

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

VUSIONGROUP SA,

Petitioner

v.

HANSHOW TECHNOLOGY CO., LTD.,

Patent Owner.

IPR2024-00963

U.S. Patent No. 10,701,321

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 10,701,321**

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
U.S. Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	GROUNDS FOR STANDING.....	1
III.	STATEMENT OF PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED.....	1
IV.	'321 PATENT OVERVIEW	3
	A. Specification.....	3
	B. Prosecution History	4
V.	LEVEL OF ORDINARY SKILL	6
VI.	CLAIM CONSTRUCTION	6
	A. “object detection system”	7
	B. “video analytics system”	8
	C. “object analysis system”	8
VII.	DETAILED EXPLANATION OF GROUNDS.....	8
	A. Ground 1A: Obvious Over Laganiere363, Laganiere817, and Laganiere461 (Claims 1-5, 7-21)	8
	1. Laganiere363.....	8
	2. Laganiere817.....	10
	3. Laganiere461	12
	4. Independent Claim 1	12
	a. 1.pre	12
	b. 1.1	13
	c. 1.2	14
	d. 1.3.1	16
	e. 1.3.2	19
	f. 1.4.1	21
	g. 1.4.2	22

TABLE OF CONTENTS **(continued)**

	Page
h. 1.5.1	23
i. 1.5.2	27
j. 1.5.3	28
5. Independent Claim 13	30
6. Claims 2, 14	31
7. Claims 3, 15	32
8. Claim 4, 16	34
9. Claims 5, 18	35
10. Claims 7, 17	36
11. Claims 8, 19	37
12. Claims 9, 20	38
13. Claims 10, 21	40
14. Claim 11	41
15. Claim 12	41
B. Ground 1B: Obviousness over Laganiere363, Laganiere817, Laganiere461, and Chen (Claims 4, 6, 16)	42
1. Chen	42
2. Motivation to Combine	43
3. Claims 4, 16	45
4. Claim 6	46
C. Ground 2: Obviousness Over Bedros (Claims 1-21)	46
1. Bedros	46
2. Independent Claim 1	50
a. 1.pre	50
b. 1.1	50

TABLE OF CONTENTS (continued)

	Page
c. 1.2	51
d. 1.3.1	57
e. 1.3.2	58
f. 1.4.1	61
g. 1.4.2	63
h. 1.5.1	64
i. 1.5.2	67
j. 1.5.3	68
3. Independent Claim 13	71
4. Claims 2, 14	72
5. Claims 3, 15	73
6. Claims 4, 6, 16	74
7. Claims 5, 18	75
8. Claims 7, 17	77
9. Claims 8, 19	78
10. Claims 9, 20	79
11. Claims 10, 21	80
12. Claim 11	81
13. Claim 12	81
VIII. DISCRETIONARY DENIAL WOULD BE INAPPROPRIATE.....	82
IX. CONCLUSION.....	83
X. MANDATORY NOTICES	83
A. Real Party in Interest	83
B. Related Matters.....	84

TABLE OF CONTENTS
(continued)

	Page
C. Notice of Counsel and Service Information.....	84
D. Power of Attorney	85

EXHIBIT LIST

Exhibit	Reference
1001	U.S. Patent No. 10,701,321 (“’321 patent” or “’321Pat.”)
1002	File History of the ’321 patent (“’321FH”)
1003	Declaration of Dr. Henry Houh in Support of Petition
1004	Dr. Henry Houh Curriculum Vitae
1005	U.S. Patent Application Publication No. 2012/0195363 (“Laganiere363”)
1006	U.S. Patent No. 8,139,817 (“Laganiere817”)
1007	U.S. Patent Application Publication No. 2011/0113461 (“Laganiere461”)
1008	U.S. Patent No. 8,634,601 (“Bedros”)
1009	U.S. Patent Application Publication No. 2014/0044348 (“Chen”)
1010	Excerpt of THE AUTHORITATIVE DICTIONARY OF IEEE STANDARDS TERMS, 7th ed. (2000)
1011	Excerpt of MERRIAM-WEBSTER’S COLLEGIATE DICTIONARY, 11th ed. (2005)
1012	Excerpt of NIST SPECIAL PUBLICATION 800-53, rev. 4 (Jan. 22, 2015)
1013	Excerpt of REVERSE ACRONYMS, INITIALISMS AND ABBREVIATIONS DICTIONARY, 10th ed., vol. 3, part 1 (Oct. 23, 1985)
1014	U.S. Patent No. 8,619,150 (“Deever ’150”)
1015	U.S. Patent No. 8,432,965 (“Deever ’965”)
1016	Excerpt of Dictionary of Contemporary Photography (1974)
1017	Excerpt of AXIS 211M Network Camera User’s Manual (Feb. 2010)
1018	Complaint, <i>Hanshow Technology Co., Ltd. v. SES-imagotag SA et al.</i> , Case No. 2:23-cv-00617 (E.D. Tex.).

Exhibit	Reference
1019	Waiver of Service, <i>Hanshow Technology Co., Ltd. v. SES-imagotag SA et al.</i> , Case No. 2:23-cv-00617 (E.D. Tex.).
1020	Unopposed Motion to Reset Trial Date and Amend Docket Control Order, <i>Hanshow Technology Co., Ltd. v. SES-imagotag SA et al.</i> , Case No. 2:23-cv-00617 (E.D. Tex.).
1021	U.S. Patent Application Publication No. 2016/0050396 (“Gali”)
1022	U.S. Patent Application Publication No. 2016/0191779 (“Yang”)
1023	U.S. Patent Application Publication No. 2013/0166711 (“Wang”)
1024	Christopher John Lindblad, <i>A Programming System for the Dynamic Manipulation of Temporally Sensitive Data</i> , (1994) (Ph.D. dissertation, Massachusetts Institute of Technology) (“Lindblad”)
1025	David James Wetherall, <i>An Interactive Programming System for Media Computation</i> (1994) (Master of Science dissertation, Massachusetts Institute of Technology) (“Wetherall”)
1026	William F. Stasior, <i>An Interactive Approach to the Identification and Extraction of Visual Events</i> (1997) (Ph.D. dissertation, Massachusetts Institute of Technology) (“Stasior”)
1027	Jia Deng et al., ImageNet: A Large-Scale Hierarchical Image Database, <i>2009 IEEE Conference on Computer Vision and Pattern Recognition</i> , (2009) (“Deng”)
1028	Alex Krizhevsky et al., ImageNet Classification with Deep Convolutional Neural Networks, <i>Comm’n of the ACM</i> , vol. 60, issue 6, pp. 84-90 (2017) (“Krizhevsky”)
1029	Andrew Tanenbaum, <i>Computer Networks</i> (3rd ed. 1996) (“Tanenbaum”)
1030	Joel F. Adam & David L. Tennenhouse, <i>The Vidboard: A Video Capture and Processing Peripheral for a Distributed Multimedia System</i> , MIT LABORATORY FOR COMPUTER SCIENCE (1994) (“Adam”)

LISTING OF CHALLENGED CLAIMS

Reference	Claim Limitation
Claim 1	
1.pre	A video surveillance system, comprising:
1.1	one or more surveillance cameras that capture image data;
1.2	an object detection system that detects objects of interest within the image data and extracts portions of frames that include the detected objects of interest and sends the extracted portions of the frames; and
1.3.1	a video analytics system that receives only the extracted portions of the frames of the image data, transmitted over a network, from the object detection system,
1.3.2	the image data received by the video analytics system corresponding to objects of interest detected by the object detection system, the video analytics system analyzing the objects of interest;
1.4.1	wherein the video analytics system receives the image data from multiple object detection systems after being transmitted over the network and
1.4.2	the video analytics system is operated by a different business entity than the one or more surveillance cameras; and
1.5.1	wherein the image data corresponding to objects of interest detected by the object detection system are ranked by the object detection system according to factors including close proximity of the objects of interest to the one or more surveillance cameras, picture quality of the objects of interest, and/or contrast of the objects of interest to create a ranking and
1.5.2	the ranking is saved in metadata with the image data and
1.5.3	an object analysis system of the video analytics system reads the rankings within the metadata and decides whether or not to analyze the image data.
Claim 2	

Reference	Claim Limitation
2	The system of claim 1, wherein the object detection system creates cropped images from the image data that include the detected objects of interest and sends the cropped images to the video analytics system.
Claim 3	
3	The system of claim 1, wherein the image data include video data and the object detection system provides frames from the video data that include the detected objects of interest and sends the frames to the video analytics system as the image data.
Claim 4	
4	The system of claim 1, wherein the video analytics system includes an object analysis system that analyzes the image data corresponding to objects of interest detected by the object detection system by comparing the image data to reference images to identify the objects of interest, by recognizing patterns in the image data, and/or by applying image processing functions, neural network functions, deep learning functions, and/or object classification functions to the image data.
Claim 5	
5	The system of claim 1, wherein the object detection system is integrated within the one or more surveillance cameras.
Claim 6	
6	The system of claim 1, wherein the image data corresponding to objects of interest detected by the object detection system are compared by the video analytics system to reference images to identify the objects of interest.
Claim 7	
7	The system of claim 6, further comprising a database that receives metadata associated with the identified objects of interest from the video analytics system.
Claim 8	

Reference	Claim Limitation
8	The system of claim 1, further comprising a gateway system that provides access to the video analytics system over a public network.
Claim 9	
9	The system of claim 8, wherein the object detection system is integrated within the gateway system.
Claim 10	
10	The system of claim 1, wherein the one or more surveillance cameras each include a gateway system that provides access to the video analytics system over a public network.
Claim 11	
11	The system of claim 1, wherein the video analytics system receives the image data from the object detection system after being transmitted over an enterprise network.
Claim 12	
12	The system of claim 1, wherein the video analytics system receives the image data from the object detection system after being transmitted over a cloud network.
Claim 13	
13.pre	A method for analyzing video in a video surveillance system, the method comprising:
13.1	capturing image data from one or more surveillance cameras;
13.2.1	detecting objects of interest within the image data via an object detection system; the object detection system extracting portions of frames that include the detected objects of interest and transmitting the extracted portions of the frames of the image data over a network,
13.2.2	the image data transmitted by the object detection system corresponding to objects of interest detected by the object detection system; and
13.3	a video analytics system, which is operated by a different business entity than the one or more surveillance cameras, only

Reference	Claim Limitation
	receiving the extracted portions of the frames of the image data from the object detection system and analyzing the objects of interest;
13.4	the video analytics system further receiving image data from other object detection systems;
13.5.1	the object detection system ranking the image data corresponding to objects of interest detected by the object detection system according to factors including close proximity of the objects of interest to the one or more surveillance cameras, picture quality of the objects of interest, and/or contrast of the objects of interest to create a ranking and
13.5.2	the ranking is saved in metadata with the image data; and
13.5.3	an object analysis system of the video analytics system reading the rankings within the metadata and deciding whether or not to analyze the image data.
Claim 14	
14	The method of claim 13, further comprising the object detection system creating cropped images from the image data that include the detected objects of interest and sending the cropped images to the video analytics system.
Claim 15	
15	The method of claim 13, further comprising the image data including video data and the object detection system providing frames from the video data that include the detected objects of interest and sending the frames to the video analytics system as the image data.
Claim 16	
16	The method of claim 13, wherein the video analytics system analyzing the image data corresponding to objects of interest detected by the object detection system comprises comparing the image data to reference images to identify the objects of interest, by recognizing patterns in the image data, and/or by applying image processing functions, neural network functions, deep

Reference	Claim Limitation
	learning functions, and/or object classification functions to the image data.
Claim 17	
17	The method of claim 16, further comprising receiving metadata associated with the identified objects of interest from the video analytics system.
Claim 18	
18	The method of claim 13, further comprising integrating the object detection system within the one or more surveillance cameras.
Claim 19	
19	The method of claim 13, further comprising providing access to the video analytics system over a public network via a gateway system.
Claim 20	
20	The method of claim 19, further comprising integrating the object detection system within the gateway system.
Claim 21	
21	The method of claim 13, further comprising the one or more surveillance cameras each including a gateway system that provides access to the video analytics system over a public network.

I. INTRODUCTION

Petitioner requests *inter partes* review (“IPR”) of claims 1-21 (“Challenged Claims”) of U.S. Patent 10,701,321. Petitioner has shown a reasonable likelihood of prevailing as to each Challenged Claim and respectfully requests the Board institute trial.

II. GROUNDS FOR STANDING

Petitioner certifies the ’321 patent is available for IPR, and Petitioner is not barred or estopped from requesting IPR. EX1018, EX1019.

III. STATEMENT OF PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED

Petitioner respectfully requests review and cancellation of the Challenged Claims¹ in view of:

¹ Petitioner does not concede any Challenged Claim satisfies other requirements for patentability that cannot be raised in IPR, including 35 U.S.C. §§101, 112.

Prior Art	
U.S. Patent Application Publication 2012/0195363 (“Laganiere363”), published August 2, 2012; prior art under at least 35 U.S.C. §102(a)(1) (EX1005)	
U.S. Patent 8,139,817 (“Laganiere817”), issued March 20, 2012; prior art under at least 35 U.S.C. §102(a)(1) (EX1006)	
U.S. Patent Application Publication 2011/0113461 (“Laganiere461”), published May 12, 2011; prior art under at least 35 U.S.C. §102(a)(1) (EX1007)	
U.S. Patent 8,634,601 (“Bedros”), issued January 21, 2014; prior art under at least 35 U.S.C. §102(a)(1) (EX1008)	
U.S. Patent Application Publication 2014/0044348 (“Chen”), published February 13, 2014; prior art under at least 35 U.S.C. §102(a)(1) (EX1009)	

Ground (35 U.S.C. §103)	Claims	References
1A	1-5, 7-21	Laganiere363, Laganiere817, Laganiere461
1B	4, 6, 16	Laganiere363, Laganiere817, Laganiere461, Chen
2	1-21	Bedros

IV. '321 PATENT OVERVIEW

A. Specification

The '321 patent is titled “System and Method for Distributed Video Analysis,” and acknowledges that existing “[v]ideo surveillance systems can increase the volume of data traffic over the networks,” causing “performance problems.” *'321Pat.*, Title, 1:47-53; *Houh*, ¶¶45-50 (explaining background of the technology). To lower transmitted data, image data processing is divided into “separate object detection and object analysis functions or systems.” *Id.*, 2:8-10, 2:15-17. “The object detection system detects objects of interest” and “only image data that includes the detected objects of interest” is sent to the local or remote video analytics systems. *Id.*, 2:11-20, 2:41-44. Annotated Figure 1 below shows object detection system 108-1 and 108-2 (yellow) connected to video analytics system 122 (purple) through network 60 (red). *Houh*, ¶39

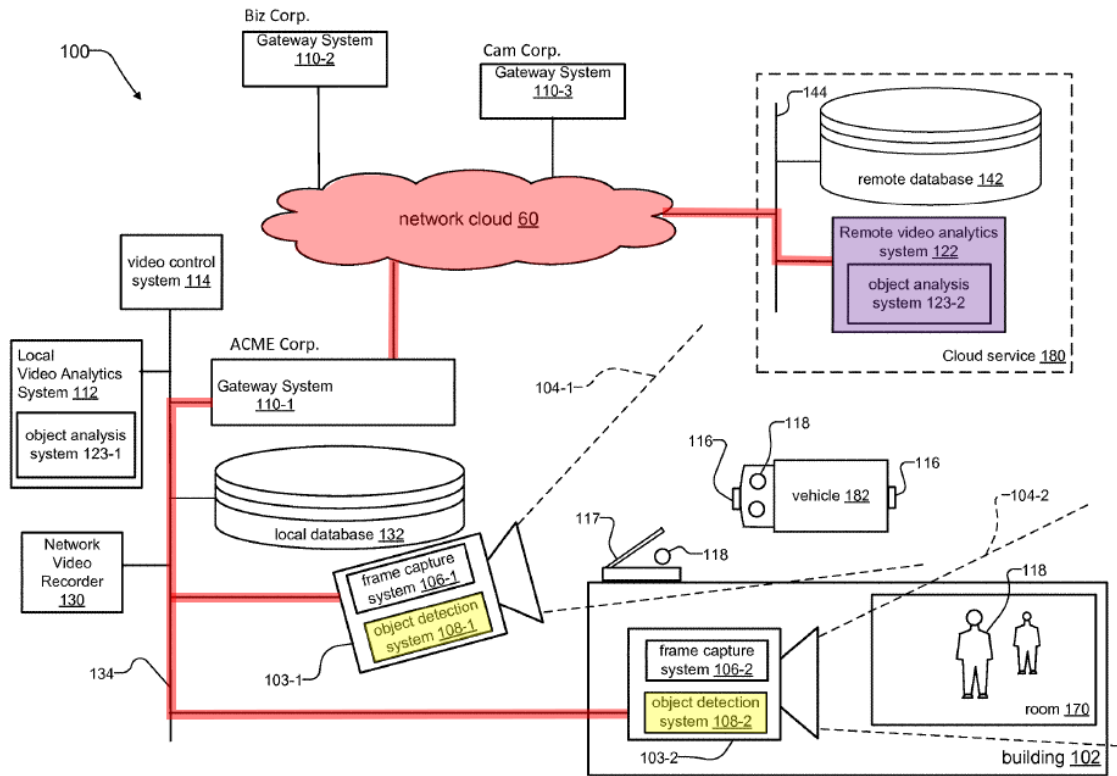


FIG. 1

To further reduce bandwidth, the object detection system can create and send image data that includes “only the detected objects of interest” to the video analytics system. *Id.*, 2:48-54; *Houh*, ¶40.

