D model. It is my opinion that one of the additional depth images of the location, as described by Newcombe, corresponds to a “second 3D representation of at least part of the surface by scanning at least part of the location,” as claimed.

225. Furthermore, Newcombe discloses two separate scanning processes, through which multiple depth images may be provided. Specifically, Newcombe discloses the following:

Firstly, a depth image of the scene to be segmented (as captured by the mobile depth camera) is received 400. If the segmentation operation is making use of the camera-tracking results, then this current depth image is received at the frame alignment engine 318. The current depth image is also known as the destination depth image. A previous depth image of at least a portion of the scene is then retrieved 402. The previous depth image is also known as the source depth image. In some examples, the previous depth image may be the preceding frame captured by the depth camera and stored at a storage device. In alternative examples, the previous depth image may be generated from the dense 3D model stored in graphics memory by requesting the previous depth image from the GPU. The GPU can generate the previous depth image from the dense 3D model stored on the graphics memory by determining the view from a virtual camera at the last known location and orientation of the real depth camera looking in the dense 3D model. The use of the dense 3D model provides a higher resolution previous depth image.

226. *Id.* at 6:49-67. It is my opinion that each of the previous and current depth images disclosed by Newcombe correspond to the first and second 3D representations recited in the claim.

227. Accordingly, it is my opinion that Newcombe teaches “providing a second 3D representation of at least part of the surface by scanning at least part of the location,” as recited by this element of claim 25.

- d. **[Claim25c]: “determining for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation, and/or determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

228. Newcombe teaches this claim element. Newcombe describes identifying, through segmentation, moving objects in the foreground of a location. It is my opinion that a POSITA at the time the '551 patent was filed would understand Newcombe to disclose this claim element in connection with the disclosed segmentation process. As described above, Newcombe discloses a previous 3D depth image and a current 3D depth image, which a POSITA would understand correspond to first and second 3D representations. *See supra*, §§ X.A.1.b-c. It is further my opinion that a POSITA at the time the '551 patent was filed would understand that all spaces within the obtained 3D depth images for the

location that *do not* correspond to a surface within the background of the location would collectively constitute an excluded volume, as recited by the claims, and more particularly would constitute an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present, according to the court's claim construction. For example, it is my opinion that a POSITA would have understood that the empty space shown in Figure 1, shown below, between camera 102 and the stationary objects in the room would constitute an excluded volume for the first 3D representation, according to the court's construction.

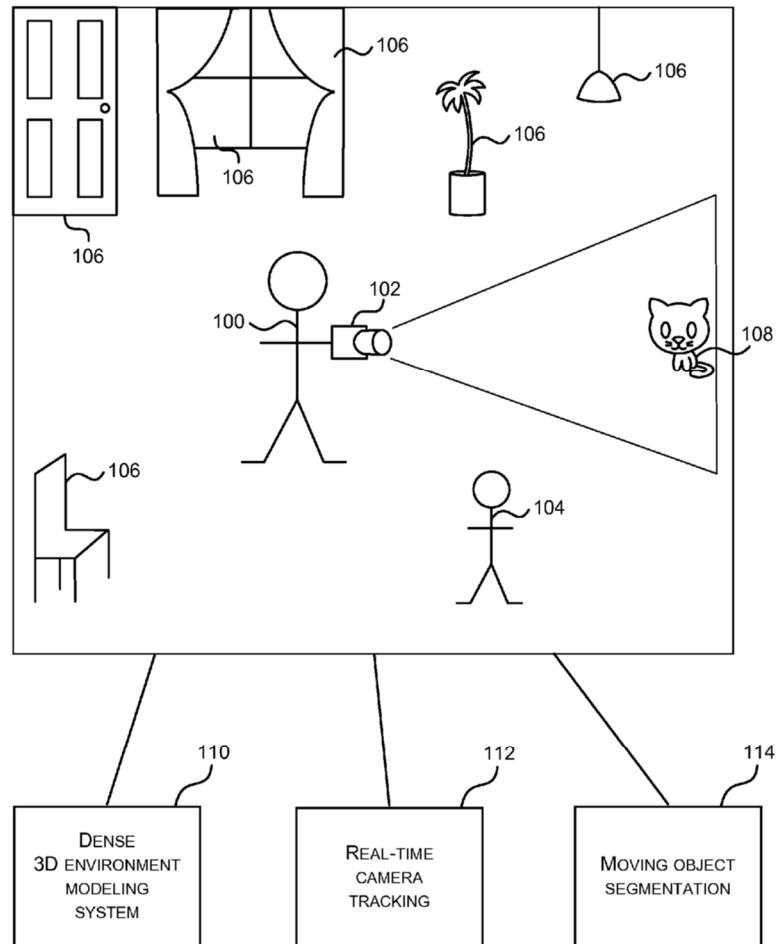


FIG. 1

229. It is my opinion that this understanding of Newcombe is supported by the reference's disclosure of the performance of compatibility tests to determine whether points in two depth images are corresponding points, which can be attributed to rigid objects, or are outliers, which can be attributed to moving objects. *See* Ex-1005 at 8:27-29. Specifically, it is my opinion that Newcombe discloses that at least the first compatibility test can be used to determine an excluded volume

where no rigid object should be present in any depth image of the location, and consequently, where any object detected is therefore a moving or “non-rigid” object at the location. *See, e.g., id.* at 9:26-29.

230. As described by Newcombe, the first compatibility test relies on defining a threshold distance for a measurement between points in two 3D depth images – i.e., a “source point” in a previous depth image and a “destination point” in a current depth image. *Id.* at 8:3-33. As Newcombe explains, “[i]f the destination point is too far away from the source point, then it is considered to be an outlier (i.e. non-corresponding)” *Id.* at 8:33-35. Newcombe further explains that “outliers” or “outlying points” refer to “points in the depth image where something has changed since the previous depth image” and “can be caused by dynamic or non-rigid objects.” *Id.* at 9:26-29.

231. This scenario is demonstrated in Figure 6, shown below, of Newcombe: “[T]he example of Fig. 6 shows the case where differences between the previous depth image 500 and the current depth image 600 (e.g. due to new or moving objects) result in the distance 604 between the source and destination points being greater than the predefined distance threshold, and hence the destination point is labeled as an outlier.” *Id.* at 8:42-48. Newcombe further explains the process as follows:

Moving object segmentation using depth images is described. In an

example, a moving object is segmented from the background of a depth image of a scene received from a mobile depth camera. A previous depth image of the scene is retrieved, and compared to the current depth image using an iterative closest point algorithm. The iterative closest point algorithm includes a determination of a set of points that correspond between the current depth image and the previous depth image. During the determination of the set of points, one or more outlying points are detected that do not correspond between the two depth images, and the image elements at these outlying points are labeled as belonging to the moving object. In examples, the iterative closest point algorithm is executed as part of an algorithm for tracking the mobile depth camera, and hence the segmentation does not add substantial additional computational complexity.

Id. at 1:49-65.

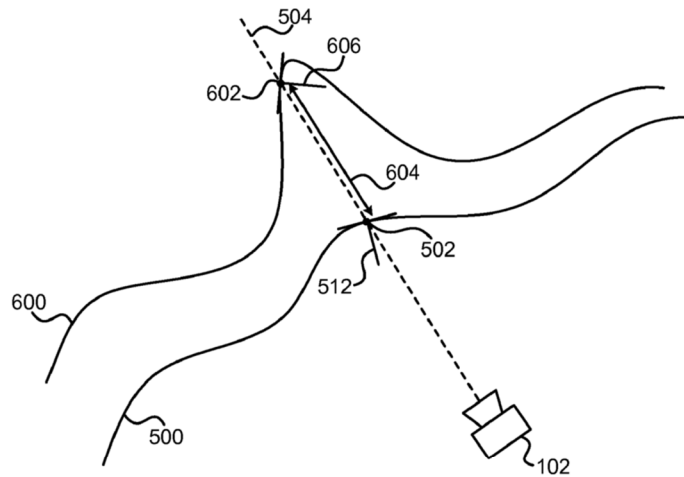


FIG. 6

232. Thus, it is my opinion that a POSITA would understand that by defining a distance threshold for the two three-dimensional depth images beyond which any point will be deemed an outlier, Newcombe teaches the determination of an excluded volume, and particularly, an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present,” as construed by the district court.

233. Further, it is my opinion that a POSITA would understand Newcombe’s disclosure of the space between the camera and the stationary objects in the

background of a location, as discussed above, as teaching an excluded volume necessarily applies equally to both a first depth image and a second depth image. Thus, it is my opinion that a POSITA would have understood that the empty space shown in the figure – i.e., the space in the scene that is not occupied by the fixed objects in the background would constitute an excluded volume for the second 3D representation, according to the court’s construction.

234. Again, as described above, this understanding of Newcombe by a POSITA is supported by Newcombe’s disclosure of a compatibility test to determine an excluded volume where no rigid object should be present in any depth image of the location, and consequently, where any object detected is therefore a moving or “non-rigid” object at the location. *See, e.g., id.* at 9:26-29. As described above, the compatibility test necessarily incorporates a first 3D representation and a second 3D representation (i.e., a previous 3D depth image and a current 3D depth image). Furthermore, it is my opinion that Newcombe teaches that the excluded volume determined by the compatibility test – i.e., the space beyond the distance threshold where any observed points will be deemed outliers corresponding to non-rigid objects – is also necessarily common to both depth images, because points in both images are required to define the threshold distance and compare the corresponding Euclidean distance between the source and destination points, as is further evidenced

visually by Newcombe's Figure 6 below. *See id.* at 8:3-48.

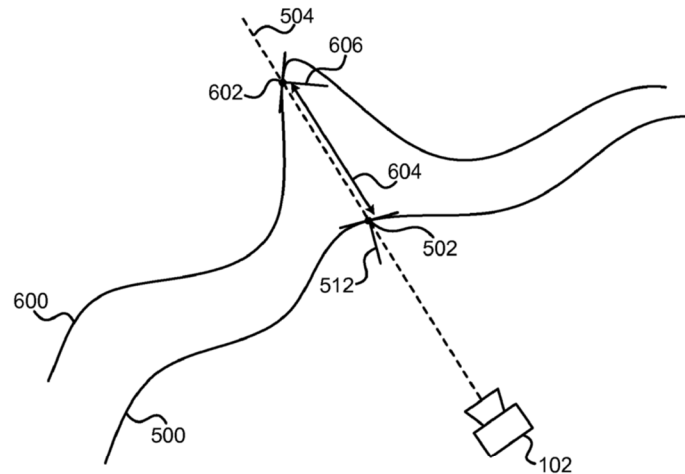


FIG. 6

235. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious, “determining for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation, and/or determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 25.

- e. **[Claim 25d]: “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, and/or if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model.”**

236. Newcombe in view of Hähnel teaches this claim element. As discussed above, a POSITA would understand Newcombe to teach a method for determining excluded volumes corresponding to previous and current 3D depth images. *See supra*, §§ X.A.1.d. As is also discussed above, Newcombe discloses that points corresponding to surfaces are identified as outliers if they are in an excluded volume (i.e., in a volume outside a predefined threshold distance), or as corresponding points if they are not. *See id.* Thus, Newcombe teaches that a surface in the first 3D representation that is located in space in the second excluded volume will be identified as an outlier. *See id.*

237. Newcombe further teaches that the outliers, which correspond to moving objects, can be “stored and provided as input to the object segmentation system.” Ex-1005 at 9:25-26. Newcombe discloses that the segmentation of the image can include “segment[ing] all the outlying points and label[ing] these as foreground objects in the depth image.” *Id.* at 9:30-32. The segmentation may be

completed using “segmentation algorithms.” *See, e.g., id.* at 10:1-4.

238. Newcombe discloses that moving objects that have been segmented “can be separately processed as desired.” *Id.* at 3:51-54. For example, Newcombe teaches that “the segmentation of moving objects enables these moving objects to be tracked separately and independently of the background. This can be achieved executing a further ICP algorithm on the portion of the depth image that was identified as a moving object by the segmentation system.” *Id.* at 10:31-36.

239. Newcombe further discloses that, in addition to “independently tracking a moving object in the depth image, an independent dense 3D model of the object can also be generated.” *Id.* at 10:57-59. Newcombe teaches that the creation of a “dense 3D model of a segmented object can be performed in a number of ways,” including a technique that “enables objects to be instantly segmented upon removal from the scene.” *Id.* at 11:3-23.

240. Additional examples of the processing of segmented objects disclosed by Newcombe include “input[ting] to an object recognition system, classification according the “behavior in the scene,” and “control[ling] further aspect of the operation of the system.” *Id.* at 11:28-51. With respect to the last example, Newcombe discloses that a segmented, moving object may be either “arranged to interact” with other objects or may be “deliberately ignored.” *Id.* at 11:51-53.

241. In my opinion, Newcombe's disclosure that segmented objects may be "separately processed as desired," as well as the disclosure of the various ways that segmented moving objects may be processed, including with respect to the "removal from the scene" and or being "deliberately ignored," would be understood by a POSITA to teach that these objects, which necessarily includes the surfaces of the objects, may be "disregarded" or otherwise removed or omitted from a virtual 3D model generated according to the methods disclosed by Newcombe.

242. However, to the extent that Newcombe may be deemed not to disclose the "disregarding" of a portion of a surface in a 3D representation, Hähnel supplies this teaching. Like Newcombe, Hähnel discloses 3D modeling processes based on simultaneous localization and mapping (SLAM). *See* Ex-1005 at 3:35-37; Ex-1007 at 1. Thus, in a context of mapping a location in a dynamic environment using 3D representation analogous to that disclosed by Newcombe, Hähnel teaches that moving objects (i.e., dynamic objects), can lead to serious errors in mapping a location. *See, e.g.*, Ex-1007 at 1. Hähnel further teaches that these moving objects can be reliably filtered out to obtain accurate models of a location. *Id.* at 1, 6; Fig. 11. It is my opinion that a POSITA would have understood Hähnel's description of filtering out of dynamic objects to correspond to "disregard[ing]" a "portion of a surface," as claimed.

243. As discussed above, Newcombe teaches that moving objects can be detected as outliers and segmented and integrated into separate 3D volumes. *Supra*, §§ X.A.1.d. Newcombe also discloses that segmented objects can be separately processed as desired. Ex-1005, at 3:51-54. Thus, it is my opinion that it would have been obvious to a POSITA to process the moving objects identified by segmentation in the depth image-based system of Newcombe with the filtering of Hähnel in order to reduce errors in the dense 3D environment model.

244. It is my opinion that a POSITA would further understand that Newcombe teaches that a surface in the second 3D representation that is located in space in the first excluded volume will be identified as an outlier. *See id.* Again, it is my opinion that a POSITA would understand Newcombe to teach that segmented, moving objects, identified as outliers in an excluded volume of a depth image, can be processed in a variety of ways. Such processing can include removal from a scene or deliberately ignoring the segmented objects. Newcombe further teaches that the identification of moving objects and subsequent processing can be performed repeatedly with respect to any pair of depth images. *See* Ex-1005 at 8:3-9:23. It is my opinion that a POSITA would therefore understand Newcombe's teachings with respect to the determination of excluded volumes and subsequent disregarding of points within those excluded volumes to be equally applicable to both first and

second 3D representations. In other words, it is my opinion that a POSITA would understand Newcombe to teach that, if a portion of a surface in a second 3D representation is located in space in the corresponding first excluded volume, the portion of the surface in the second 3D representation should be disregarded, just as a POSITA would understand Newcombe to teach, as described above, that if a portion of a surface in a first 3D representation is located in space in the corresponding second excluded volume, the portion of the surface in the first 3D representation should be disregarded.

245. As described above, it is also my opinion that a POSITA would have found it obvious to combine Hähnel's filtering process with Newcombe's process of identifying and segmenting moving objects in a scene. As further described above, Newcombe discloses that its segmenting process can be applied repeatedly to multiple depth images to obtain multiple separate 3D volumes with respect to the multiple depth images. Ex-1005 at 8:3-9:23. In my opinion, a POSITA therefore would have been motivated and found it obvious to apply Hähnel's filtering process repeatedly to multiple segmented objects in the separate 3D volumes disclosed by Newcombe's iterative process, in order to reduce any errors in location mapping caused by the segmented moving objects. Accordingly, Newcombe in view of Hähnel teaches this element.

246. Accordingly, it is my opinion that Newcombe in view of Hähnel teaches, and therefore renders obvious “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, and/or if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model,” as recited by this element of claim 25.

B. Claims 2-5, 18, and 19 Are Rendered Obvious By Newcombe in View of Hähnel, Further in View of McQueston

247. It is my opinion that Newcombe in view of Hähnel, further in view of McQueston teaches, and therefore renders obvious, every element of claims 2-5, 18, and 19 of the '551 patent.

