

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent of: Meadow et al.

U.S. Patent No.: 8,078,396

Attorney Docket No.: 19473-0325IP1

Issue Date: December 13, 2011

Appl. Serial No.: 12/035,423

Filing Date: February 21, 2008

Title: METHODS FOR AND APPARATUS FOR GENERATING A CONTINUUM OF
THREE DIMENSIONAL IMAGE DATA

Mail Stop Patent Board

Patent Trial and Appeal Board

U.S. Patent and Trademark Office

P.O. Box 1450

Alexandria, VA 22313-1450

**PETITION FOR *INTER PARTES* REVIEW OF UNITED STATES
PATENT NO. 8,078,396 PURSUANT TO 35 U.S.C. §§ 311–319, 37 C.F.R. § 42**

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	MANDATORY NOTICES UNDER 37 C.F.R. § 42.8.....	2
	A. Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1).....	2
	B. Related Matters Under 37 C.F.R. § 42.8(b)(2)	2
	C. Lead And Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3).....	2
	D. Service Information	2
III.	PAYMENT OF FEES – 37 C.F.R. § 42.103	3
IV.	REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104	3
	A. Grounds for Standing Under 37 C.F.R. § 42.104(a)	3
	B. Challenge Under 37 C.F.R. § 42.104(b) and Relief Requested	3
V.	SUMMARY OF THE '396 PATENT	5
	A. Brief Description.....	5
	B. Summary of the Original Prosecution	6
	C. The '396 Patent's Priority Claim	7
	D. Claim Construction under 37 C.F.R. §§ 42.104(b)(3)	9
VI.	THERE IS A REASONABLE LIKELIHOOD THAT AT LEAST ONE CLAIM OF THE '369 PATENT IS UNPATENTABLE	11
	A. Ground 1 [Anticipation under § 102 by Fruh] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, and 8-13.....	12
	B. Ground 2 [Obviousness under § 103 by Fruh] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, and 8-13.....	18
	C. Ground 3 [Obviousness under § 103 by Fruh in view of Di Bernardo] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, 8-13, and 16	19
	D. Ground 4 [Anticipation under § 102 by Akbarzadeh] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, and 8-10	23
	E. Ground 5 [Obviousness under § 103 by Akbarzadeh] sets forth a reasonable likelihood to prevail on at least Claims 1, 2, 5, and 8-10.27	
	F. Ground 6 [Obviousness under § 103 by Akbarzadeh in view of Di Bernardo] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, 8-13, and 16.....	27

VII.	[GROUND 1 CLAIM CHART] – Anticipation of Claims 1, 2, 5, 8-13 under § 102 by Fruh	28
VIII.	[GROUND 3 CLAIM CHART] – Obviousness of Claims 1, 2, 5, 8-13, and 16 under § 103 by Fruh in view of Di Bernardo	37
IX.	[GROUND 4 CLAIM CHART] – Anticipation of Claims 1, 2, 5, and 8-10 under § 102 by Akbarzadeh	43
X.	[GROUND 6 CLAIM CHART] – Obviousness of Claims 1, 2, 5, 8-13, and 16 under § 103 by Akbarzadeh in view of Di Bernardo	50
XI.	CONCLUSION	56

EXHIBITS

GOOGLE1001	Declaration of Dr. Fuchs
GOOGLE1002	U.S. Patent No. 8,078,396 (“the ’396 patent”)
GOOGLE1003	Select Prosecution History of the ’396 patent (Serial No. 12/035,423)
GOOGLE1004	“An Automated Method for Large-Scale, Ground-Based City Model Acquisition” by Früh et al. (October 2004) (“Fruh”)
GOOGLE1005	U.S. Patent No. 6,895,126 to Di Bernardo et al. (“Di Bernardo”)
GOOGLE1006	“Towards Urban 3D Reconstruction From Video” by Akbarzadeh et al. (June 14, 2006) (“Akbarzadeh”)
GOOGLE1007	U.S. Patent 7,929,800 to Meadows et al. (Application No. 11/702,708)
GOOGLE1008	U.S. Publication No. 2008/0189031 to Meadows et al. (Application No. 11/702,707)
GOOGLE1009	Exhibit C to VRE’s Disclosure of Asserted Claims and Infringement Contentions for U.S. Patent No. 8,078,396
GOOGLE1010	U.S. Patent 7,389,181 to Meadows et al. (Application No. 11/216,465)

Google Inc. (“Petitioner”) petitions for *Inter Partes* Review (“IPR”) under 35 U.S.C. §§ 311–319 and 37 C.F.R. § 42 of claims 1, 2, 5, 8-13, and 16 (“the challenged claims”) of U.S. Patent 8,078,396 (“the ’396 patent”). Below, Petitioner demonstrates there is a reasonable likelihood of prevailing in its challenge of at least one of the challenged claims.

I. INTRODUCTION

The ’396 patent claims a method for generating a “continuum of image data.” See, e.g., GOOGLE1002 at 1:63-65. Generally speaking, the purported invention of the ’396 patent generates a continuum of 3D object representations based on point cloud data. *Id.* at 1:57-59. The claimed invention, however, was not new by the ’396 patent’s filing date or by the 2005 date to which it improperly claims priority. Rather, the ’396 patent was granted without full consideration of highly pertinent prior art. For example, the Fruh paper and the Akbarzadeh paper each discloses generating a continuum of 3D object representations based on point cloud data. Specifically, Fruh discloses generating textured 3D city models using two 2D laser scanners and a camera. GOOGLE1004 at Abstract. Similarly, Akbarzadeh discloses reconstructing a geo-registered 3D model of a scene from multiple video streams. GOOGLE1006 at Abstract.

In sum, if the Office had been aware of Fruh or Akbarzadeh (alone or in combination with Di Bernardo), the ’396 patent never would have issued. Petitioner therefore requests the Board to institute *Inter Partes* Review of the challenged claims on the grounds set for the below.

II. MANDATORY NOTICES UNDER 37 C.F.R § 42.8

A. Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1)

Google Inc. is the real party-in-interest.

B. Related Matters Under 37 C.F.R. § 42.8(b)(2)

Petitioner is not aware of any reexamination certificate for the '396 patent. The Patent Owner, a non-practicing entity under the name Visual Real Estate, Inc. ("VRE"), filed a complaint alleging infringement of the '396 patent in lawsuit against the Petitioner (*Visual Real Estate, Inc., v. Google, Inc.*, M.D. Fla., Case No. 3:14-cv-00274-TJC-JRK). The complaint was filed on March 10, 2014. Petitioner has also petitioned for *Inter Partes* Review of two other patents at issue in the above-noted litigation: U.S. Patent 7,389,181 and U.S. Patent 7,929,800.

C. Lead And Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3)

Petitioner provides the following designation of counsel.

LEAD COUNSEL	BACK-UP COUNSEL
John C. Phillips, Reg. No. 35,322 3200 RBC Plaza, 60 South Sixth Street Minneapolis, MN 55402 Tel: 858-678-4304 / Fax 858-678-5099	Michael T. Hawkins, Reg. No. 57,867 3200 RBC Plaza, 60 South Sixth Street Minneapolis, MN 55402 Tel: 612-337-2569 / Fax 612-288-9696

D. Service Information

Please address all correspondence and service to the address of both counsel listed above. Petitioner also consents to electronic service by email at IPR19473-0325IP1@fr.com (referencing No. 19473-0325IP1 and cc'ing phillips@fr.com and hawkins@fr.com).

III. PAYMENT OF FEES – 37 C.F.R. § 42.103

Petitioner authorizes the Patent and Trademark Office to charge Deposit Account No. 06-1050 for the petition fee set in 37 C.F.R. § 42.15(a) and for any other required fees.

IV. REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104

A. Grounds for Standing Under 37 C.F.R. § 42.104(a)

Petitioner certifies that the '396 patent is available for IPR and that Petitioner is not barred or estopped from requesting IPR.

B. Challenge Under 37 C.F.R. § 42.104(b) and Relief Requested

Petitioner requests IPR of claims 1, 2, 5, 8-13, and 16 of the '396 patent on the grounds listed in the table below. In support, this Petition includes claim charts for each of these grounds and a supporting evidentiary declaration of Dr. Henry Fuchs (GOOGLE1001).

Ground	Claims	Basis for Rejection
Ground 1	1, 2, 5, and 8-13	Anticipated under § 102 by Fruh (GOOGLE1004)
Ground 2	1, 2, 5, and 8-13	Obvious under § 103 by Fruh
Ground 3	1, 2, 5, 8, 9, 11-13, and 16	Obvious under § 103 by Fruh in view of U.S. Patent No. 6,895,126 to Di Bernardo (GOOGLE1005)
Ground 4	1, 2, 5, and 8-10	Anticipated under § 102 by Akbarzadeh (GOOGLE1006)
Ground 5	1, 2, 5, and 8-10	Obvious under § 103 by Akbarzadeh

Ground 6	1, 2, 5, 8-13, and 16	Obvious under § 103 by Akbarzadeh in view of Di Bernardo
-----------------	----------------------------------	--

The '396 patent is a continuation-in-part of App. No. 11/216,465 (GOOGLE1010) filed August 31, 2005, a continuation-in-part of App. No. 11/702,707 (GOOGLE1008) filed February 6, 2007, and a continuation-in-part of App. No. 11/702,708 (GOOGLE1007) filed February 6, 2007. As explained below, however, the '465 application, the '707 application, and the '708 application have different specifications and do not provide sufficient support for all of the elements of claim 1 of the '396 patent. Accordingly, at most, claim 1 and its dependent claims are entitled only to the benefit of February 21, 2008—the filing date of the '396 patent.

Regardless of the priority date of the '396 patent, both Fruh and Di Bernardo (Grounds 1-3) are prior art under 35 U.S.C. § 102(a) or 102(b). Namely, Fruh was published on October 1, 2004 and thus is prior art to the '396 patent under at least 35 U.S.C. § 102(a) (and is prior art under § 102(b) when the February 21, 2008, priority date is properly applied). Di Bernardo was filed on January 11, 2001, and issued on May 17, 2005, and thus is prior art to the '396 patent under at least 35 U.S.C. § 102(a) (and is prior art under § 102(b) when the February 21, 2008, priority date is properly applied). Akbarzadeh (Grounds 4-7) was published June 14, 2006, and thus is prior art to the '396 patent under 35 U.S.C. § 102(a) when the February 21, 2008, priority date is properly applied. None of these references were considered by the Examiner during prosecution of the '396 patent.

V. SUMMARY OF THE '396 PATENT

A. Brief Description

The '396 patent claims a method for creating a continuum of image data by receiving two or more 2D image data sets of a subject captured from disparate points on a continuum, generating a composite image of the subject from aligned portions of two or more of the 2D image data sets, receiving data descriptive of a location of features in the image data sets, tracking the position of at least one feature in two or more of the images, generating a point cloud array for the feature, converting the point cloud array into a polygon based model, and associating a location for the polygon based model relative to the portions of the image data sets and based on the location of the features in the image data sets. See GOOGLE1002 at 2:32-54.

In an embodiment described in the '396 patent, image data is captured from disparate points along a continuum. *Id.* at 3:14-15. The continuum includes the path of a vehicle carrying a digital camera. *Id.* at 3:15-18. The image data sets capture the subject matter such that they overlap each other, which allows for features present in one image data set to be present in a second image data set. *Id.* at 5:23-29. The location of a feature is registered and the relative position of the feature in each image data set is tracked. *Id.* at 5:43-45.

A composite image is created by aligning portions of data from more than one image data set. *Id.* at 7:5-10. In particular, vertical slices of image data are aligned in a horizontal

plane to form the composite image. *Id.* at 7:50-57.

Point cloud arrays are generated which represent various features (e.g., a house, a car, or a tree) in the image data sets. *Id.* at 5:50-52. Point cloud arrays are converted to 3D polygon based models of the features present in the image data sets. *Id.* at 6:50-52. The image data derived from image data sets corresponding to the 3D models can be “sprayed over” the 3D polygon models. *Id.* at 6:60-63.