B. Prosecution History

U.S. Patent Application 14/634,428 was filed Feb. 27, 2015, without a claim to earlier priority. '321*Pat.*, Cover. After extensive prosecution history, the '321 patent issued on June 30, 2020. *Houh*, ¶¶41-44.

Over five years, the Examiner repeatedly rejected all pending claims because various prior art references independently taught surveillance cameras that detect

objects of interest within the image data and transmit the image data over a network.

Id., 79-87, 112-21, 154-62, 183-93, 245-52, 272-81, 320-28.

The Examiner noted Chen (EX1009) disclosed an object detection system that ranks image data according to the factors listed in pending claims 4 and 17.² *Id.*, 83-84, 116, 159-60, 187, 249, 276, 324-25 (citing Chen ¶34). The Examiner also noted Chen disclosed “comparing the image data to reference images to identify the objects of interest.” *Id.* The applicant never contested the Examiner’s Chen-based rejections. *See id.* Rather, in response to the October 11, 2019, office action, applicant amended claims 4 and 17 to recite limitations including “creat[ing],” “sav[ing],” and “read[ing]” the ranking to “decide[.]” whether to analyze the image data. *Id.*, 344-47. The Examiner incorporated as-amended claims 4 and 17 into independent claims 1 and 13 in an Examiner’s amendment and allowed the application without providing reasons for allowance. *Id.*, 361-68.

Laganiere363, Laganiere817, Laganiere416, and Bedros were not cited in the ’321 file history. Nor was the cited art related to those references.

² As-filed claims 4 and 17 were substantially incorporated as limitations 1.5.1 and 13.5.1 into the as-issued independent claims 1 and 13.

V. LEVEL OF ORDINARY SKILL

A person of ordinary skill in the art (“POSITA”) in February 2015 would have had a bachelor’s degree in electrical engineering, computer engineering, computer science, or a related field, and two or more years of work experience or research related to computer vision, image analytics or automated object recognition methods and systems. *Houh*, ¶¶18-19. Additional education in a relevant field or relevant industry experience may compensate for a deficit in one of the other aspects of the requirements stated above. *Id.*

Dr. Houh was at least a POSITA as of February 2015. *Houh*, ¶¶7-17, 20-21.

VI. CLAIM CONSTRUCTION

Petitioner is unaware of any “prior claim construction determination” related to the ’321 patent. 37 C.F.R. §42.100(b). No formal claim constructions are necessary in this proceeding because “claim terms need only be construed ‘to the extent necessary to resolve the controversy.’”³ *Wellman, Inc. v. Eastman Chem. Co.*, 642 F.3d 1355, 1361 (Fed. Cir. 2011).

37 C.F.R. §42.104(b)(3) requires a petition to include, for any means-plus-function limitations, “the specific portions of the specification that describe the

³ Petitioner reserves all rights to raise claim construction arguments in other forums or to address construction of any terms PO argues in this proceeding.

structure, material, or acts corresponding to each claimed function.” No claim terms in the Challenged Claims use the term “means,” and as such, there is a rebuttable presumption that §112(f) does not apply. *Williamson v. Citrix Online, LLC*, 792 F.3d 1339, 1348 (Fed. Cir. 2015). If the Board finds the presumption is overcome for any of the “system” limitations, Petitioner contends that the specification recites the processors and hardware identified in the tables below but lacks disclosure of sufficient algorithms to perform the claimed functions, rendering each such term indefinite. *Houh*, ¶¶51-52. Nevertheless, because the prior art identified herein discloses equivalent structure for each such term, “[i]ndefiniteness ‘does not necessarily preclude the Board from addressing the patentability of the claims on section 102 and 103 grounds.’” *Sony Elecs. Inc. v. Jawbone Innovations, LLC*, IPR2023-01352, Paper 7, 47 (PTAB Mar. 6, 2024) (quoting *Samsung Elecs. Am., Inc. v. Prisia Eng’g Corp.*, 948 F.3d 1342, 1353 (Fed. Cir. 2020)); *Intel Corp. v. Qualcomm Inc.*, 21 F.4th 801, 813 (Fed. Cir. 2021).

A. “object detection system”

<u>Claims:</u>	1.2, 13.2.1; 1.5.1, 13.5.1; 1.5.2, 13.5.2; 2, 14; 3, 15
<u>Structure:</u>	Processor with memory in a local analytics system, local camera, or local gateway, <i>e.g.</i> , ’321 <i>Pat.</i> , Figs. 2, 4, 6, 7, 1:64-2:2, 2:50-53, 3:1-4, 3:32-34, 4:45-49, 6:9-11, 6:19-23, 6:29-36, 6:41-43, 7:65-67, 9:3-15, 11:1-8, but lacking disclosure of sufficient algorithms to perform the claimed functions.

B. “video analytics system”

<u>Claims:</u>	1.3.1, 13.3; 1.3.2, 13.3; 1.4.1, 13.4; 4, 16; 6
<u>Structure:</u>	Processor with memory in a remote analytics system, remote gateway, or cloud, <i>e.g.</i> , '321 <i>Pat.</i> , Figs. 2, 4, 6, 7, 1:64-2:2, 2:21-36, 2:59-63, 3:14-22, 3:36-38, 4:50-52, 6:52-7:3, 7:19-41, 9:6-31, 10:24-45, 11:14-30, but lacking disclosure of sufficient algorithms to perform the claimed functions.

C. “object analysis system”

<u>Claims:</u>	1.5.3, 13.5.3; 4
<u>Structure:</u>	Processor with memory in a remote analytics system, remote gateway, or cloud, <i>e.g.</i> , '321 <i>Pat.</i> , Figs. 2, 4, 6, 7, 1:64-2:2, 2:21-36, 3:20-22, 6:38-40, 6:52-7:3, 7:19-41, 9:16-37, 10:24-45, 11:14-35, but lacking disclosure of sufficient algorithms to perform the claimed functions.

VII. DETAILED EXPLANATION OF GROUNDS

A. Ground 1A: Obvious Over Laganriere363, Laganriere817, and Laganriere461 (Claims 1-5, 7-21)

1. Laganriere363

Laganriere363 (EX1005) is analogous art to the '321 patent. *Houh*, ¶¶55-60.

Laganriere363 discloses a video analytics system that pre-processes video data “at the source end to reduce the bandwidth requirement for transmitting the video data.” *Laganriere363*, Abstract, ¶¶2-7. Laganriere363’s system includes multiple video sources 102/116 (*e.g.*, cameras) (yellow) connected to server 108 (purple) over WAN 106 (red) through optional gateway 104:

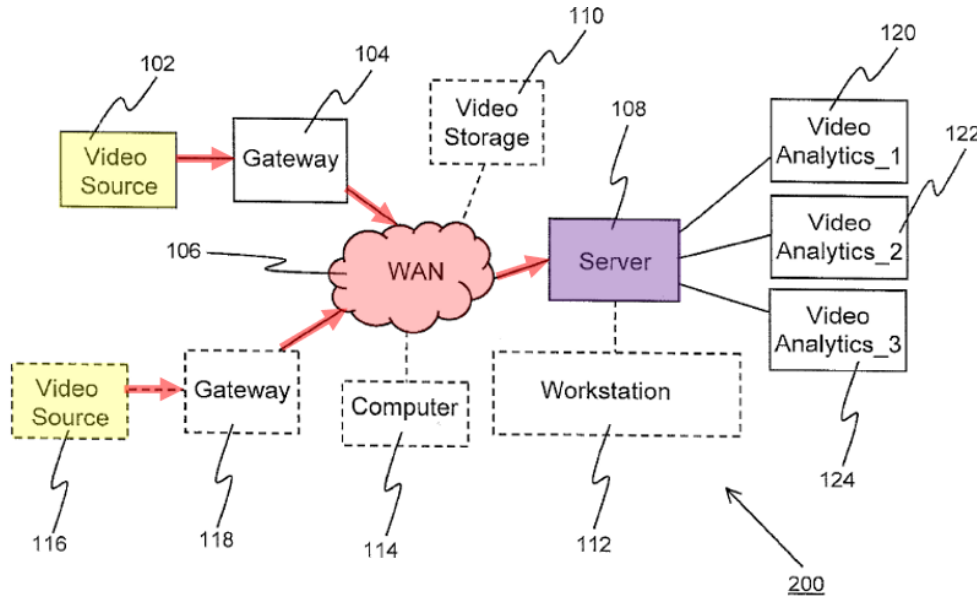


Fig. 2

Id., Fig. 2 (annotated), ¶¶29-33; *see id.*, Fig. 1, ¶¶25-28; *Houh*, ¶¶56-57. Video data is captured and “pre-processed...using video analytics” at the video sources to produce a “data stream that requires an amount of bandwidth below a bandwidth limit of WAN 106.” *Laganieri*363, ¶34. The pre-processing operation provides the benefit of “reduc[ing] the bandwidth requirement for transmitting the video data” (*id.*) and can include “cropping regions of the video data that relate other than the event of interest” (*id.*, ¶21). For example, data relating to license plates or faces of individual can be preprocessed. *Id.*, ¶¶30, 41; *Houh*, ¶58.

The pre-processed data is then transmitted from video source 102 to central server 108. *Id.*, ¶35. The transmitted data is further processed at central server 108 or at “one of the video analytics engines 120-124.” *Id.*, ¶35, Fig. 5 (below). The

engines may process the data to perform facial recognition and/or reading license plate characters. *Id.*, ¶30.

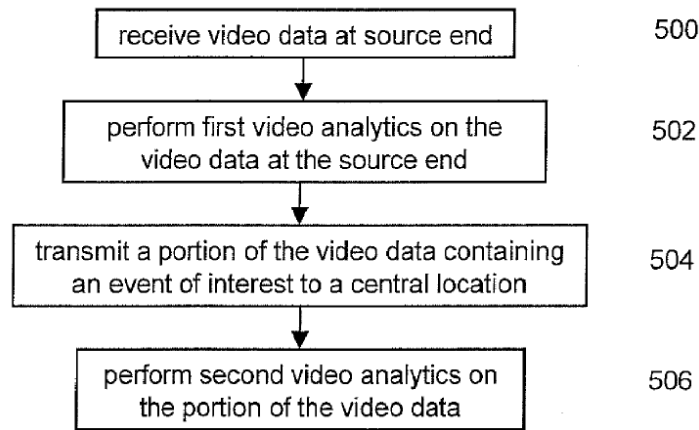


Fig. 5

2. Laganiere817

Laganiere817 is analogous art to the '321 patent and shares inventorship with Laganiere363. *Houh*, ¶¶61-64.

Laganiere817 discloses a system and method to create a ranked queue of data, a “face log,” from video and/or image data captured by a camera for subsequent facial recognition analysis. *Laganiere817*, 2:20-50. Figure 1 depicts Laganiere817’s system. *Id.*, 3:55-56.

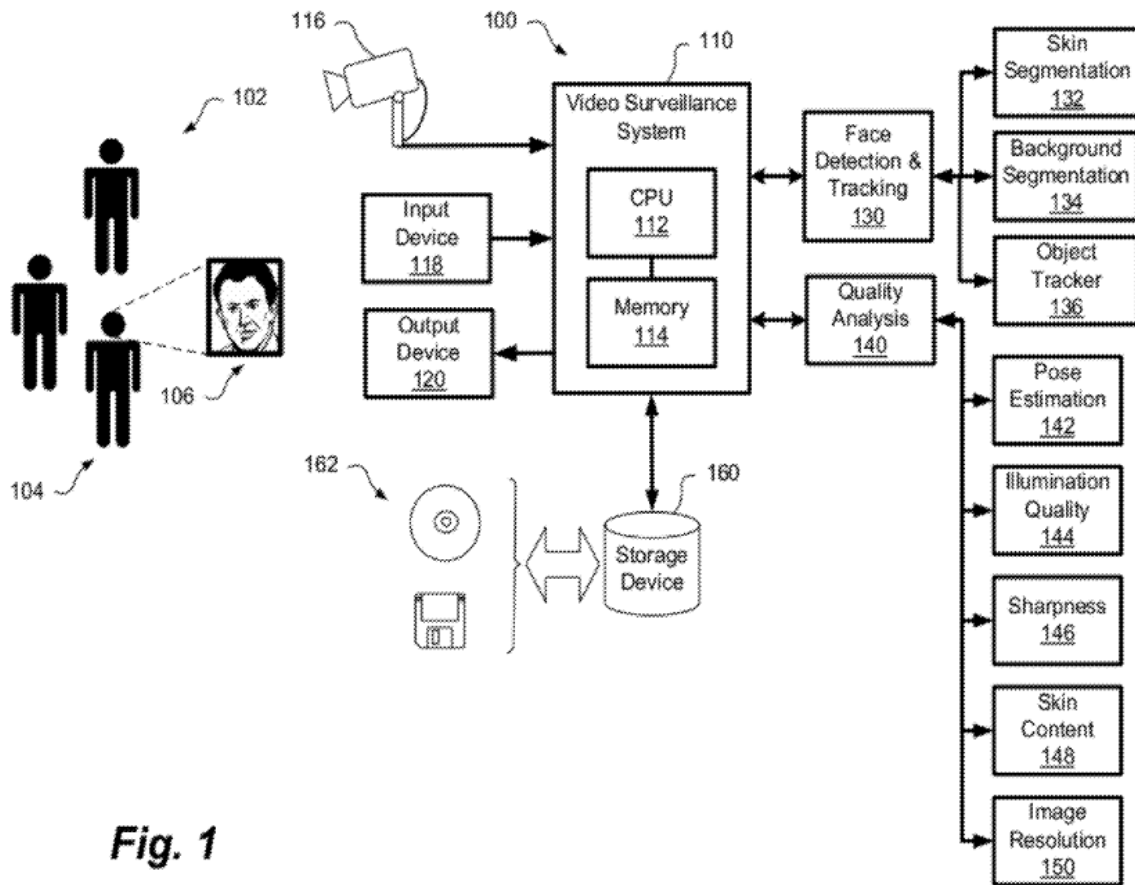


Fig. 1

Video cameras 116 capture video or still images, and image characteristics, “such as resolution, pose, illumination, skin content and sharpness,” are used to select images. *Id.*, 3:35-38, 55-56. These characteristics are scored and incorporated to create an “overall quality score of an image.” *Id.*, 3:38-42. The quality scores are then used to create a face log (*id.*, 7:55-10:32), containing only images of an individual of sufficient quality to perform facial recognition (*id.*, 7:48-51, 9:61-63).