248. As discussed above in Section X.A, it is my opinion that Newcombe in view of Hähnel teaches all the elements of claims 1, 6-12, 15-17, and 20-25. Newcombe in view of Hähnel identifies the problem of moving bodies obstructing a scanning volume and interfering with 3D modeling of a dense background environment.

249. McQueston describes a similar, if not identical, problem albeit in the context of intraoral imaging, where dental instruments or other objects may interfere

with the 3D imaging of teeth by obstructing a scanning volume between a scanning device and the tooth surfaces. Ex-1008 at 1:9-23. As with Newcombe in view of Hähnel, McQueston teaches using laser range finding for distance measurements underlying the three-dimensional tooth modeling. Ex-1008 at 4:65-5:26; 5:47-63. Likewise, as with the depth imaging described by Newcombe, McQueston teaches that imaging techniques may be incorporated into the three-dimensional modeling of the teeth. Ex-1008 at 3:31-61.

250. It is my opinion that a POSITA would be motivated to apply the imaging and modeling techniques and components of Newcombe in view of Hähnel to the intraoral environment of McQueston in order to achieve the stated goal of McQueston—minimizing the interference of dental objects with the three-dimensional model of the teeth being imaged or measured. It is my opinion that a POSITA would understand that the depth imaging and object segmentation of Newcombe (Ex-1005), could be integrated into a dental scanning apparatus, like that described by McQueston in order to address the same problem of removing any influence or interference of moving bodies from three-dimensional modeling of stationary objects. Thus, it is my opinion that Newcombe in view of Hähnel, further in view of McQueston renders claims 2-5, 18, and 19 obvious, as discussed in more detail below.

1. Claim 2 Is Rendered Obvious By Newcombe in View of Hähnel, Further in View of McQueston

251. As discussed above, McQueston describes obtaining a three-dimensional model of teeth located in an intra-oral environment. Ex-1008 at 1:9-23. For the reasons stated above, it is my opinion that it would have been obvious to a POSITA to apply the optical measurement techniques of Newcombe in view of Hähnel to the intra-oral environment of McQueston in order to measure tooth surfaces without the interference of moving bodies and thereby obtain a more accurate three-dimensional model of the tooth surfaces. Therefore, it is my opinion that Newcombe in view of Hähnel, further in view of McQueston renders obvious, and therefore invalid, that the “rigid object is a patient’s set of teeth, and the location is the mouth of the patient,” as claimed.

2. Claims 3-5, 18, and 19 Are Rendered Obvious By Newcombe in View of Hähnel, Further in View of McQueston

252. As discussed above, McQueston teaches that the movable objects that may interfere with the three-dimensional modeling of the teeth may be a number of example dental objects, including soft tissue, intraoral structures, and dental instruments. Ex-1008 at 4:4-46. McQueston also describes that a finger may interfere with the tooth surface measurement. *See, e.g., id.* at 1:9-23. Therefore, it is my opinion that Newcombe in view of Hähnel, further in view of McQueston

renders obvious, and therefore invalid, the claimed “movable object” being “soft tissue,” “dentist’s instrument,” “finger,” “tongue,” or “dental suction device,” as recited in claims 3-5, 18, and 19.

C. Claims 1, 6-12, 15-17, and 20-25 Are Rendered Obvious by Herbst

253. It is my opinion that Herbst teaches, and therefore renders obvious, every element of claims 1, 6-12, 15-17, and 20-25 of the ’551 patent.

1. Claim 1 Is Rendered Obvious by Herbst

254. As explained below, it is my opinion that Herbst teaches, and therefore renders obvious, every element of claim 1.

a. [Claim 1 pre]: “A method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the method comprises:”

255. As discussed above, in section VIII.A, for the purposes of this declaration, I have treated the preamble of claim 1 as limiting. It is my opinion that Herbst teaches this claim element.

256. Herbst teaches “a method to determine movable parts of a scene based on differencing with another scene.” Ex-1006 at 7. Herbst further teaches that “[t]he inputs to our system are two or more RGB-D videos representing separate visits to the same location. From each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*. Objects can move between visits but are assumed static

during each visit.” *Id.* at 2. In my opinion, Herbst’s disclosed method of “determin[ing] movable parts of a scene” corresponds to the claimed method for “detecting a movable object in a location.” Similarly, it is my opinion that the “static” objects disclosed by Herbst correspond to the “rigid object” recited in this claim element. I further believe that Herbst discloses scanning these static objects by means of a 3D scanner for generating a virtual 3D model of the rigid object, as recited by this element. Specifically, Herbst teaches the use of RGB-D cameras and RGB-D Mapping, which Herbst further discloses to be “a SLAM technique for generating dense 3-D maps.” *Id.* at 2. RGB-D cameras capture color in the form of light in red, green, and blue wavelengths (RGB) and also capture depth information (D) using a 3D depth sensor. The color and depth information captured by the camera can be used to create 3D models. As Herbst explains, “We use the PrimeSensor, an *RGB-D* camera that provides dense registered color and depth measurements, to detect objects by means of their movement in 3D between two maps of the same location.” *Id.* at 1. RGB-D cameras and their abilities to scan objects and locations to thereby capture 3D models were well-known to POSITAs at the time of filing of the ’551 patent. Furthermore, SLAM stands for “Simultaneous Localization And Mapping,” which is a technique for generating and updating virtual 3D models, and more specifically 3D map models. In both my

opinion and as evidenced by Herbst's discussion of relevant prior art references utilizing the SLAM technique, a POSITA at the time would have known and understood SLAM as a technique for generating and updating 3D models of objects, including both rigid and moveable objects, in a scene or location. *See id.* at 1.

257. Accordingly, it is my opinion that Herbst teaches "detecting a movable object in a location when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object," as recited by the preamble of claim 1.

b. [Claim 1a]: "providing a first 3D representation of at least part of a surface by scanning at least part of the location;"

258. Herbst teaches this claim element. As discussed above, Herbst describes scanning objects in a location using RGB-D cameras and SLAM to create and update 3D maps. Ex-1006 at 1. Herbst teaches that "[t]he inputs to our system are two or more RGB-D videos representing separate visits to the same location." He also teaches that "[f]rom each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*." He further teaches that "[i]n each experiment, the RGB-D camera was carried by hand to generate partial coverage of the scene." In my opinion, a POSITA at the time the '551 patent was filed would have understood that 3-D reconstruction from the first visit by Herbst to the location constitutes a first

3D representation, as recited by this element and that “partial coverage of the scene” constitutes “part of a surface”, as recited by this element. Herbst teaches scanning and detecting surfaces within a location or scene. Specifically, Herbst states that its “algorithm aligns the frames of each RGB-D video using SLAM, then reconstructs each scene as a dense set of surface elements.” *Id.* at 1. Herbst further states that the RGB-D cameras “provide a dense matrix of pixels, each measuring the color and distance of the closest surface in the pixel direction.” *Id.* at 2. Figure 6(a) of Herbst, shown below, demonstrates such a first 3D reconstruction or representation of surfaces, including surfaces corresponding to a table and objects on the table, obtained from scanning a location.



Fig. 6(a)

259. Accordingly, it is my opinion that Herbst teaches “providing a first 3D representation of at least part of a surface of the location,” as recited in this element of claim 1.

c. [Claim 1b]: “providing a second 3D representation of at least part of the surface by scanning at least part of the location;”

260. Herbst teaches this claim element. As discussed in the preceding section above, Herbst discloses “providing a first 3D representation of at least part of a surface of the location.” Herbst further discloses that this process of providing 3D representations is iterative and performed two or more times with respect to a particular scene or location. Specifically, Herbst states that “[t]he inputs to [its] system are two or more RGB-D videos representing separate visits to the same location.” Ex-1006 at 2. Herbst further discloses that its method relies on “information in the difference between two maps made at different times.” *Id.* at 1. As discussed in the preceding section, he also teaches that “[f]rom each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*.” He further teaches that “[i]n each experiment, the RGB-D camera was carried by hand to generate partial coverage of the scene.” It is my opinion that the 3-D reconstruction from the second visit among multiple visits taught by Herbst for partial coverage of

the location corresponds to “a second 3D representation of at least part of a surface of the location,” as recited by this claim element. For example, Figure 6(d) of Herbst, reproduced below, shows a second 3D reconstruction or representation of the same tabletop scene shown in Figure 6(a) above, but with different objects in different locations on the table within the scene.

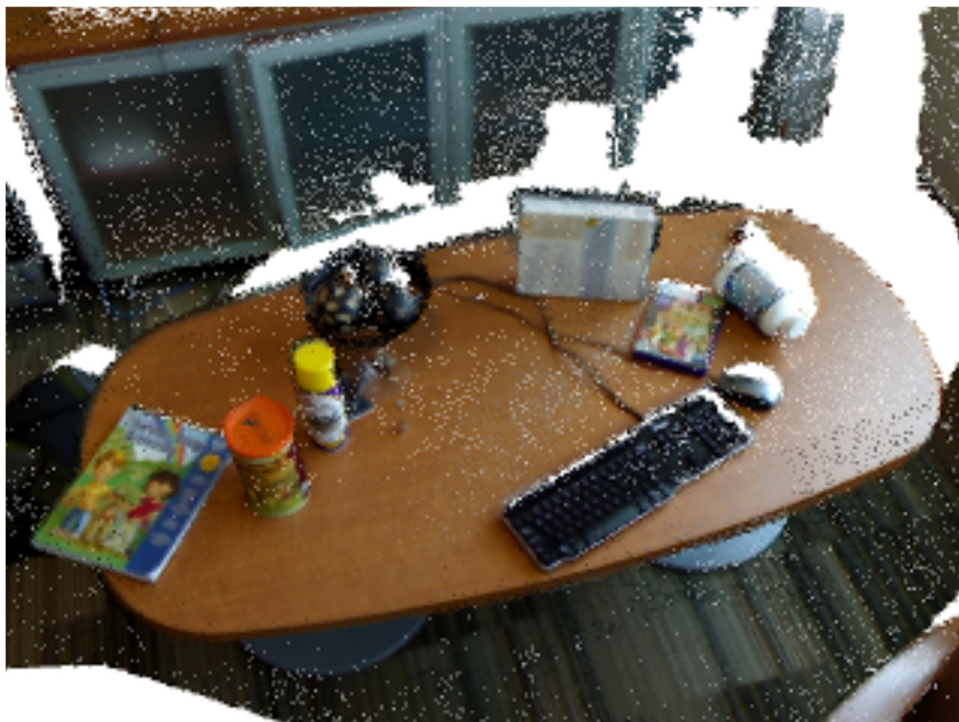


Fig. 6(d)

261. Accordingly, it is my opinion that Herbst teaches “providing a second 3D representation of at least part of the surface by scanning at least part of the location,” as recited by this element of claim 1.

- d. [Claim 1c]: “determining for the first 3D representation a first excluded volume in space where**

no surface can be present in both the first 3D representation and the second 3D representation;”

262. Herbst teaches this claim element. As discussed above, I understand that the district court has construed “excluded volume,” as used in this claim element, to mean an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present.” Specifically, it is my opinion that Herbst teaches the determination of an excluded volume through the use of the probabilistic model disclosed by the reference.

263. Herbst teaches a probabilistic model for obtaining pixel measurements corresponding to surfaces observed in a scene by an RGB-D camera. *See* Ex-1006 at 2-3. The model incorporates surface information related to depth, color, and orientation of each pixel. *Id.* at 3. Herbst’s model considers two cases: one in which no information about a scene or surfaces is known and one in which such information is known. *Id.* at 3; Fig. 3. In the first case, Herbst’s model is able to calculate the probability that a surface within a scene is at a particular depth measurement (along with additional information related to color and orientation of the surface). *Id.* at 3; Fig. 3. Similarly, in the second case in which scene and surface information are already known, Herbst’s model is able to calculate a depth measurement of a surface relative to an expected known depth of the surface (again, along with corresponding

color and orientation information). *Id.* at 3-4; Fig. 3. By aligning and comparing two different 3D representations of a location and applying these probabilistic models to the surfaces observed in the scene in each 3D representation, Herbst's method can determine a probability that any of the surfaces observed in either 3D representation moved relative to its location in or absence from the other 3D representation. *Id.* at 4-5. Accordingly, with respect to each of a first 3D representation and a second 3D representation, Herbst's model is able to predict the depth, and therefore the location, of all surfaces within the volume of the scene scanned by the RGB-D camera.

264. Herbst further teaches that “[f]ollowing global registration, we convert points to surfels [14] to reduce map size and smooth evidence from multiple frames. Each 3-D point from each frame is assigned to a surfel, either an existing one or a newly created one if there is no existing surfel close to the point. Fig. 1 shows two reconstructions of a location.” *Id.* at 2. It is my opinion that a POSITA at the time the '551 patent was filed would understand that Herbst's disclosure of the reduction of the size of the map (i.e. 3D representation) of the scanning visit, by assigning surfels to 3D points detected in frames of that scanning visit, is teaching that the empty space determined accordingly between the scanner and the surface, where no 3D point on any surface is detected, is excluded from the 3D surfel representation of

the corresponding scan visit. Therefore, a POSITA would further understand that the empty space for each scanning visit would constitute an excluded volume for the 3D representation of that scanning visit.

265. It is my opinion that a POSITA would therefore understand Herbst to teach that its probabilistic model is used to determine not only the location of surfaces but also the empty space in a location. The scanned 3D space where no 3D point is present does not contribute to the 3D surface representation and would constitute an excluded volume in the 3D space being scanned, as recited by the claims, and more particularly would constitute an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present, according to the court's claim construction.

266. For example, it is my opinion that a POSITA would have understood that the empty space shown in Figure 6(a) below – i.e., the space in the scene that is not occupied by the table, the objects on the table, or any other surface – would constitute an excluded volume for the first 3D representation, according to the court's construction.



Fig. 6(a)

267. Accordingly, it is my opinion that Herbst teaches “determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 1.

- e. **[Claim 1d]: “determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

268. Herbst teaches this claim element. My analysis and discussion of Herbst’s probabilistic model in the preceding section as teaching an excluded volume necessarily applies equally to both the probabilistic model’s application to a

first 3D representation and to a second 3D representation. Thus, with respect to Figure 6(d) below, which as discussed above shows a second 3D representation according to Herbst, it is my opinion that a POSITA would have understood that the empty space shown in the figure – i.e., the space in the scene that is not occupied by the table, the objects on the table, or any other surface – would constitute an excluded volume for the second 3D representation, according to the court’s construction.

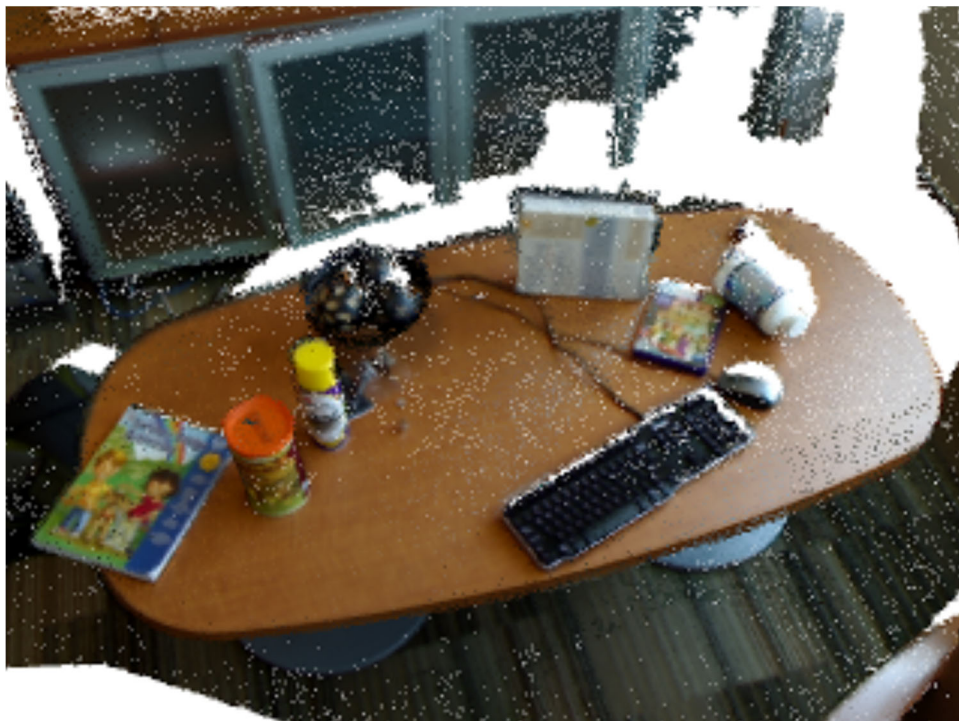


Fig. 6(d)

269. Accordingly, it is my opinion that Herbst teaches “determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as

recited in claim 1.