Positional data of the camera capturing image data sets is recorded for each disparate point along the continuum from which an image data set is captured. *Id.* at 3:22-25. Positional data can include Cartesian coordinates or geospatial data such as latitude and longitude coordinates. *Id.* at 3:26-29, 3:46-48. Positional data can be related to each polygon model and used to position the polygon model in relation to other image data and polygon models. *Id.* at 6:63-65.

B. Summary of the Original Prosecution

The '396 patent issued from a patent application filed on February 21, 2008. Claim 1 was allowed in July 2011 after a brief examination involving prior art references that did not provide an accurate picture of the state of the art prior to August 31, 2005.

The first office action provided no prior art citation against any element of any claim, and instead rejected claim 1 under 35 U.S.C. § 112 only. GOOGLE1003 at Non-final Rejection dated September 22, 2010, p. 2. Dependent claims 2-25 were rejected for incorporating the deficiencies of claim 1. *Id.* In response, the Applicant amended claim 1 by deleting

the phrase “three dimensional” from the preamble of the claim to address the § 112 rejection. *Id.* at Amendment dated May 18, 2011. In addition, the Applicant amended claim 2 to include “wherein the polygon based model comprises at least one three dimensional polygon based model” and explained “it is claimed that a polygon based model may be generated by conversion of a point cloud array to a polygon based model, and that the polygon based model may be three dimensional.” *Id.* The sole independent claim was never subject to a prior art rejection, and the Applicant never commented on any cited prior art reference in comparison to any specific claim element during prosecution.

The Examiner issued a notice of allowance on July 13, 2011, and the '396 patent subsequently issued on November 22, 2011. A certificate of correction was issued on October 30, 2012, inserting the priority claim to App. No. 11/216,465, App. No. 11/702,707, and App. No. 11/702,708. The Applicant filed another request for certificate of correction on July 31, 2013, clarifying the “continuation-in-part application” relationship to the priority applications. *Id.* at Request for Certificate of Correction dated July 31, 2013.

As described below, more pertinent prior art never considered by the Examiner expressly disclosed the features of claims 1, 2, 5, 8-13, and 16 of the '396 patent. As such, a more meaningful review of the '396 patent is necessary and just here.

C. The '396 Patent's Priority Claim

As shown in the Patent Owner's Request for a Certificate of Correction, the '396 patent is a continuation-in-part of each of the '465 application (GOOGLE1010) filed August 31,

2005, the '707 application (GOOGLE1008) filed February 6, 2007, and the '708 application (GOOGLE1007) also filed February 6, 2007:

The present application claims priority to pending patent application S/N 11/216,465, filed, Aug 31, 2005 and entitled, "Apparatus and Method for Producing Video Drive-By Data Corresponding to a Geographic Location," as a Continuation in Part Patent Application, the contents of which are relied upon and incorporated by reference, and also pending application 11/ 702,708 filed February 6, 2007 and entitled Methods and Apparatus for Generating a Continuum of Image Data, as a Continuation in Part Patent Application, the contents of which are each relied upon and incorporated by reference; and also pending application 11/ 702,707 filed February 6, 2007 and entitled "Methods and Apparatus for Presenting a Continuum of Image Data", as a Continuation in Part Patent Application; the contents of which are each relied upon and incorporated by reference.

Id. at Request for Certificate of Correction dated July 31, 2013. For a patent claim to receive the benefit of the filing date of a prior application, the prior application must satisfy the written description and enablement requirements of 35 U.S.C. § 112, ¶ 1, with respect to the patented claim. See 35 U.S.C. § 120. The prior applications here (the '465 application, the '707 application, and the '708 application) do not provide written description support for all of the elements of independent claim 1 of the '396 patent. Accordingly, at most, claim 1 and its dependent claims are entitled only to the filing date of February 21, 2008—the filing date of the '396 patent.

Specifically, claim 1 of the '396 patent recites "generating a point cloud array for the at least one of the plurality of features" and "converting the point cloud array to a polygon based model." See *a/so* GOOGLE1002 at 2:41-47. In contrast, the earlier-filed '465 and '708 applications do not contain a written description of generating point clouds or pro-

cessing the point clouds to generate polygon based models. While the '707 application describes polygons overlaid on images, it also does not contain a written description of generating point clouds or processing the point clouds to generate polygon based models. See GOOGLE1008 at ¶ [0026], [0028], [0033], [0035], [0038]-[0041], [0044], [0046], [0060]; see also GOOGLE1001 at ¶ 17.

For at least these reasons, claim 1 and its dependent claims are not entitled to the filing date of the '465 application, the '707 application, or the '708 application, but instead are entitled to the filing date of February 21, 2008. GOOGLE1001 at ¶ 17.

D. Claim Construction under 37 C.F.R. §§ 42.104(b)(3)

Claims are to be given their “broadest reasonable construction in light of the specification.” 37 C.F.R. § 42.100(b). The constructions are intended to aid in this proceeding, and should not be understood as waiving any arguments concerning indefiniteness or claim breadth that may be raised in any litigation, which requires different construction standards.¹

“a continuum of image data” – two or more images captured along a path.

GOOGLE1001 at ¶ 19. This interpretation is consistent with the broadest reasonable interpretation of these terms and with the specification of the '396 patent. See GOOGLE1002 at 3:14-18 (“According to the present invention, image data is captured from disparate points

¹ The Patent Owner’s preliminary infringement contentions are attached as Exhibit GOOGLE1009 for the Board’s reference.

along a continuum. Referring now to FIG. 1, in some embodiments the continuum 100 includes the path of a vehicle carrying a digital camera, or other image data capturing device.”); *see also* GOOGLE1007 at 2:5-11.

“point cloud array” – a set of points corresponding to locations of features.

GOOGLE1001 at ¶ 20. This interpretation is consistent with the broadest reasonable interpretation of these terms and with the specification of the ’396 patent. See GOOGLE1002 at 2:41-44 (“In general, point clouds are generated utilizing the location of feature points present in multiple image sets in combination with the position and orientation of a camera capturing the image data set.”).

“polygon based model” - model representing 3D structure of objects using polygons. GOOGLE1001 at ¶ 21. This interpretation is consistent with the broadest reasonable interpretation of these terms and with the specification of the ’396 patent. See GOOGLE1002 at 2:44-46 (“The point clouds can in turn be processed to generate three-dimensional polygon based models representative of objects”); 6:50-53 (“point cloud arrays can be converted to three-dimensional polygon based models of artifacts present in the image data sets. The three-dimensional polygon based models are based upon physical attributes.”).

“spraying image data onto the polygon based model” – mapping image data on to the polygon based model. GOOGLE1001 at ¶ 22. This interpretation is consistent with the broadest reasonable interpretation of these terms and with the specification of the ’707

application to which the '396 patent claims priority. See GOOGLE2012 at ¶ [0060]; see also GOOGLE1001 at ¶ 22.

“associate the composite image with a particular portion of the subject” – associate the composite image with a particular parcel of a geographic area. GOOGLE1001 at ¶ 23. This interpretation is consistent with the broadest reasonable interpretation of these terms and with the specification of the '396 patent claims priority. See GOOGLE1002 at 1:33-34 (“Photographs . . . are easily associated with a particular real estate parcel”), 1:45-50 (“correlation of particular portions of a motion picture with a particular real estate parcel” “to locate and view a particular real estate parcel in the context of its neighborhood setting”), 7:42-44 (“correlates a ribbon of geographic image data with geospatial designations to facilitate identification of a particular geographic area”).

VI. THERE IS A REASONABLE LIKELIHOOD THAT AT LEAST ONE CLAIM OF THE '369 PATENT IS UNPATENTABLE

As detailed below, claims 1, 2, 5, 8-13, and 16 of the '396 patent are anticipated or rendered obvious by at least one reference or combination of references. The seven Grounds for rejection of claims 1, 2, 5, 8-13, and 16 are not cumulative or redundant, but instead rely upon different primary references that individually assert unique benefits to the user and additionally address the dependent claims in different ways (including different statutory grounds (§ 102 and § 103)). See GOOGLE1001 at ¶¶ 24-66.

For example, the Fruh paper, on which Grounds 1-3 are based, describes a method for ground-based acquisition of large-scale 3D city models using a camera and two 2D laser

scanners. GOOGLE1004 at Abstract. By contrast, the Akbarzadeh paper, on which Grounds 4-6 are based, describes a method for 3D reconstruction of urban scenes from multiple video streams. Akbarzadeh at Abstract. The various Grounds of rejection differ in the way in which they address the challenged claims of the '396 patent in numerous other ways, as detailed below.

A. Ground 1 [Anticipation under § 102 by Fruh] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, and 8-13

Referring to Ground 1 (charted below), Fruh discloses all elements of, and thus anticipates, claims 1, 2, 5, and 8-13. Similar to the preferred system of the '396 patent, Fruh discloses using a digital camera and ground-based laser scanners to continuously obtain geometry and texture (i.e., image) data and then generate 3D city models. GOOGLE1004 at Abstract and p. 6. This is depicted in Figure 1 of Fruh. See GOOGLE1001 at ¶ 26.

The sideways looking digital camera captures multiple images as the vehicle travels on a path. GOOGLE1004 at p. 6 and 7. This is depicted in Figure 5 of Fruh. See GOOGLE1001 at ¶ 26. Accordingly, regarding elements [1.P] and [1.1], Fruh discloses creating a continuum of image data and receiving multiple 2D image data sets of a subject, where each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features. GOOGLE1001 at ¶ 26.

Fruh discloses acquiring 3D geometry and 2D texture image data using laser scanners and a digital camera. GOOGLE1004 at p. 6. Fruh discloses that “due to its regular structure, the processed point cloud can easily be transformed into a triangular mesh by

connecting adjacent vertices.” GOOGLE1004 at p. 16. Fruh further discloses that the camera and laser scanners are synchronized and calibrated in the 3D coordinate system so that for every arbitrary 3D point, the corresponding image coordinate can be computed and the texture can be mapped onto each mesh triangle to result in textured façade models of the buildings that the acquisition vehicle drove by. *Id.* A POSITA would have recognized that texturing a 3D model includes extracting a single vertical strip of a camera image corresponding to a regular structure of the point cloud and aligning vertical strips in adjacent images corresponding to adjacent vertices. GOOGLE1001 at ¶ 27. Accordingly, Fruh discloses generating a composite image of the subject from portions of two or more of the multiple 2D image data sets, wherein the portions are aligned in a dimension consistent with the continuum (as recited in element [1.2] below), and wherein the composite image comprises a single vertical slice of each of the two or more images (as recited in claim 11 below). GOOGLE1001 at ¶ 27.

Regarding element [1.3], Fruh discloses acquiring successive ground-based horizontal scans and matching the scans to an aerial photograph or a global map. GOOGLE1004 at Abstract. Fruh also discloses that GPS may be used for estimating position of a scanner. *Id.* at p. 6. Thus, Fruh discloses receiving data descriptive of a location of the plurality of features. GOOGLE1001 at ¶ 28.

Fruh discloses that the camera and laser scanners are synchronized and calibrated in the 3D coordinate system, and an image coordinate is computed for every 3D point.

GOOGLE1004 at p. 16. Fruh also discloses matching successive horizontal scans with each other to determine an estimate of the vehicle's motion and to determine a pose of successive scans and camera images. *Id.* at Abstract and p. 6. A POSITA would have understood that doing so would include tracking the position of objects in the successive horizontal scans and that tracking the position of objects in the successive horizontal scans results in tracking the position of the same objects in successive 2D images captured by the camera corresponding to the successive horizontal scans because the camera and laser scanners are synchronized and calibrated in the 3D coordinate system. GOOGLE1001 at ¶ 29. Therefore, regarding element [1.4], Fruh discloses tracking the position of at least one of the plurality of features in two or more of the 2D images. *Id.*

Regarding element [1.5], Fruh discloses generating a point cloud array for the at least one of the plurality of features. GOOGLE1001 at ¶ 30. For example, Fruh discloses generating a 3D point cloud for a city block from vertical laser scans. GOOGLE1004 at p. 16 and 19. The structured point cloud for a city block is shown in Figure 19. See GOOGLE1001 at ¶ 30.

Fruh discloses transforming the point cloud into a triangular mesh by connecting adjacent vertices, as shown in Figure 21. *Id.* at p. 16 and 20; see GOOGLE1001 at ¶ 31. Thus, regarding element [1.6], Fruh discloses converting the point cloud array to a polygon based model. GOOGLE1001 at ¶ 31.