3. Laganiere461

Laganiere461 is analogous art to the '321 patent and shares inventorship with Laganiere363 and Laganiere817. *Houh*, ¶¶65-66. Laganiere461's system is similar to the system depicted in Laganiere363. *Houh*, ¶¶67-68.

Laganiere461's system "reduc[es] the amount of space that is required to store video data, based on selective deletion of portions of video data that are stored in a video storage device." *Laganiere461*, ¶2. In Laganiere461, a central server performs deletion of images. *Id.*, ¶¶24-25.

4. Independent Claim 1

a. 1.pre

If limiting, Laganiere363 teaches 1.pre. *Houh*, ¶69. Laganiere363 seeks to overcome limitations with existing "video surveillance systems." *Laganiere363*, ¶¶2-7. Laganiere363 discloses systems with video sources 102 "for monitoring a known field of view (FOV)" such as "a parking lot, an entry/exit point of a building, and an automated teller machine (ATM)." *Id.*, ¶25; *id.*, Figs. 1, 2. The systems also include video analytics engines 120-124 that can detect and read license plates and "perform[] facial recognition." *Id.*, ¶30. Laganiere363 suggests these and other video analytics engines are needed by "different monitoring or surveillance applications." *Id.* Accordingly, POSITAs would have understood Laganiere363 discloses "a video surveillance system." *Houh*, ¶¶70-71.

b. 1.1

Laganiere363 teaches 1.1. *Houh*, ¶72. Each source end video source 102/116 in Figures 1-2 is a “network IP **camera** with onboard video analytics capabilities,” or “a ‘dumb’ IP **camera** or an analogue **video camera**.”⁴ *Laganiere363*, ¶¶25, 28, 33.

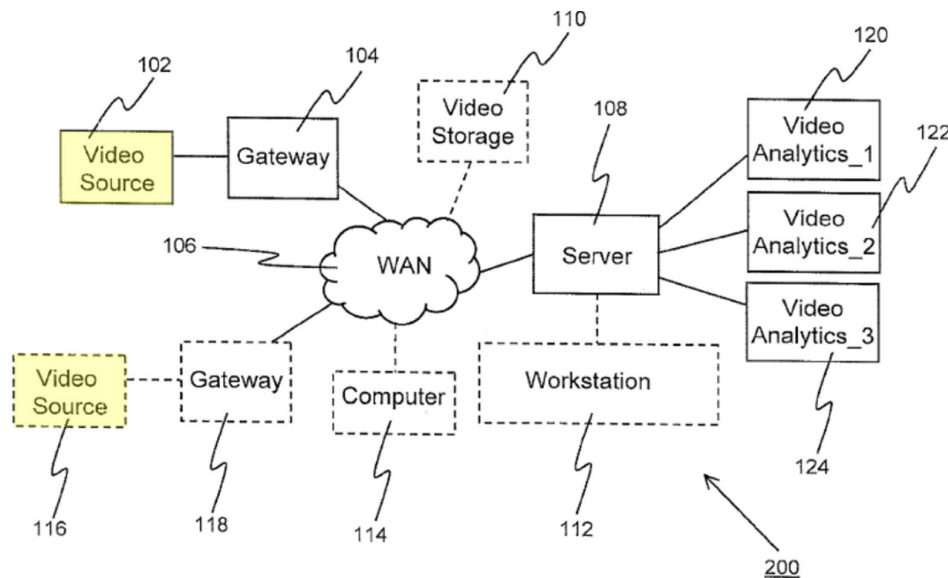


Fig. 2

Id., Fig. 2; *id.*, Fig. 1. POSITAs would have understood Laganiere363’s source-end video cameras to be “*surveillance cameras*” because they monitor parking lots, building entries, or ATMs and capture “video data relating to an event of interest.” *Id.*, ¶¶25, 34; *Houh*, ¶72.

⁴ Unless otherwise noted, all emphasis in any quoted material has been added.

Laganiere363 describes how a “video camera captures video data at a known frame rate.” Laganiere, '363, ¶34. Laganiere363 describes analyzing the video frames to read license plates and detect faces. *Id.*, ¶¶30, 41 (discussing selection of “a frame with a suitable license plate **image**”). POSITAs would have understood video frames captured by the video camera are image data. *Houh*, ¶¶73-75.

c. 1.2

Laganiere363 teaches 1.2. *Houh*, ¶76. Laganiere363 discloses an object detection system at the source end. Figure 5 (annotated below) shows a method of analyzing video data using the “systems shown in Fig. 1 and Fig. 2.” *Laganiere363*, ¶34. After receiving video data, “[a]t 502 **first video analytics** is performed on the video data **using a first processor disposed at the source end**, the first video analytics for **detecting the event of interest in a portion of the video data.**” *Id.*, ¶40.

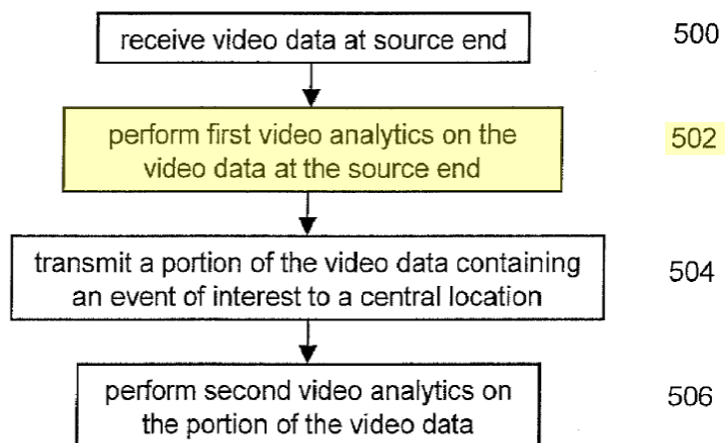


Fig. 5

As an example, “the first video analytics locates a vehicle license plate within video data captured using a video camera that monitors a parking lot entrance.” *Id.*, ¶41. The license plate is “*an object of interest within the image data*” captured by the surveillance camera. *Houh*, ¶¶77-79.

Laganiere363 further teaches “*extract[ing] portions of frames that include the detected objects of interest*” and “*send[ing] the extracted portions of the frames.*” In the license plate example, “the identified license plate is **cropped**, such that **only the region of video data that contains the identified license plate is transmitted** to the central location via WAN 106.” *Laganiere363*, ¶41; *id.*, ¶21 (describing “cropping” as an example of “pre-processing”). The transmitting process is performed at Figure 5’s operation 504 (blue, below). *Id.*, ¶40.

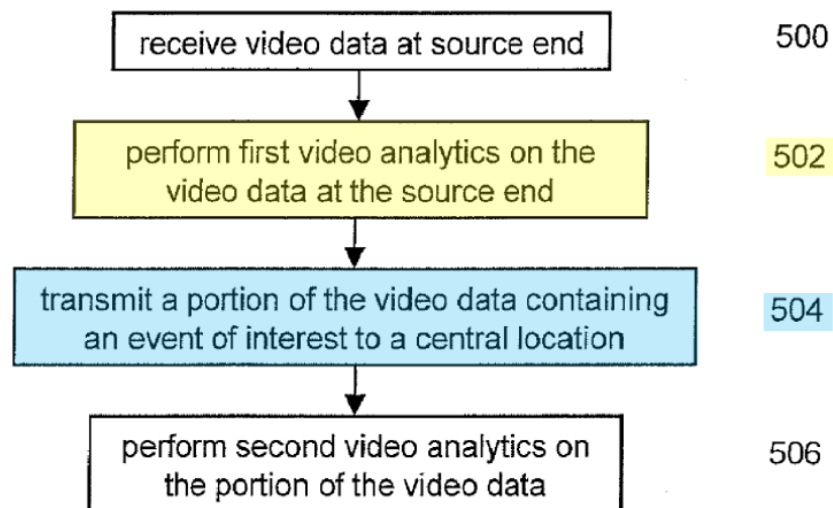


Fig. 5

Laganiere363 describes the same cropping and transmitting steps as the '321 patent to obtain the same benefit—bandwidth reduction. *Houh*, ¶¶80-82.

Laganiere '363 teaches the same or equivalent structure to that described by the '321 patent for performing the claimed functions. *Houh*, ¶82. In Laganiere363, “first video analytics” of object detection, extraction, and transmission steps are performed at the source end camera “using a first processor.” *Laganiere363*, ¶¶10, 25, 40-41. POSITAs would have understood firmware or software programming were required to perform these functions. *Houh*, ¶82.

Accordingly, Laganiere363 teaches a processor executing software disposed at a source end (“*object detection system*”) that detects an event of interest in a portion of the video data (“*objects of interest within the image data*”) and crops (“*extracts portions of frames that include the detected objects of interest*”) and transmits the cropped portion of data (“*sends the extracted portions of the frames*”). *Houh*, ¶83.

d. 1.3.1

Laganiere363 teaches 1.3.1. *Houh*, ¶84. As discussed in §VII.A.4.c, in Laganiere363’s operation 504, “**only the region of video data** that contains the identified license plate **is transmitted to the central location** via WAN 106.” *Laganiere363*, ¶¶10, 41. The video data is transmitted to the central location from a source end (e.g., video source 102/116). *Id.*, ¶¶25-29, 40-41.

In Laganiere363, a processor at a central location (“*video analytics system*”) receives the region of video data containing the identified license plate (“*image data...corresponding to objects of interest detected by the object detection system*”) from the source end (e.g., video source 102/116) (“*object detection system*”). *Id.*, ¶¶40-41; *Houh*, ¶85. The transmission of the region of video data is performed over WAN 106 (“*network*”). *Laganiere363*, ¶¶25, 40-41; *Houh*, ¶85.

Figure 5’s method is “described...with reference to the systems shown in FIG. 1 and in FIG. 2.” *Laganiere363*, ¶40. Central server 108 depicted in Figures 1-2 receives “pre-processed video data...via WAN” and the region of video data containing the identified license plate is an example of pre-processed video data. *Id.*, Abstract, ¶¶21, 40-41; *Houh*, ¶¶86-87. Indeed, “video source 102 is in communication with a central server 108 via gateway 104 and Wide Area Network (WAN) 106.” *Laganiere363*, ¶¶25, 28. Annotated Figure 1 below shows video data transmission from video sources 102/116 through WAN 106 (red) to central server 108 (purple).

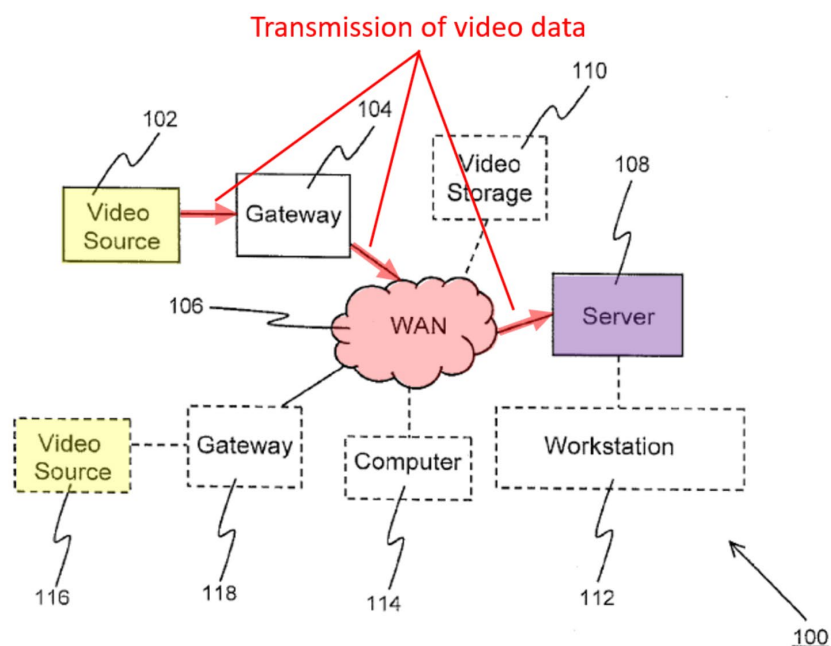


Fig. 1

Annotated Figure 2 below shows video data transmission from video sources 102/116 through WAN 106 (red) to central server 108 (purple) and to video analytics engines 120-124 (orange). *Laganiere* 363, ¶¶29-30.

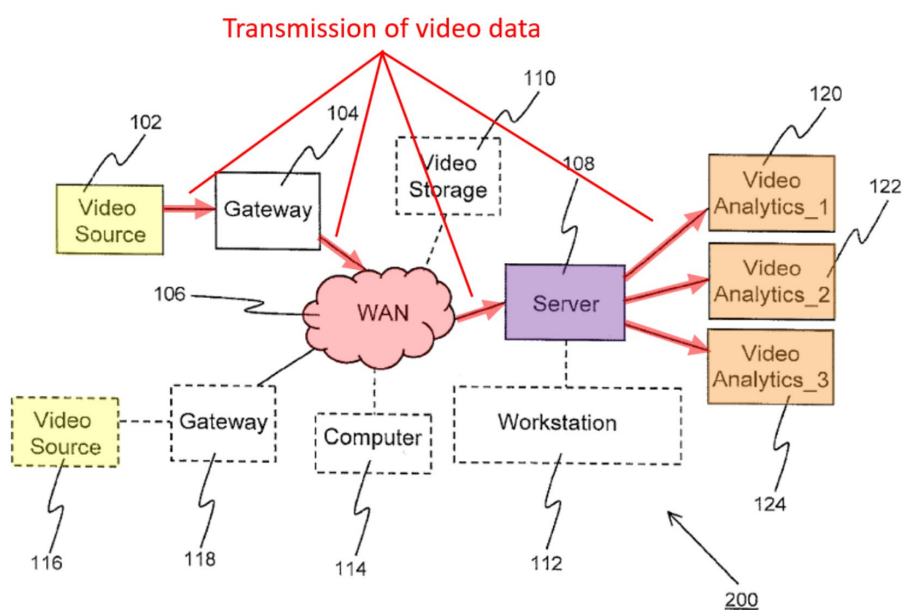


Fig. 2

Laganier363 teaches the video analytics system that receives the extracted portions of the image data may be central server 108 (i.e., a central location) which is equivalent to the '321 patent's disclosure that the video analytics system can be a node on an enterprise network. *Houh*, ¶88. Laganier363 thus teaches the same or equivalent structure to that described by the '321 patent for performing the claimed functions. Laganier363 discloses “second video analytics” performs receiving and analyzing the image data at a central location (central server 108) alone or with video analytics engines 120-124 using a second processor. *Id.*, ¶88; *Laganier363*, ¶¶10, 28-30, 40. POSITAs would have understood these systems required associated programming for receiving the image data. *Houh*, ¶¶88-89.

e. 1.3.2

Laganier363 teaches 1.3.2. *Houh*, ¶90. At operation 506 (below, purple), a **“second video analytics is performed on the portion of the video data** using the second processor at the central location, the second video analytics **for determining information relating to the event of interest.”** *Laganier363*, ¶40.