- f. **[Claim 1e]: “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, and/or”**

270. Herbst teaches this claim element. As discussed above, Herbst teaches a method for comparing two or more 3D representations of a scene to determine if objects within the scene have moved between the two representations. Herbst further teaches that once all of the moveable objects have been detected, the two 3D representations can be merged into a single virtual 3D model from which all of the moveable objects are removed. *See* Ex-1006 at 5. Herbst teaches a probabilistic “differencing” approach relying, in particular, on depth differences for each surfel in a 3D representation to determine if a surfel belongs to a moving object or to the static background. *Id.* at 4-5. Herbst further discloses that a Bayesian framework is used to determine the probability of a surfel belonging to a moving object or to the static background, by first aligning the first representation with the second representation to find the corresponding surfels, so that an “expected” value of the surfel depth in the second representation can be computed based on the first representation. *Id.* The probability of a surfel in the second representation being a moving one or not is then determined as a function of the difference between its

measured depth and its expected depth value (determined from the first representation). *Id.* Herbst also teaches a regularization approach is used so that the probabilistic model is not overly sensitive to small differences due to measurement or alignment errors, i.e., only large probabilities contribute to the identification of moving surfels. *Id.*

271. In my opinion, a POSITA at the time the '551 patent was filed would have understood that the difference between the measured surfel depth in the second representation and its expected value computed from the first representation, which Herbst uses to determine the conditional probabilities could result only in three possibilities: (1) measured depth is smaller than the expected depth (difference is negative, leading to a high probability of motion for Herbst), (2) measured depth is larger than the expected depth (difference is positive, leading again to a high probability of motion for Herbst), (3) measured depth is approximately the same as the expected value (no significant difference, so the difference is within measurement and alignment error, i.e. low probability of motion). A POSITA would have understood that the first case occurs when a surfel from the first 3D representation is located in the second excluded volume. Herbst further discloses that such surfel could be identified with high probability and therefore removed from the merged 3D model in SLAM. Since a collection of surfels in the first 3D

representation would define a portion of the surface in the first representation, it is my opinion that this claim element is made obvious by the teachings of Herbst.

272. For example, Herbst shows a first 3D representation of a tabletop scene in Figure 6(a) and a second 3D representation of the same tabletop scene in Figure 6(d). Both figures are reproduced below.



Fig. 6(a) (annotated)

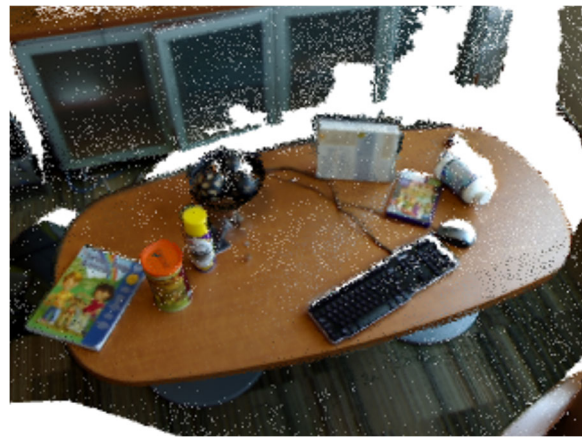


Fig. 6(d)

As discussed above, a POSITA would understand Herbst to teach that all empty space in these respective 3D representations that is not occupied by a surface would collectively constitute an excluded volume for each 3D representation. Thus, when comparing the two representations according to Herbst's probabilistic model, those surfaces in the first 3D representation that are in the excluded volume of the second 3D representation (i.e., the second excluded volume) are determined to have moved. Those two objects are identified in the left figure above with **red** arrows.

273. Furthermore, in Figure 7(b) of Herbst below, which illustrates a virtual

3D model of the scene created by merging the two 3D representations of Figures 6(a) and 6(d), the moved objects in the first 3D representation – i.e., those objects that are in the second excluded volume – have been removed or, in the language of the claim, “disregarded.” In other words, the two moved objects are not included in the virtual 3D model, but instead shown separately over the virtual 3D model of the scene.



Fig. 7(b)

274. Accordingly, it is my opinion that Herbst teaches “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model,” as recited by this element of claim 1.

- g. **[Claim 1f]: “if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the**

second 3D representation is disregarded in the generation of the virtual 3D model.”

275. Herbst teaches this claim element. My analysis of Herbst’s teachings in the preceding section apply equally to this claim element. *See supra*, § X.C.1.f. As described in the preceding section, a POSITA at the time the ’551 patent was filed would have understood that the difference between the measured surfel depth in the second representation and its expected value computed from the first representation, which Herbst uses to determine the conditional probabilities could result only in three possibilities: (1) measured depth is smaller than the expected depth (difference is negative, leading to a high probability for Herbst), (2) measured depth is larger than the expected depth (difference is positive, leading again to a high probability for Herbst), (3) measured depth is approximately the same as the expected value (no significant difference, so the difference is within measurement and alignment error, i.e. low probability of motion). *Id.* A POSITA would have understood that the second case occurs when a surfel from the second 3D representation is located in the first excluded volume. As described in the preceding section, Herbst further discloses that such surfels could be identified with high probability and therefore removed from the merged 3D model in SLAM. Since a collections of surfels in the second 3D representation would define a portion of the surface in the second representation, it is my opinion that this claim element is made

obvious by the teachings of Herbst.

276. Referring again to Figures 6(a) and 6(d) of Herbst, reproduced below, as stated above a POSITA would understand Herbst to teach that all empty space in these respective 3D representations that is not occupied by a surface would collectively constitute an excluded volume for each 3D representation.

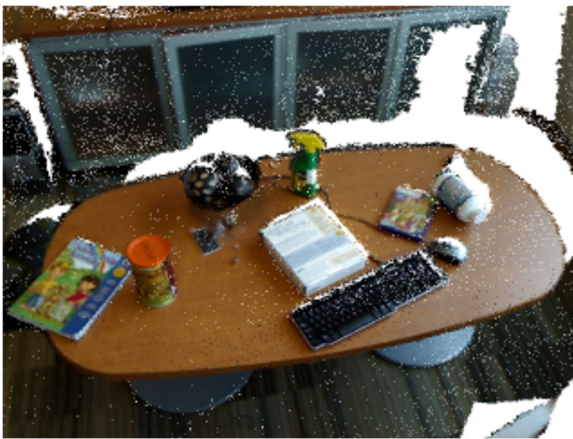


Fig. 6(a)

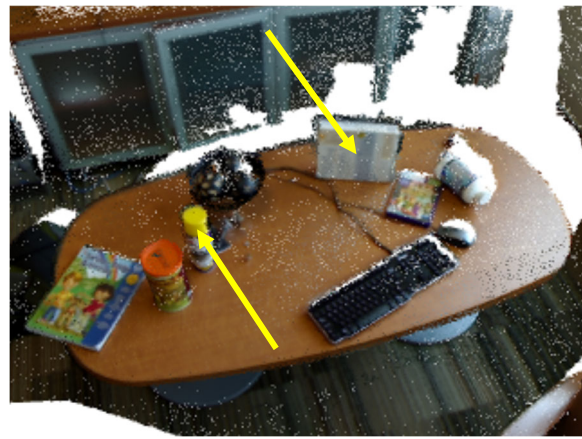


Fig. 6(d) (annotated)

Thus, when comparing the two representations according to Herbst's probabilistic model, those surfaces in the second 3D representation that are in the excluded volume of the first 3D representation (i.e., the first excluded volume) are determined to have moved. Those two objects are identified in the right figure above with **yellow** arrows.

277. Thus, again when considering the merged virtual 3D model of Figure 7(b) of Herbst, shown below, the moved objects in the second 3D representation – i.e., those objects that are in the first excluded volume – have been

removed or, in the language of the claim, “disregarded.” In other words, the two moved objects are not included in the virtual 3D model, but instead shown separately over the virtual 3D model of the scene.



Fig. 7(b)

278. Accordingly, it is my opinion that Herbst teaches “if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model,” as recited by this element of claim 1.

2. Claim 6 Is Rendered Obvious by Herbst

279. As explained below, it is my opinion that Herbst renders every element of claim 6 obvious.

- a. [Claim 6]: “The method according to claim 1, wherein at least part of the surface captured in the first representation and at least part of the surface captured in the second representation are overlapping**

the same surface part on the rigid object.”

280. Herbst teaches this claim element. As discussed above, Herbst teaches providing first and second 3D representations of at least part of a surface by scanning at least part of the location. *Supra*, §§ X.C.1.b-c. Figures 1(a) and 6(a) of Herbst, reproduced below, show examples of the first 3D representations disclosed by Herbst, while Figures 1(b) and 6(d), also reproduced below, show examples of the second 3D representations disclosed by Herbst.



Fig. 1(a)



Fig. 1(b)



Fig. 6(a)

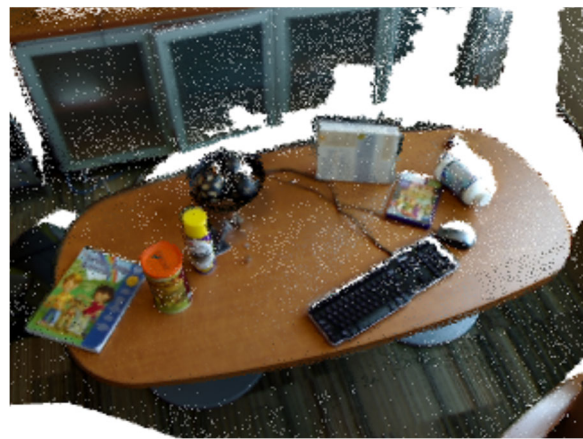


Fig. 6(d)

281. As is seen in the images above, at least a part of the surface captured in the first 3D representation of Herbst in Fig. 1(a) and at least part of the surface captured in the second 3D representation of Herbst in Fig. 1(b) are overlapping the same surface part on the rigid object (e.g., the table shown in Figs. 1(a) and 1(b)). Similarly, at least a part of the surface captured in the first 3D representation of Herbst in Fig. 6(a) and at least part of the surface captured in the second 3D representation of Herbst in Fig. 6(d) are overlapping the same surface part on the rigid object (e.g., the table shown in Figs. 6(a) and 6(d)).

282. Accordingly, it is my opinion that Herbst teaches “wherein at least part of the surface captured in the first representation and at least part of the surface captured in the second representation are overlapping the same surface part on the rigid object,” as recited in claim 6.

3. Claim 7 Is Rendered Obvious by Herbst

283. As explained below, it is my opinion that Herbst renders every element of claim 7 obvious.

- a. **[Claim 7]: “The method according to claim 1, wherein the first representation of at least part of the surface is defined as the first representation of at least a first part of the surface, and the second representation of at least part of the surface is defined as the second representation of at least a second part of the surface.”**

284. Herbst teaches this claim element. As discussed above, Herbst teaches providing first and second 3D representations of at least part of a surface by scanning at least part of the location. *Supra*, §§ X.C.1.b-c. Figures 1(a) and 6(a) of Herbst, reproduced below, show examples of the first 3D representations disclosed by Herbst, while Figures 1(b) and 6(d), also reproduced below, show examples of the second 3D representations disclosed by Herbst.

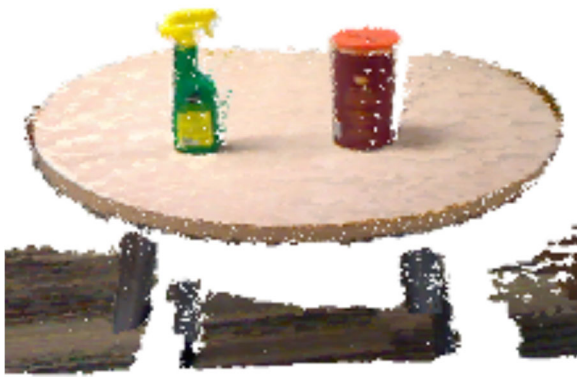


Fig. 1(a)



Fig. 1(b)



Fig. 6(a)

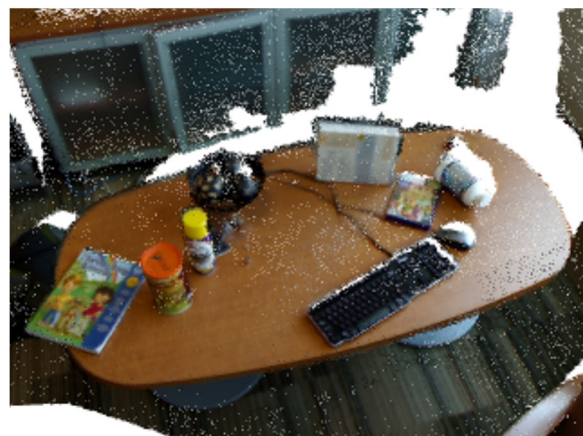


Fig. 6(d)

285. As is seen in the images above, the first representation of at least part

of the surface is defined as the first representation of at least a first part of the surface, as shown in the first 3D representation of Herbst in Fig. 1(a), and the second representation of at least part of the surface is defined as the second representation of at least a second part of the surface, as shown in the second 3D representation of Herbst in Fig. 1(b). Similarly, the first representation of at least part of the surface is defined as the first representation of at least a first part of the surface, as shown in the first 3D representation of Herbst in Fig. 6(a), and the second representation of at least part of the surface is defined as the second representation of at least a second part of the surface, as shown in the second 3D representation of Herbst in Fig. 6(d).

286. Accordingly, it is my opinion that Herbst teaches “wherein the first representation of at least part of the surface is defined as the first representation of at least a first part of the surface, and the second representation of at least part of the surface is defined as the second representation of at least a second part of the surface,” as recited in claim 7.

4. Claim 8 Is Rendered Obvious by Herbst

287. As explained below, it is my opinion that Herbst renders every element of claim 8 obvious.

- a. **[Claim 8] “The method according to claim 1, wherein the method comprises determining a first scan volume in space related to the first representation of at least part of the surface, and determining a second scan**

**volume in space related to the second representation
of at least part of the surface.”**

288. Herbst teaches this claim element. Herbst discloses that “[w]e use the PrimeSensor, an RGB-D camera that provides dense registered color and depth measurements, to detect objects by means of their movement in 3-D between two maps of the same location.” Ex-1006 at 1. In my opinion the disclosure of using an RGB-D sensor such as the PrimeSensor would make Claim 8 obvious, since a POSITA at the time the ’551 patent was filed would have understood that such sensors comprise of an RGB camera and an infrared laser to project a pattern in the scene (typically random spots). A POSITA would have known that an RGB camera would have a field of view, defining the extent of a volume in space in front of the aperture of the camera that is pictured by the camera, and a depth of field defining a range of depth within the field of view that is in focus. Also, a POSITA would have known that laser pattern projection would have a depth range depending on the laser strength and the environment ambient infrared light. Therefore, it is my opinion that a POSITA would have understood that the combination of field of view, depth of field, and the laser range associated with an RGB-D sensor would define a volume of space in front of the sensor where the scanner could effectively capture 3D point clouds. Therefore, as the user moves the sensor’s position and orientation, a new scan volume is defined in space, where the RGB-D scanner can effectively produce

3D point clouds.

289. On the other hand, Herbst teaches providing first and second 3D representations of at least part of a surface by scanning at least part of the location. *See supra*, §§ X.C.1.b-c. Therefore, Herbst determines a scan volume related to each 3D representation for part of the surface since as discussed above, it determines the reduction of the scan volume into surfels and the associated excluded volume, i.e., the empty space between part of the surface being scanned and the sensor. Figures 1(a) and 6(a) of Herbst, reproduced below, show examples of the first 3D representations disclosed by Herbst, while Figures 1(b) and 6(d), also reproduced below, show examples of the second 3D representations disclosed by Herbst.



Fig. 1(a)



Fig. 1(b)



Fig. 6(a)

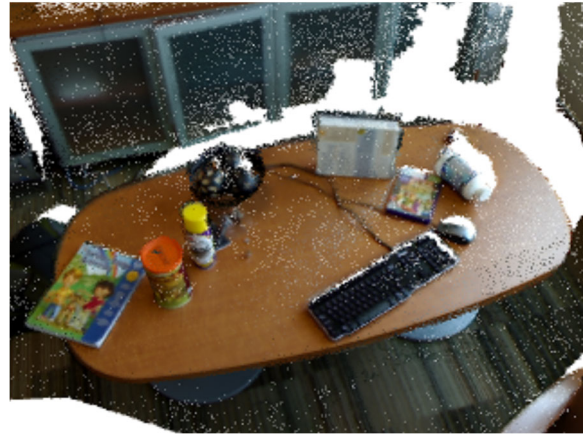


Fig. 6(d)

290. As seen in the images above, a scan volume in space related to each of the first and second 3D representations of at least part of the surface is taught by Herbst.

291. Accordingly, it is my opinion that Herbst teaches “wherein the method comprises determining a first scan volume in space related to the first representation of at least part of the surface, and determining a second scan volume in space related to the second representation of at least part of the surface,” as recited in claim 8.