Fruh further discloses determining the pose of successive laser scans and camera

images in a global coordinate system using an aerial photograph or a global map, calculating 3D coordinates of vertical scan points in the global coordinate system, and computing the corresponding image coordinate for every arbitrary 3D point to create textured façade models of the buildings that the acquisition vehicle drove by. GOOGLE1004 at Abstract and p. 6, 16, and 20. Fruh further discloses that the façade models have been brought into perfect registration with either an aerial photo or a global map. *Id.* A POSITA would have understood that in Fruh, the locations of features in the façade models is associated with locations in the 2D camera images, and for a façade model to be constructed properly, the images and the associated horizontal laser scans would have to be registered to the aerial photo or the global map based on the locations of the features. GOOGLE1001 at ¶ 32. Thus, regarding element [1.7], Fruh discloses associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features. GOOGLE1001 at ¶ 32.

Fruh discloses acquiring texture data using a digital camera. GOOGLE1004 at p. 6. Fruh's system computes the corresponding image coordinate for every arbitrary 3D point and maps the texture onto each mesh triangle to create textured models of the buildings that the acquisition vehicle drove by. GOOGLE1004 at p. 6 and 16. Thus, regarding claim 2 and element [5.2] in claim 5, Fruh discloses "spraying" image data onto the polygon based model, wherein the polygon based model comprises at least one 3D polygon based model

and wherein the image data sprayed onto the polygon based model is based upon 2D image data sets and the location of the polygon based model and “spraying” image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon 2D image data sets and the location of each respective polygon based model.” GOOGLE1001 at ¶ 33.

Regarding element [5.1] of claim 5, Fruh discloses continuously acquiring 3D geometry and texture data of an entire city. GOOGLE1004 at p. 6. Textured façade models of the buildings that the acquisition vehicle drove by are created from the acquired data. *Id.* at p. 16. Façade models for a geographic area, e.g., the downtown Berkeley, California area, are aligned with each other to create large-scale 3D city models. *Id.* at Abstract and p. 20. Figure 22 of Fruh shows the alignment of two independently acquired façade models. GOOGLE1001 at ¶ 34. Thus, Fruh discloses repeating steps a) through f) multiple times each from disparate points along the continuum. GOOGLE1001 at ¶ 34.

Regarding claims 8 and 9, Fruh discloses using a Cartesian world coordinate system for the vertical scan points and that GPS can be and is often used for position estimates in outdoor environments. GOOGLE1004 at p. 6, 8, and 16. Thus, Fruh discloses that the data descriptive of a location comprises a Cartesian coordinate and the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate. GOOGLE1001 at ¶ 35, 36.

Regarding claim 10, Fruh discloses associating metadata with a polygon based

model. Fruh discloses that the textured façade models have been brought into perfect registration with either an aerial photo or a global map. GOOGLE1004 at p. 16. A POSITA would have understood that registering the textured façade models with an aerial photo or a global map provides the claimed associating metadata with the polygon based model. GOOGLE1001 at ¶ 37.

Regarding claim 12, Fruh discloses recording positional data descriptive of a respective location of each of the multiple 2D image data sets of a subject (element [12.1]) and associating the composite image with a particular portion of the subject based upon the positional data (element [12.2]). GOOGLE1001 at ¶ 38. For example, Fruh discloses acquiring 3D geometry and 2D texture image data using laser scanners and a digital camera. GOOGLE1004 at p. 6. Fruh discloses matching horizontal laser scans to an aerial photograph or global map to determine the final global pose of the acquisition vehicle. *Id.* at Abstract. Fruh further discloses that the camera and laser scanners are synchronized and calibrated in the 3D coordinate system so that for every arbitrary 3D point, the corresponding image coordinate can be computed and the texture can be mapped onto each mesh triangle to result in textured façade models of the buildings that the acquisition vehicle drove by. *Id.* at p. 16. A POSITA would have recognized that matching horizontal laser scans to an aerial photograph or a global map would result in recording of positional data descriptive of a respective location of each of the multiple 2D image data sets of a subject and that the com-

posite texture image is associated with a particular portion of a subject based on the matching of the horizontal laser scans with the aerial photo or the global map. GOOGLE1001 at ¶ 38.

Regarding claim 13, Fruh discloses the subject comprises a geographic area and the positional data comprises a geospatial designation. GOOGLE1001 at ¶ 38. For example, Fruh discloses acquiring large-scale 3D city models where localization of the acquisition vehicle can be based on GPS. GOOGLE1004 at Abstract and p. 6.

For at least these reasons, and the additional explanations described in the chart below and the accompanying declaration by Dr. Fuchs, there is a reasonable likelihood that claims 1, 2, 5, 8, 9, and 11-13 are anticipated by Fruh. See GOOGLE1001 at ¶ 24-39.

B. Ground 2 [Obviousness under § 103 by Fruh] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, and 8-13

Referring to Ground 2, as demonstrated above in Section VI.A and below in Section VII, Fruh discloses each element of claims 1, 2, 5, 8, and 9. However, to the extent that any limitations are considered not to be anticipated by Fruh alone, such limitations are also rendered obvious by Fruh, as a POSITA would have found such limitations obvious in Fruh's disclosure in light of the ordinary knowledge of skilled artisan and the state of the art known before 2008. GOOGLE1001 at ¶ 39. Additionally, to the extent that any variances in claim scope are not necessarily shown by the support included in the below claim charts, such variances would have been obvious to a POSITA based on the level of ordinary skill in the pertinent art as it existed at the purported time of the invention for the same reasons cited in

the accompanying expert declaration of Dr. Fuchs. GOOGLE1001 at ¶ 24-66.

C. Ground 3 [Obviousness under § 103 by Fruh in view of Di Bernardo] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, 8-13, and 16

Referring to Ground 3 (charted below), to the extent that Fruh is interpreted as not expressly disclosing any of elements [1.P] or [1.2] of claim 1 (or any of dependent claims 11-13 and 16), these elements were commonly practiced in similar image processing systems well before the priority date of the '396 patent. GOOGLE1001 at ¶ 40-50. For example, regarding element [1.P], Di Bernardo also discloses a method for creating a continuum of image data. GOOGLE1001 at ¶ 42. In particular, Di Bernardo discloses synthesizing images of a geographic location to generate composite images of the location. GOOGLE1005 at 2:16-21. The geographic location may be a particular street in a geographic area with the composite images providing a view of the objects on each side of the street. *Id.*

To form a “composite image” as recited in element [1.2] of claim 1, Di Bernardo teaches the conventional practice at the time in which image data from selected image frames is extracted and combined. *Id.* at 5:67-6:15. The image data extracted from each image frame is an optical column that consists of a vertical set of pixels. *Id.* The composite image is created on a column-by-column basis by extracting the corresponding optical column from each image frame. *Id.* The columns of pixels are stacked side by side to generate a single image forming the composite image. *Id.* at 9:34-43. For example, in Figure 2 of Di Bernardo (see GOOGLE1001 at ¶ 43), portions of images captured at time t_1 , t_i , and t_N

are combined into the wider image 40. Accordingly, Di Bernardo discloses generating a composite image of the subject from portions of two or more of the multiple 2D image data sets, wherein the portions are aligned in a dimension consistent with the continuum (as recited in element [1.2] below), and wherein the composite image comprises a single vertical slice of each of the two or more images (as recited in claim 11 below). GOOGLE1001 at ¶ 43.

Regarding claims 12 and 13, Di Bernardo teaches that an image recording device moves along a path recording images of objects along the path while a GPS receiver or inertial navigation system provides position information of the image recording device as the images are acquired. GOOGLE1005 at 2:22-28. The GPS receiver computes latitude and longitude coordinates indicating the position of the recording device during the recording of a particular image frame. *Id.* at 5:17-23. The image and position information is provided to the computer to associate each image with the position information. *Id.* at 2:22-28. Di Bernardo also teaches tracking the GPS position of landmarks in captured images. *Id.* at 6:60-65, 7:6-12. Thus, Di Bernardo discloses recording positional data descriptive of a respective location of each of the multiple 2D image data sets of a subject (as recited in element [12.1] below). GOOGLE1001 at ¶ 44.

Using the position information, Di Bernardo's computer segments the trajectory taken by the recording device into street segments and associates each street segment with identifying information about the street segment, such as a street name and number range.

GOOGLE1005 at 6:38-49. The computer generates a series of composite images depicting a portion of each street segment and stores each composite image in an image database along with the identifying information of the street segment with which it is associated. *Id.* Thus, Di Bernardo discloses associating the composite image with a particular portion of the subject based upon the positional data (as recited in element [12.2] below), wherein the subject comprises a geographic area and the positional data comprises a geospatial designation (as recited in claim 13 below). GOOGLE1001 at ¶ 45.

As Figure 16 (see GOOGLE1001 at ¶ 46) shows, Di Bernardo discloses that icons are displayed on top of objects in the composite image. GOOGLE1005 at 12:54-58. When an icon is passed over or an object is selected by the user's cursor, these icons display or link to additional information associated with the selected object in the composite image based on the object's location. GOOGLE1005 at 10:22-33, 12:61-67. Di Bernardo discloses overlaying the metadata or link over the composite image in that information associated with a particular location "is preferably displayed each time the user terminal's cursor or pointing device is passed above the icon." GOOGLE1005 at 12:62-64. Accordingly, regarding claim 16, Di Bernardo discloses overlaying metadata on the composite image descriptive of the composite image. GOOGLE1001 at ¶ 46, 47. To the extent that such limitations are considered not to be expressly disclosed by Di Bernardo, such limitations are also rendered obvious by Di Bernardo, as a POSITA would have found the particular placement of metadata in relation to the composite image to be an obvious matter of design choice.

GOOGLE1001 at ¶ 47.

There are a number of reasons that would have prompted a POSITA to combine the teachings of Di Bernardo with Fruh's system for creating 3D models of a geographic area. First, a POSITA would have been prompted to modify Fruh's system to incorporate Di Bernardo's teachings regarding overlaying metadata so that the resulting system would provide context to the objects being depicted (as suggested by Di Bernardo). GOOGLE1001 at ¶ 48; GOOGLE1005 at Abstract, 2:6-8.

Second, a POSITA would have been prompted to modify Fruh's system with the above-mentioned teachings of Di Bernardo because doing so would be merely the use of known techniques (e.g., generating a composite image of a geographic area, recording positional data descriptive of the geographic area, or both) to improve similar devices (e.g., Fruh's system for generating 3D models from 2D scans and images) in the same way. GOOGLE1001 at ¶ 49; *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 417 (2007). Here, both Fruh and Di Bernardo disclose similar systems for generating visualizations from multiple 2D images. Not only would a POSITA have recognized that applying Di Bernardo's suggestions to Fruh's system would have led to predictable results, but the POSITA would have understood that such a modification would not significantly alter or hinder the functions performed by the system of Fruh, but would instead provide the additional benefits conferred by practicing the functions described in Di Bernardo). GOOGLE1001 at ¶ 49.

For at least these reasons and the additional explanations described in the chart below, there is a reasonable likelihood that claims 1, 2, 5, 8-13, and 16 of the '396 patent are rendered obvious by Fruh in view of Di Bernardo. Moreover, to the extent the Patent Owner argues that any element of claims 1, 2, 5, 8-13, and 16 is not expressly disclosed by the combination of Fruh and Di Bernardo, such element would be rendered obvious in view of the disclosures of Fruh and Di Bernardo and in light of the knowledge of a POSITA. GOOGLE1001 at ¶ 50.

D. Ground 4 [Anticipation under § 102 by Akbarzadeh] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, and 8-10

Referring to Ground 4 (charted below), Akbarzadeh discloses all elements of, and thus anticipates, claims 1, 2, 5, and 8-10. Similar to the preferred system of the '396 patent, Akbarzadeh discloses a method for reconstructing 3D models of urban scenes from multiple video streams captured by an on-vehicle acquisition system. GOOGLE1006 at Abstract and p.2 and 5. Figure 2 of Akbarzadeh (see *also* GOOGLE1001 at ¶ 53) shows the dense reconstruction of a city block. GOOGLE1006 at FIG.2. Thus, Akbarzadeh's method for creating a continuum of image data (as recited in element [1.P] below) is like the preferred method of the '396 patent in that the 3D models created from the multiple video streams captured by an on-vehicle acquisition system provides the claimed continuum of image data. GOOGLE1001 at ¶ 53.