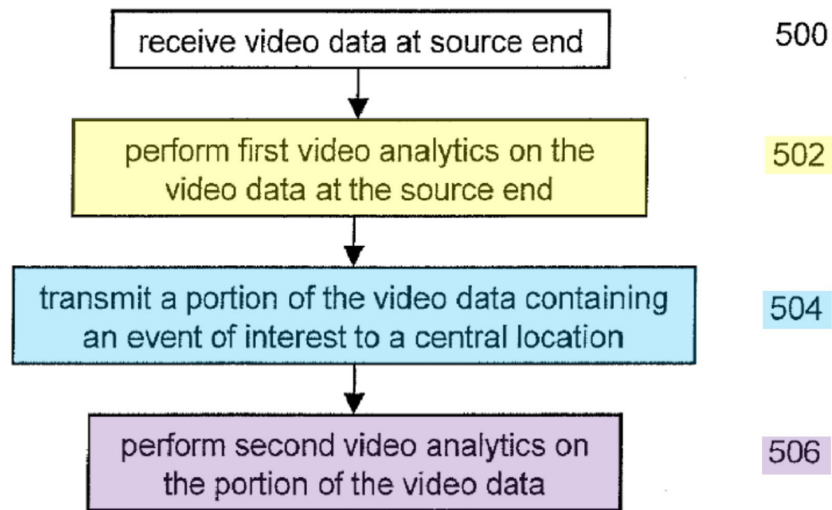


Fig. 5

As discussed in §VII.A.4.d, central server 108 (“*video analytics system*”) receives from video sources 102/116 cropped video data containing an event of interest. *Id.*, ¶¶40-41; *Houh*, ¶91. Central server 108 performs a “**second video analytics** [] for determining the license plate number” after it receives “only the region of video data that contains the identified license plate” (“*analyzing the objects of interest*”). *Laganiere363*, ¶¶40-41; *Houh*, ¶91. Alternatively, *Laganiere363* discloses central server 108 may further provide the video data to video analytics engines 120-124, performing various analysis functions such as detecting and reading a license plate and performing facial recognition. *Laganiere363*, ¶¶29-30; *Houh*, ¶91. In such case, both central server 108 and video analytics engines 120-124 form the “video analytics system.” *Id.*

Laganiere363 teaches the same or equivalent structure to that described by the '321 patent for performing the claimed function. Laganiere363's video analytics system that analyzes objects of interest may be central server 108 which utilizes video analytics engines and is equivalent to the '321 patent's disclosure that the video analytics system can be a node on an enterprise network. *Houh*, ¶92. POSITAs would have understood such systems would have required associated programming for analyzing the objects of interest. *Id.*

Accordingly, Laganiere363 teaches central server 108 alone or with video analytics engines 120-124 (“*video analytics system*”) receiving only the region of video data that contains the identified license plate (“*only the extracted portions of the frames of the image data*”) that “*correspond[] to objects of interest detected by the object detection system*”) over WAN 106 (“*transmitted over a network*”). *Houh*, ¶93. Laganiere363 further teaches video analytics engines 120-124 (“*video analytics system*”) perform video analytics processes such as reading license plate characters (“*analyzing the objects of interest*”). *Id.*

f. 1.4.1

Laganiere363 teaches 1.4.1. *Houh*, ¶94. Video sources 102/116 can each be associated with different clients and both “transmit video data over WAN 106 to a shared central processing facility, e.g., central server 108, which is capable of

performing different video analytics processing according to the individual needs of each client.” *Laganiere363*, ¶28; *Houh*, ¶94.

Accordingly, *Laganiere363* teaches central server 108 (“*video analytics system*”) receiving only the region of video data containing the identified license plate (“*image data*”) over WAN 106 (“*transmitted over the network*”). *Houh*, ¶95.

g. 1.4.2

Laganiere363 teaches 1.4.2. *Houh*, ¶96. For example, *Laganiere363* explains “some video analytics processing may be ‘**farmed out**’ to third parties if specialized video analytics engines are required” (*Laganiere363*, ¶23) and “some or all of the video analytics engines 120-124 are fee-per-use-based or subscription based” (*id.*, ¶30).

Laganiere363 further describes various benefits of “moving the processing infrastructure away from the client’s local network and ‘into the cloud,’” such as lower costs, more flexibility, and parallel processing. *Laganiere363*, ¶¶23-24; *Houh*, ¶97. Moreover, it was well-known at the relevant time that cloud services and local surveillance cameras were operated by different entities. *Houh*, ¶98.

It would have been obvious to POSITAs for central server 108 and video analytics engines 120-124 to be “*operated by a different business entity.*” *Id.*

h. 1.5.1

The Laganieri363-Laganieri817 combination teaches 1.5.1. *Houh*, ¶99. In the license plate example, Laganieri363 reduces bandwidth by selecting particular frames “with a **suitable license plate image**” and transmitting only the “one or a few **frames having a suitable quality** relative to the other frames.” *Laganieri363*, ¶41. POSITAs would have recognized this process as equally applicable to Laganieri363’s facial recognition. *Id.*, ¶30; *Houh*, ¶¶100-101. Laganieri363 further includes metadata about the pre-processing along with the video data to be processed to inform the central server how to handle the data. *Laganieri363*, ¶37. However, Laganieri363 does not describe how frames having a suitable quality are selected, what is included in the metadata, or how the central server uses the metadata. POSITAs would have thus looked to other references, such as Laganieri817, that address these teachings. *Houh*, ¶102.

Laganieri817 discloses creating a “face log” from video and/or image data captured by a camera for subsequent facial recognition analysis. *Laganieri817*, 2:20-50. Laganieri817 explains concise face logs “ensure adequate images are provided to identify an individual yet not unnecessarily consume processing or storage resources by providing excessive numbers of images.” *Id.*, 2:51-55. Moreover, Laganieri817 explains the “face log need not contain every instance of every face as it only takes one good quality image to identify an individual.” *Id.*,

2:55-57. Accordingly, “to reduce the number of faces added to a face log...only high-quality image(s) for each individual [are] stored reducing capacity requirements and operator intervention or judgment.” *Id.*, 2:59-63.

In Laganriere817, several criteria are used to select face images from the video data, “includ[ing] image characteristics such as resolution, pose, illumination, skin content and sharpness.” *Id.*, 3:35-38. These characteristics (“*factors*”) are scored (“*ranked*”) and incorporated into and used to create an “overall quality score of an image” (“*to create a ranking*”). *Id.*, 3:38-42. POSITAs would have understood resolution, illumination, and sharpness all relate to the face image’s picture quality. *Houh*, ¶¶103-104. Once these scores are calculated, they are combined to create a single quality score. *Laganriere817*, 4:52-7:9 (determining scores S1-6).

Laganriere817 explains the quality score (“*ranking*”) is used to create a face log, as detailed in Figures 7-10. *Id.*, 7:55-10:32. The face log contains only images of sufficient quality to perform facial recognition because images below a certain quality threshold are deleted from the log. *Id.*, 7:48-51, 9:61-63. Figure 11 shows an example of the face log. *Id.*, 10:34-35, Fig. 11.

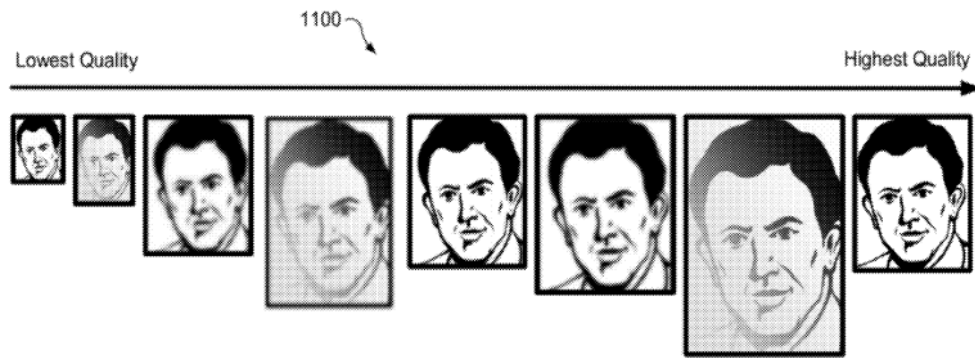


Fig. 11

Each face image in the face log is stored with “metadata with the quality score associated with the images.” *Id.*, 8:51-54. “Once the face log is created and is associated with an individual further purging or reduction of the face log can be performed by removing lower quality images as required therefore only keeping the best quality images.” *Id.*, 10:50-54.

As explained above, Laganier363 detects objects in image and video data (e.g., license plates or faces) and sends only high-quality image data to the central server for further processing. *Laganier363*, ¶¶29-30, 41. Furthermore, Laganier363 performs license plate and/or facial recognition on image data sent to the central server. POSITAs would have understood the high-quality image data related to any of license plates (e.g., vehicles) or faces (e.g., people). *Id.*, ¶30; *Houh*, ¶¶105-106. POSITAs would have understood the face log in Laganier817 was a suitable method of accomplishing Laganier363’s goals of reducing

bandwidth by transmitting only high-quality object images for further processing. *Houh*, ¶107.

POSITAs would have been motivated to use Laganriere817's face log in Laganriere363 to reduce the amount of data sent to Laganriere363's central server and to ensure only high-quality images are used for subsequent license plate or facial recognition. *Laganriere363*, ¶41; *Houh*, ¶108. Laganriere817 and Laganriere363 share inventors, applicants, content, and field of use. *Houh*, ¶108. Moreover, they also both seek to reduce the amount of data transmitted over networks without reducing the ability to perform subsequent recognition. *Laganriere363*, ¶¶34, 38; *Laganriere817*, 2:51-63; *Houh*, ¶108. Thus, POSITAs would have understood Laganriere363 and Laganriere817 were combinable and would have been motivated to combine them. *Houh*, ¶108.

POSITAs would have also been motivated to combine them because this implementation would have required nothing more than the application of known, conventional techniques (using image quality to create a group of images to further process) to yield predictable results (reducing bandwidth of transmitted data while ensuring adequate data for subsequent recognition) as described in Laganriere363 and Laganriere817. *Laganriere363*, ¶¶34, 38, 41; *Laganriere817*, 2:51-63; *Houh*, ¶109; *KSR Int'l Co., v. Teleflex, Inc.*, 550 U.S. 398, 416 (2007); *ZUP, LLC v. Nash Mfg, Inc.*, 896 F.3d 1365, 1371-73 (Fed. Cir. 2018) (upholding invalidity for obviousness

over combination of two patents to the same inventor directed to solving the same problem because the “motivation to combine may be found explicitly or implicitly in...‘interrelated teachings of multiple patents’”).

Accordingly, the Laganiere363-Laganiere817 combination teaches images containing faces (“*the image data corresponding to objects of interest detected by the object detection system*”) are scored (“*are ranked by the object detection system*”) according to image characteristics such as resolution, pose, illumination, skin content and sharpness (“*according to factors including close proximity of the objects of interest to the one or more surveillance cameras, picture quality of the objects of interest, and/or contrast of the objects of interest*”) to create an overall quality score of an image (“*to create a ranking*”). *Houh*, ¶110.

i. 1.5.2

The Laganiere363-Laganiere817 combination teaches 1.5.2. *Houh*, ¶111. As explained above, in Laganiere817, each face image in the face log is stored with “metadata with the quality score associated with the images.” *Laganiere817*, 8:51-54. POSITAs would have been motivated to combine these references’ teachings. §VII.A.4.h.

Accordingly, the Laganiere363-Laganiere817 combination teaches the metadata with quality scores associated with the images (“*the ranking saved in metadata*”) is stored with the face images (“*image data*”). *Houh*, ¶112.

j. 1.5.3

The Laganier363-Laganier817 combination teaches 1.5.3. *Houh*, ¶113.

Following transmission of the image data (i.e., face log) Laganier363's central server reads the data and performs operations on it based on the metadata. Laganier363, ¶37. Laganier817 discloses higher quality images are better for performing facial recognition. Laganier817, 1:24-26; *see Laganier363*, ¶41 (discussing selection of "suitable" license plate images for downstream reading). POSITAs would have found it obvious to combine these references' teachings for the central server to start with the highest quality images when performing facial recognition. *Houh*, ¶114. In so doing, the central server would read the rankings within the metadata to decide whether or not to analyze the image data with the same structure. *Id.*

Additionally, Laganier817 discloses purging lower quality images from the face log once a face (or license plate) has been recognized. *Laganier817*, 10:50-54. By purging lower quality images from the face log, the central server would also decide not to analyze those images based (at least in part) on their quality rankings because they would not have been analyzed to perform the recognition. *Houh*, ¶115.

If PO argues the Laganier363-Laganier817 combination does not teach the central server purges images from the face log, POSITAs would have found it obvious to reduce the number of images in the central server because in a related

application filed on the same day (Laganiere461) as Laganiere363, Laganiere461 explains for a similar system “the central server controls a process for selectively erasing portions of video data that are stored in the video storage device, based on a result of video analytics processing of the video data.” *Laganiere461*, ¶10.

Laganiere461 shares inventorship with both Laganiere363 and Laganiere817. *Houh*, ¶¶116-117. Laganiere461 describes a system similar to Laganiere363 “for reducing the amount of space that is required to store video data, based on selective deletion of portions of video data that are stored in a video storage device.” *Laganiere461*, ¶2; compare *Laganiere461*, Fig. 2, with *Laganiere363*, Fig. 1. “The selective deletion is performed under the control of a central server based on the results of video analytics processing of the video data.” *Laganiere461*, ¶2.

In Laganiere461, “under control of the central server 108, video analytics is performed on the portion of the video data, for identifying a subset of the portion of the video data that contains predetermined information relating to the event of interest.” *Id.*, ¶25; see *id.*, ¶24. Laganiere461 further explains “under control of the central server 108, portions of the retrievably stored portion of the video data, which are other than the identified subset of the portion of the video data, are deleted from the memory storage device 118.” *Id.*, ¶25.

POSITAs would have been motivated to use Laganiere461’s processing at the central server to further reduce the amount of data the central server must process,

thereby speeding up the recognition process. *Houh*, ¶¶118-119. Moreover, each of the Laganieri references seeks to reduce the amount of data required for the central server to process. *Laganieri461*, ¶28; *Laganieri363*, ¶¶34, 38; *Laganieri817*, 2:51-63; *Houh*, ¶119. POSITAs would have also been motivated to combine *Laganieri461* with *Laganieri363* and *Laganieri817* because this implementation would have required nothing more than the application of known, conventional techniques (reducing data to be processed at a server based on metadata indicating quality) to yield predictable results (processing less data) as described in *Laganieri461*, *Laganieri363*, and *Laganieri817*. *Laganieri461*, ¶¶9, 28; *Laganieri363*, ¶¶34, 38, 41; *Laganieri817*, 2:51-63; *Houh*, ¶¶119-120; *KSR*, 550 U.S. at 416; *ZUP, LLC*, 896 F.3d at 1371-73.

Accordingly, the *Laganieri363*-*Laganieri817* combination or that combination with *Laganieri461* teaches central server 108 (“*object analysis system of the video analytics system*”) that reads the metadata with quality scores associated with the images (“*reads the ranking within the metadata*”) and purges lower quality images from the face log (“*decides whether or not to analyze the image data*”).

5. Independent Claim 13

The combination teaches Claim 13 for substantially the same reasons as claim 1 (§VII.A.4) because the claims are substantively identical, and the combination further teaches steps of a method “for analyzing video in a video surveillance system”

described above. *Laganiere363*, Abstract; *Laganiere817*, Abstract; *Laganiere461*, Abstract; *compare* 1.pre-1.5.3, with 13.pre-13.5.3; *Houh*, ¶121.

Claim 13.pre is substantially similar to 1.pre, except it recites a “method” instead of a “system.” *Houh*, ¶122. As discussed in §VII.A.4, the combination describes the recited method because it discloses steps to perform the video analysis of claim 13. §VII.A.4; *Houh*, ¶122.

The table below correlates the reasons why the combination teaches the limitations of claim 1 with limitations 13.1-13.5.3. *Houh*, ¶123.

Claim 13		
Limitation	Substantially Similar Limitation	Reference
13.1	1.1	§VII.A.4.b
13.2.1	1.2	§VII.A.4.c
13.2.2	1.3.2	§VII.A.4.e
13.3	1.3.1, 1.4.2	§§VII.A.4.d, VII.A.4.g
13.4	1.4.1	§VII.A.4.f
13.5.1-13.5.3	1.5.1-1.5.3	§VII.A.4.h

6. Claims 2, 14

Claims 2 and 14 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶124. *Laganiere363* further teaches claim 2’s additional limitation. *Id.*

Laganiere363 crops image data to the object of interest. §VII.A.4.c; *Laganiere363*, ¶41; *Houh*, ¶125. Cropping regions of the video data is an example of “pre-processing.” *Laganiere363*, ¶21.

Laganiere363 teaches video sources 102 (“*object detection system*”) creating the region of frame that contains the event of interest (“*creates cropped image from the image data that include the detected objects of interest*”) and transmitting the region of video data that contains the event of interest to the central location (“*sends the cropped images to the video analytics system*”). *Houh*, ¶126.