5. Claim 9 Is Rendered Obvious by Herbst

292. As explained below, it is my opinion that Herbst renders every element of claim 9 obvious.

- a. **[Claim 9]: “The method according to claim 1, wherein the scan volume is defined by focusing optics in the 3D scanner and the distance to the surface which is captured.”**

293. Herbst teaches this claim element. As discussed above, Herbst discloses a scan volume with respect to each of the first and second 3D representations. *Supra*, § X.C.4.a. Furthermore, as discussed above, Herbst discloses the use of an RGB-D camera for 3D scanning. In my opinion, a POSITA at the time the '551 patent was filed would know that RGB-D cameras use electro-optical sensors with associated focusing optics for capturing depth information. Therefore, it is my opinion that a POSITA would know that the scan volume created by an RGB-D camera when scanning a surface to create 3D representations would depend on the field of view and the depth of field of the electro-optical sensor, both of which depend on the focusing optics and the distance to the surface captured by the depth camera.

294. Accordingly, it is my opinion that Herbst teaches “wherein the scan volume is defined by focusing optics in the 3D scanner and the distance to the surface which is captured,” as recited in claim 9.

6. Claim 10 Is Rendered Obvious by Herbst

295. As explained below, it is my opinion that Herbst renders every element of claim 10 obvious.

- a. **[Claim 10] “The method according to claim 1, wherein the first scan volume related to the first representation of at least part of the surface is the volume in space between the focusing optics of the 3D**

scanner and the surface captured in the first representation; and the second scan volume related to the second representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the second representation.”

296. Herbst teaches this claim element. As discussed above with respect to claim 8, Herbst describes determining a first and second scan volume related to the first and second representations of at least a first and second part of the surface. *Supra*, § X.C.4.a. As discussed with respect to claim 9, it is my opinion that the RGB-D camera taught by Herbst would necessarily have focusing optics that a POSITA would understand to define a scan volume between the focusing optics and the captured surface. *Supra*, § X.C.5.a.

297. Accordingly, it is my opinion that Herbst teaches “wherein the first scan volume related to the first representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the first representation; and the second scan volume related to the second representation of at least part of the surface is the volume in space between the focusing optics of the 3D scanner and the surface captured in the second representation,” as recited by claim 10.

7. Claim 11 Is Rendered Obvious by Herbst

298. As explained below, it is my opinion that Herbst renders every element

of claim 11 obvious.

- a. **[Claim 11]: “The method according to claim 1, wherein the first excluded volume and second excluded volume in space where no surface can be present corresponds to the first scan volume and the second scan volume, respectively.”**

299. Herbst teaches this claim element. As discussed above with respect to the excluded volume and scan volume elements of claims 8-10, a POSITA would understand the scanned depths that do not correspond to a surface to be excluded volumes where the surface of the stationary body cannot be present. *Supra*, §§ X.C.4.a-6.a. Further, as is shown in Figures 1(a)-1(b) and 6(a), 6(d) below, these scanned depths correspond to the scan volumes, including, for example, those volumes between the RGB-D camera and the surface.



Fig. 1(a)



Fig. 1(b)



Fig. 6(a)

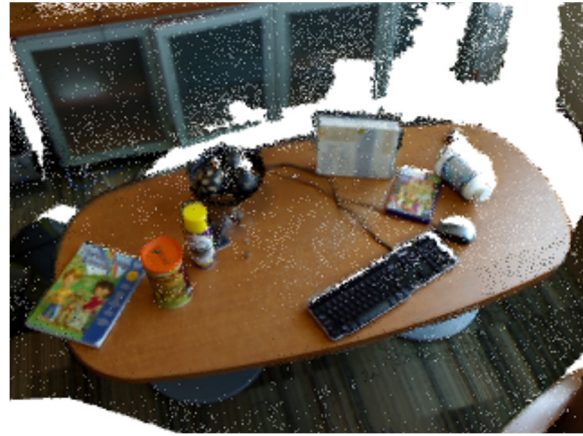


Fig. 6(d)

300. Accordingly, it is my opinion that Herbst teaches “wherein the first excluded volume and second excluded volume in space where no surface can be present corresponds to the first scan volume and the second scan volume, respectively,” as recited by claim 11.

8. Claim 12 Is Rendered Obvious by Herbst

301. As explained below, it is my opinion that Herbst renders every element of claim 12 obvious.

- a. [Claim 12]: “The method according to claim 1, wherein the volume of the 3D scanner itself is defined as an excluded volume.”**

302. To the extent that Herbst does not explicitly describe that the 3D scanner is defined as an excluded volume, it is my opinion that this claim element would be obvious to a POSITA in light of the teachings of Herbst. A POSITA would recognize that the RGB-D camera of Herbst would necessarily be a part of “an

excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present,” per the district court’s construction, because it is impossible for the RGB-D camera to be included in any scan volume. Therefore, it is my opinion that Herbst renders claim 12 obvious.

9. Claim 15 Is Rendered Obvious by Herbst

303. As explained below, it is my opinion that Herbst renders every element of claim 15 obvious.

- a. [Claim 15]: “The method according to claim 1, wherein the first representation of at least part of a surface is a first subscan of at least part of the location, and the second representation of at least part of the surface is a second subscan of at least part of the location.”**

304. It is my opinion that a POSITA would understand that Herbst describes the elements of this claim. Although Herbst does not explicitly use the term “subscan” with respect to the first and second 3D representations it teaches, as shown, for example, in Figures 1(a)-1(b) and 6(a), 6(d) below, a POSITA would understand that these multiple representations are subscans of at least a part of the location depicted in the scene and that together these subscans will be used to generate a virtual 3D model. *See, e.g.*, Ex-1001 at 3:44-49 (“Typically when scanning an object, such as a set of teeth, more representations or subscans may be provided . . . during

a complete scanning process.”). Herbst further discloses that “[i]n each experiment, the RGB-D camera was carried by hand to generate partial coverage of the scene.” Ex-1006 at 5. In my opinion, a POSITA would have understood each section of “partial coverage of the scene” generated by the camera to constitute multiple, separate subscans.

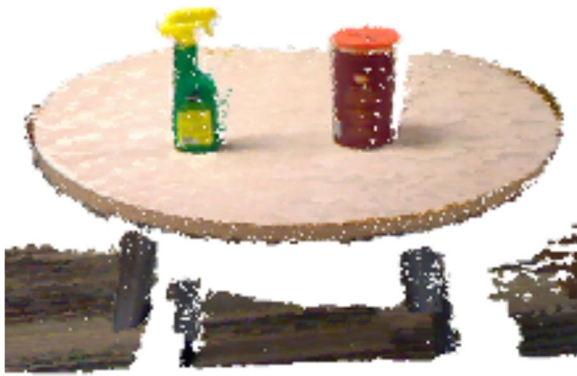


Fig. 1(a)



Fig. 1(b)



Fig. 6(a)

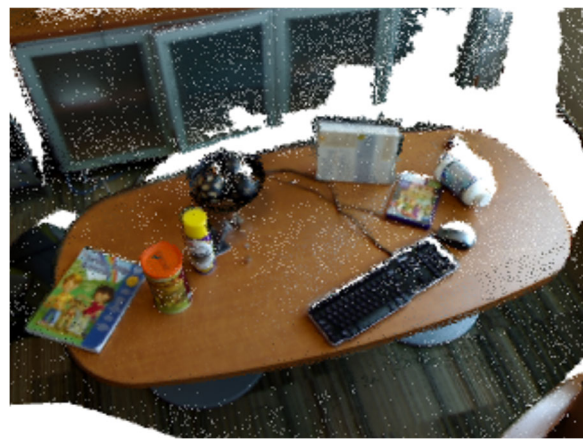


Fig. 6(d)

305. Accordingly, it is my opinion that Herbst renders obvious “wherein the first representation of at least part of a surface is a first subscan of at least part of the

location, and the second representation of at least part of the surface is a second subscan of at least part of the location,” as recited by claim 15.

10. Claim 16 Is Rendered Obvious by Herbst

306. As explained below, it is my opinion that Herbst renders every element of claim 16 obvious.

- a. **[Claim 16]: “The method according to claim 1, wherein the first representation of at least part of a surface is a provisional virtual 3D model comprising the subscans of the location acquired already, and the second representation of at least part of the surface is a second subscan of at least part of the location.”**

307. It is my opinion that a POSITA would understand that Herbst renders obvious the elements of claim 16. As discussed with respect to claim 15 above, a POSITA would understand that the first and second 3D representations of Herbst, as shown, for example, in Figures 1(a)-1(b) and 6(a), 6(d) above, can be considered subscans of the final model of the tabletop scene (e.g., as shown in Fig. 7) because the model is not complete until both those scanning models are used to finalize the three-dimensional model. *See, supra*, § X.C.9.a; Ex-1006 at 2 (“The inputs to our system are two or more RGB-D videos . . .”).

308. Accordingly, it is my opinion that Herbst renders obvious “wherein the first representation of at least part of a surface is a provisional virtual 3D model comprising the subscans of the location acquired already, and the second

representation of at least part of the surface is a second subscan of at least part of the location,” as recited by claim 16.

11. Claim 17 Is Rendered Obvious by Herbst

309. As explained below, it is my opinion that Herbst renders every element of claim 17 obvious.

- a. **[Claim 17]: “The method according to claim 1, wherein acquired subscans of the location are adapted to be added to the provisional virtual 3D model concurrently with the acquisition of the subscans.”**

310. It is my opinion that a POSITA would find the elements of this claim obvious in light of Herbst’s teachings. Herbst teaches that “We build on RGB-D Mapping [10], a SLAM technique for generating dense 3-D maps.” Ex-1006 at 2. Claim 17 would be rendered obvious by the above teaching of Herbst, since a POSITA would understand that SLAM methods are designed to update a 3D model while tracking the scanner position and orientation. A POSITA would have understood that adapting subscans “to be added to the provisional virtual 3D model concurrently with the acquisition of the subscans.” would directly correspond to the definition of methods in the prior arts known as simultaneous localization and mapping (SLAM). It would have been obvious to a POSITA to update a real-time provisional model of the tabletop scenes of Herbst in order to obtain a preview of the final model.

311. Accordingly, it is my opinion that Herbst, in combination with the knowledge and creativity of a POSITA renders obvious, “wherein acquired subscans of the location are adapted to be added to the provisional virtual 3D model concurrently with the acquisition of the subscans,” as recited by claim 17.

12. Claim 20 Is Rendered Obvious by Herbst

312. As explained below, it is my opinion that Herbst renders every element of claim 20 obvious.

- a. **[Claim 20 pre]: “The method according to claim 1, where in the method includes:”**

[Claim 20a]: “disregarding the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model, and”

313. Herbst teaches this claim element. As discussed above with respect to second excluded volume element of claim 1, the method described by Herbst removes, and therefore disregards, any surface measurements corresponding to moving objects in the second excluded volume of the second representation. *Supra*, § X.C.1.f.

- b. **[Claim 20b]: “disregarding the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model.”**

314. Herbst teaches this claim element. As discussed above with respect to

first excluded volume element of claim 1, the method described by Herbst removes, and therefore disregards, any surface measurements corresponding to moving objects in the first excluded volume of the first representation. *Supra*, § X.C.1.g.

315. Accordingly, it is my opinion that Herbst renders obvious “disregarding the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model, and disregarding the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model,” as recited by claim 20.

13. Claim 21 Is Rendered Obvious by Herbst

316. As explained below, it is my opinion that Herbst renders every element of claim 21 obvious.

a. [Claim 21 pre]: “The method according to claim 1, wherein:”

[Claim 21a]: “the first excluded volume is determined to be a space where no surface of the rigid object is present, and”

317. Herbst teaches this claim element. As discussed above with respect to first excluded volume element of claim 1, Herbst discloses the determination of a first excluded volume that contains empty space and no rigid objects. *See supra*, § X.C.1.d. Further, the method described by Herbst removes, and therefore disregards,

any surface measurements corresponding to moving objects in the first excluded volume of the first representation. *Supra*, § X.C.1.g.

- b. **[Claim 21b]: “the second excluded volume is determined to be a space where no surface of the rigid object is present.”**

318. Herbst teaches this claim element. As discussed above with respect to second excluded volume element of claim 1, Herbst discloses the determination of a second excluded volume that contains empty space and no rigid objects. *See supra*, § X.C.1.e. Further, the method described by Herbst removes, and therefore disregards, any surface measurements corresponding to moving objects in the second excluded volume of the second representation. *Supra*, § X.C.1.f.

319. Accordingly, it is my opinion that Herbst renders obvious, “wherein the first excluded volume is determined to be a space where no surface of the rigid object is present, and the second excluded volume is determined to be a space where no surface of the rigid object is present,” as recited by claim 21.

14. Claim 22 Is Rendered Obvious by Herbst

320. As explained below, it is my opinion that Herbst renders every element of claim 22 obvious.

- a. **[Claim 22]: “A nontransitory computer readable medium encoded with a computer program product comprising program code for causing a data processing system to perform the method of claim 1,**

when said program code is executed on the data processing system.”

321. Herbst teaches this claim element. Herbst teaches that the steps of its algorithm and method are all performed on a “system” that comprises a “machine” that has “16 hardware threads.” *See, e.g.*, Ex-1006 at 6-7. In my opinion, a POSITA would therefore understand Herbst to teach or render obvious a nontransitory computer readable medium encoded with a computer program product comprising program code for causing a data processing system to perform the method disclosed by Herbst.

322. Accordingly, Herbst renders obvious, “a nontransitory computer readable medium encoded with a computer program product comprising program code for causing a data processing system to perform the method of claim 1, when said program code is executed on the data processing system,” as recited by claim 22.

15. Claim 23 Is Rendered Obvious by Herbst

323. As explained below, it is my opinion that Herbst teaches, and therefore renders obvious, every element of claim 23.

- a. **[Claim 23 pre]: “A system for detecting a movable object in a location when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the system comprises a hardware processor configured**

to:”

324. As discussed above, in section VIII.A, for the purposes of this declaration, I have treated the preamble of claim 23 as limiting. It is my opinion that Herbst teaches this claim element.

325. Herbst teaches “a method to determine movable parts of a scene based on differencing with another scene.” Ex-1006 at 7. Herbst further teaches that “[t]he inputs to our system are two or more RGB-D videos representing separate visits to the same location. From each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*. Objects can move between visits but are assumed static during each visit.” *Id.* at 2. In my opinion, Herbst’s disclosed method of “determin[ing] movable parts of a scene” corresponds to the claimed method for “detecting a movable object in a location.” Similarly, it is my opinion that the “static” objects disclosed by Herbst correspond to the “rigid object” recited in this claim element. I further believe that Herbst discloses scanning these static objects by means of a 3D scanner for generating a virtual 3D model of the rigid object, as recited by this element. Specifically, Herbst teaches the use of RGB-D cameras and RGB-D Mapping, which Herbst further discloses to be “a SLAM technique for generating dense 3-D maps.” *Id.* at 2. RGB-D cameras capture color in the form of light in red, green, and blue wavelengths (RGB) and also capture depth information

(D) using a 3D depth sensor. The color and depth information captured by the camera can be used to create 3D images and videos. As Herbst explains, “[w]e use the PrimeSensor, an *RGB-D* camera that provides dense registered color and depth measurements, to detect objects by means of their movement in 3D between two maps of the same location.” *Id.* at 1. RGB-D cameras and their abilities to scan objects and locations to thereby capture 3D images and videos were well-known to POSITAs at the time of filing of the ’551 patent. Furthermore, SLAM stands for “Simultaneous Localization And Mapping,” which is a technique for generating virtual 3D models, and more specifically 3D map models. In both my opinion and as evidenced by Herbst’s discussion of relevant prior art references utilizing the SLAM technique, a POSITA at the time would have known and understood SLAM as a technique for generating 3D models of objects, including both rigid and moveable objects, in a scene or location. *See id.* at 1.

326. Furthermore, Herbst teaches that its methods may all be performed on a computer using a program, as discussed with respect to the previous claim above. *See, e.g.*, Ex-1006 at 6-7. It is my opinion that a POSITA would understand Herbst to therefore disclose that its methods may be implemented on a computer system comprising a hardware processor.