Akbarzadeh discloses collecting multiple video streams using a multi-camera system

while driving through streets. GOOGLE1006 at Abstract and p. 1 and 3. Accordingly, regarding element [1.1], Akbarzadeh discloses receiving multiple 2D image data sets of a subject, where each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features. GOOGLE1001 at ¶ 54.

Akbarzadeh also discloses generating ground-level 3D models of urban scenes from video data captured by a multi-camera system. GOOGLE1006 at p. 1. To generate the 3D models, Akbarzadeh generates depth maps for each video frame by performing stereo matching on the input images and fuses the depth maps to enforce consistency between them. *Id.* at p. 3. Accordingly, regarding element [1.2], Fruh discloses generating a composite image of the subject from portions of two or more of the multiple 2D image data sets, wherein the portions are aligned in a dimension consistent with the continuum. GOOGLE1001 at ¶ 55.

Akbarzadeh discloses collecting GPS and INS measurements to determine the coordinates of features. GOOGLE1006 at 4. Akbarzadeh also discloses a “2D feature tracker” (see portion of Figure 3)(see GOOGLE1001 at ¶ 57) that determines correspondence between 3D feature points in consecutive image frames. *Id.* at 3-4. See GOOGLE1001 at ¶ 57. Thus, Akbarzadeh discloses receiving data descriptive of a location of the plurality of features (as recited in element [1.3] below), tracking the position of at least one of the plurality of features in two or more of the 2D images (as recited in element [1.4] below), wherein

the data descriptive of a location comprises a Cartesian coordinate (as recited in claim 8 below), and wherein the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate (as recited in claim 9 below). GOOGLE1001 at ¶ 56, 57.

Regarding elements [1.5] and [1.6], Akbarzadeh discloses generating a point cloud array for the at least one of the plurality of features and converting the point cloud to a polygon based model. GOOGLE1001 at ¶ 58. For example, Akbarzadeh discloses determining 2D-2D point correspondences in consecutive video frames by tracking 1000 feature points in a 1024 x 768 image. GOOGLE1006 at p. 3. Akbarzadeh also discloses producing a depth map for each video frame that provides a depth estimate for each point in a frame. *Id.* at p. 4 and 5. A triangulated, texture-mapped, 3D model of a scene is generated from the depth maps. *Id.* at 5.

Akbarzadeh discloses that the 3D reconstruction of urban scenes are geo-registered. GOOGLE1006 at Abstract. To determine geo-registered coordinates of the features in the 3D model, Akbarzadeh uses geo-registered poses of the cameras estimated from the video and INS/GPS data. *Id.* at p. 4. The '396 patent explicitly describes metadata of images as including geospatial data from a GPS receiver unit. GOOGLE1002 at 5:17-22. Thus, Akbarzadeh discloses associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features (as recited in element [1.7] below) and associating metadata with a polygon based model (as recited in claim 10 below). GOOGLE1001 at ¶ 59.

Akbarzadeh discloses generating a 3D model of a scene in geo-registered coordinates and providing texture for the surfaces of the 3D model using the video data.

GOOGLE1006 at p. 2 and 5. Thus, regarding claim 2 and element [5.2] in claim 5, Akbarzadeh discloses “spraying” image data onto the polygon based model, wherein the polygon based model comprises at least one 3D polygon based model and wherein the image data sprayed onto the polygon based model is based upon 2D image data sets and the location of the polygon based model and “spraying” image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon 2D image data sets and the location of each respective polygon based model.” GOOGLE1001 at ¶ 60.

Regarding element [5.1] of claim 5, Akbarzadeh discloses capturing and processing large amounts of video data and reconstructing 3D models of urban scenes from several hours of video data. GOOGLE1006 at p. 1. Akbarzadeh discloses a fully automatic system that operates in continuous mode to post-process large video datasets. *Id.* at p. 3. In particular, Akbarzadeh discloses “an approach for fully automatic 3D reconstruction of urban scenes from several hours of video data captured by a multi-camera system.” *Id.* Thus, Akbarzadeh discloses repeating steps a) through f) multiple times each from disparate points along the continuum. GOOGLE1001 at ¶ 61.

For at least these reasons, and the additional explanations described in the chart below and the accompanying Fuchs declaration, there is a reasonable likelihood that claims 1,

2, 5, and 8-10 of the '396 patent are anticipated by Akbarzadeh. See GOOGLE1001 at ¶ 51-62.

E. Ground 5 [Obviousness under § 103 by Akbarzadeh] sets forth a reasonable likelihood to prevail on at least Claims 1, 2, 5, and 8-10

Referring to Ground 5, as demonstrated above in Section VI.D and below in Section IX, Akbarzadeh discloses each element of claims 1, 2, 5, and 8-10. However, to the extent that any limitations are considered not to be anticipated by Akbarzadeh alone, such limitations are also rendered obvious by Akbarzadeh, as a POSITA would have found such limitations obvious in Akbarzadeh's disclosure in light of the ordinary knowledge of skilled artisan and the state of the art known before 2008. GOOGLE1001 at ¶ 61. Additionally, to the extent that any variances in claim scope are not necessarily shown by the support included in the below claim charts, such variances would have been obvious to a POSITA based on the level of ordinary skill in the pertinent art as it existed at the purported time of the invention for the same reasons cited in the accompanying expert declaration of Dr. Fuchs. GOOGLE1001 at ¶ 24-66.

F. Ground 6 [Obviousness under § 103 by Akbarzadeh in view of Di Bernardo] sets forth a reasonable likelihood to prevail on at least one of Claims 1, 2, 5, 8-13, and 16

Referring to Ground 6 (charted below), Akbarzadeh in combination with Di Bernardo renders obvious claims 1, 2, 5, 8-13, and 16. To the extent that Akbarzadeh is considered not to expressly disclose any of elements [1.P.] or [1.2] of claim 1 (or any of dependent

claims 11, 12, 13, and 16), these elements were commonly practiced in similar image processing systems well before the priority date of the '396 patent. GOOGLE1001 at ¶ 63-66. For example, Di Bernardo discloses the features of elements [1.P.] and [1.2] of claim 1 and dependent claims 11, 12, 13, and 16, as described above in Section VI.C.

Reasons similar to those described in Section VI.C would have prompted a POSITA to combine the teachings of Di Bernardo with Akbarzadeh's system for creating 3D images of a scene. GOOGLE1001 at ¶ 65.

For at least these reasons and the additional explanations described in the chart below, there is a reasonable likelihood that claims 1, 2, 5, 8, 9, 11-13, and 16 of the '396 patent are rendered obvious by Akbarzadeh in view of Di Bernardo. Moreover, to the extent the Patent Owner argues that any element of claims 1, 2, 5, 8, 9, 11-13, and 16 is not expressly disclosed by the combination of Akbarzadeh and Di Bernardo, such element would be rendered obvious in view of the disclosures of Akbarzadeh and Di Bernardo and in light of the knowledge of a POSITA. GOOGLE1001 at ¶ 66.

VII. [GROUND 1 CLAIM CHART] – Anticipation of Claims 1, 2, 5, 8-13 under § 102 by Fruh

U.S. Pat. 8,078,396	Fruh
[1.P] A method for creating a continuum of image data, the method comprising:	Fruh discloses a method for creating a continuum of image data. GOOGLE1001 at ¶ 26. For example, Fruh discloses “an automated method for fast, ground-based acquisition of large-scale 3D city models . Our experimental set up consists of a truck equipped with one camera and two fast, inexpensive 2D laser scanners , being driven on city streets under normal traffic conditions.” GOOGLE1004 at Abstract. “Specifically, the final global pose is obtained by

	<p>utilizing an aerial photograph or a Digital Surface Model as a global map, to which the ground-based horizontal laser scans are matched. A fairly accurate, textured 3D cof the downtown Berkeley area has been acquired in a matter of minutes.” <i>Id.</i> “In this paper, we propose “drive-by scanning” as a method that is capable of rapidly acquiring 3D geometry and texture data of an entire city at the ground level by using a configuration of two fast, inexpensive 2D laser scanners and a digital camera. ... This approach has the advantage that data can be acquired continuously, rather than in a stop-and-go fashion.” <i>Id.</i> at p. 6.</p>
<p>[1.1] a) receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features;</p>	<p>Fruh discloses receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features. GOOGLE1001 at ¶ 26. For example, Fruh discloses “Our experimental set up consists of a truck equipped with one camera and two fast, inexpensive 2D laser scanners, being driven on city streets under normal traffic conditions. One scanner is mounted vertically to capture building facades. ...” GOOGLE1004 at Abstract. “A fairly accurate, textured 3D cof the downtown Berkeley area has been acquired in a matter of minutes.” <i>Id.</i> “In this paper, we propose “drive-by scanning” as a method that is capable of rapidly acquiring 3D geometry and texture data of an entire city at the ground level by using a configuration of two fast, inexpensive 2D laser scanners and a digital camera. This data acquisition system is mounted on a truck moving at normal speed on public roads.” GOOGLE1004 at p. 6. “The data acquisition system is mounted on a truck and consists of two parts: a sensor module and a processing unit (Früh and Zakhori, 2001a). ... the sensor module consists of two SICK 2D laser scanners, a digital camera, and a heading sensor.” GOOGLE1004 at p. 7.</p>
<p>[1.2] b) generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are</p>	<p>Fruh discloses generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum. GOOGLE1001 at ¶ 27. For example, Fruh discloses “acquiring 3D geometry and texture data of an entire city at the ground level by using</p>

<p>aligned in a dimension consistent with the continuum;</p>	<p>a configuration of two fast, inexpensive 2D laser scanners and a digital camera.” GOOGLE1004 at p. 6. “[D]ue to its regular structure, the processed point cloud can easily be transformed into a triangular mesh by connecting adjacent vertices. Furthermore, camera and laser scanners are synchronized by trigger signals and are mounted in a rigid configuration on the sensor platform. We calibrate the camera before our measurements and determine the transformation between its coordinate system and the laser coordinate system. Therefore, for every arbitrary 3D point, we can compute the corresponding image coordinate and map the texture onto each mesh triangle. The result is textured façade models of the buildings that the acquisition vehicle drove by. Since these façade models have been brought into perfect registration with either aerial photo or DSM, they can eventually be merged with models derived from this same airborne data.” <i>Id.</i> at p. 16.</p>
<p>[1.3] c) receiving data descriptive of a location of the plurality of features;</p>	<p>Fruh discloses receiving data descriptive of a location of the plurality of features. GOOGLE1001 at ¶ 28. For example, Fruh discloses “determin[ing] an estimate of the vehicle’s motion, and relative motion estimates are concatenated to form an initial path. ... the final global pose is obtained by utilizing an aerial photograph or a Digital Surface Model as a global map, to which the ground-based horizontal laser scans are matched.” GOOGLE1004 at Abstract. “Zhao and Shibasaki (1999) ... localization is based on the Global Positioning System (GPS). GPS is by far the most common source of global position estimates in outdoor environments.” <i>Id.</i> at p. 6. “[I]t is necessary to determine the pose of successive laser scans and camera images in a global coordinate system with centimeter and sub-degree accuracy in order to reconstruct a consistent model.” <i>Id.</i> “The vertical scanner detects the shape of the building facades as we drive by, and therefore we refer to our method as drive-by scanning; the horizontal scanner operates in a plane parallel to the ground and is used for pose estimation as described in this paper. The camera’s line of sight is the intersection between the orthogonal scanning planes. All sensors are synchronized with each other</p>