Claim 14 is substantially similar to claim 2, except it recites a “method” instead of a “system.” *Houh*, ¶127. For the reasons discussed above for claim 2, Laganiere363 teaches claim 14’s additional limitation. *Id.*

7. Claims 3, 15

Claims 3 and 15 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶128. Laganiere363 further teaches claims 3’s additional limitation. *Id.*

POSITAs would have understood “*the image data include video data*” to mean the image data is included in video data as frames of video data. *Houh*, ¶129. Laganiere363 teaches video sources 102 (“*object detection system*”) extracting and transmitting the cropped frame of data (“*portions of the frames from the video data*”) that contains the event of interest (“*detected objects of interest*”) to the central location (“*video analytics system*”) for further processing. §VII.A.4.c.

POSITAs reading claim 3 in light of claim 1 would have understood the recited “*provides frames from the video data that include the detected objects of*

interest” as referring back to claim 1’s recitation of “extracted portions of the frames.” *Houh*, ¶130. Under this interpretation, Laganier363 teaches this limitation.

If “*frames*” is interpreted as full frames, claim 3 would be indefinite. Claim 1 of the ’321 patent, from which claim 3 depends, requires that the “*video analytics system that receives only the extracted portions of the frames of the image data.*” POSITAs would have thus understood that if **only** portions of frames can be received, then it would be impossible for full frames to be received by the video analytics system. Regardless, Laganier363 sends full frames of data to the central server for further processing. *Laganier363*, ¶¶21, 38 (sending frames at lower resolution or rate); *Houh*, ¶131. Thus, Laganier363 teaches claim 3’s additional limitation, even under this interpretation.

Claim 15 is substantially similar to claim 3, except that it recites a “*method*” instead of a “*system.*” *Houh*, ¶132. For the reasons discussed above for claim 3, Laganier363 teaches claim 15’s additional limitation. *Id.*

8. Claim 4,⁵ 16

Claims 4 and 16 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶¶133-134. Laganriere363 further teaches claims 4's additional limitation. *Id.*

The Laganriere363-Laganriere817 combination teaches central server 108 (“*video analytics system*”) including analytics engines (“*object analysis systems*”) that analyzes portions of video data containing the event of interest detected by the video source (“*image data corresponding to objects of interest detected by the object detection system*”). §VII.A.4.f.

Claim 4 contains three parallel clauses beginning with “*by*” (i.e., “*by comparing*,” “*by recognizing*,” and “*by applying*”) that are connected with an “*and/or*” conjunction. Laganriere363 teaches at least one of these three parallel clauses. *Houh*, ¶135. For example, Laganriere363 discloses each of the analytics engines 120-124 (“*object analysis systems*”) “performs a different video analytics function” (“*applying image processing functions*”) on the data frames. Laganriere363, ¶30. POSITAs would have regarded the video analytics functions such as facial recognition and reading license plate numbers as image processing

⁵ For the purpose of this IPR, Petitioner assumes “*an object analysis system*” in claims 1 and 4 are the same.

functions. *Houh*, ¶135. Accordingly, Laganieri363 teaches claim 4’s additional limitation.

Claim 16 is substantially similar to claim 4, except it recites a “*method*” instead of a “*system*.”⁶ *Houh*, ¶¶136-137. For the reasons discussed above for claim 4, Laganieri363 teaches claim 16’s additional limitation. *Id.*

9. Claims 5, 18

Claims 5 and 18 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶138. Laganieri363 further teaches claims 5’s additional limitation. *Id.*

Laganieri363 teaches multiple source ends including video sources 102 (“*one or more surveillance cameras*”) that contains a processor (“*object detection system*”) for “detecting the event of interest in a portion of the video data.” §VII.A.4.c; *Laganieri363*, ¶40.

Accordingly, Laganieri363 teaches a processor (“*object detection system*”) integrated in video source 102 (“*one or more surveillance cameras*”). *Houh*, ¶¶139-140.

⁶ Claim 16 appears to have typographical errors, in that it recites “by recognizing” and “by applying” in its recited methods. POSITAs would have realized “by” was an artifact of transposing claim 4 into method claim format. *Houh*, ¶136.

Claim 18 is substantially similar to claim 5, except it recites a “*method*” instead of a “*system*.” *Houh*, ¶141. For the reasons discussed above for claim 5, Laganier363 teaches claim 18’s additional limitation. *Id.*

10. Claims 7, 17

Claims 7 and 17 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶142. Laganier363 further teaches claims 7’s additional limitation. *Id.*

In Laganier363, the event of interest (e.g., a vehicle license plate within video data) teaches “*identified objects of interest*.” §VII.A.4.c. Further, in Laganier363, the “second video analytics is performed for determining the **license plate number**” and “[u]pon completion some or all of the frames are stored in **non-volatile memory**.” *Laganier363*, ¶41. POSITAs would have recognized a “license plate number” is metadata because it is data describing the license plate image. *Houh*, ¶143. “[T]he necessary frames and **the result are stored** in association one with another as evidence and conclusion.” *Laganier363*, ¶41. POSITAs would have been motivated to utilize a database to store the image data and result because databases were widely used for persistent storage of application data. *Houh*, ¶143. Laganier363’s central server 108 accesses a database. *Laganier363*, ¶¶35, 38. POSITAs would have known such a database was suitable for associating the image

data and result (e.g., license plate number) as Laganier363 uses tables and/or relations between database tables. *Houh*, ¶¶143-144.

Accordingly, Laganier363 teaches a database that receives the license plate number (“*receives metadata*”) associated with a vehicle license plate (“*identified objects of interest*”) from the central server 108 (“*video analytics system*”).

Claim 17 is substantially similar to claim 7, except it recites a “*method*” instead of a “*system*.” *Houh*, ¶145. For the reasons discussed above for claim 7, Laganier363 teaches claim 17’s additional limitation. *Id*

11. Claims 8, 19

Claims 8 and 19 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶146. Laganier363 further teaches claims 8’s additional limitation. *Id*.

Laganier363’s system includes a “video source 102 [] in communication with a **central server 108 via gateway 104 and Wide Area Network (WAN) 106.**” §VII.A.4.d; *Laganier363*, ¶25. Laganier363 explains “[s]ome non-limiting examples of a WAN include: a computer network such as the **Internet of the World Wide Web.**” *Id.*, ¶21. POSITAs would have understood the Internet is an example of a public network because it is open to all data traffic. *Houh*, ¶147. Figure 2 shows gateway 104 (blue) in communication with central server 108 (purple) through WAN 106 (red):

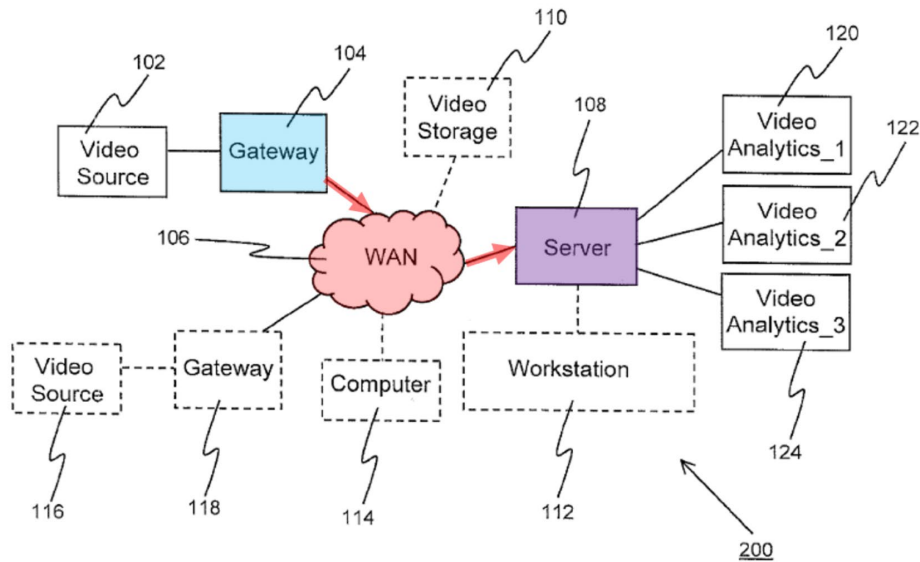


Fig. 2

Accordingly, Laganieri363 teaches gateway 104 (“gateway system”) in communication with central server 108 (“video analytics system”) over the Internet (“public network”). *Houh*, ¶148.

Claim 19 is substantially similar to claim 8, except it recites a “method” instead of a “system.” *Houh*, ¶149. For the reasons discussed above for claim 8, Laganieri363 teaches claim 19’s additional limitation. *Id.*

12. Claims 9, 20

Claims 9 and 20 depend from claims 8 and 19, respectively, which as discussed above, the combination teaches. *Houh*, ¶150. Laganieri363 further teaches claims 9’s additional limitation. *Id.*

Laganieri363 discloses “[a]lternatively, the video source 102 is a ‘dumb’ IP camera or an analogue video camera, coupled with a not illustrated video encoder

and/or a local video analytics engine.” *Laganriere363*, ¶¶25, 29. POSITAs would have recognized under such configuration, the processor (“*object detection system*”) disposed at the source end for detecting the event of interest would be in the “video encoder and/or a local video analytics engine.” *Id.*; *Houh*, ¶151. *Laganriere363* does not limit the processor to be within video source 102; rather, the processor for performing the initial pre-processing of video data is “disposed at the source end.” *Laganriere363*, ¶¶10, 40. POSITAs would have been motivated to integrate the source end processing into gateway 104 because both are computers. *Houh*, ¶151. POSITAs would have understood consolidation would save cost and eliminate unnecessary hardware. *Id.* POSITAs would have been further motivated to integrate the processor in gateway 104 to take advantage of the known video encoding capabilities of gateway systems. *Id.* (citing *Wang*, ¶39).⁷

Claim 20 is substantially similar to claim 9, except it recites a “*method*” instead of a “*system*.” *Houh*, ¶152. For the reasons discussed above for claim 9, *Laganriere363* teaches claim 20’s additional limitation. *Id.*

⁷ Deever ’150, Deever ’965, Gali, Yang, and Wang are cited in this Petition to demonstrate POSITA’s knowledge and/or as evidence to support motivation to combine. *KSR*, 550 U.S. at 420. These exhibits are not part of the unpatentability grounds.

13. Claims 10, 21

Claims 10 and 21 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶153. Laganriere363 further teaches claims 10's additional limitation. *Id.*

In Laganriere363, “[o]ptionally, the video source 102 connects to WAN 106 without the gateway 104.” *Laganriere363*, ¶¶28, 33. Accordingly, POSITAs would have understood video source 102 incorporates a gateway for accessing the network. *Houh*, ¶154. Laganriere363 also discloses “video source 102 is a **network IP camera** with onboard video analytics capabilities, such as for instance **an AXIS 211M Network Camera**.” *Laganriere363*, ¶¶25, 29. By the time of the alleged invention, network-connectable cameras were known in the art. *Houh*, ¶154. Indeed, the AXIS 211M Network Camera could directly connect to a LAN and to the Internet. *Houh*, ¶154; EX1017 (AXIS 211M manual), 8, 25.

Accordingly, Laganriere363 teaches video source 102 (“*one or more surveillance cameras*”) includes a gateway (“*gateway system*”) that provides access to the Internet (“*public network*”). *Houh*, ¶155.

Claim 21 is substantially similar to claim 10, except it recites a “*method*” instead of a “*system*.” *Houh*, ¶156. For the reasons discussed above for claim 10, Laganriere363 teaches claim 21's additional limitation. *Id.*

14. Claim 11

Claim 11 depends from claim 1, which as discussed above, the combination teaches. *Houh*, ¶157. Laganier363 further teaches claim 11's additional limitation. *Id.*

Laganier363's central server 108 receives image data from the source end over a WAN. §VII.A.4.f. POSITAs would have known a WAN is a group of computer devices that typically share common resources within a limited geographical area for a small to medium size business (e.g., enterprise). *Houh*, ¶158. POSITAs would have thus understood an enterprise network is an example of WAN consisting of such geographically limited group of computers, and it connects devices within an entity. *Id.*

Accordingly, Laganier363 teaches video analytics engines 120-124 (“*video analytics system*”) receives video data (“*image data*”) from processor of video source 102 (“*object detection system*”) through a WAN (“*enterprise network*”). *Houh*, ¶159.

15. Claim 12

Claim 12 depends from claim 1, which as discussed above, the combination teaches. *Houh*, ¶160. Laganier363 further teaches claim 12's additional limitation. *Id.*

The '321 patent discloses “**network cloud 60 is preferably a public/service provider network such as the Internet.**”⁸ '321Pat., 5:13-14. Likewise, Laganier363 teaches “[c]loud computing is a general term for anything that involves delivering hosted services over the Internet.” *Laganier363*, ¶23. Laganier363 further discloses “[a]ccording to at least some of the described embodiments, the pre-processed video data is enhanced to be more useful when subsequent **video analytics processing is performed ‘in the cloud.’**” *Id.*, ¶24. Therefore, Laganier363 discloses central server 108 and/or video analytics engines 120-124 (“*video analytics system*”) perform video analytics “in the cloud,” and POSITAs would have understood Laganier363’s video analytics system receives the image data through a cloud network. *Houh*, ¶¶161-162. Image data is received from the processor of video source 102 (“*object detection system*”). §VII.A.14.

B. Ground 1B: Obviousness over Laganier363, Laganier817, Laganier461, and Chen (Claims 4, 6, 16)

1. Chen

Chen is analogous art to the '321 patent. *Houh*, ¶164. Chen is directed to a system and method “for assessing the quality of an image.” *Chen*, Abstract. Chen’s

⁸ Although the '321 patent uses the phrase “network cloud,” POSITAs would have understood that to also refer to a cloud network. *Houh*, ¶161.

system includes surveillance camera 502, computer 500, and external storage 506, as depicted in Figure 5 below.

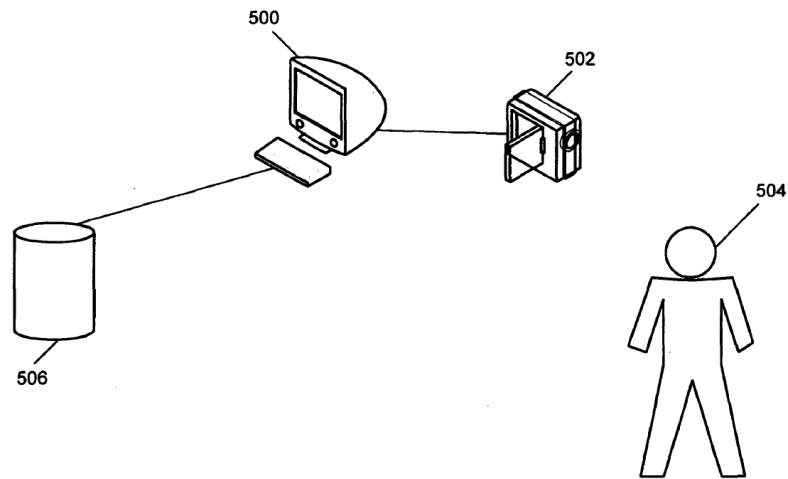


Fig. 5

Chen's Figure 6 describes a "face recognition system" to determine "the identity of the face." *Chen*, ¶¶90-92. The workflow begins with assessing a set of captured face images for image quality at 604. *Id.* At operation 610, "a most suitable face matching method is selected" based on "how well aligned the face images are" with gallery images 612. *Chen*, ¶91. Known face matching methods are selected based on the alignment. *Id.*; *Houh*, ¶¶165-167.