327. Accordingly, it is my opinion that Herbst teaches “a system for

detecting a movable object in a location when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the system comprises a hardware processor” as recited by the preamble of claim 23.

b. [Claim 23a]: “provide a first 3D representation of at least part of a surface by scanning at least part of the location;”

328. Herbst teaches this claim element. As discussed above, Herbst describes scanning objects in a location using RGB-D cameras and SLAM to create and update 3D maps. Ex-1006 at 1. Herbst teaches that “[t]he inputs to our system are two or more RGB-D videos representing separate visits to the same location.” He also teaches that “[f]rom each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*.” He further teaches that “[i]n each experiment, the RGB-D camera was carried by hand to generate partial coverage of the scene.” In my opinion, a POSITA at the time the ’551 patent was filed would have understood that 3-D reconstruction from the first visit by Herbst to the location constitutes a first 3D representation, as recited by this element and that “partial coverage of the scene” constitutes “part of a surface”, as recited by this element. Herbst teaches scanning and detecting surfaces within a location or scene. Specifically, Herbst states that its “algorithm aligns the frames of each RGB-D video using SLAM, then reconstructs

each scene as a dense set of surface elements.” *Id.* at 1. Herbst further states that the RGB-D cameras “provide a dense matrix of pixels, each measuring the color and distance of the closest surface in the pixel direction.” *Id.* at 2. Figure 6(a) of Herbst, shown below, demonstrates such a first 3D reconstruction or representation of surfaces, including surfaces corresponding to a table and objects on the table, obtained from scanning a location.



Fig. 6(a)

329. Accordingly, it is my opinion that Herbst teaches “a hardware processor configured to provide a first 3D representation of at least part of a surface of the location,” as recited in this element of claim 23.

c. [Claim 23b]: “provide a second 3D representation of at least part of the surface by scanning at least part of the location;”

330. Herbst teaches this claim element. As discussed in the preceding section above, Herbst discloses “providing a first 3D representation of at least part of a surface of the location.” Herbst further discloses that this process of providing 3D representations is iterative and performed two or more times with respect to a particular scene or location. Specifically, Herbst states that “[t]he inputs to [its] system are two or more RGB-D videos representing separate visits to the same location.” Ex-1006 at 2. Herbst further discloses that its method relies on “information in the difference between two maps made at different times.” *Id.* at 1. As discussed in the preceding section, he also teaches that “[f]rom each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*.” He further teaches that “[i]n each experiment, the RGB-D camera was carried by hand to generate partial coverage of the scene.” It is my opinion that the 3-D reconstruction from the second visit among multiple visits taught by Herbst for partial coverage of the location corresponds to “a second 3D representation of at least part of a surface of the location,” as recited by this claim element. For example, Figure 6(d) of Herbst, reproduced below, shows a second 3D reconstruction or representation of the same tabletop scene shown in Figure 6(a) above, but with different objects in different

locations on the table within the scene.



Fig. 6(d)

331. Accordingly, it is my opinion that Herbst teaches “a hardware processor configured to ... provide a second 3D representation of at least part of the surface by scanning at least part of the location,” as recited by this element of claim 23.

- d. **[Claim 23c]: “determine for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

332. Herbst teaches this claim element. As discussed above, I understand that the district court has construed “excluded volume,” as used in this claim element,

to mean an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present.” Specifically, it is my opinion that Herbst teaches the determination of an excluded volume through the use of the probabilistic model disclosed by the reference.

333. Herbst teaches a probabilistic model for obtaining pixel measurements corresponding to surfaces observed in a scene by an RGB-D camera. *See* Ex-1006 at 2-3. The model incorporates surface information related to depth, color, and orientation of each pixel. *Id.* at 3. Herbst’s model considers two cases: one in which no information about a scene or surfaces is known and one in which such information is known. *Id.* at 3; Fig. 3. In the first case, Herbst’s model is able to calculate the probability that a surface within a scene is at a particular depth measurement (along with additional information related to color and orientation of the surface). *Id.* at 3; Fig. 3. Similarly, in the second case in which scene and surface information are already known, Herbst’s model is able to calculate a depth measurement of a surface relative to an expected known depth of the surface (again, along with corresponding color and orientation information). *Id.* at 3-4; Fig. 3. By aligning and comparing two different 3D representations of a location and applying these probabilistic models to the surfaces observed in the scene in each 3D representation, Herbst’s

method can determine a probability that any of the surfaces observed in either 3D representation moved relative to its location in or absence from the other 3D representation. *Id.* at 4-5. Accordingly, with respect to each of a first 3D representation and a second 3D representation, Herbst's model is able to predict the depth, and therefore the location, of all surfaces within the volume of the scene scanned by the RGB-D camera.

334. Herbst further teaches that “[f]ollowing global registration, we convert points to surfels [14] to reduce map size and smooth evidence from multiple frames. Each 3-D point from each frame is assigned to a surfel, either an existing one or a newly created one if there is no existing surfel close to the point. Fig. 1 shows two reconstructions of a location.” *Id.* at 2.

335. It is my opinion that a POSITA at the time the '551 patent was filed would understand that Herbst's disclosure of the reduction of the size of the map (i.e. 3D representation) of the scanning visit, by assigning surfels to 3D points detected in frames of that scanning visit, is teaching that the empty space determined accordingly between the scanner and the surface, where no 3D point on any surface is detected, is excluded from the 3D surfel representation of the corresponding scan visit. Therefore, a POSITA would further understand that the empty space for each scanning visit would constitute an excluded volume for the 3D representation of that

scanning visit.

336. It is my opinion that a POSITA would therefore understand Herbst to teach that its probabilistic model is used to determine not only the location of surfaces but also the empty space in a location. The scanned 3D space where no 3D point is present does not contribute to the 3D surface representation and would constitute an excluded volume in the 3D space being scanned, as recited by the claims, and more particularly would constitute an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present, according to the court's claim construction.

337. For example, it is my opinion that a POSITA would have understood that the empty space shown in Figure 6(a) below – i.e., the space in the scene that is not occupied by the table, the objects on the table, or any other surface – would constitute an excluded volume for the first 3D representation, according to the court's construction.



Fig. 6(a)

338. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious, “a hardware processor configured to ... determine for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 23.

- e. **[Claim 23d]: “determine for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

339. Herbst teaches this claim element. My analysis and discussion of Herbst’s probabilistic model in the preceding section as teaching an excluded

volume necessarily applies equally to both the probabilistic model's application to a first 3D representation and to a second 3D representation. Thus, with respect to Figure 6(d) below, which as discussed above shows a second 3D representation according to Herbst, it is my opinion that a POSITA would have understood that the empty space shown in the figure – i.e., the space in the scene that is not occupied by the table, the objects on the table, or any other surface – would constitute an excluded volume for the second 3D representation, according to the court's construction.

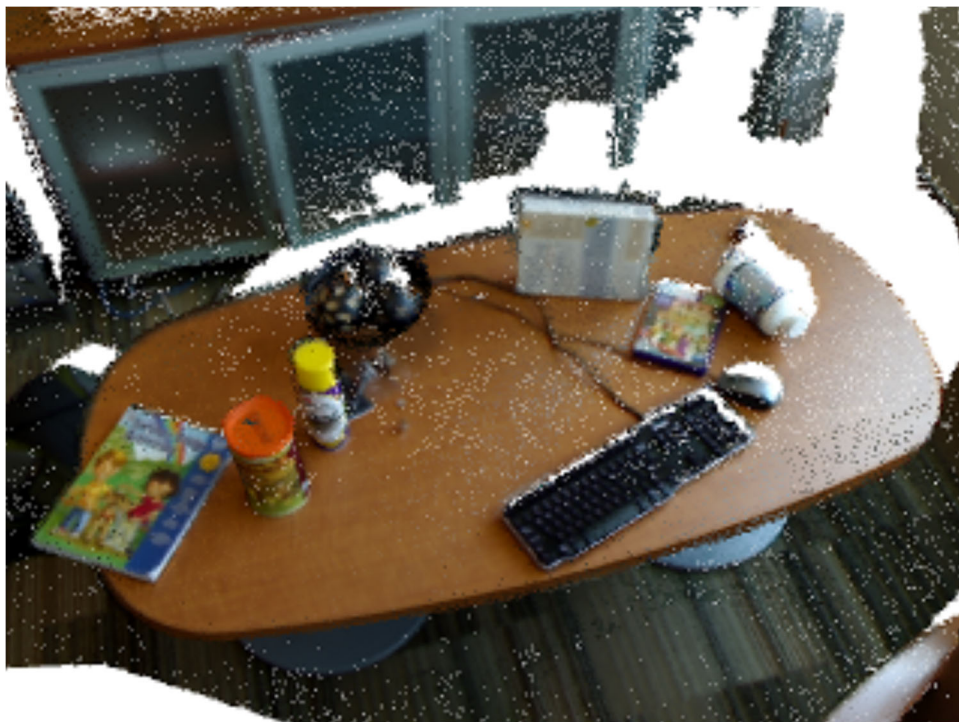


Fig. 6(d)

340. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious, “a hardware processor configured to ... determine for the second 3D

representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 23.

- f. **[Claim 23e]: “disregard the portion of the surface in the first 3D representation in the generation of the virtual 3D model, if a portion of the surface in the first 3D representation is located in space in the second excluded volume, and/or”**

341. Herbst teaches this claim element. As discussed above, Herbst teaches a method for comparing two or more 3D representations of a scene to determine if objects within the scene have moved between the two representations. Herbst further teaches that once all of the moveable objects have been detected, the two 3D representations can be merged into a single virtual 3D model from which all of the moveable objects are removed. *See* Ex-1006 at 5. Herbst teaches a probabilistic “differencing” approach relying, in particular, on depth differences for each surfel in a 3D representation to determine if a surfel belongs to a moving object or to the static background. *Id.* at 4-5. Herbst further discloses that a Bayesian framework is used to determine the probability of a surfel belonging to a moving object or to the static background, by first aligning the first representation with the second representation to find the corresponding surfels, so that an “expected” value of the surfel depth in the second representation can be computed based on the first

representation. *Id.* The probability of a surfel in the second representation being a moving one or not is then determined as a function of the difference between its measured depth and its expected depth value (determined from the first representation). *Id.* Herbst teaches also a regularization approach is used so that the probabilistic model is not overly sensitive to small differences due to measurement or alignment errors, i.e., only large probabilities contribute to the identification of moving surfels. *Id.*

342. In my opinion, a POSITA at the time the '551 patent was filed would have understood that the difference between the measured surfel depth in the second representation and its expected value computed from the first representation, which Herbst uses to determine the conditional probabilities could result only in three possibilities: (1) measured depth is smaller than the expected depth (difference is negative, leading to a high probability of motion for Herbst), (2) measured depth is larger than the expected depth (difference is positive, leading again to a high probability of motion for Herbst), (3) measured depth is approximately the same as the expected value (no significant difference, so the difference is within measurement and alignment error, i.e. low probability of motion). A POSITA would have understood that the first case occurs when a surfel from the first 3D representation is located in the second excluded volume. Herbst further discloses

that such surfel could be identified with high probability and therefore removed from the merged 3D model in SLAM. Since a collection of surfels in the first 3D representation would define a portion of the surface in the first representation, it is my opinion that this claim element is made obvious by the teachings of Herbst.

343. For example, Herbst shows a first 3D representation of a tabletop scene in Figure 6(a) and a second 3D representation of the same tabletop scene in Figure 6(d). Both figures are reproduced below.

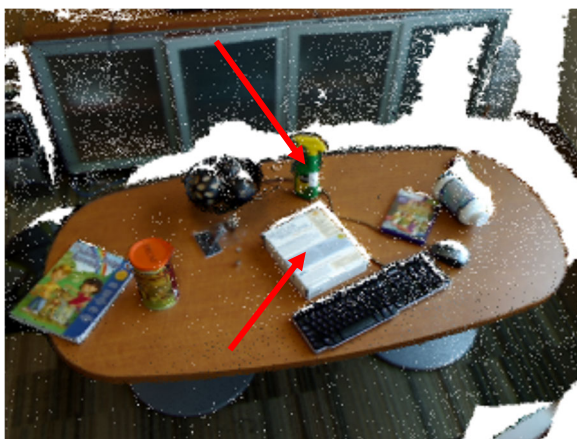


Fig. 6(a) (annotated)



Fig. 6(d)

As discussed above, a POSITA would understand Herbst to teach that all empty space in these respective 3D representations that is not occupied by a surface would collectively constitute an excluded volume for each 3D representation. Thus, when comparing the two representations according to Herbst's probabilistic model, those surfaces in the first 3D representation that are in the excluded volume of the second 3D representation (i.e., the second excluded volume) are determined to have moved.

Those two objects are identified in the left figure above with **red** arrows.

344. Furthermore, in Figure 7(b) of Herbst below, which illustrates a virtual 3D model of the scene created by merging the two 3D representations of Figures 6(a) and 6(d), the moved objects in the first 3D representation – i.e., those objects that are in the second excluded volume – have been removed or, in the language of the claim, “disregarded.” In other words, the two moved objects are not included in the virtual 3D model, but instead shown separately over the virtual 3D model of the scene.



Fig. 7(b)

345. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious “a hardware processor configured to ... disregard the portion of the surface in the first 3D representation in the generation of the virtual 3D model, if a portion

of the surface in the first 3D representation is located in space in the second excluded volume,” as recited by this element of claim 23.

- g. [Claim 23f]: “disregard the portion of the surface in the second 3D representation in the generation of the virtual 3D model, if a portion of the surface in the second 3D representation is located in space in the first excluded volume.”**

346. Herbst teaches this claim element. My analysis of Herbst’s teachings in the preceding section apply equally to this claim element. *See supra*, § X.C.15.f. As described in the preceding section, a POSITA at the time the ’551 patent was filed would have understood that the difference between the measured surfel depth in the second representation and its expected value computed from the first representation, which Herbst uses to determine the conditional probabilities could result only in three possibilities: (1) measured depth is smaller than the expected depth (difference is negative, leading to a high probability for Herbst), (2) measured depth is larger than the expected depth (difference is positive, leading again to a high probability for Herbst), (3) measured depth is approximately the same as the expected value (no significant difference, so the difference is within measurement and alignment error, i.e. low probability of motion). *Id.* A POSITA would have understood that the second case occurs when a surfel from the second 3D representation is located in the first excluded volume. As described in the preceding

section, Herbst further discloses that such surfels could be identified with high probability and therefore removed from the merged 3D model in SLAM. Since a collection of surfels in the second 3D representation would define a portion of the surface in the second representation, it is my opinion that this claim element is made obvious by the teachings of Herbst.

347. Referring again to Figures 6(a) and 6(d) of Herbst, reproduced below, as stated above a POSITA would understand Herbst to teach that all empty space in these respective 3D representations that is not occupied by a surface would collectively constitute an excluded volume for each 3D representation.



Fig. 6(a)



Fig. 6(d) (annotated)

Thus, when comparing the two representations according to Herbst's probabilistic model, those surfaces in the second 3D representation that are in the excluded volume of the first 3D representation (i.e., the first excluded volume) are determined to have moved. Those two objects are identified in the right figure above with

yellow arrows.

348. Thus, again when considering the merged virtual 3D model of Figure 7(b) of Herbst, shown below, the moved objects in the second 3D representation – i.e., those objects that are in the first excluded volume – have been removed or, in the language of the claim, “disregarded.” In other words, the two moved objects are not included in the virtual 3D model, but instead shown separately over the virtual 3D model of the scene.



Fig. 7(b)

349. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious, “a hardware processor configured to ... disregard the portion of the surface in the second 3D representation in the generation of the virtual 3D model, if a portion of the surface in the second 3D representation is located in space in the first excluded

volume,” as recited by this element of claim 23.

16. Claim 24 Is Rendered Obvious by Herbst

350. As explained below, it is my opinion that Herbst renders every element of claim 24 obvious.

- a. **[Claim 24 Pre]: “The system according to claim 23, where in the hardware processor is configured to:”**

[Claim 24a]: “disregard the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model, and”

351. Herbst teaches this claim element. As discussed above with respect to second excluded volume element of claim 23, the method described by Herbst removes, and therefore disregards, any surface measurements corresponding to moving objects in the second excluded volume of the second representation. *Supra*, § X.C.15.e.

352. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious “a hardware processor configured to ... disregard the portion of the surface in the first 3D representation that is located in space in the second excluded volume in the generation of the virtual 3D model,” as recited by this element of claim 24.

- b. **[Claim 24b]: “disregard the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model.”**

353. Herbst teaches this claim element. As discussed above with respect to first excluded volume element of claim 23, the method described by Herbst removes, and therefore disregards, any surface measurements corresponding to moving objects in the first excluded volume of the first representation. *Supra*, § X.C.15.d.

354. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious, “a hardware processor configured to ... disregard the portion of the surface in the second 3D representation that is located in space in the first excluded volume in the generation of the virtual 3D model,” as recited by this element of claim 24.

17. Claim 25 Is Rendered Obvious by Herbst

355. As explained below, it is my opinion that Herbst teaches, and therefore renders obvious, every element of claim 25.

- a. [Claim 25 pre]: “A method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object, wherein the method comprises:”**

356. As discussed above, in section VIII.A, for the purposes of this declaration, I have treated the preamble of claim 25 as limiting. It is my opinion that Herbst teaches this claim element.