	<p>and acquire data at prespecified times. Figure 2 shows a picture of the truck with rack and equipment.” <i>Id.</i> at p. 7. “In this section, we derive a global edge map either from an aerial photo or from a DSM, and ... use the airborne edge maps as an occupancy grid for global localization.” <i>Id.</i> at p. 10; see also pp. 10-16 (§§ 4.1-4.3, 5).</p>
<p>[1.4] d) tracking the position of at least one of the plurality of features in two or more of the two-dimensional images;</p>	<p>Fruh discloses tracking the position of at least one of the plurality of features in two or more of the two-dimensional images. GOOGLE1001 at ¶ 29. For example, Fruh discloses “Successive horizontal scans are matched with each other in order to determine an estimate of the vehicle’s motion.” GOOGLE1004 at Abstract. “More specifically, it is necessary to determine the pose of successive laser scans and camera images in a global coordinate system.” <i>Id.</i> at p. 6. “[D]ue to its regular structure, the processed point cloud can easily be transformed into a triangular mesh by connecting adjacent vertices. Furthermore, camera and laser scanners are synchronized by trigger signals and are mounted in a rigid configuration on the sensor platform. We calibrate the camera before our measurements and determine the transformation between its coordinate system and the laser coordinate system. Therefore, for every arbitrary 3D point, we can compute the corresponding image coordinate and map the texture onto each mesh triangle. The result is textured façade models of the buildings that the acquisition vehicle drove by. Since these façade models have been brought into perfect registration with either aerial photo or DSM, they can eventually be merged with models derived from this same airborne data.” <i>Id.</i> at p. 16.</p>
<p>[1.5] e) generating a point cloud array for the at least one of the plurality of features;</p>	<p>Fruh discloses generating a point cloud array for the at least one of the plurality of features. GOOGLE1001 at ¶ 30. For example, Fruh discloses “Once the pose of the vehicle and thus the laser scanners is known, the generation of a 3D point cloud is straightforward. We calculate the 3D coordinates of the vertical scan points.” GOOGLE1004 at p. 16. “Using the accurate pose information, we transform the vertical laser scans into global coordinates and obtain a structured point cloud for a city block as shown in Fig. 19, with a detailed view shown in Fig. 20.” <i>Id.</i> at p. 19; see GOOGLE1001 at ¶ 30.</p>

<p>[1.6] f) converting the point cloud array to a polygon based model; and</p>	<p>Fruh discloses converting the point cloud array to a polygon based model. GOOGLE1001 at ¶ 31. For example, Fruh discloses “ground-based acquisition of large-scale 3D city models.” GOOGLE1004 at Abstract. “[T]he processed point cloud can easily be transformed into a triangular mesh by connecting adjacent vertices.” <i>Id.</i> at p. 16. “Triangulating the structured point cloud by connecting adjacent vertices, we obtain a surface mesh as shown in Fig. 21.” <i>Id.</i> at p. 20; see GOOGLE1001 at ¶ 31.</p>
<p>[1.7] g) associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features.</p>	<p>Fruh discloses associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features. GOOGLE1001 at ¶ 32. For example, Fruh discloses “the final global pose is obtained by utilizing an aerial photograph or a Digital Surface Model as a global map, to which the ground-based horizontal laser scans are matched. ... it is necessary to determine the pose of successive laser scans and camera images in a global coordinate system with centimeter and sub-degree accuracy in order to reconstruct a consistent model.” Fruh at p. 6. “Once the pose of the vehicle and thus the laser scanners is known, the generation of a 3D point cloud is straightforward. We calculate the 3D coordinates of the vertical scan points by applying a coordinate transformation from the local to the world coordinate system. The structure of the resulting point cloud is given by scan number and angle, and therefore each vertex has defined neighbors.” <i>Id.</i> at p. 16. “[F]or every arbitrary 3D point, we can compute the corresponding image coordinate and map the texture onto each mesh triangle. ... these facade models have been brought into perfect registration with either aerial photo or DSM, they can eventually be merged with models derived from this same airborne data.” <i>Id.</i> “Using the accurate pose information, we transform the vertical laser scans into global coordinates and obtain a structured point cloud for a city block as shown in Fig. 19, with a detailed view shown in Fig. 20.” <i>Id.</i> at p. 19.</p>
<p>[2] The method of claim 1 additionally comprising the step of spraying image data</p>	<p>Fruh discloses spraying image data onto the polygon based model, wherein the polygon based model comprises at least one three dimensional polygon based model and wherein</p>

<p>onto the polygon based model, wherein the polygon based model comprises at least one three dimensional polygon based model and wherein the image data sprayed onto the polygon based model is based upon two-dimensional image data sets and the location of the polygon based model.</p>	<p>the image data sprayed onto the polygon based model is based upon two-dimensional image data sets and the location of the polygon based model. GOOGLE1001 at ¶ 33. For example, Fruh discloses “A fairly accurate, textured 3D cof the downtown Berkeley area has been acquired in a matter of minutes.” GOOGLE1004 at Abstract. “In this paper, we propose ‘drive-by scanning’ as a method that is capable of rapidly acquiring 3D geometry and texture data of an entire city at the ground level by using a configuration of two fast, inexpensive 2D laser scanners and a digital camera.” <i>Id.</i> at p. 6. “We calibrate the camera before our measurements and determine the transformation between its coordinate system and the laser coordinate system. Therefore, for every arbitrary 3D point, we can compute the corresponding image coordinate and map the texture onto each mesh triangle. The result is textured facade models of the buildings that the acquisition vehicle drove by.” <i>Id.</i> at p. 16; see <i>Id.</i> at Figure 23.</p>
<p>[5.P] The method of claim 1 additionally comprising the steps of:</p>	<p>Fruh discloses the method of claim 1. See analysis of claim 1, <i>supra</i>.</p>
<p>[5.1] repeating steps a) through f) multiple times each from disparate points along the continuum; and</p>	<p>Fruh discloses repeating steps a) through f) multiple times each from disparate points along the continuum. GOOGLE1001 at ¶ 34. For example, Fruh discloses “fast, ground-based acquisition of large-scale 3D city models. ... Successive horizontal scans are matched with each other.” GOOGLE1004 at Abstract. “In this paper, we propose ‘drive-by scanning’ as a method that is capable of rapidly acquiring 3D geometry and texture data of an entire city at the ground level ... data can be acquired continuously ... it is necessary to determine the pose of successive laser scans and camera images in a global coordinate system with centimeter and sub-degree accuracy in order to reconstruct a consistent model.” <i>Id.</i> at p. 6. “Once the pose of the vehicle and thus the laser scanners is known, the generation of a 3D point cloud is straightforward. ... Since these facade models have been brought into perfect registration with either aerial photo or DSM, they can eventually be merged with models derived from this same air-</p>

	<p>borne data.” <i>Id.</i> at p. 16. “Using the accurate pose information, we transform the vertical laser scans into global coordinates and obtain a structured point cloud for a city block as shown in Fig. 19, with a detailed view shown in Fig. 20.” <i>Id.</i> at p. 19. “Triangulating the structured point cloud by connecting adjacent vertices, we obtain a surface mesh as shown in Fig. 21. As we have discussed before, our global pose estimation aligns scans from different passes consistently with each other; thus, the resulting facade models align as well. The two facade faces shown in Fig. 22 were taken in two different passes, with several hundred meters driving distance between the two acquisitions. Nevertheless, they align with each other up to a small gap.” <i>Id.</i> at p. 20; see GOOGLE1001 at ¶ 34.</p>
[5.2] spraying image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon two-dimensional image data sets and the location of each respective polygon based model.	<p>Fruh discloses spraying image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon two-dimensional image data sets and the location of each respective polygon based model. GOOGLE1001 at ¶ 33; see claim 2, <i>supra</i> (incorporated here).</p>
[8] The method of claim 1 wherein the data descriptive of a location comprises a Cartesian coordinate.	<p>Fruh discloses the data descriptive of a location comprises a Cartesian coordinate. GOOGLE1001 at ¶ 35. For example, Fruh discloses “a Cartesian world coordinate system [x, y, z] where x, y defines the geographical Figure 3. Two scans before matching (left) and after matching (right) ground plane and z the altitude as shown in Fig. 1. We also define a truck coordinate system [u, v] which is implied by the horizontal laser scanner. In a flat environment, a 2D pose of the truck and its local coordinate system can be entirely described by the two coordinates x, y and the yaw orientation angle θ; in this case, the u-v coordinate system is assumed to be parallel to the x-y plane as shown in Fig. 1.” <i>Id.</i> at p. 8; see also pp. 10-16 (§§ 4.1-4.3, 5, 6).</p>
[9] The method of claim 8 wherein the data descriptive of a location of the plurality	<p>Fruh discloses the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate. GOOGLE1001 at ¶ 36. For example, Fruh discloses “localization is based on the Global Positioning System (GPS).</p>

of features comprises a latitude and longitude coordinate.	GPS is by far the most common source of global position estimates in outdoor environments.” GOOGLE1004 at p. 6. “Additionally, in our current implementation, the pose correction is limited to urban areas, since it is dependent on features visible in the scans and the global map. In this sense, it is complementary to GPS, which is more accurate in countryside areas than in urban canyons. If desired, it is straightforward to extend our approach to rural areas by including GPS information during the MCL perception phase.” <i>Id.</i> at p. 21.
[10] The method of claim 1 additionally comprising the step of associating metadata with a polygon based model.	Fruh discloses associating metadata with a polygon based model. GOOGLE1001 at ¶ 37. For example, Fruh discloses “for every arbitrary 3D point, we can compute the corresponding image coordinate and map the texture onto each mesh triangle. The result is textured façade models of the buildings that the acquisition vehicle drove by. Since these façade models have been brought into perfect registration with either aerial photo or DSM, they can eventually be merged with models derived from this same airborne data.” <i>Id.</i> at p. 16.
[11] The method of claim 1, wherein the composite image comprises a single vertical slice of each of the two or more images.	Fruh discloses the composite image comprises a single vertical slice of each of the two or more images. GOOGLE1001 at ¶ 27; see claim [1.2], <i>supra</i> (incorporated here).
[12.P] The method of claim 11, additionally comprising the steps of:	Fruh discloses the method of claim 11. See analysis of claim 11, <i>supra</i> .
[12.1] recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject: and	Fruh discloses recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject. GOOGLE1001 at ¶ 38. For example, Fruh discloses “the final global pose is obtained by utilizing an aerial photograph or a Digital Surface Model as a global map, to which the ground-based horizontal laser scans are matched.” GOOGLE1004 at Abstract. “Zhao and Shibasaki ... localization is based on the Global Positioning System (GPS). GPS is by far the most common source of global position estimates in outdoor environments.” GOOGLE1004 at p. 6. “[I]t is necessary to determine the pose of successive laser scans and camera images in a

	<p>global coordinate system with centimeter and sub-degree accuracy in order to reconstruct a consistent model.” Id. at p. 6. “[W]e use an aerial image or an airborne DSM to precisely reconstruct the path of the acquisition vehicle in offline computations: First, relative position changes are computed with centimeter accuracy by matching successive horizontal laser scans against each other, and are concatenated to reconstruct an initial path estimate. ... Monte-Carlo-Localization (MCL) is then used in conjunction with an airborne map to correct global pose. Our approach is to match features observed in both the laser scans and the aerial image or the DSM, whereby the airborne data can be regarded as a global map onto which the ground-based scan points have to be registered.” Id. at p. 6. “[I]t is straightforward to extend our approach to rural areas by including GPS information during the MCL perception phase.” Id. at p. 21.</p>
<p>[12.2] associating the composite image with a particular portion of the subject based upon the positional data.</p>	<p>Fruh discloses associating the composite image with a particular portion of the subject based upon the positional data. GOOGLE1001 at ¶ 38. For example, Fruh discloses “fast, ground-based acquisition of large-scale 3D city models. ... Specifically, the final global pose is obtained by utilizing an aerial photograph or a Digital Surface Model as a global map, to which the ground-based horizontal laser scans are matched.” GOOGLE1004 at Abstract. “Zhao and Shibasaki ... localization is based on the Global Positioning System (GPS).” Id. at p. 6. “[I]t is necessary to determine the pose of successive laser scans and camera images in a global coordinate system with centimeter and sub-degree accuracy in order to reconstruct a consistent model.” Id. at p. 6. “We calculate the 3D coordinates of the vertical scan points by applying a coordinate transformation from the local to the world coordinate system. The structure of the resulting point cloud is given by scan number and angle, and therefore each vertex has defined neighbors.” Id. at p. 16. “We calibrate the camera before our measurements and determine the transformation between its coordinate system and the laser coordinate system. Therefore, for every arbitrary 3D point, we can compute the corresponding image coordinate and map the texture onto each mesh triangle. ... Since these facade models have been brought into perfect registration</p>