2. Motivation to Combine

Laganiere363 discloses performing pre-processing at the source end and sending the result to the central server for further analysis. §VII.A.4.c; *Laganiere363*, ¶¶34, 38. Central server 108 reads the metadata and "perform[s] the

additional video analytics processing of the pre-processed data.” *Id.*, ¶37. The additional video analytics can be performed by multiple different video analytics engines 120-124, each of which “performs a different video analytics function.” *Laganriere363*, ¶30. Specifically, “Video_Analytics_3 124 performs facial recognition.” *Id.* However, *Laganriere363* does not expressly disclose how a person is identified through the facial recognition process. POSITAs would have thus looked to other references for details.

POSITAs would have looked to *Chen*, which is in the same field of endeavor as *Laganriere363* and is directed specifically to performing a facial recognition process on a set of facial images. *Houh*, ¶¶168-169. *Chen* and *Laganriere363* both disclose pre-processing video data and selecting high-quality images to form a subset of images for subsequent facial recognition. *Chen*, ¶¶90-92; *Laganriere363*, ¶¶34, 41. *Chen* further discloses comparing received images with gallery images and using known face matching methods to perform the recognition. *Chen*, ¶91.

POSITAs would have been motivated to use *Chen*’s facial recognition algorithms (e.g., workflow 600) in *Laganriere363*’s video analytics engine 124 to reliably perform a facial recognition process and to successfully determine the identity of the person. *Houh*, ¶170.

POSITAs would have also been motivated to combine *Chen* with *Laganriere363*’s video analytics engines because it would have required nothing

more than the application of known, conventional techniques (using facial recognition processes) to yield predictable results (to reliably and efficiently determine identity of person) as described in Laganriere363 and Chen. *Laganriere363*, ¶¶30, 36; *Chen*, ¶¶90-92; *Houh*, ¶171; *KSR*, 550 U.S. at 416.

POSITAs would have further been motivated to combine Chen with Laganriere363's video analytics engines because it would have been applying a known technique (comparing received images to gallery images to select the most suitable facial matching method) to Laganriere363's known device (video analytics engine) ready for improvement to yield predictable results (reliably and efficiently performing a facial recognition process). *Laganriere363*, ¶¶30, 36; *Chen*, ¶¶90-92; *Houh*, ¶172; *KSR*, 550 U.S. at 416.

3. Claims 4,⁹ 16

Claims 4 and 16 depend from claims 1 and 13, respectively, which as discussed above, the combination teaches. *Houh*, ¶¶163, 173. Laganriere363 discloses each of the analytics engines 120-124 (“*object analysis systems*”) “performs a different video analytics function” (“*applying image processing functions*”) on the data frames. §VII.A.8; *Laganriere363*, ¶30; *Houh*, ¶173.

⁹ See note 5.

Chen compares a received subset of face images 608 to gallery images 612 to determine how well the faces align across the two sets of images. *Chen*, ¶91. Based on that determination, a most suitable face-matching process is selected to determine the identity of the person at operation 620. *Id.*, ¶¶91-92. Therefore, Chen discloses comparing a subset of face images 608 (“*comparing the image data*”) to gallery images (“*reference images*”) to determine the identity of the face (“*identify the objects of interest*”). *Houh*, ¶¶174-175.

4. Claim 6

Claim 6 depends from claim 1, which as discussed above, the combination teaches. *Houh*, ¶176. Chen discloses comparing a subset of face images 608 (“*comparing the image data*”) to gallery images (“*reference images*”) to determine the identity of the face (“*identify the objects of interest*”). §VII.B.3; *Houh*, ¶¶176-177.

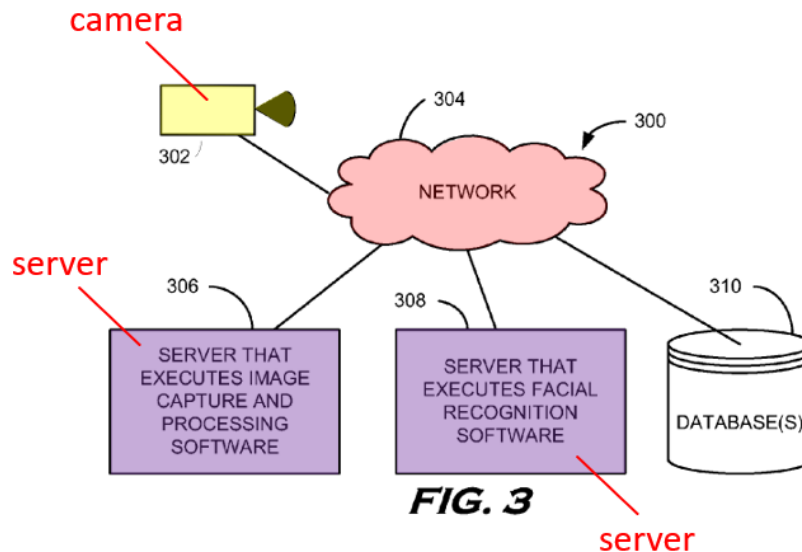
C. Ground 2: Obviousness Over Bedros (Claims 1-21)

1. Bedros

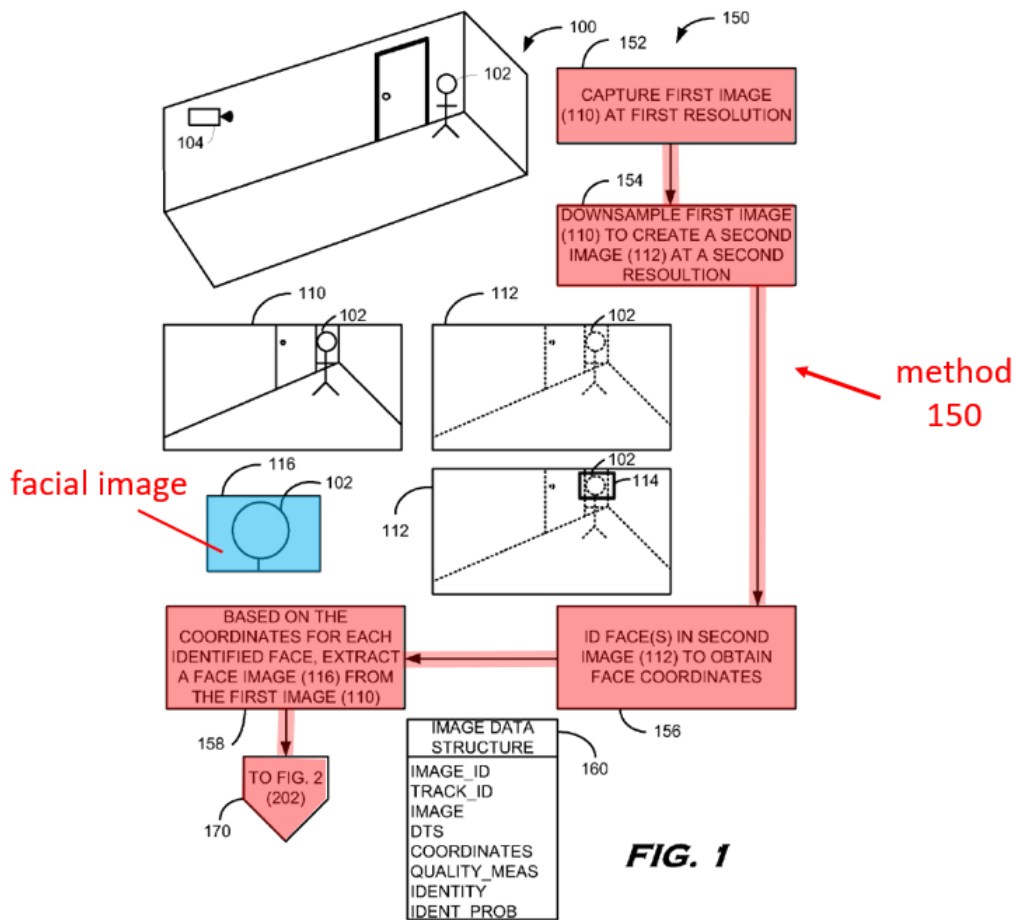
Bedros (EX1008) is analogous art to the ’321 patent. *Houh*, ¶¶178-179.

Bedros is directed to systems and methods that “capture images, locate individuals, and provide images” to a “[surveillance-based high-resolution] facial recognition process.” *Bedros*, Title, Abstract, 1:6-26, 1:43-57; *Houh*, ¶179-182. Bedros’s system includes a network-connected camera and “one or more computing

devices” (*id.*, 8:17-20) that “execute[] image capturing and processing software” (server 306) and “execute facial recognition software” (server 308) (*id.*, 8:13-49). Figure 3 shows system 300, including camera 302 (yellow) connected to servers 306 and 308 (purple) through network 304 (red). *Id.*, Fig. 3 (annotated), 8:13-49.



Bedros discloses methods 150 and 200 implemented by system 300. *Id.*, 8:14-16. Method 150 acquires and processes images before “submit[ting] to a facial recognition process.” *Id.*, 2:40-43; *see id.*, 2:61-4:29. The extracted facial image only contains the face of a person. *Id.*, Fig. 1, 3:42-46; *Houh*, ¶180. Figure 1 shows method 150 (red) and facial image 116 (blue). *Id.*, Fig. 1 (annotated), 2:36-43.



Method 200 is optionally performed following method 150 and describes “filtering and prioritizing facial images 116 for submission to the facial recognition program.” *Id.*, 4:23-29, 6:23-26. “[A] quality measure is applied [] to the facial image 116” by evaluating various features of facial image 116, and “a composite score may be assigned” to facial image 116. *Id.*, 6:4-16. Figure 2 shows processing elements 222, 224, and 226 (blue) prioritizing facial images based on quality scores and determining which facial images to submit first into image queue. *Id.*, 6:39-7:25.

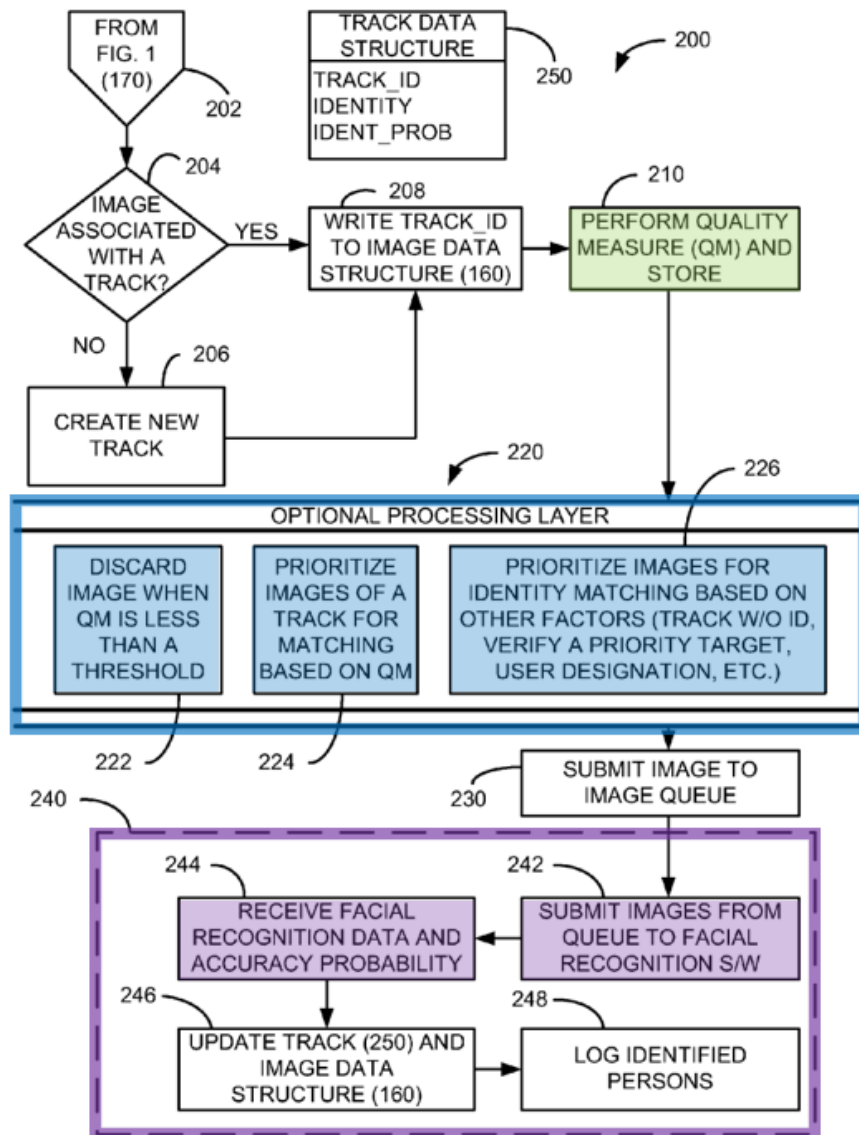


FIG. 2

Bedros explains submitting the image to the queue “may include removing and reordering” images based on “processing outcomes.” *Id.*, 7:21-25. Images are submitted from the queue and the facial recognition program. *Id.*, 7:26-30.

2. Independent Claim 1

a. 1.pre

If limiting, Bedros teaches 1.pre. *Houh*, ¶183-186. Bedros’s “systems, methods, and software” capture and provide images to a “[surveillance-based high-resolution] facial recognition process.” *Bedros*, Title, Abstract, 1:6-26, 1:43-57. Bedros’s “system” includes “camera 302 that is connected to a network 304” (*id.*, 8:17-18) and “one or more computing devices that operate server process and database 310 systems” (*id.*, 8:18-20).

b. 1.1

Bedros teaches 1.1. *Houh*, ¶187. Camera 104 monitoring scene 100 “may be a digital video camera or digital still **image camera** with suitable optics to **capture images of scene** as needed.” *Bedros*, 2:44-50. Specifically, method 150 (Figure 1) begins with “capturing 152 a first image 110 acquired by a camera 104 that monitors a scene 100.” *Id.*, 2:44-46. POSITAs would have understood captured images of the scene include “*image data*.” *Houh*, ¶187-189; ’321*Pat.*, 5:29-34. Moreover, Bedros explains captured images are stored in an “image data structure.” *Bedros*, 2:40-43, 3:56-58; *Houh*, ¶187. POSITAs would have understood Bedros’s camera 104 to be a “*surveillance camera*” because it monitors scenes of interest, such as “the presence of a person 102” in a room. *Bedros*, 2:44-46, Fig. 1; *Houh*, ¶190.

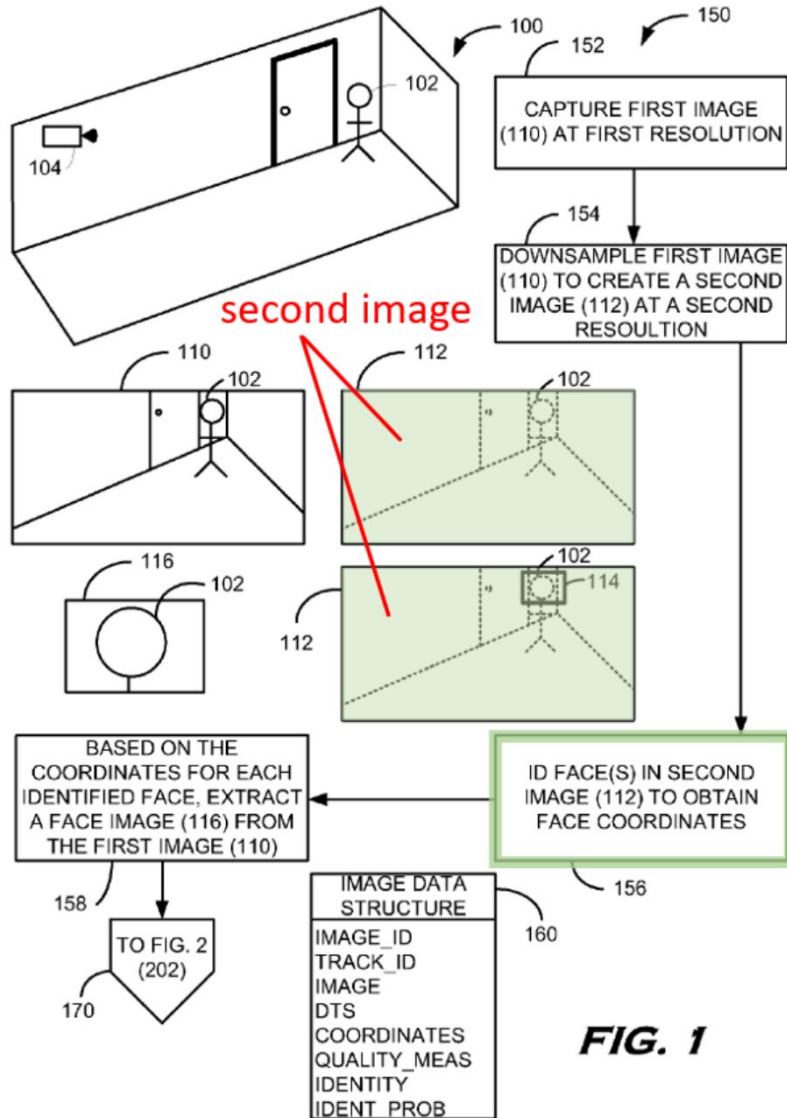
Accordingly, Bedros teaches at least one camera 104 capturing images of scene (“*one or more surveillance cameras that capture image data*”). *Houh*, ¶191.

c. 1.2

Bedros teaches 1.2. *Houh*, ¶192. Bedros’s method 150 includes detecting an individual’s face within a captured image (“*detects objects of interest within the image data*”), extracting a face image from the captured image (“*extracts portions of frames that include the detected objects of interest*”), and sending the face image to the facial recognition program (“*sends the extracted portions of the frames*”). *Bedros*, 2:36-43, 3:19-55, 7:19-25; *Houh*, ¶192.