357. Herbst teaches “a method to determine movable parts of a scene based on differencing with another scene.” Ex-1006 at 7. Herbst further teaches that “[t]he

inputs to our system are two or more RGB-D videos representing separate visits to the same location. From each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*. Objects can move between visits but are assumed static during each visit.” *Id.* at 2. In my opinion, Herbst’s disclosed method of “determin[ing] movable parts of a scene” corresponds to the claimed method for “detecting a movable object in a location.” Similarly, it is my opinion that the “static” objects disclosed by Herbst correspond to the “rigid object” recited in this claim element. I further believe that Herbst discloses scanning these static objects by means of a 3D scanner for generating a virtual 3D model of the rigid object, as recited by this element. Specifically, Herbst teaches the use of RGB-D cameras and RGB-D Mapping, which Herbst further discloses to be “a SLAM technique for generating dense 3-D maps.” *Id.* at 2. RGB-D cameras capture color in the form of light in red, green, and blue wavelengths (RGB) and also capture depth information (D) using a 3D depth sensor. The color and depth information captured by the camera can be used to create 3D images and videos. As Herbst explains, “[w]e use the PrimeSensor, an *RGB-D* camera that provides dense registered color and depth measurements, to detect objects by means of their movement in 3D between two maps of the same location.” *Id.* at 1. RGB-D cameras and their abilities to scan objects and locations to thereby capture 3D images and videos were well-known to

POSITAs at the time of filing of the '551 patent. Furthermore, SLAM stands for “Simultaneous Localization And Mapping,” which is a technique for generating virtual 3D models, and more specifically 3D map models. In both my opinion and as evidenced by Herbst’s discussion of relevant prior art references utilizing the SLAM technique, a POSITA at the time would have known and understood SLAM as a technique for generating 3D models of objects, including both rigid and moveable objects, in a scene or location. *See id.* at 1.

358. Accordingly, it is my opinion that Herbst teaches “method for detecting a movable object in a location, when scanning a rigid object in the location by means of a 3D scanner for generating a virtual 3D model of the rigid object,” as recited by the preamble of claim 25.

b. [Claim 25a]: “providing a first 3D representation of at least part of a surface by scanning at least part of the location;”

359. Herbst teaches this claim element. As discussed above, Herbst describes scanning objects in a location using RGB-D cameras and SLAM to create and update 3D maps. Ex-1006 at 1. Herbst teaches that “[t]he inputs to our system are two or more RGB-D videos representing separate visits to the same location.” He also teaches that “[f]rom each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*.” He further teaches that “[i]n each experiment, the

RGB-D camera was carried by hand to generate partial coverage of the scene.” In my opinion, a POSITA at the time the ’551 patent was filed would have understood that 3-D reconstruction from the first visit by Herbst to the location constitutes a first 3D representation, as recited by this element and that “partial coverage of the scene” constitutes “part of a surface”, as recited by this element. Herbst teaches scanning and detecting surfaces within a location or scene. Specifically, Herbst states that its “algorithm aligns the frames of each RGB-D video using SLAM, then reconstructs each scene as a dense set of surface elements.” *Id.* at 1. Herbst further states that the RGB-D cameras “provide a dense matrix of pixels, each measuring the color and distance of the closest surface in the pixel direction.” *Id.* at 2. Figure 6(a) of Herbst, shown below, demonstrates such a first 3D reconstruction or representation of surfaces, including surfaces corresponding to a table and objects on the table, obtained from scanning a location.



Fig. 6(a)

360. Accordingly, it is my opinion that Herbst teaches “providing a first 3D representation of at least part of a surface of the location,” as recited in this element of claim 25.

c. [Claim 25b]: “providing a second 3D representation of at least part of the surface by scanning at least part of the location;”

361. Herbst teaches this claim element. As discussed in the preceding section above, Herbst discloses “providing a first 3D representation of at least part of a surface of the location.” Herbst further discloses that this process of providing 3D representations is iterative and performed two or more times with respect to a

particular scene or location. Specifically, Herbst states that “[t]he inputs to [its] system are two or more RGB-D videos representing separate visits to the same location.” Ex-1006 at 2. Herbst further discloses that its method relies on “information in the difference between two maps made at different times.” *Id.* at 1. As discussed in the preceding section, he also teaches that “[f]rom each sequence we generate a geometrically consistent 3-D reconstruction, or *scene*.” He further teaches that “[i]n each experiment, the RGB-D camera was carried by hand to generate partial coverage of the scene.” It is my opinion that the 3-D reconstruction from the second visit among multiple visits taught by Herbst for partial coverage of the location corresponds to “a second 3D representation of at least part of a surface of the location,” as recited by this claim element. For example, Figure 6(d) of Herbst, reproduced below, shows a second 3D reconstruction or representation of the same tabletop scene shown in Figure 6(a) above, but with different objects in different locations on the table within the scene.

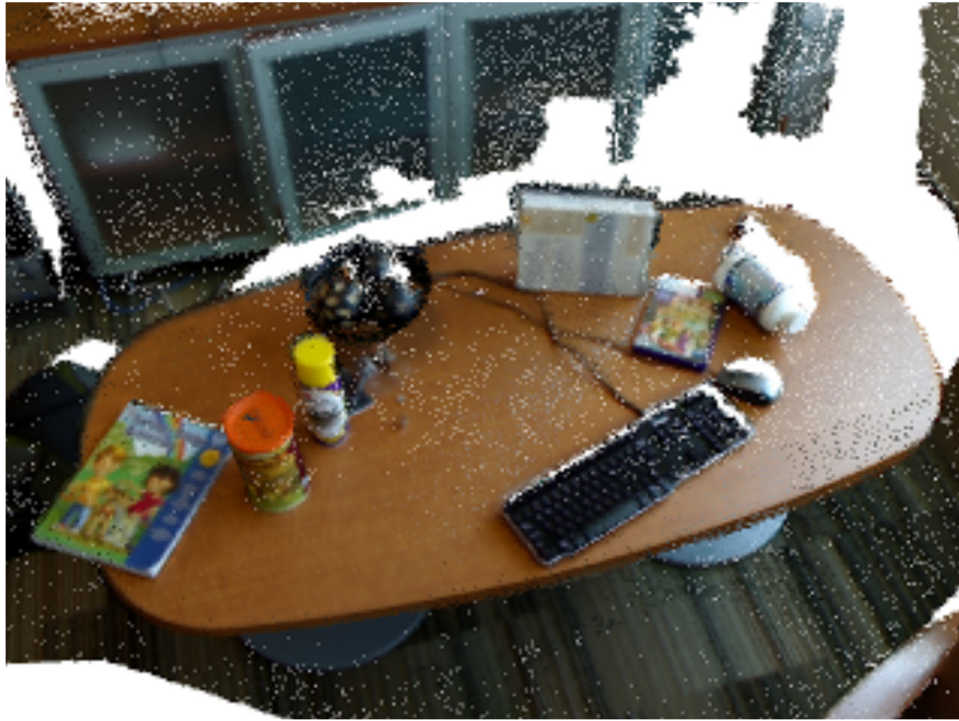


Fig. 6(d)

362. Accordingly, it is my opinion that Herbst teaches “providing a second 3D representation of at least part of the surface by scanning at least part of the location,” as recited by this element of claim 1.

- d. **[Claim25c]: “determining for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation, and/or determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation;”**

363. Herbst teaches this claim element. As discussed above, I understand that the district court has construed “excluded volume,” as used in this claim element,

to mean an “excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present.” Specifically, it is my opinion that Herbst teaches the determination of an excluded volume through the use of the probabilistic model disclosed by the reference.

364. Herbst teaches a probabilistic model for obtaining pixel measurements corresponding to surfaces observed in a scene by an RGB-D camera. *See* Ex-1006 at 2-3. The model incorporates surface information related to depth, color, and orientation of each pixel. *Id.* at 3. Herbst’s model considers two cases: one in which no information about a scene or surfaces is known and one in which such information is known. *Id.* at 3; Fig. 3. In the first case, Herbst’s model is able to calculate the probability that a surface within a scene is at a particular depth measurement (along with additional information related to color and orientation of the surface). *Id.* at 3; Fig. 3. Similarly, in the second case in which scene and surface information are already known, Herbst’s model is able to calculate a depth measurement of a surface relative to an expected known depth of the surface (again, along with corresponding color and orientation information). *Id.* at 3-4; Fig. 3. By aligning and comparing two different 3D representations of a location and applying these probabilistic models to the surfaces observed in the scene in each 3D representation, Herbst’s

method can determine a probability that any of the surfaces observed in either 3D representation moved relative to its location in or absence from the other 3D representation. *Id.* at 4-5. Accordingly, with respect to each of a first 3D representation and a second 3D representation, Herbst's model is able to predict the depth, and therefore the location, of all surfaces within the volume of the scene scanned by the RGB-D camera.

365. Herbst further teaches that “[f]ollowing global registration, we convert points to surfels [14] to reduce map size and smooth evidence from multiple frames. Each 3-D point from each frame is assigned to a surfel, either an existing one or a newly created one if there is no existing surfel close to the point. Fig. 1 shows two reconstructions of a location.” *Id.* at 2. It is my opinion that a POSITA at the time the '551 patent was filed would understand that Herbst's disclosure of the reduction of the size of the map (i.e. 3D representation) of the scanning visit, by assigning surfels to 3D points detected in frames of that scanning visit, is teaching that the empty space determined accordingly between the scanner and the surface, where no 3D point on any surface is detected, is excluded from the 3D surfel representation of the corresponding scan visit. Therefore, a POSITA would further understand that the empty space for each scanning visit would constitute an excluded volume for the 3D representation of that scanning visit.

366. It is my opinion that a POSITA would therefore understand Herbst to teach that its probabilistic model is used to determine not only the location of surfaces but also the empty space in a location. The scanned 3D space where no 3D point is present does not contribute to the 3D surface representation and would constitute an excluded volume in the 3D space being scanned, as recited by the claims, and more particularly would constitute an excluded volume in space that is common to both the first and second 3D representations where no surface from either of the 3D representations should be present, according to the court's claim construction.

367. For example, it is my opinion that a POSITA would have understood that the empty space shown in Figure 6(a) below – i.e., the space in the scene that is not occupied by the table, the objects on the table, or any other surface – would constitute an excluded volume for the first 3D representation, according to the court's construction.



Fig. 6(a)

368. It is further my opinion that a POSITA would understand that these teachings of Herbst necessarily apply equally to both the probabilistic model's application to a first 3D representation and to a second 3D representation. Thus, with respect to Figure 6(d) below, which as discussed above shows a second 3D representation according to Herbst, it is my opinion that a POSITA would have understood that the empty space shown in the figure – i.e., the space in the scene that is not occupied by the table, the objects on the table, or any other surface – would constitute an excluded volume for the second 3D representation, according to the court's construction.

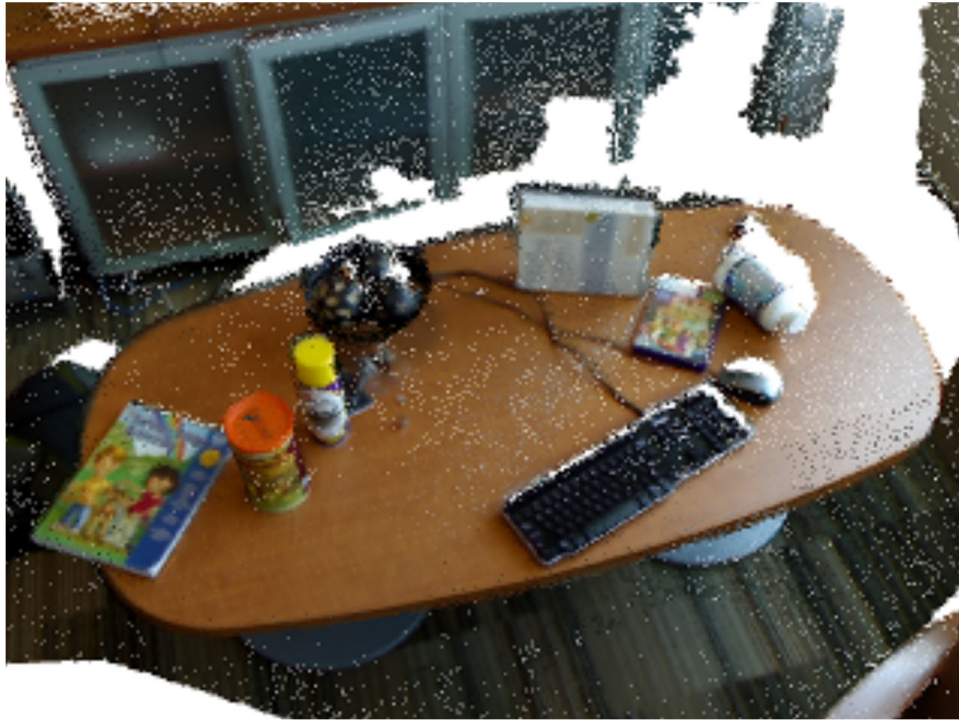


Fig. 6(d)

369. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious, “determining for the first 3D representation a first excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation, and/or determining for the second 3D representation a second excluded volume in space where no surface can be present in both the first 3D representation and the second 3D representation,” as recited in claim 25.

- e. **[Claim 25d]: “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, and/or if a portion of the surface in the second 3D representation is located in**

space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model.”

370. Herbst teaches this claim element. As discussed above, Herbst teaches a method for comparing two or more 3D representations of a scene to determine if objects within the scene have moved between the two representations. Herbst further teaches that once all of the moveable objects have been detected, the two 3D representations can be merged into a single virtual 3D model from which all of the moveable objects are removed. *See* Ex-1006 at 5. Herbst teaches a probabilistic “differencing” approach relying, in particular, on depth differences for each surfel in a 3D representation to determine if a surfel belongs to a moving object or to the static background. *Id.* at 4-5. Herbst further discloses that a Bayesian framework is used to determine the probability of a surfel belonging to a moving object or to the static background, by first aligning the first representation with the second representation to find the corresponding surfels, so that an “expected” value of the surfel depth in the second representation can be computed based on the first representation. *Id.* The probability of a surfel in the second representation being a moving one or not is then determined as a function of the difference between its measured depth and its expected depth value (determined from the first representation). *Id.* Herbst teaches also a regularization approach is used so that the

probabilistic model is not overly sensitive to small differences due to measurement or alignment errors, i.e., only large probabilities contribute to the identification of moving surfels. *Id.*

371. In my opinion, a POSITA at the time the '551 patent was filed would have understood that the difference between the measured surfel depth in the second representation and its expected value computed from the first representation, which Herbst uses to determine the conditional probabilities could result only in three possibilities: (1) measured depth is smaller than the expected depth (difference is negative, leading to a high probability of motion for Herbst), (2) measured depth is larger than the expected depth (difference is positive, leading again to a high probability of motion for Herbst), (3) measured depth is approximately the same as the expected value (no significant difference, so the difference is within measurement and alignment error, i.e. low probability of motion). A POSITA would have understood that the first case occurs when a surfel from the first 3D representation is located in the second excluded volume. Herbst further discloses that such surfel could be identified with high probability and therefore removed from the merged 3D model in SLAM. Since a collection of surfels in the first 3D representation would define a portion of the surface in the first representation, it is my opinion that this claim element is made obvious by the teachings of Herbst.

372. For example, Herbst shows a first 3D representation of a tabletop scene in Figure 6(a) and a second 3D representation of the same tabletop scene in Figure 6(d). Both figures are reproduced below.



Fig. 6(a) (annotated)

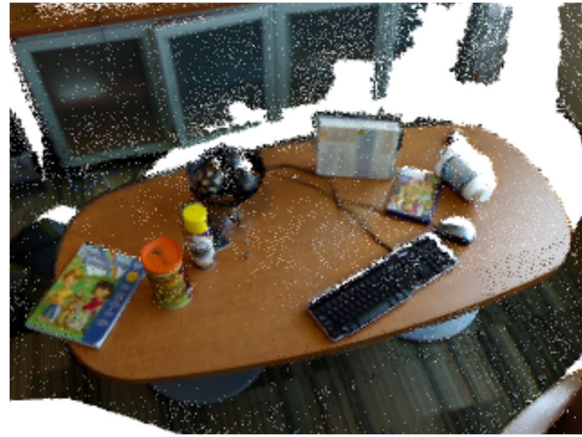


Fig. 6(d)

As discussed above, a POSITA would understand Herbst to teach that all empty space in these respective 3D representations that is not occupied by a surface would collectively constitute an excluded volume for each 3D representation. Thus, when comparing the two representations according to Herbst's probabilistic model, those surfaces in the first 3D representation that are in the excluded volume of the second 3D representation (i.e., the second excluded volume) are determined to have moved. Those two objects are identified in the left figure above with **red** arrows.