	with either aerial photo or DSM, they can eventually be merged with models derived from this same airborne data.” Id.
[13] The method of claim 12 wherein the subject comprises a geographic area and the positional data comprises a geospatial designation.	Fruh discloses the subject comprises a geographic area and the positional data comprises a geospatial designation. GOOGLE1001 at ¶ 38. For example, Fruh discloses “fast, ground-based acquisition of large-scale 3D city models. ... A fairly accurate, textured 3D of the downtown Berkeley area has been acquired.” GOOGLE1004 at Abstract. “More specifically, it is necessary to determine the pose of successive laser scans and camera images in a global coordinate system with centimeter and sub-degree accuracy in order to reconstruct a consistent model.” Id. “[W]e use an aerial image or an airborne DSM to precisely reconstruct the path of the acquisition vehicle in offline computations: First, relative position changes are computed with centimeter accuracy by matching successive horizontal laser scans against each other, and are concatenated to reconstruct an initial path estimate. ... Monte-Carlo-Localization (MCL) is then used in conjunction with an airborne map to correct global pose. Our approach is to match features observed in both the laser scans and the aerial image or the DSM.” Id. “We first introduce a Cartesian world coordinate system [x, y, z] where x, y defines the geographical Figure 3. Two scans before matching (left) and after matching (right) ground plane and z the altitude as shown in Fig. 1. We also define a truck coordinate system [u, v] which is implied by the horizontal laser scanner. In a flat environment, a 2D pose of the truck and its local coordinate system can be entirely described by the two coordinates x, y and the yaw orientation angle θ ; in this case, the u-v coordinate system is assumed to be parallel to the x-y plane as shown in Fig. 1.” Id. at p. 8. “[I]t is straightforward to extend our approach to rural areas by including GPS information during the MCL perception phase.” Id. at p. 21.

VIII. [GROUND 3 CLAIM CHART] – Obviousness of Claims 1, 2, 5, 8-13, and 16 under § 103 by Fruh in view of Di Bernardo

U.S. Pat. 8,078,396	Fruh in view of Di Bernardo
[1.P] A method for creating a	Fruh discloses a method for creating a continuum of image

<p>continuum of image data, the method comprising:</p>	<p>data. See Section VII, claim [1.P], <i>supra</i> (incorporated here).</p> <p>Similarly, Di Bernardo also discloses an analogous system that performs a method for creating a continuum of image data. GOOGLE1001 at ¶ 42. For example, Di Bernardo discloses “a computer-implemented system and method for synthesizing images of a geographic location to generate composite images of the location. The geographic location may be a particular street in a geographic area with the composite images providing a view of the objects on each side of the street.” GOOGLE1005 at 2:16-21. “The composite images are created by synthesizing individual image frames acquired by a video camera moving through the location and filming the objects in its view. The composite images may depict an urban scene including the streets and structures of an entire city, state, or country.” <i>Id.</i> at 3:40-48.</p>
<p>[1.1] a) receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features;</p>	<p>Fruh discloses receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features. See Section VII, claim [1.1], <i>supra</i> (incorporated here).</p>
<p>[1.2] b) generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum;</p>	<p>Fruh discloses generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum. See Section VII, claim [1.2], <i>supra</i> (incorporated here).</p> <p>Similarly, Di Bernardo also discloses generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum. GOOGLE1001 at ¶ 43. In particular, Di Bernardo discloses “synthesizing images of a geographic location to generate composite images of the location.” GOOGLE1005 at 2:16-18. “Image data from each selected image frame 42 is then</p>

	<p>extracted and combined to form the composite image. Preferably, the image data extracted from each image frame is an optical column that consists of a vertical set of pixels. The composite image is preferably created on a column-by-column basis by extracting the corresponding optical column from each image frame.” <i>Id.</i> at 5:67-6:15; see <i>Id.</i> at FIG. 2; see GOOGLE1001 at ¶ 43. “[T]he computer 28, in step 138, identifies an associated image frame and extracts a column of RGB pixel values from the frame corresponding to the optical rays 46 originating from the fictitious camera 44. In step 140, the column of RGB pixel values are stacked side by side to generate a single image bitmap forming the composite image.” <i>Id.</i> at 9:34-43. For the reasons described above in Section VI.C, a POSITA would have been prompted to combine Fruh and Di Bernardo. GOOGLE1001 at ¶ 48-50.</p>
[1.3] c) receiving data descriptive of a location of the plurality of features;	Fruh discloses receiving data descriptive of a location of the plurality of features. See Section VII, claim [1.3], <i>supra</i> (incorporated here).
[1.4] d) tracking the position of at least one of the plurality of features in two or more of the two-dimensional images;	Fruh discloses tracking the position of at least one of the plurality of features in two or more of the two-dimensional images. See Section VII, claim [1.4], <i>supra</i> (incorporated here).
[1.5] e) generating a point cloud array for the at least one of the plurality of features;	Fruh discloses generating a point cloud array for the at least one of the plurality of features. See Section VII, claim [1.5], <i>supra</i> (incorporated here).
[1.6] f) converting the point cloud array to a polygon based model; and	Fruh discloses converting the point cloud array to a polygon based model. See Section VII, claim [1.6], <i>supra</i> (incorporated here).
[1.7] g) associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features.	Fruh discloses associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features. See Section VII, claim [1.7], <i>supra</i> (incorporated here).
[2] The method of claim 1 additionally comprising the step of spraying image data onto the polygon based	Fruh discloses spraying image data onto the polygon based model, wherein the polygon based model comprises at least one three dimensional polygon based model and wherein the image data sprayed onto the polygon based model is

model, wherein the polygon based model comprises at least one three dimensional polygon based model and wherein the image data sprayed onto the polygon based model is based upon two-dimensional image data sets and the location of the polygon based model.	based upon two-dimensional image data sets and the location of the polygon based model. See Section VII, claim [2], <i>supra</i> (incorporated here).
[8] The method of claim 1 wherein the data descriptive of a location comprises a Cartesian coordinate.	Fruh discloses the data descriptive of a location comprises a Cartesian coordinate. See Section VII, claim [8], <i>supra</i> (incorporated here).
[9] The method of claim 8 wherein the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate.	Fruh discloses the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate. See Section VII, claim [9], <i>supra</i> (incorporated here).
[11] The method of claim 1, wherein the composite image comprises a single vertical slice of each of the two or more images.	<p>Fruh discloses the composite image comprises a single vertical slice of each of the two or more images. See Section VII, claim [11], <i>supra</i> (incorporated here).</p> <p>Di Bernardo also discloses the composite image comprises a single vertical slice of each of the two or more images. GOOGLE1001 at ¶ 43; See claim [1.2], <i>supra</i> (incorporated here).</p>
[12.P] The method of claim 11, additionally comprising the steps of:	Fruh in combination with Di Bernardo discloses the method of claim 11. See analysis of claim 1, <i>supra</i> .
[12.1] recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject: and	<p>Fruh discloses recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject. See Section VII, claim [12.1], <i>supra</i> (incorporated here).</p> <p>Di Bernardo also discloses recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject. GOOGLE1001 at ¶ 44. In particular, Di Bernardo discloses “an image recording</p>

	<p>device moves along a path recording images of objects along the path. A GPS receiver and/or inertial navigation system provides position information of the image recording device as the images are being acquired."</p> <p>GOOGLE1005 at 2:22-28. "The data acquisition and processing system further includes a GPS receiver 16 for receiving position information from a set of GPS satellites 18 as the cameras 10 move along the trajectory. An inertial navigation system 20 including one or more accelerometers and/or gyroscopes also provides position information to the data acquisition and processing system." <i>Id.</i> at 3:62-4:1.</p> <p>"As the camera 10 moves along the path and records the objects in its view, the GPS receiver 16 computes latitude and longitude coordinates using the information received from the set of GPS satellites 18 at selected time intervals (e.g. one sample every two seconds). The latitude and longitude coordinates indicate the position of the camera 10 during the recording of a particular image frame." <i>Id.</i> at 5:17-23. "In step 60, the camera 10 acquires a series of images of a particular geographic location. At the same time, the GPS receiver 16 and/or inertial navigation system 20 acquires the position of the camera 10 while the images are being acquired." <i>Id.</i> at 6:26-34.</p>
<p>[12.2] associating the composite image with a particular portion of the subject based upon the positional data.</p>	<p>Fruh discloses associating the composite image with a particular portion of the subject based upon the positional data. See Section VII, claim [12.2], <i>supra</i> (incorporated here).</p> <p>Di Bernardo also discloses associating the composite image with a particular portion of the subject based upon the positional data. GOOGLE1001 at ¶ 45. In particular, Di Bernardo discloses "The composite image 40 is . . . associated with an identifier identifying the particular geographic location depicted in the image."</p> <p>GOOGLE1005 at 6:15-19. "[I]f the camera traverses through various streets, the computer 28 segments the trajectory into multiple straight street segments and associates each street segment with a street name and number range. In step 66, the computer 28 generates a series of composite images depicting a portion of each segment, and in step</p>

	<p>68, stores each composite image in the image database 32 along with the identifying information of the segment with which it is associated.” <i>Id.</i> at 6:38-49. “FIG. 14 is an illustration of an exemplary image coordinates table 200 for associating the composite images with the street segments in the street segments table 170. The image coordinates table 200 includes a plurality of composite image records where each record includes a segment ID 202 of the street segment being depicted in the composite image. In addition, each composite image record includes information of the side of the street segment 204 being depicted. For example, the side may be described as even or odd based on the street numbers on the side of the street being depicted. Each composite image entry also includes a distance from the segment origin to the center Xc of the composite image 206 indicating the position along the street segment for which the image was computed. The distance information is used to retrieve an appropriate composite image for each position on the street segment.” <i>Id.</i> at 11:19-35.</p>
<p>[13] The method of claim 12 wherein the subject comprises a geographic area and the positional data comprises a geospatial designation.</p>	<p>Fruh discloses the subject comprises a geographic area and the positional data comprises a geospatial designation. See Section VII, claim [13], <i>supra</i> (incorporated here).</p> <p>Di Bernardo also discloses that the subject comprises a geographic area and the positional data comprises a geospatial designation. GOOGLE1001 at ¶ 45. For example, Di Bernardo discloses that “synthesizing images of a geographic location to generate composite images of the location. The geographic location may be a particular street in a geographic area with the composite images providing a view of the objects on each side of the street.” <i>Id.</i> at 2:14-21. “The composite images may depict an urban scene including the streets and structures of an entire city, state, or country. The composite images may also depict other locales such as a zoo, national park, or the inside of a museum.” GOOGLE1005 at 3:46-49. “As the camera 10 moves along the path and records the objects in its view, the GPS receiver 16 computes latitude and longitude coordinates using the information received from the set of GPS satellites 18 at selected time intervals (e.g. one sample</p>

	every two seconds). The latitude and longitude coordinates indicate the position of the camera 10 during the recording of a particular image frame.” GOOGLE1005 at 5:17-23.
[16] The method of claim 12 wherein the method additionally comprises the steps of overlaying metadata on the composite image descriptive of the composite image.	To the extent that Fruh does not expressly disclose this “overlaying” step, such features were commonly used in other methods for creating a continuum of image data. In one example, Di Bernardo discloses overlaying metadata on the composite image descriptive of the composite image. GOOGLE1001 at ¶ 46, 47. In particular, Di Bernardo discloses that the software is additionally operative to overlay metadata on the composite image descriptive of the composite image. GOOGLE1001 at ¶ 37. For example, Di Bernardo discloses “The user may obtain information about the objects being visualized in the composite image by actuating one of the information icons 234 above the image of a particular object.” In displaying the information icons 234, a range of street addresses for the currently displayed image is computed. The listings in the object information database with street numbers that fall inside the computed range are then selected and associated with the information icons 234 displayed on top of the image of the object. If the objects are business establishments, the information displayed upon actuating the information icons 234 may include the name, address, and phone number 236 of the establishment. This information is preferably displayed each time the user terminal's cursor or pointing device is passed above the icon. GOOGLE1005 at 12:49-64; see GOOGLE1001 at ¶ 46. For the reasons described above in Section VI.C, a POSITA would have been prompted to implement Di Bernardo's teachings with Fruh's system. GOOGLE1001 at ¶¶ 48, 49.