In operation 152 of method 150, “a first image 110 [is] acquired by a camera 104 that monitors a scene 100.” *Bedros*, 2:44-46. POSITAs would have understood image 110 constitutes a frame of image data at least because image 110 is obtained through “an image or video stream received from the camera 104.” *Id.*, 2:61-62, 4:8-11, 4:20-22; *Houh*, ¶193. Method 150 proceeds to “downsampling 154” where the resolution of first image 110 is downsized to form second image 112. *Bedros*, 3:8-27. Therefore, first and second images 110 and 112 contain the same content (e.g., both depict the same image data) but have different resolutions. *Houh*, ¶193. “[T]he image 112 is [] processed to **locate 156 one or more faces present in the image 112**, if any, and to obtain face coordinates within the image 112.” *Id.*, 3:28-31. Accordingly, Bedros discloses locating a face in an image (“*detects objects of*

interest within the image data”). Houh, ¶193. Figure 1 below shows second image 112 (green shade) and capture operation 156 (green box).



System 300 implements method 150. *Bedros*, 8:13-16, Fig. 3. In particular, server 306 of system 300 “executes image capturing and processing software, such as software that implements one or both of the methods 150 and 200.” *Id.*, 8:20-22. “[S]erver 306 receives the images or video stream in a first resolution, including a

first image,” “downsample[s] an entirety of the first image to create a second image at a second resolution,” and “search[es] the second image to identify faces present in the second image.” *Id.*, 8:32-42, Fig. 3 (annotated).

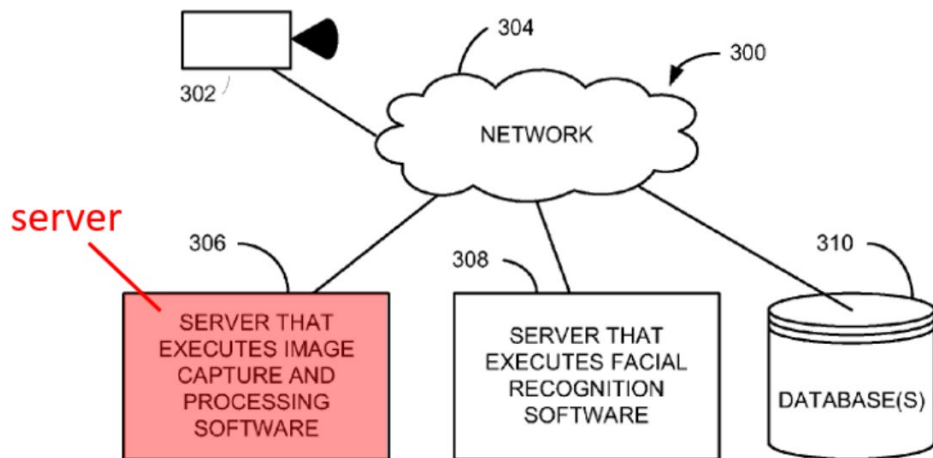
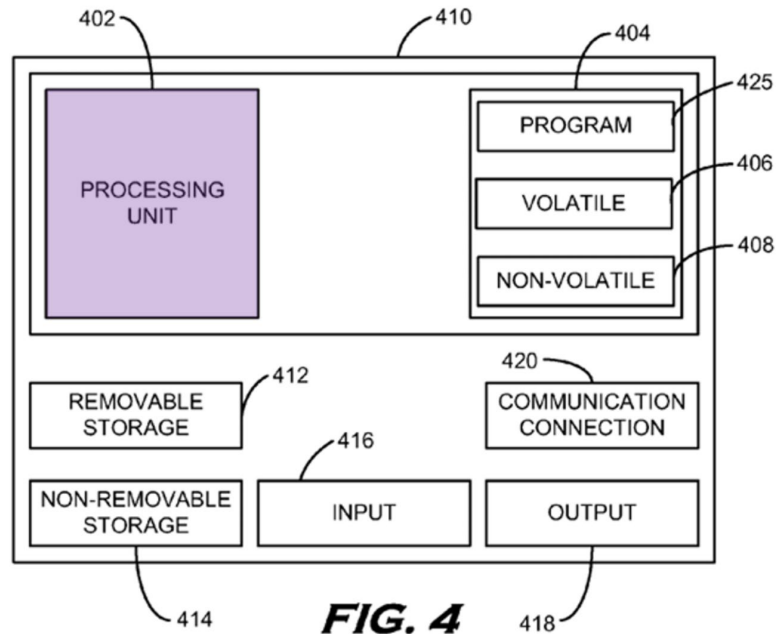


FIG. 3

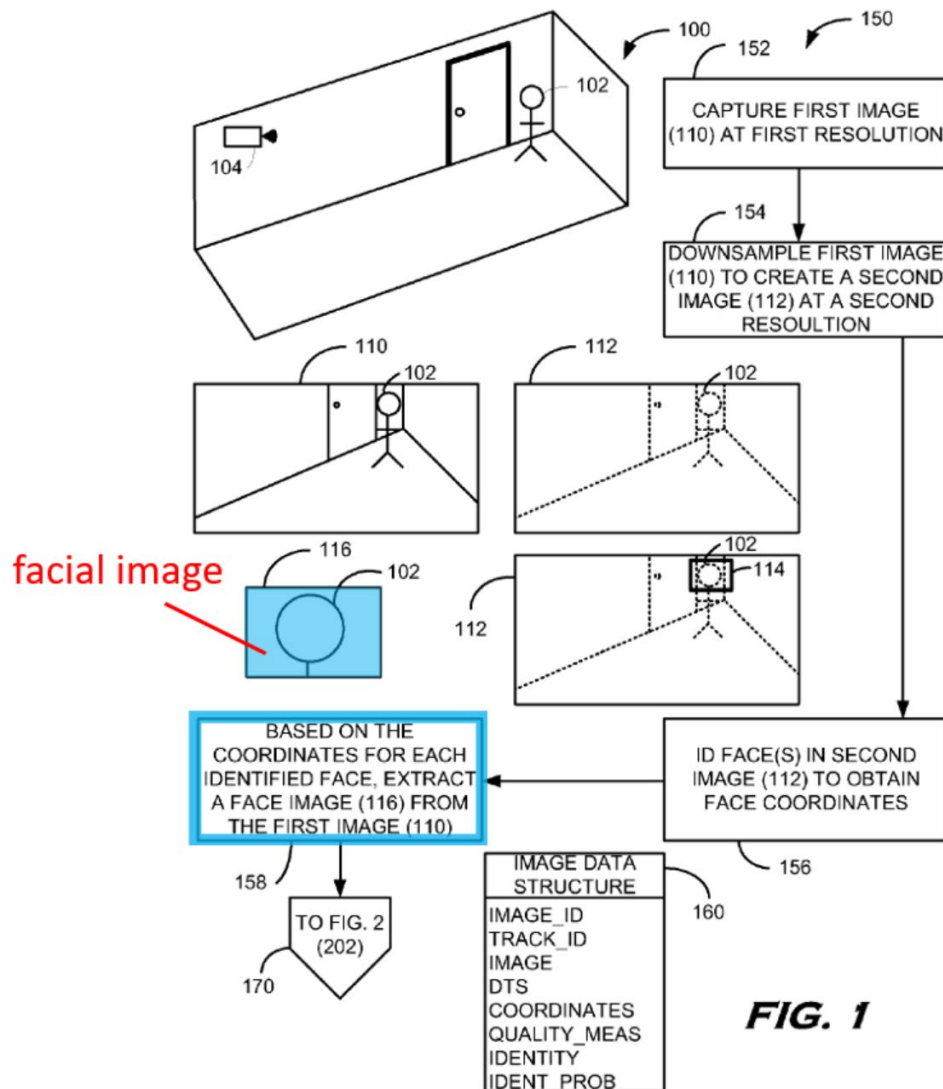
Furthermore, Bedros’s server 306 may be a computing device, which includes a processor (purple) and memory computer-readable medium for executing instructions. *Id.*, 8:50-9:2, Fig. 4 (annotated). The processor executes instructions stored on the computer-readable medium to perform the functions of methods 150 and/or 200. *Id.*, 9:26-34; *Houh*, ¶¶194-195. Accordingly, Bedros teaches server 306 (“*object detection system*”) locating one or more faces (“*detects objects of interest*”) present in the image 112 (“*within the image data*”). *Houh*, ¶195.



Bedros further discloses extracting person 102's face from first image 110.¹⁰ *Houh*, ¶196. After faces are identified in operation 156, “method 150 continues by **extracting 158 a face image 116** from the first image 110” such that “face image 116 includes the face of the person 102 include[d] in the scene 100 and the captured image 110.” *Bedros*, 3:42-46. Facial image 116 is a portion of first image 110 because only box 114 containing the person's face is kept in facial image 116 while the rest of second image 112 is excluded. *Id.*, Fig. 1; *Houh*, ¶¶196-197. Annotated

¹⁰ If PO argues identifying the face in image 112 does not “detect[] objects of interest within the image data,” because image 110 is the image data, POSITAs would have recognized extracting face image 116 from image 110 would have regardless detected the face in the image data. *Houh*, ¶196.

Figure 1 below shows facial image 116 (blue shade) and extracting operation 158 (blue box). Method 150 is implemented using server 306 depicted in Figure 3. *Bedros*, 8:13-42.



Accordingly, *Bedros* teaches server 306 (“*object detection system*”) extracts face image 116 from first image 110 (“*extracts portions of frames*”) that includes person 102’s face (“*that include the detected objects of interest*”). *Houh*, ¶198.

Bedros sends facial image 116 to a facial recognition program. *Bedros*, 4:23-29, 7:26-39; *Houh*, ¶199. Specifically, Bedros explains method 200 “**includes submitting 230 the facial image 116** to the facial image queue.” *Bedros*, 7:19-21. Method 200 is a continuation of method 150 and is implemented using Figure 3’s system. *Id.*, 8:46-49 (explaining server 306 (of Figure 3) software causes the **facial images 116 to be submitted** to server 308, executing the facial recognition software); *see id.*, 4:24-26, 8:13-20.

Accordingly, Bedros teaches submitting facial image 116 to the facial recognition software (“*sends the extracted portions of the frames*”). *Houh*, ¶200.

Bedros’s extraction processes operate in the same manner as the processes of creating and sending “cropped images” described in the ’321 patent. *Houh*, ¶201. Bedros’s extraction process removes the portion of image 110 outside of box 114, leaving only a region of the frame containing the identified face of person 102. *Bedros*, 3:42-46, Fig. 1. In the same way, “object detection system 108-1 [of the ’321 patent] can create cropped frames” that “include the detected objects of interest.” ’321*Pat.*, 9:3-5; *Houh*, ¶201.

Bedros teaches the same or equivalent structure to that described by the ’321 patent for performing the claimed functions. *Houh*, ¶202. Bedros’ server 306 including a processor and memory is equivalent to the local video analytics system 112 of the ’321 patent. *Bedros*, 8:50-9:11; *Houh*, ¶202. POSITAs would have

understood these computers would have required the same type of software programming to perform the recited detection, extracting, and sending functions. *Bedros*, 9:30-34; *Houh*, ¶202.

Accordingly, *Bedros* teaches server 306 (“*object detection system*”) locating one or more faces (“*detects objects of interest*”) present in image 112 (“*within the image data*”), extracting face image 116 from first image 110 (“*extracts portions of frames*”) that includes person 102’s face (“*that include the detected objects of interest*”), and submitting facial image 116 to the facial recognition software (“*sends the extracted portions of the frames*”). *Houh*, ¶203.

d. 1.3.1

Bedros teaches 1.3.1. *Houh*, ¶204. Facial images 116 (“*extracted portions of the frames of the image data*”) contain person 102’s face (“*objects of interest detected by the object detection system*”), and “server 306 software may cause the facial images 116 to be submit[ed]...to the server 308 that executes the facial recognition software.” §VII.C.2.c; *Bedros*, 8:46-49; *see id.*, 3:19-41; *Houh*, ¶204. POSITAs would have understood facial image 116 is submitted from server 306 (“*object detection system*”) to server 308 (“*video analytics system*”) through network 304 (“*network*”) because servers 306 and 308 are connected through network 304 and “facial images are submitted to...the facial recognition program **over a network.**” *Bedros*, Fig. 3, 11:1-3; *Houh*, ¶204.

Bedros further explains server 308 may be a computing device as depicted in Figure 4, annotated above, which includes a processor and memory or computer-readable medium for executing instructions. *Id.*, 8:50-9:2. The processor executes instructions stored on the computer-readable medium to perform the functions of methods 150 and/or 200. *Id.*, 9:26-34; *Houh*, ¶205.

Bedros teaches the same or equivalent structure to that described by the '321 patent for performing the claimed functions. Bedros's server 308 including a processor and memory is at least equivalent to the '321 patent's disclosure of the remote analytics system 122, which can be a node on a network. *Bedros*, 8:50-9:11; *Houh*, ¶206. POSITAs would have understood these systems required associated programming for receiving the image data. *Bedros*, 9:30-34; *Houh*, ¶206.

Accordingly, Bedros teaches server 308 (*"a video analytics system"*) receiving facial image 116 (*"the extracted portions of the frames of the image data"*), transmitted over network 304 (*"network"*), from server 306 (*"object detection system"*), the facial image 116 containing person 102's face (*"the image data received by the video analytics system corresponding to objects of interest detected by the object detection system"*). *Houh*, ¶207.

e. 1.3.2

Bedros teaches 1.3.2. *Houh*, ¶208. As discussed in §VII.C.2.c, server 306 submits each facial image 116 to the server 308, where the facial recognition

software is executed. *Bedros*, 8:46-49; *see id.*, 3:19-41. Specifically, facial image 116 is submitted “to the facial image queue” and subsequently submitted to “queue processing 240,” which includes “submitting 242 images from the queue to the facial recognition program and **receiving 244 facial recognition data** in response.” *Id.*, 7:19-21, 7:27-29, Fig. 2 (annotated below showing operations 240 (purple box), and 242/244 (purple shaded)). *Bedros*’s server 308 executes the facial recognition software to find “**an identity of a person** of the facial image 116” and provide “data indicative of [the] identity” as facial recognition data in operation 244. *Id.*, 7:30-35, 8:46-49, Fig. 3; *Houh*, ¶208.