373. Furthermore, in Figure 7(b) of Herbst below, which illustrates a virtual 3D model of the scene created by merging the two 3D representations of Figures 6(a) and 6(d), the moved objects in the first 3D representation – i.e., those objects that

are in the second excluded volume – have been removed or, in the language of the claim, “disregarded.” In other words, the two moved objects are not included in the virtual 3D model, but instead shown separately over the virtual 3D model of the scene.



Fig. 7(b)

374. As described above, a POSITA at the time the '551 patent was filed would have further understood that the difference between the measured surfel depth in the second representation and its expected value computed from the first representation, which Herbst uses to determine the conditional probabilities could result only in three possibilities: (1) measured depth is smaller than the expected depth (difference is negative, leading to a high probability for Herbst), (2) measured depth is larger than the expected depth (difference is positive, leading again to a high probability for Herbst), (3) measured depth is approximately the same as the

expected value (no significant difference, so the difference is within measurement and alignment error, i.e. low probability of motion). A POSITA would have understood that the second case occurs when a surfel from the second 3D representation is located in the first excluded volume. As described in the preceding section, Herbst further discloses that such surfels could be identified with high probability and therefore removed from the merged 3D model in SLAM. Since a collection of surfels in the second 3D representation would define a portion of the surface in the second representation, it is my opinion that this claim element is made obvious by the teachings of Herbst.

375. Referring again to Figures 6(a) and 6(d) of Herbst, reproduced below, as stated above a POSITA would understand Herbst to teach that all empty space in these respective 3D representations that is not occupied by a surface would collectively constitute an excluded volume for each 3D representation.



Fig. 6(a)



Fig. 6(d) (annotated)

376. Thus, when comparing the two representations according to Herbst's probabilistic model, those surfaces in the second 3D representation that are in the excluded volume of the first 3D representation (i.e., the first excluded volume) are determined to have moved. Those two objects are identified in the right figure above with yellow arrows.

377. Thus, again when considering the merged virtual 3D model of Figure 7(b) of Herbst, shown below, the moved objects in the second 3D representation – i.e., those objects that are in the first excluded volume – have been removed or, in the language of the claim, “disregarded.” In other words, the two moved objects are not included in the virtual 3D model, but instead shown separately over the virtual 3D model of the scene.



Fig. 7(b)

378. Accordingly, it is my opinion that Herbst teaches, and therefore renders obvious, “if a portion of the surface in the first 3D representation is located in space in the second excluded volume, the portion of the surface in the first 3D representation is disregarded in the generation of the virtual 3D model, and/or if a portion of the surface in the second 3D representation is located in space in the first excluded volume, the portion of the surface in the second 3D representation is disregarded in the generation of the virtual 3D model,” as recited in claim 25.

D. Claims 2-5, 18, and 19 Are Rendered Obvious By Herbst in View of McQueston

379. It is my opinion that Herbst in view of McQueston teaches, and therefore renders obvious, every element of claims 2-5, 18, and 19 of the '551 patent.

380. As discussed above in Section X.C, it is my opinion that Herbst teaches all the elements of claims 1, 6-12, 15-17, and 20-25. Herbst identifies the problem of moving bodies obstructing a scanning volume and interfering with 3D modeling of an environment.

381. McQueston describes a similar, if not identical, problem albeit in the context of intraoral imaging, where dental instruments or other objects may interfere with the 3D imaging of teeth by obstructing a scanning volume between a scanning device and the tooth surfaces. Ex-1008 at 1:9-23. As with Herbst, McQueston teaches using laser range finding for distance measurements underlying the three-

dimensional tooth modeling. Ex-1008 at 4:65-5:26; 5:47-63. Likewise, as with the RGB-D depth imaging described by Herbst, McQueston teaches that imaging techniques may be incorporated into the three-dimensional modeling of the teeth. Ex-1008 at 3:31-61.

382. It is my opinion that a POSITA would be motivated to apply the RGB-D imaging techniques and components of Herbst to the intraoral environment of McQueston in order to achieve the stated goal of McQueston — minimizing the interference of dental objects with the three-dimensional model of the teeth being imaged or measured. It is my opinion that a POSITA would understand that the imaging and modeling techniques and devices of Herbst (Ex-1006), could be integrated into a dental scanning apparatus, like that described by McQueston, in order to address the same problem of removing any influence or interference of moving bodies from three-dimensional modeling of stationary objects, in this case teeth. Thus, it is my opinion that Herbst in view of McQueston renders claims 2-5, 18, and 19 obvious, as discussed in more detail below.

1. Claim 2 Is Rendered Obvious By Herbst in View of McQueston

383. As discussed above, McQueston describes obtaining a three-dimensional model of teeth located in an intra-oral environment. Ex-1008 at 1:9-23 For the reasons stated above, it is my opinion that it would have been obvious to a

POSITA to apply the modeling techniques of Herbst to the intra-oral environment of McQueston in order to measure tooth surfaces without the interference of moving bodies and thereby obtain a more accurate three-dimensional model of the tooth surfaces. Therefore, it is my opinion that Herbst in view of McQueston renders obvious, and therefore invalid, that the “rigid object is a patient’s set of teeth, and the location is the mouth of the patient,” as claimed.

2. Claims 3-5, 18, and 19 Are Rendered Obvious By Herbst in View of McQueston

384. As discussed above, McQueston teaches that the movable objects that may interfere with the three-dimensional modeling of the teeth may be a number of example dental objects, including soft tissue, intraoral structures, and dental instruments. Ex-1008 at 4:4-46. McQueston also describes that a finger may interfere with the tooth surface measurement. *See, e.g., id.* at 1:9-23. Therefore, it is my opinion that Herbst in view of McQueston renders obvious, and therefore invalid, the claimed “movable object” being “soft tissue,” “dentist’s instrument,” “finger,” “tongue,” or “dental suction device,” as recited in claims 3-5, 18, and 19.

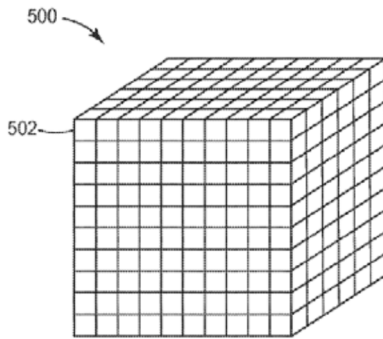
XI. DEMONSTRATION OF WRITTEN DESCRIPTION SUPPORT FOR MCQUESTON’S EFFECTIVE FILING DATE OF FEBRUARY 22, 2011

385. McQueston issued on January 26, 2016, and claims priority to U.S. Provisional Application No. 61/445,499 (Ex-1009 (McQueston Provisional)), which

was filed February 22, 2011. Below I provide my opinion regarding the McQueston Provisional's written description support for claim 1 of McQueston.

386. I have reviewed the McQueston Provisional and have compared its disclosure against claim 1 of McQueston. I note that claim 1 of McQueston is nearly identical to claim 1 of the McQueston provisional, and further incorporates a modified version of the limitation of claim 5 of the McQueston provisional. Based on my review of the documents, McQueston and the McQueston provisional appear to contain nearly identical written specifications. I did not identify any substantive differences in the disclosures of the two documents. Thus, in my opinion and based on my review of McQueston and the McQueston Provisional, the McQueston Provisional contains a written description of the claimed inventions in such full, clear, concise, and exact terms, to enable any ordinarily skilled artisan to practice the invention claimed in claim 1 of McQueston. A more detailed analysis of the support found in the McQueston Provisional for each element of claim 1 of McQueston is included below.

Elements of Claim 1 of McQueston	Support in McQueston Provisional
A method comprising:	"In one aspect, a method disclosed herein includes storing a spatial matrix in a memory, the spatial matrix including information for a plurality of voxels in a mesh of a scan space; obtaining three-dimensional data from an object within the scan space using a scanner; identifying a one of the plurality of voxels in the mesh that is unoccupied;

	storing an indication in the spatial matrix for a modified processing of data from the one of the plurality of voxels, thereby identifying a prohibited volume within the scan space; and processing a portion of the three-dimensional data that is not in the prohibited volume according to an unmodified processing and processing a second portion of the three-dimensional data that is in the prohibited volume according to the modified processing.” (McQueston Provisional, [0004].)
storing a spatial matrix in a memory, the spatial matrix including information for a plurality of voxels in a mesh of a scan space;	“In one aspect, a method disclosed herein includes storing a spatial matrix in a memory, the spatial matrix including information for a plurality of voxels in a mesh of a scan space;” (McQueston Provisional, [0004].) “Fig. 5 illustrates a spatial matrix for use in space carving. In general, the spatial matrix 500 may include a number of cells 502, each storing information related to a voxel in a scan volume.” (McQueston Provisional, [0055].)  <i>FIG. 5</i> (McQueston Provisional, FIG. 5.) “The spatial matrix 500 may include any memory or combination of memories capable of storing and retrieving information according to voxels within the scan space, and all such variations that would be apparent to one of ordinary skill in the art are intended to fall within the scope of this invention. The spatial matrix 500 may define a mesh of voxels that together form the scan

	volume or scan space in which data is processed as generally described below.” (McQueston Provisional, [0056].) “As shown in step 702, the process 700 may begin with storing a spatial matrix in a memory, such as the spatial matrix described above that includes information for a plurality of voxels in a mesh of a scan space.” (McQueston Provisional, [0060].)
obtaining three-dimensional data from an object within scan space using a scanner;	“In one aspect, a method disclosed herein includes ... obtaining three-dimensional data from an object within the scan space using a scanner; ...” (McQueston Provisional, [0004].) “Fig. 1 shows an image capture system. In general, the system 100 may include a scanner 102 that captures images of a subject 104 within an image plane 106, and forwards the images to a computer 108, which may include a display 110 and one or more user input devices such as a mouse 112 or a keyboard 114.” (McQueston Provisional, [0024].) “The scanner 102 may include any camera or camera system suitable for capturing images from which a three-dimensional point cloud may be recovered. For example, the scanner 102 may employ a multi-aperture system as disclosed, for example, in U.S. Pat. Pub. No. 20040155975 to Hart et al., the entire contents of which is incorporated herein by reference. While Hart discloses one multi-aperture system, it will be appreciated that any multi-aperture system suitable for reconstructing a three-dimensional point cloud from a number of two-dimensional images may similarly be employed, including systems with moving apertures, fixed apertures, and/or electromechanically shuttered apertures. In one multi-aperture embodiment, the scanner 102 may include a plurality of apertures including a center aperture positioned along a center optical axis of a lens and any

	associated imaging hardware. The scanner 102 may also, or instead, include a stereoscopic, triscopic or other multicamera or other configuration in which a number of cameras or optical paths are maintained in fixed or moving relation to one another to obtain two-dimensional images of an object from a number of slightly different perspectives. The scanner 102 may include suitable processing for deriving a three-dimensional point cloud from an image set or a number of image sets, or each two-dimensional image set may be transmitted to an external processor such as contained in the computer 108 described below. In other embodiments, the scanner 102 may employ structured light, laser scanning, direct ranging (e.g., time of flight in a known direction), or any other technology suitable for acquiring three-dimensional data, or two-dimensional data that can be resolved into three-dimensional data.” (McQueston Provisional, [0025].) “In one embodiment, the scanner 102 is a handheld, freely positionable probe having at least one user input device, such as a button, lever, dial, thumb wheel, switch, or the like, for user control of the image capture system 100 such as starting and stopping scans. In an embodiment, the scanner 102 may be shaped and sized for dental scanning. More particularly, the scanner may be shaped and sized for intraoral scanning and data capture, such as by insertion into a mouth of an imaging subject and passing the imaging plane 106 over one or more intraoral structures at a suitable distance to acquire surface data from teeth, gums, and so forth. The scanner 102 may, through such a continuous acquisition process, capture a point cloud of surface data having sufficient spatial resolution and accuracy to prepare dental objects such as prosthetics, hardware, appliances, and the like therefrom, either directly or through a variety of intermediate processing steps. In other embodiments, surface data may be acquired from a dental model such as
--	---

	a dental prosthetic, to ensure proper fitting using a previous scan of corresponding dentition, such as a tooth surface prepared for the prosthetic.” (McQueston Provisional, [0026].) “In an example of general operation of a three-dimensional image capture system 100, the scanner 102 may acquire two-dimensional image sets at a video rate while the scanner 102 is passed over a surface of the subject. The two-dimensional image sets may be forwarded to the computer 108 for derivation of three-dimensional point clouds. The three-dimensional data for each newly acquired two-dimensional image set may be derived and fitted or "stitched" to existing three-dimensional data using a number of different techniques. One useful example of such a technique is described in commonly-owned U.S. App. No. 11/270,135, filed on November 9, 2005, the entire contents of which is incorporated herein by reference. However, it will be appreciated that this example is not limiting, and that the principles described herein may be applied to a wide range of three-dimensional image capture systems. It will also be understood that terms such as "video" or "video rate" imply a wide range of possible frame rates associated with such video. While most modem video formats employ a frame rate of 25 to 30 frames per second, early video employed frame rates as low as 8 frames per second, and movies of the early 1900's varied from 12 to 18 frames per second. In addition, it is common for specialized imaging equipment to employ a rate adapted to the computational demands of particular imaging and rendering techniques, and some video systems operate with frame rates anywhere from 4 frames per second (for computationally extensive imaging systems) to 100 frames per second or higher (for high-speed video systems). As used herein, the terms video rate and frame rate should be interpreted broadly. Notwithstanding this broad meaning, it is noted that
--	--

	useful and visually pleasing three-dimensional imaging systems have been constructed according to the foregoing with frame rates of at least ten frames per second, frame rates of at least twenty frames per second, and frame rates between 25 and 30 frames per second.” (McQueston Provisional, [0032].) “As shown in step 704, three-dimensional data may be obtained from an object within a scan space using, e.g., any of the scanners described above.” (McQueston Provisional, [0061].)
identifying a one of the plurality of voxels in the mesh that is unoccupied:	“In one aspect, a method disclosed herein includes ... identifying a one of the plurality of voxels in the mesh that is unoccupied; ...” (McQueston Provisional, [0004].) “Identifying the one of the plurality of voxels may include identifying a plurality of points on a ray between the scanner and the object that are unoccupied. The method may include providing a spatial margin of one or more voxels adjacent to a surface of the object for which processing cannot be prohibited. The spatial margin may be a multiple of a standard deviation of a z-axis measurement for the scanner. The modified processing may include detecting a point in the prohibited volume and displaying the point with a decay time. The modified processing may include omitting any points in the prohibited volume from a surface mesh of the object created with the three-dimensional data. The mesh may divide the scan space into a spatial grid substantially identical to a second spatial grid used to create the surface mesh. The modified processing may include omitting any points in the prohibited volume from a high-accuracy refinement of the three-dimensional data.” (McQueston Provisional, [0005].) “Fig. 6 illustrates a ray trace used to identify unoccupied space. During a scan such as any of the scans described above, a scanner 602 such as any of the scanners

	described above may be passed over an object 604 within a scan space 606 to acquire three-dimensional data. While depicted two-dimensionally, it will be understood that the scan space 606 would generally be a three-dimensional volume. Given a position of the scanner 602 and the object 604, one or more rays - straight lines passing through the scan space 606 - may be determined such as a first ray 608 that intersects a surface of the object 604 and a second ray 610 that does not intersect a surface of the object 604. With the scan space 606 divided into a volumetric mesh according to the spatial matrix described above, all of the voxels that each ray 608, 610 passes through can be determined, and more particularly, each unoccupied one of the voxels in the mesh can be identified. Each unoccupied one of the voxels may be flagged in the spatial matrix as prohibited using, e.g., a binary flag or any other suitable identifier. In this manner, voxels that are known to be unoccupied at some time during the scan can be marked for modified processing and differentially processed during a scan, all as generally contemplated below.” (McQueston Provisional, [0057].) “Space carving is generally known in the art as a technique for identifying unoccupied volume, and any corresponding techniques or tools may be adapted to identify ‘prohibited’ voxels as contemplated herein. For example, commercially available software tools such as Voxel Space (Novalogic) and/or 3D-DDA support ray tracing through voxel spaces and may be used with space carving as described herein.” (McQueston Provisional, [0058].) “As shown in step 706, the process 700 may include identifying at least one of the plurality of voxels in the scan space that is unoccupied. This may be accomplished, for example, using any suitable ray tracing software or algorithm. A variety of commercially available software tools (e.g. 3D DDA) and general algorithms (e.g.,
--	---