IX. [GROUND 4 CLAIM CHART] – Anticipation of Claims 1, 2, 5, and 8-10 under § 102 by Akbarzadeh

U.S. Pat. 8,078,396	Akbarzadeh
[1.P] A method for creating a continuum of image data, the method comprising:	Akbarzadeh discloses a method for creating a continuum of image data. GOOGLE1001 at ¶ 53. For example, Akbarzadeh discloses “a processing pipeline for automatic geo-registered 3D reconstruction of urban scenes from video.

	<p>The system collects multiple video streams, as well as GPS and INS measurements in order to place the reconstructed models in georegistered coordinates.”</p> <p>GOOGLE1006 at Abstract. “The reconstruction step is divided into multi-view stereo, which produces depth-maps from multiple views with a single reference view, and depth-map fusion, which resolves conflicts between multiple depth maps and derives a coherent surface description. The dense reconstruction stage also provides texture for the surfaces using the video input.” <i>Id.</i> at p. 2. “The on-vehicle video acquisition system consists of two main sub-systems - an 8-camera digital video recorder (DVR) and an Applanix INS/GPS (model POS LV) navigation system. The DVR streams the raw images to disk, and the Applanix system tracks position and orientation so the 3D models produced in post-processing can be created in a common geo-registered coordinate system.” <i>Id.</i> at p. 5; see GOOGLE1001 at ¶ 53.</p>
<p>[1.1] a) receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features;</p>	<p>Akbarzadeh discloses receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features.</p> <p>GOOGLE1001 at ¶ 54. For example, Akbarzadeh discloses “The system collects multiple video streams, as well as GPS and INS measurements in order to place the reconstructed models in georegistered coordinates.”</p> <p>GOOGLE1006 at Abstract. “A system that automatically generates texturemapped, ground-level 3D models should be capable of capturing large amounts of data while driving through the streets and of processing these data efficiently. In this paper, we introduce an approach for fully automatic 3D reconstruction of urban scenes from several hours of video data captured by a multi-camera system.” <i>Id.</i> at p. 1. “The emphasis in our project is on developing a fully automatic system that is able to operate in continuous mode without the luxury of capturing data from selected viewpoints since capturing is performed from a moving vehicle constrained to the vantage points of urban streets.” <i>Id.</i> at p. 3.</p>

<p>[1.2] b) generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum;</p>	<p>Akbarzadeh discloses generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum. GOOGLE1001 at ¶ 55. For example, Akbarzadeh discloses “A system that automatically generates texture-mapped, ground-level 3D models should be capable of capturing large amounts of data while driving through the streets and of processing these data efficiently. In this paper, we introduce an approach for fully automatic 3D reconstruction of urban scenes from several hours of video data captured by a multi-camera system.” GOOGLE1006 at p. 1. “Once the camera poses have been computed, we use them together with the video frames to perform stereo matching on the input images. This leads to a depth map for each frame. These depth maps are later fused to enforce consistency between them.” <i>Id.</i> at p. 3.</p>
<p>[1.3] c) receiving data descriptive of a location of the plurality of features;</p>	<p>Akbarzadeh discloses receiving data descriptive of a location of the plurality of features. GOOGLE1001 at ¶ 56. For example, Akbarzadeh discloses “determining 2D-2D point correspondences in consecutive video frames. Then, we use the relative camera geometry of the internally calibrated cameras to establish a Euclidean space for the cameras. The INS/GPS information is used to compute the camera position in the geo-spatial coordinate system. ... Once the camera poses have been computed, we use them together with the video frames to perform stereo matching on the input images. ... To establish 2D feature correspondences between consecutive video frames we track features with a hierarchical KLT tracker [22]. ... we can also use a detect and match tracker similar to [24].” GOOGLE1006 at p. 3. “To determine geo-registered coordinates of the features in the 3D model, we employ the INS/GPS data. The INS/GPS measurement system is outfitted with a GPS receiver, gyroscopes, and accelerometers. It delivers highly accurate measurements of the position and orientation of the vehicle on which the cameras are mounted.” <i>Id.</i> at p. 4. “We are currently developing a second approach which overcomes these limitations by fusing geo-location measurements and tracked 2D features either</p>

	using a Kalman filter or through bundle adjustment.” <i>Id.</i> at p. 4; see GOOGLE1001 at ¶ 57.
[1.4] d) tracking the position of at least one of the plurality of features in two or more of the two-dimensional images;	Akbarzadeh discloses tracking the position of at least one of the plurality of features in two or more of the two-dimensional images. GOOGLE1001 at ¶ 57.; see element [1.3], <i>supra</i> (incorporated here).
[1.5] e) generating a point cloud array for the at least one of the plurality of features;	Akbarzadeh discloses generating a point cloud array for the at least one of the plurality of features. GOOGLE1001 at ¶ 58. For example, Akbarzadeh discloses “determining 2D-2D point correspondences in consecutive video frames. ... To establish 2D feature correspondences between consecutive video frames we track features with a hierarchical KLT tracker [22]. To achieve real-time tracking with video frame rate we use an implementation of the hierarchical KLT tracker on the GPU [23]. It needs on average 30ms to track 1000 feature points in a 1024 X 768 image on an ATI X 1900 graphics card.” GOOGLE1006 at p. 3. “[A] plane is swept through space in steps along a predefined direction, typically parallel to the optical axis. At each position of the plane , all views are projected on it. If a point on the plane is at the correct depth, then, according to the brightness constancy assumption, all pixels that project on it should have consistent intensities. We measure this consistency by summing the absolute intensity differences in square aggregation windows defined in the reference image, for which the depth map is computed.” <i>Id.</i> at p. 4. “[P]oints are visible in large number of frames, which provide multiple depth estimates for each point.” <i>Id.</i> at p. 5.
[1.6] f) converting the point cloud array to a polygon based model; and	Akbarzadeh discloses converting the point cloud array to a polygon based model. GOOGLE1001 at ¶ 58. For example, Akbarzadeh discloses “The multi-view stereo module takes as input the camera poses and images from a single video stream and produces a depth map for each frame. It uses a plane-sweep algorithm of Collins [13], which is an efficient multi-image matching technique. Conceptually, a plane is swept through space in steps along a predefined direction, typically parallel to the optical axis. At each position of the plane, all views are projected on it. If a point on the plane is at the correct depth, then, according to the brightness constancy assumption, all pixels that project on it should have

	<p>consistent intensities. We measure this consistency by summing the absolute intensity differences in square aggregation windows defined in the reference image, for which the depth map is computed.” GOOGLE1006 at p. 4. “[T]he final mesh can be generated rapidly by triangulating the depth estimates of adjacent pixels.” <i>Id.</i> at p. 5. “[P]oints are visible in large number of frames, which provide multiple depth estimates for each point.” <i>Id.</i> at p. 5. “From the fused stereo depth maps, we generate a triangulated, texture-mapped, 3D model of the scene in geo-registered coordinates.” <i>Id.</i></p>
<p>[1.7] g) associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features.</p>	<p>Akbarzadeh discloses associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features. GOOGLE1001 at ¶ 59. For example, Akbarzadeh discloses “geo-registered 3D reconstruction of urban scenes from video. The system collects multiple video streams, as well as GPS and INS measurements in order to place the reconstructed models in georegistered coordinates.” GOOGLE1006 at Abstract. “To determine geo-registered coordinates of the features in the 3D model, we employ the INS/GPS data. The INS/GPS measurement system is outfitted with a GPS receiver, gyroscopes, and accelerometers. It delivers highly accurate measurements of the position and orientation of the vehicle on which the cameras are mounted.” <i>Id.</i> at p. 4. “From the fused stereo depth maps, we generate a triangulated, texture-mapped, 3D model of the scene in geo-registered coordinates.” <i>Id.</i> at p. 5. “The on-vehicle video acquisition system consists of two main sub-systems - an 8-camera digital video recorder (DVR) and an Applanix INS/GPS (model POS LV) navigation system. The DVR streams the raw images to disk, and the Applanix system tracks position and orientation so the 3D models produced in post-processing can be created in a common geo-registered coordinate system.” <i>Id.</i> at p. 5.</p>
<p>[2] The method of claim 1 additionally comprising the step of spraying image data onto the polygon based model, wherein the polygon</p>	<p>Akbarzadeh discloses spraying image data onto the polygon based model, wherein the polygon based model comprises at least one three dimensional polygon based model and wherein the image data sprayed onto the polygon based model is based upon two-dimensional image data sets and</p>

<p>based model comprises at least one three dimensional polygon based model and wherein the image data sprayed onto the polygon based model is based upon two-dimensional image data sets and the location of the polygon based model.</p>	<p>the location of the polygon based model. GOOGLE1001 at ¶ 60. For example, Akbarzadeh discloses “dense reconstruction during which a texture-mapped, 3D model of the urban scene is computed from the video data. ... In dense reconstruction, the surfaces of the buildings, ground and other structures are estimated using multi-view stereo techniques. The goal of this step is to provide accurate surfaces wherever possible even in the presence of ambiguous or little surface texture, occlusion or specularly. The reconstruction step is divided into multi-view stereo, which produces depth-maps from multiple views with a single reference view, and depth-map fusion, which resolves conflicts between multiple depth maps and derives a coherent surface description. The dense reconstruction stage also provides texture for the surfaces using the video input.” GOOGLE1006 at p. 2. “From the fused stereo depth maps, we generate a triangulated, texture-mapped, 3D model of the scene in geo-registered coordinates.” <i>Id.</i> at p. 5.</p>
<p>[5.P] The method of claim 1 additionally comprising the steps of:</p>	<p>Akbarzadeh discloses the method of claim 1. See analysis of claim 1, <i>supra</i>.</p>
<p>[5.1] repeating steps a) through f) multiple times each from disparate points along the continuum; and</p>	<p>Akbarzadeh discloses repeating steps a) through f) multiple times each from disparate points along the continuum. GOOGLE1001 at ¶ 61. For example, Akbarzadeh discloses “A system that automatically generates texture-mapped, ground-level 3D models should be capable of capturing large amounts of data while driving through the streets and of processing these data efficiently. In this paper, we introduce an approach for fully automatic 3D reconstruction of urban scenes from several hours of video data captured by a multi-camera system. The goal is an automatic system for processing very large amounts of video data acquired in an unconstrained manner.” <i>Id.</i> at p. 1. “The emphasis in our project is on developing a fully automatic system that is able to operate in continuous mode. ... Our system design is also driven by the performance goal of being able to post-process the large video datasets in a time equal to the acquisition time.” <i>Id.</i> at p. 3. “We approach this by determining 2D-2D point correspondences in consecutive video frames. ... Once the</p>

	<p>camera poses have been computed, we use them together with the video frames to perform stereo matching on the input images. This leads to a depth map for each frame. These depth maps are later fused to enforce consistency between them.” <i>Id.</i> at p. 3. “From the fused stereo depth maps, we generate a triangulated, texture-mapped, 3D model of the scene in geo-registered coordinates.” <i>Id.</i> at p. 5.</p>
[5.2] spraying image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon two-dimensional image data sets and the location of each respective polygon based model.	<p>Akbarzadeh discloses spraying image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon two-dimensional image data sets and the location of each respective polygon based model. GOOGLE1001 at ¶ 60; see claim 2, <i>supra</i> (incorporated here).</p>
[8] The method of claim 1 wherein the data descriptive of a location comprises a Cartesian coordinate.	<p>Akbarzadeh discloses the data descriptive of a location comprises a Cartesian coordinate. GOOGLE1001 at ¶ 56; see claim [1.3], <i>supra</i> (incorporated here).</p>
[9] The method of claim 8 wherein the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate.	<p>Akbarzadeh discloses the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate. GOOGLE1001 at ¶ 56; see claim [1.3], <i>supra</i> (incorporated here).</p>
[10] The method of claim 1 additionally comprising the step of associating metadata with a polygon based model.	<p>Akbarzadeh discloses associating metadata with a polygon based model. GOOGLE1001 at ¶ 59. For example, Akbarzadeh discloses “geo-registered 3D reconstruction of urban scenes from video. The system collects multiple video streams, as well as GPS and INS measurements in order to place the reconstructed models in georegistered coordinates.” GOOGLE1006 at Abstract. “To determine geo-registered coordinates of the features in the 3D model, we employ the INS/GPS data. The INS/GPS measurement system is outfitted with a GPS receiver, gyroscopes, and accelerometers. It delivers highly accurate measurements of the position and orientation of the vehicle on which the cameras are mounted.” <i>Id.</i> at p. 4. “The DVR streams the raw</p>

	images to disk, and the Applanix system tracks position and orientation so the 3D models produced in post-processing can be created in a common geo-registered coordinate system. ” <i>Id.</i> at p. 5-6.
--	--