	Bresenham's line algorithm) are known in the art for traversing voxels in a space and identifying what voxels contain a ray, any of which may be usefully adapted to identifying unoccupied voxels in a scan space as contemplated herein. In a ray tracing embodiment, unoccupied space may, for example, include voxels behind or beyond a known surface of an object, or voxels between a scanner and the surface, as well as combinations of these. In another aspect, other information may also or instead be used to determine whether a voxel should be forbidden or permitted. For example, for a voxel to be forbidden, the process 700 may provide a threshold for a minimum number of adjacent voxels that are already identified as forbidden. In another aspect, the number of rays passing through a voxel may be counted and used, e.g., as a minimum threshold for identifying a voxel as forbidden. For occluded surfaces of the scanned object, the process 700 may provide an offset or step away from the border of the occlusion in a field of view. More generally, any suitable techniques for characterizing a voxel as containing or not containing the surface of interest may be suitably employed to identify prohibited space as contemplated herein." (McQueston Provisional, [0062].)
storing an indication in the spatial matrix for a modified processing of data from the one of the plurality of voxels, thereby identifying a prohibited volume within the scan space; and	"In one aspect, a method disclosed herein includes ... storing an indication in the spatial matrix for a modified processing of data from the one of the plurality of voxels, thereby identifying a prohibited volume within the scan space; ..." (McQueston Provisional, [0004].) "As shown in step 708, the process 700 may include storing an indication for modified processing of the one of the plurality of voxels in the spatial matrix 710. After completing this step for one or more of the plurality of voxels, the spatial matrix identifies a prohibited volume that includes one or more voxels within the scan space. It will be understood that identifying the prohibited volume may be a cumulative process that is repeated as each new

	frame of three-dimensional data is repeated for a scan such as the sequential scans described above. Thus for example, each frame of data may provide different information on surfaces of the object that are occluded, that have changed, or that have disappeared, and this information may be generally used to update the spatial matrix in any suitable way. In general, it is contemplated that every voxel of the matrix is a permitted voxel until a reliable inference can be made that the voxel should be prohibited. However, this is not strictly required, and variations to this initial condition may be made without departing from the scope of this disclosure.” (McQueston Provisional, [0064].)
processing a portion of the three-dimensional data that is not in the prohibited volume according to an unmodified processing and processing a second portion of the three-dimensional data that is in the prohibited volume according to the modified processing, wherein the modified processing includes detecting a point in the prohibited volume and displaying the point for a decay time with the second portion of the three-dimensional data.	“In one aspect, a method disclosed herein includes ... and processing a portion of the three-dimensional data that is not in the prohibited volume according to an unmodified processing and processing a second portion of the three-dimensional data that is in the prohibited volume according to the modified processing.” (McQueston Provisional, [0004].) “As shown in step 716, the process 700 may include processing a portion of the three-dimensional data that is not in the prohibited volume - or alternatively stated, three-dimensional data that is in the permitted volume of the scan space -- according to an unmodified or normal processing. This may include any processing associated with the acquisition of three-dimensional data such as derivation of a camera path for a scanner, feature-based or other derivation of individual three-dimensional data points from frames of video or image data, registration of three-dimensional surfaces and so forth. In general, it is contemplated that any three-dimensional process that does not use space carving or the like may be performed in an ‘unmodified’ processing step as that term is used herein, including without limitation acquisition, display, and refinement (e.g., optimization, error minimization, etc.) of a three-dimensional surface representation based

	on, e.g., a sequence of frames of image data or the like from a three-dimensional scanner.” (McQueston Provisional, [0070].) “As shown in step 718, the process 700 may include processing a portion of the three-dimensional data that is in the prohibited volume according to the modified processing. Modified processing may include any processing suitable for data in the ‘prohibited’ volume.” (McQueston Provisional, [0071].) “Modified processing may, for example, include selective use of three-dimensional data points in a three-dimensional reconstruction according to whether they are in prohibited or permitted volumes of the scan space. For example, a reconstruction process may treat data within the prohibited volume as outliers that are excluded from processing. Or data in the prohibited volume may be de-weighted or otherwise deemphasized in the reconstruction process. This data exclusion or de-weighting may be performed in a real time reconstruction process while three-dimensional data is being displayed, or in a post-processing refinement step, or some combination of these.” (McQueston Provisional, [0072].) “The modified processing may include detecting a point in the prohibited volume and displaying the point with a decay time.” (McQueston Provisional, [0005].) “In another aspect, modified processing may include controlling how data in the prohibited volume is displayed. For example, data within the prohibited volume may be excluded from display, or may be displayed temporarily using a brief decay time such as 0.5 seconds or any other short interval.” (McQueston Provisional, [0073].)
--	---

387. Based on my review of McQueston and the McQueston Provisional, I have also determined that the McQueston provisional provides support for all of the subject matter of McQueston upon which I have relied in my analysis in this declaration. Again, I have observed that the relevant disclosures of the two documents, as summarized in the chart below, are nearly identical.

Disclosures of McQueston Cited in Support of The Analysis in this Declaration	Corresponding Disclosures in McQueston Provisional
McQueston, Ex-1008 at 1:9-23: “A wide range of techniques exist for acquiring three-dimensional data from a subject. One broad category of techniques employs continuous scanning in which incremental three-dimensional data is acquired and assembled into a full three-dimensional model. In this latter category, regardless of the particular scanning technology, the calculation and display of three-dimensional data can become encumbered by transient obstructions within a scanning volume. For example, in a dental scan, a patient's tongue or other soft tissue may temporarily obstruct or move adjacent to a scanned surface. A technician may similarly obstruct a scanner's view of the scanned surface with a finger, tongue depressor, or other instrument. There remains a need for techniques to exclude such outliers from a three-dimensional scanning process.”	McQueston Provisional, ¶ [0002]: “A wide range of techniques exist for acquiring three-dimensional data from a subject. One broad category of techniques employs continuous scanning in which incremental three-dimensional data is acquired and assembled into a full three-dimensional model. In this latter category, regardless of the particular scanning technology, the calculation and display of three-dimensional data can become encumbered by transient obstructions within a scanning volume. For example, in a dental scan, a patient's tongue or other soft tissue may temporarily obstruct or move adjacent to a scanned surface. A technician may similarly obstruct a scanner's view of the scanned surface with a finger, tongue depressor, or other instrument. There remains a need for techniques to exclude such outliers from a three-dimensional scanning process.”

McQueston, Ex-1008 at 3:31-61: “In the following description, the term “image” generally refers to a two-dimensional set of pixels forming a two-dimensional view of a subject within an image plane. The term “image set” generally refers to a set of related two-dimensional images that might be resolved into three-dimensional data. The term “point cloud” generally refers to a three-dimensional set of points forming a three-dimensional view of the subject, such as might be reconstructed from a number of two-dimensional views. In a three-dimensional image capture system, a number of such point clouds may also be registered and combined into an aggregate point cloud constructed from images captured by, for example, a moving camera. Thus it will be understood that pixels generally refer to two-dimensional data and points generally refer to three-dimensional data, unless another meaning is specifically indicated or otherwise clear from the context. The terms “three-dimensional surface representation”, “digital surface representation”, “three-dimensional surface map”, and the like, as used herein, are intended to refer to any three-dimensional surface map of an object, such as a point cloud of surface data, a set of two-dimensional	McQueston Provisional, ¶¶ [0018]-[0020]: “In the following description, the term "image" generally refers to a two-dimensional set of pixels forming a two-dimensional view of a subject within an image plane. The term "image set" generally refers to a set of related two-dimensional images that might be resolved into three-dimensional data. The term "point cloud" generally refers to a three-dimensional set of points forming a three-dimensional view of the subject, such as might be reconstructed from a number of two-dimensional views. In a three-dimensional image capture system, a number of such point clouds may also be registered and combined into an aggregate point cloud constructed from images captured by, for example, a moving camera. Thus it will be understood that pixels generally refer to two-dimensional data and points generally refer to three-dimensional data, unless another meaning is specifically indicated or otherwise clear from the context. The terms "three-dimensional surface representation", "digital surface representation", "three-dimensional surface map", and the like, as used herein, are intended to refer to any three-dimensional surface map of an object, such as a point cloud of surface data, a set of two-dimensional
--	--

polygons, a particle system, or any other renderable primitives or other data representing all or some of the surface of an object, as might be obtained through the capture and/or processing of three-dimensional scan data, unless a different meaning is explicitly provided or otherwise clear from the context. A “three-dimensional representation” may include any of the three-dimensional surface representations described above, as well as volumetric, vector, and/or other representations, unless a different meaning is explicitly provided or otherwise clear from the context.”	polygons, a particle system, or any other renderable primitives or other data representing all or some of the surface of an object, as might be obtained through the capture and/or processing of three-dimensional scan data, unless a different meaning is explicitly provided or otherwise clear from the context. A "three-dimensional representation" may include any of the three-dimensional surface representations described above, as well as volumetric, vector, and/or other representations, unless a different meaning is explicitly provided or otherwise clear from the context.”
McQueston, Ex-1008 at 4:4-46: “The term “dental object”, as used herein, is intended to refer broadly to subject matter specific to dentistry. This may include intraoral structures such as dentition, and more typically human dentition, such as individual teeth, quadrants, full arches, pairs of arches which may be separate or in various types of occlusion, soft tissue, and the like, as well bones and any other supporting or surrounding structures. As used herein, the term “intraoral structures” refers to both natural structures within a mouth as described above and artificial structures such as any of the dental objects described below that might be	McQueston Provisional, ¶ [0022]: “The term "dental object", as used herein, is intended to refer broadly to subject matter specific to dentistry. This may include intraoral structures such as dentition, and more typically human dentition, such as individual teeth, quadrants, full arches, pairs of arches which may be separate or in various types of occlusion, soft tissue, and the like, as well bones and any other supporting or surrounding structures. As used herein, the term "intraoral structures" refers to both natural structures within a mouth as described above and artificial structures such as any of the dental objects described below that might be

present in the mouth. Dental objects may include “restorations”, which may be generally understood to include components that restore the structure or function of existing dentition, such as crowns, bridges, veneers, inlays, onlays, amalgams, composites, and various substructures such as copings and the like, as well as temporary restorations for use while a permanent restoration is being fabricated. Dental objects may also include a “prosthesis” that replaces dentition with removable or permanent structures, such as dentures, partial dentures, implants, retained dentures, and the like. Dental objects may also include “appliances” used to correct, align, or otherwise temporarily or permanently adjust dentition, such as removable orthodontic appliances, surgical stents, bruxism appliances, snore guards, indirect bracket placement appliances, and the like. Dental objects may also include “hardware” affixed to dentition for an extended period, such as implant fixtures, implant abutments, orthodontic brackets, and other orthodontic components. Dental objects may also include “interim components” of dental manufacture which may not be intraoral structures as described above, such as dental models (full and/or partial), wax-ups, investment molds, and the like, as well as trays, bases, dies, and other components employed in the fabrication of restorations, prostheses,	present in the mouth. Dental objects may include "restorations", which may be generally understood to include components that restore the structure or function of existing dentition, such as crowns, bridges, veneers, inlays, onlays, amalgams, composites, and various substructures such as copings and the like, as well as temporary restorations for use while a permanent restoration is being fabricated. Dental objects may also include a "prosthesis" that replaces dentition with removable or permanent structures, such as dentures, partial dentures, implants, retained dentures, and the like. Dental objects may also include "appliances" used to correct, align, or otherwise temporarily or permanently adjust dentition, such as removable orthodontic appliances, surgical stents, bruxism appliances, snore guards, indirect bracket placement appliances, and the like. Dental objects may also include "hardware" affixed to dentition for an extended period, such as implant fixtures, implant abutments, orthodontic brackets, and other orthodontic components. Dental objects may also include "interim components" of dental manufacture which may not be intraoral structures as described above, such as dental models (full and/or partial), wax-ups, investment molds, and the like, as well as trays, bases, dies, and other components employed in the fabrication of restorations, prostheses,
--	--

and the like. Dental objects may be categorized in a number of ways. For example, dental objects may be categorized as natural dental objects such as the teeth, bone, and other intraoral structures described above and artificial dental objects such as the restorations, prostheses, appliances, hardware, and interim components of dental manufacture as described above. Similarly, dental objects may be categorized intraoral dental objects (which may be natural or artificial) and extraoral dental objects (which are generally but not exclusively artificial).”	and the like. Dental objects may be categorized in a number of ways. For example, dental objects may be categorized as natural dental objects such as the teeth, bone, and other intraoral structures described above and artificial dental objects such as the restorations, prostheses, appliances, hardware, and interim components of dental manufacture as described above. Similarly, dental objects may be categorized intraoral dental objects (which may be natural or artificial) and extraoral dental objects (which are generally but not exclusively artificial).”
McQueston, Ex-1008 at 4:65-5:26: “The scanner 102 may include any camera or camera system suitable for capturing images from which a three-dimensional point cloud may be recovered. For example, the scanner 102 may employ a multi-aperture system as disclosed, for example, in U.S. Pat. Pub. No. 20040155975 to Hart et al. While Hart discloses one multi-aperture system, it will be appreciated that any multi-aperture system suitable for reconstructing a three-dimensional point cloud from a number of two-dimensional images may similarly be employed, including systems with moving apertures, fixed apertures, and/or electro-mechanically shuttered apertures. In one multi-aperture embodiment, the scanner 102 may include a plurality of	McQueston Provisional, ¶ [0025]: “The scanner 102 may include any camera or camera system suitable for capturing images from which a three-dimensional point cloud may be recovered. For example, the scanner 102 may employ a multi-aperture system as disclosed, for example, in U.S. Pat. Pub. No. 20040155975 to Hart et al., the entire contents of which is incorporated herein by reference. While Hart discloses one multi-aperture system, it will be appreciated that any multi-aperture system suitable for reconstructing a three-dimensional point cloud from a number of two-dimensional images may similarly be employed, including systems with moving apertures, fixed apertures, and/or electromechanically shuttered apertures. In one multi-aperture

apertures including a center aperture positioned along a center optical axis of a lens and any associated imaging hardware. The scanner 102 may also, or instead, include a stereoscopic, triscopic or other multi-camera or other configuration in which a number of cameras or optical paths are maintained in fixed or moving relation to one another to obtain two-dimensional images of an object from a number of slightly different perspectives. The scanner 102 may include suitable processing for deriving a three-dimensional point cloud from an image set or a number of image sets, or each two-dimensional image set may be transmitted to an external processor such as contained in the computer 108 described below. In other embodiments, the scanner 102 may employ structured light, laser scanning, direct ranging (for example, time of flight in a known direction), or any other technology suitable for acquiring three-dimensional data, or two-dimensional data that can be resolved into three-dimensional data.”	embodiment, the scanner 102 may include a plurality of apertures including a center aperture positioned along a center optical axis of a lens and any associated imaging hardware. The scanner 102 may also, or instead, include a stereoscopic, triscopic or other multi-camera or other configuration in which a number of cameras or optical paths are maintained in fixed or moving relation to one another to obtain two-dimensional images of an object from a number of slightly different perspectives. The scanner 102 may include suitable processing for deriving a three-dimensional point cloud from an image set or a number of image sets, or each two-dimensional image set may be transmitted to an external processor such as contained in the computer 108 described below. In other embodiments, the scanner 102 may employ structured light, laser scanning, direct ranging (e.g., time of flight in a known direction), or any other technology suitable for acquiring three-dimensional data, or two-dimensional data that can be resolved into three-dimensional data.”
McQueston, Ex-1008 at 5:47-63: “Although not shown in FIG. 1, it will be appreciated that a number of supplemental lighting systems may be usefully employed during image capture. For example, environmental illumination may be enhanced with one	McQueston Provisional, ¶ [0027]: “Although not shown in Fig. 1, it will be appreciated that a number of supplemental lighting systems may be usefully employed during image capture. For example, environmental illumination may be enhanced with one

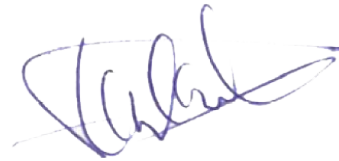
or more spotlights illuminating the subject 104 to speed image acquisition and improve depth of field (or spatial resolution depth). The scanner 102 may also, or instead, include a strobe, flash, or other light source to supplement illumination of the subject 104 during image acquisition. In other structured light systems, the illumination may be in the form of an array of laser beams that form a pattern on a surface, which pattern may be employed to recover three-dimensional data from the surface. Other systems employ a single laser beam along with directional information to gather point-by-point direction and range information. All such imaging systems may be usefully employed to acquire three-dimensional data as described herein.”	or more spotlights illuminating the subject 104 to speed image acquisition and improve depth of field (or spatial resolution depth). The scanner 102 may also, or instead, include a strobe, flash, or other light source to supplement illumination of the subject 104 during image acquisition. In other structured light systems, the illumination may be in the form of an array of laser beams that form a pattern on a surface, which pattern may be employed to recover three-dimensional data from the surface. Other systems employ a single laser beam along with directional information to gather point-by-point direction and range information. All such imaging systems may be usefully employed to acquire three-dimensional data as described herein.”
--	---

XII. CONCLUSION

388. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: August 17, 2023

By: _____



Hassan Foroosh, Ph.D.