X. [GROUND 6 CLAIM CHART] – Obviousness of Claims 1, 2, 5, 8-13, and 16 under § 103 by Akbarzadeh in view of Di Bernardo

U.S. Pat. 8,078,396	Akbarzadeh in view of Di Bernardo
[1.P] A method for creating a continuum of image data, the method comprising:	<p>Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).</p> <p>Di Bernardo also discloses a method for creating a continuum of image data. See Section VIII, claim [1.P], <i>supra</i> (incorporated here).</p>
[1.1] a) receiving multiple two-dimensional image data sets of a subject, wherein each image data set is captured from a disparate point on a continuum and each image data set comprises a plurality of features;	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[1.2] b) generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum;	<p>Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).</p> <p>To the extent that Akbarzadeh does not expressly disclose all aspects of this “generating” step, such features were commonly used in other methods for creating a continuum of image data. In one example, Di Bernardo discloses generating a composite image of the subject from portions of two or more of the multiple two-dimensional image data sets, wherein the portions are aligned in a dimension consistent with the continuum. See Section VIII, claim [1.2], <i>supra</i> (incorporated here). For the reasons described above in Section VI.F, a POSITA would have been prompted to implement Di Bernardo’s teachings with Akbarzadeh’s system. GOOGLE1001 at ¶ 65.</p>

[1.3] c) receiving data descriptive of a location of the plurality of features;	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[1.4] d) tracking the position of at least one of the plurality of features in two or more of the two-dimensional images;	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[1.5] e) generating a point cloud array for the at least one of the plurality of features;	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[1.6] f) converting the point cloud array to a polygon based model; and	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[1.7] g) associating a location for the polygon based model relative to the portions of the image data sets and based upon the location of the plurality of features.	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[2] The method of claim 1 additionally comprising the step of spraying image data onto the polygon based model, wherein the polygon based model comprises at least one three dimensional polygon based model and wherein the image data sprayed onto the polygon based model is based upon two-dimensional image data sets and the location of the polygon based model.	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[5.P] The method of claim 1 additionally comprising the steps of:	Akbarzadeh in view of Di Bernardo discloses the method of claim 1. See analysis of claim 1, <i>supra</i> .
[5.1] repeating steps a) through f) multiple times each from disparate points	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).

along the continuum; and	
[5.2] spraying image data onto the each polygon based model, wherein the image data sprayed onto the each polygon based model is based upon two-dimensional image data sets and the location of each respective polygon based model.	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[8] The method of claim 1 wherein the data descriptive of a location comprises a Cartesian coordinate.	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[9] The method of claim 8 wherein the data descriptive of a location of the plurality of features comprises a latitude and longitude coordinate.	Akbarzadeh discloses a method for creating a continuum of image data. See Section IX, claim [1.P], <i>supra</i> (incorporated here).
[10] The method of claim 1 additionally comprising the step of associating metadata with a polygon based model.	Akbarzadeh discloses associating metadata with a polygon based model. See Section IX, claim 10, <i>supra</i> (incorporated here).
[11] The method of claim 1, wherein the composite image comprises a single vertical slice of each of the two or more images.	To the extent that Akbarzadeh does not expressly disclose the composite image comprises a single vertical slice of each of the two or more images, such features were commonly used in other methods for creating a continuum of image data. In one example, Di Bernardo discloses the composite image comprises a single vertical slice of each of the two or more images, and thus the resulting combination of Akbarzadeh in view of Di Bernardo (described above in the analysis of claim 1) would provide this element. See Section VIII, claim [1.2], <i>supra</i> (incorporated here). For the reasons described above in Section VI.F, a POSITA would have been prompted to implement Di Bernardo's teachings with Akbarzadeh's system. GOOGLE1001 at ¶ 65.
[12.P] The method of claim 11, additionally comprising the steps of:	Akbarzadeh in view of Di Bernardo discloses the method of claim 11. See analysis of claim 11, <i>supra</i> .

<p>[12.1] recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject: and</p>	<p>Akbarzadeh discloses recording positional data descriptive of a respective location of each of the multiple two-dimensional image data sets of a subject. GOOGLE1001 at ¶ 64. For example, Akbarzadeh discloses “The system collects multiple video streams, as well as GPS and INS measurements in order to place the reconstructed models in georegistered coordinates.” Akbarzadeh at Abstract; p. 2. “[T]he geo-registered poses of the cameras are estimated from the video and the INS/GPS data In sparse reconstruction the trajectory of the camera is estimated from the video data using structure from motion techniques. . . . To this end, the INS/GPS data are post-processed to obtain a filtered precise trajectory of the vehicle, which is called Smoothed Best Estimated Trajectory (SBET). . . . provide reliable estimates of the camera trajectories.” <i>Id.</i> at p. 2. “The processing pipeline begins by estimating a geo-registered camera pose for each frame of the videos. . . . The INS/GPS information is used to compute the camera position in the geo-spatial coordinate system.” <i>Id.</i> at p. 3. “[W]e employ the INS/GPS data. The INS/GPS measurement system is outfitted with a GPS receiver.”. <i>Id.</i> at p. 3-4. “With each video frame recorded, one of the DVR processor modules also records a GPS-timestamp message from the Applix navigation system. . . . In post-processing, this GPS timestamp is correlated with post-processed INS/GPS data to provide a smoothed best estimate of the position and orientation.” <i>Id.</i> at p. 5-6.</p>
<p>[12.2] associating the composite image with a particular portion of the subject based upon the positional data.</p>	<p>Akbarzadeh discloses associating the composite image with a particular portion of the subject based upon the positional data. GOOGLE1001 at ¶ 64. For example, Akbarzadeh discloses “geo-registered 3D reconstruction of urban scenes from video. The system collects multiple video streams, as well as GPS and INS measurements in order to place the reconstructed models in georegistered coordinates.” GOOGLE1006 at Abstract. “[T]he geo-registered poses of the cameras are estimated from the video and the INS/GPS data . . . a texture-mapped, 3D model of the urban scene is computed from the video data.” <i>Id.</i> at p. 2. “The processing pipeline begins by</p>

	<p>estimating a geo-registered camera pose for each frame of the videos. ... The INS/GPS information is used to compute the camera position in the geo-spatial coordinate system.” <i>Id.</i> at p. 3. “To determine geo-registered coordinates of the features in the 3D model, we employ the INS/GPS data. The INS/GPS measurement system is outfitted with a GPS receiver, gyroscopes, and accelerometers. It delivers highly accurate measurements of the position and orientation of the vehicle on which the cameras are mounted.” <i>Id.</i> at p. 3-4. “The DVR streams the raw images to disk, and the Applanix system tracks position and orientation so the 3D models produced in post-processing can be created in a common geo-registered coordinate system.” <i>Id.</i> at p. 5. “With each video frame recorded, one of the DVR processor modules also records a GPS-timestamp message from the Applanix navigation system. These message events are also synchronized to the CCD exposure by means of an external TTL signal output by one of the cameras. In post-processing, this GPS timestamp is correlated with post-processed INS/GPS data to provide a smoothed best estimate of the position and orientation of the navigation system and of each camera.” <i>Id.</i> at p. 5-6.</p> <p>To the extent that Akbarzadeh does not expressly disclose all aspects of this “associating” step, such features were commonly used in other methods for creating a continuum of image data. In one example, Di Bernardo discloses associating the composite image with a particular portion of the subject based upon the positional data. See Section VIII, claim [12.2], <i>supra</i> (incorporated here). For the reasons described above in Section VI.F, a POSITA would have been prompted to implement Di Bernardo’s teachings with Akbarzadeh’s system. GOOGLE1001 at ¶ 65.</p>
<p>[13] The method of claim 12 wherein the subject comprises a geographic area and the positional data comprises a geospatial designation.</p>	<p>Akbarzadeh discloses the subject comprises a geographic area and the positional data comprises a geospatial designation. GOOGLE1001 at ¶ 64. For example, Akbarzadeh discloses “geo-registered 3D reconstruction of urban scenes from video. The system collects multiple video streams, as well as GPS and INS measurements in</p>

	<p>order to place the reconstructed models in georegistered coordinates.” GOOGLE1006 at Abstract. “Detailed, 3D models of cities are usually made from aerial data, in the form of range or passive images combined with other modalities, such as measurements from a Global Positioning System (GPS).” <i>Id.</i> at p. 1. “In this paper, we introduce an approach for fully automatic 3D reconstruction of urban scenes from several hours of video data captured by a multi-camera system.” <i>Id.</i> “Processing entails the following steps: sparse reconstruction during which the georegistered poses of the cameras are estimated from the video and the INS/GPS data; and dense reconstruction during which a texture-mapped, 3D model of the urban scene is computed from the video data and the results of the sparse step.” <i>Id.</i> at p. 2. “To this end, the INS/GPS data are post-processed to obtain a filtered precise trajectory of the vehicle, which is called Smoothed Best Estimated Trajectory (SBET).” <i>Id.</i> “The INS/GPS information is used to compute the camera position in the geo-spatial coordinate system.” <i>Id.</i> at p. 3. “To determine georegistered coordinates of the features in the 3D model, we employ the INS/GPS data. The INS/GPS measurement system is outfitted with a GPS receiver, gyroscopes, and accelerometers. It delivers highly accurate measurements of the position and orientation of the vehicle on which the cameras are mounted.” <i>Id.</i> at p. 3-4. “In post-processing, this GPS timestamp is correlated with post-processed INS/GPS data to provide a smoothed best estimate of the position and orientation of the navigation system and of each camera.” <i>Id.</i> at p. 5-6.</p>
<p>[16] The method of claim 12 wherein the method additionally comprises the steps of overlaying metadata on the composite image descriptive of the composite image.</p>	<p>To the extent that Akbarzadeh does not expressly disclose this “overlaying” step, such features were commonly used in other methods for creating a continuum of image data. In one example, Di Bernardo discloses overlaying metadata on the composite image descriptive of the composite image. See Section VIII, claim [16], <i>supra</i> (incorporated here). For the reasons described above in Section VI.F, a POSITA would have been prompted to implement Di Bernardo’s teachings with Akbarzadeh’s system. GOOGLE1001 at ¶ 65.</p>

XI. CONCLUSION

Claims 1, 2, 5, 8-13, and 16 of the '396 patent are invalid over the prior art pursuant to Grounds 1-7 set forth above. Accordingly, Petitioner requests *inter partes* review of claims 1, 2, 5, 8-13, and 16.

Respectfully submitted,

Dated: August 20, 2014

/John C. Phillips/
John C. Phillips, Reg. No. 35,322

Dated: August 20, 2014

/Michael T. Hawkins/
Michael T. Hawkins, Reg. No. 57,867

(Trial No. _____)

Both above signers of:
Fish & Richardson P.C.
3200 RBC Plaza
60 South Sixth Street
Minneapolis, MN 55402
Attorneys for Petitioner

CERTIFICATE OF SERVICE

Pursuant to 37 CFR §§ 42.6(e)(4)(i) *et seq.* and 42.105(b), the undersigned certifies that on August 20, 2014, a complete and entire copy of this Petition for *Inter Partes* Review and all supporting exhibits were provided by Express Mail, cost prepaid, to the Patent Owner by serving the correspondence address of record as follows:

Ideation Law, PLLC
P.O. Box 936
Vails Gate NY 12584

/Diana Bradley/

Diana Bradley
Fish & Richardson P.C.
60 South Sixth Street, Suite 3200
Minneapolis, MN 55402
(858) 678-5667