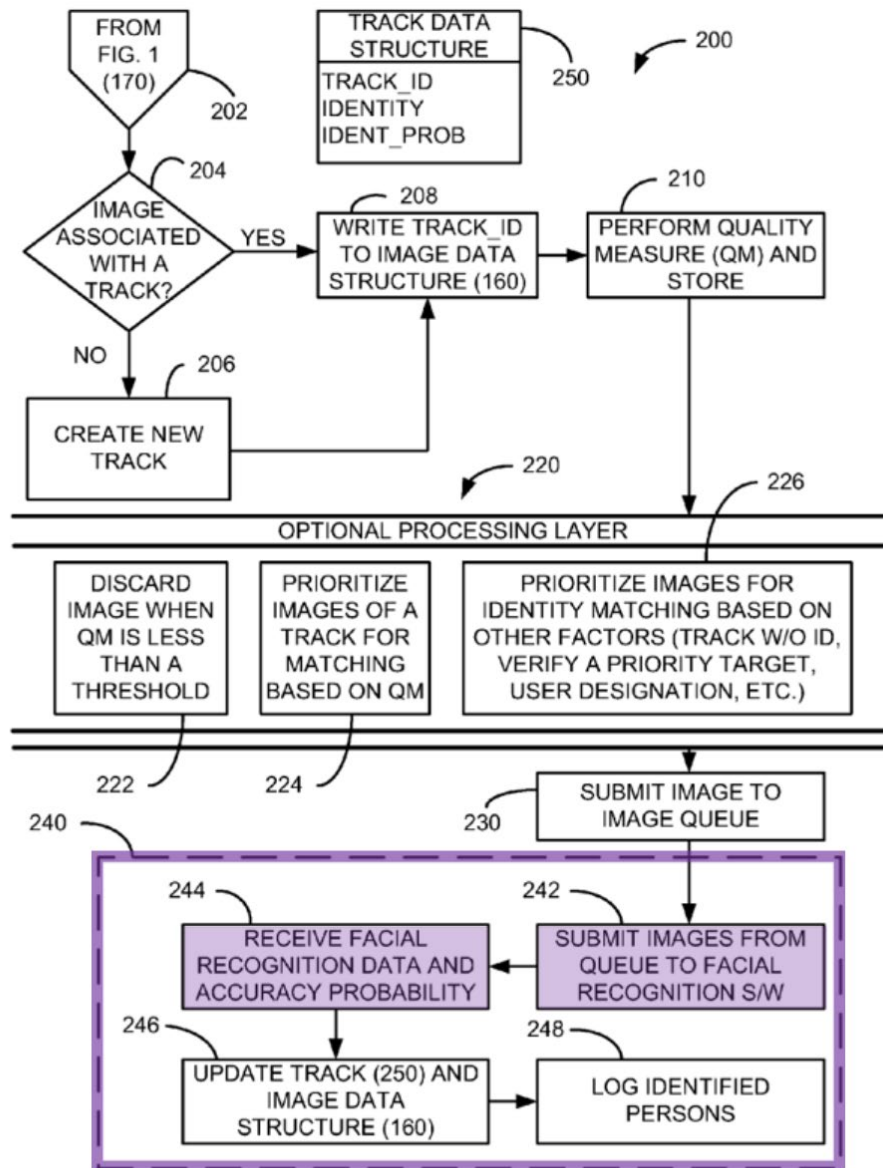
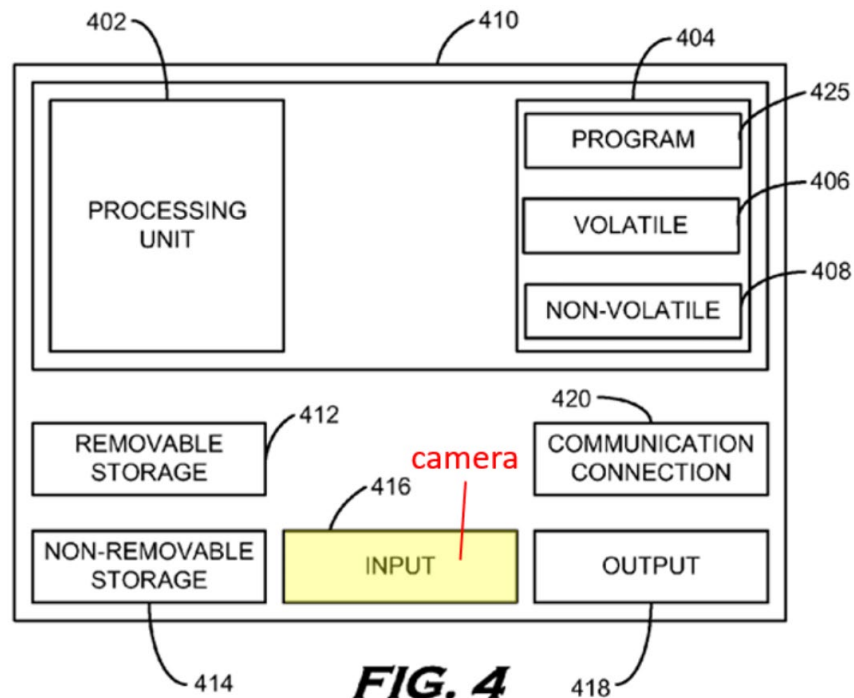


FIG. 2

Accordingly, Bedros teaches server 308 (“*video analytics system*”) executing the facial recognition software on facial image 116 to find person 102’s identity (“*analyzing the objects of interest*”). Houh, ¶209.

f. 1.4.1

Bedros teaches 1.4.1. *Houh*, ¶210. Bedros discloses server 308 (“*video analytics system*”) receives facial image 116 (“*image data*”) over a network from server 306 (“*object detection system*”). §VII.C.2.d. Bedros further discloses the computing device depicted in Figure 4 (annotated below) “may be...server[] 306.” *Bedros*, 8:55-56. Each of the computing devices can be a “computer 410,” which includes an “**input 416 [that] may include a camera** that connects directly to the computer 410.” *Id.*, 9:9-17, Fig. 4 (input 416 annotated in yellow).



If PO argues, or the Board finds, Bedros does not disclose or suggest this limitation, POSITAs would have found it obvious to utilize multiple computers 410 with connected cameras to send image data to server 308, implementing a facial

recognition system that monitors multiple locations. For example, Bedros explains prior art systems were used to identify “individuals and their locations,” but that data processing was prohibitive because they used both high- and low-resolution cameras to implement facial recognition. *Bedros*, 1:6-23, 1:43-57. However, POSITAs would have been motivated to utilize Bedros’ multiple computers with cameras to simultaneously monitor several scenes of interest (e.g., hallways and work areas) associated with different levels of human movement (e.g., remain or do not remain in the same space for extended periods), thus maintaining the ability to monitor more than one location. *Bedros*, 5:3-11, 5:26-33; *Houh*, ¶211. By monitoring multiple locations, a POSITA would have recognized the surveillance system’s usefulness is increased. *Houh*, ¶211.

POSITAs would have reasonably expected success in implementing multiple computers (i.e., servers 306) because doing so represents nothing more than the combination of known elements used according to their ordinary and expected operations. *Houh*, ¶212. This combination was well within the skill of POSITAs and would have required nothing more than the application of known, conventional techniques (utilizing more than one computer to send image data to a centralized analysis system). *Id.* Indeed, Bedros explains server 308 may be a “service” utilized by the facial recognition system over the Internet. *Bedros*, 8:27-31. POSITAs would

have thus been motivated to send multiple different image data from multiple locations to server 308 to cost-effectively utilize the service. *Houh*, ¶212.

Accordingly, Bedros teaches or suggests server 308 (“*video analytics system*”) receiving facial image 116 (“*image data*”) from multiple servers 306 (“*object detection system*”) over network 304 (“*transmitted over the network*”). *Houh*, ¶213.

g. 1.4.2

Bedros teaches 1.4.2. *Houh*, ¶¶214-215. Bedros’s server 308 is a “*video analytics system*.” §VII.C.2.d. Further, Bedros’s “facial recognition software that executes on the one or more servers 308 may be **accessible as a service over a private network, such as a virtual private network (VPN), or public network, such as the Internet.**” *Bedros*, 8:27-31. Accordingly, POSITAs would have understood Bedros’s system could be used, and would have found it obvious to use Bedros’s system, where the surveillance cameras are owned and operated by one business entity (e.g., a store or restaurant) that relies on a facial recognition service operated by a different business entity. *Houh*, ¶214. With this arrangement, POSITAs would have recognized the business using the security cameras would avoid having to create and maintain its own database with reference images of individuals to enable facial recognition. *Id.* By the time of the alleged invention, POSITAs would have been aware of the concept of software as a service (SaaS) and how it could be applied here to outsource facial recognition. *Id.*

h. 1.5.1

Bedros teaches 1.5.1. Method 200 includes an “optional processing layer 220” that can “filter or prioritize facial images 116 for submission to a facial recognition program.” *Bedros*, 6:23-26. Optional processing layer 220 filters and/or prioritizes facial images 116 (“*image data corresponding to objects of interest detected by the object detection system*”) **before** they are being submitted to an image queue at operation 230 and **before** being processed for facial recognition at operation 240. *Bedros*, 6:59-64, 7:19-30, Fig. 2; *Houh*, ¶216. Therefore, the prioritization process on facial image 116 is performed in server 306 (“*object detection system*”) before the images are submitted to server 308 (“*video analytics system*”) for facial recognition. *Houh*, ¶216.

Optional processing layer 220 (blue box) includes at least three processing elements 222, 224, and 226 (blue shaded) for prioritizing facial images. *Bedros*, 6:37-7:18, Fig. 2 (annotated below). Before optional processing layer 220 receives facial images 116, “a quality measure is applied [at] 210 to the facial image 116.” *Bedros*, 6:4-5. “The quality measure evaluates various features of the facial image 116 such as **image sharpness, an amount of present texture in the image**, and image brightness.” *Id.*, 6:6-9. POSITAs would have understood image sharpness and the amount of present texture in the image are “*factors*” referring to the picture quality of the image. *Houh*, ¶217.

Bedros further discloses “a score maybe assigned to each factor” and/or “a composite score may be assigned” to facial image 116. *Bedros*, 6:13-16; *Houh*, ¶218. Bedros explains “[a] best quality measure may be a high value or a low value depending on the particular embodiment.” *Bedros*, 6:20-22. Therefore, POSITAs would have understood Bedros to disclose or suggest facial images 116 are “*ranked*” based on the factors because the images are assigned different quality measure scores based on various factors, including image sharpness and picture quality. *Houh*, ¶218.

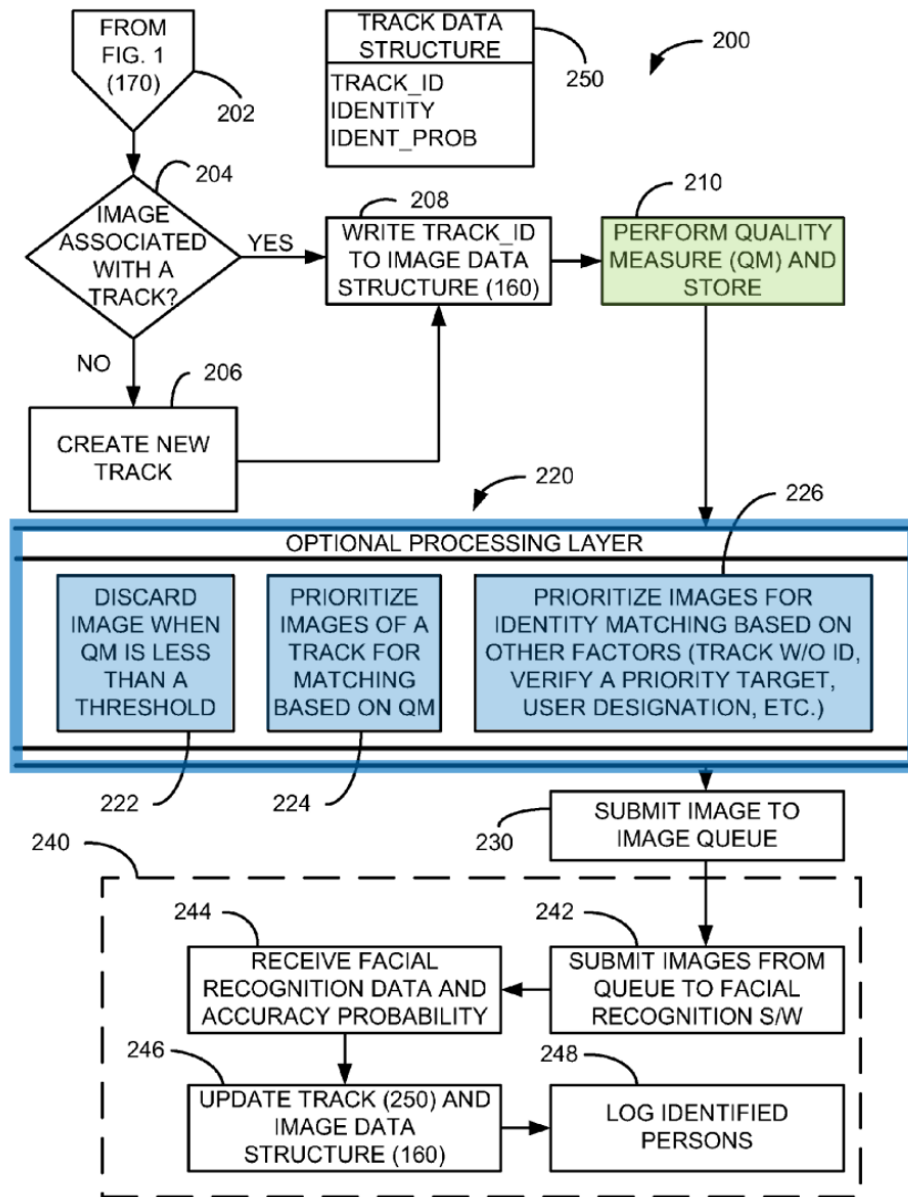


FIG. 2

Bedros further discloses processing element 222 discards facial images 116 having quality measure “less than a threshold value” (*Bedros*, 6:39-55) and processing element 224 prioritizes facial images 116 (“to create a ranking”) “based on the quality measure” and “facial images 116 placed in the queue may have a

priority assigned to them” (*id.*, 6:56-7:3). The priority determines which facial images 116 are submitted first into image queue at operation 230. *Id.*; *Houh*, ¶219.

Accordingly, Bedros teaches facial images 116 (“*the image data corresponding to objects of interest detected by the object detection system*”) are given quality measure scores (“*are ranked by the object detection system*”) based on factors including image sharpness and an amount of present texture in the image (“*according to factors including close proximity of the objects of interest to the one or more surveillance cameras, picture quality of the objects of interest, and/or contrast of the objects of interest to create a ranking*”) and/or prioritized based on the quality measure of the images by optional processing layer 220 of server 306 (“*to create a ranking*”). *Houh*, ¶220.

i. 1.5.2

Bedros teaches 1.5.2. Facial images 116 are ranked and prioritized by assigning quality measure scores based on the scores for each feature/factor. §VII.C.2.h; *Houh*, ¶221. Bedros explains “the quality measure score or scores...**may then be stored, such as to a record of the image data structure 160 of the facial image 116.**” *Bedros*, 6:15-19. “The quality measure scores may be numeric, alphabetical, or some other data type.” *Bedros*, 6:19-22. POSITAs would have understood image data structure 160, including facial image 116, includes metadata of facial image 116 (e.g., quality measure scores “QUALITY_MEAS”)

because the quality measure scores describe the properties of the image data. *Houh*, ¶221. Metadata would have been well-known to a POSITA as “data that describes other data.” EX1010, 688; EX1011, 779; EX1012, B14; *Houh*, ¶221. Moreover, it was common practice at the relevant time to include metadata with image data to provide context for the image. *Houh*, ¶221 (citing Deever ’150, 6:10-15, 8:48-56, 13:37-56 and explaining it was well-known to store metadata in an image file describing features of the image).

Accordingly, Bedros teaches the quality measures of facial images 116 (“*ranking*”) are stored to image data structure 160 with facial image 116 (“*metadata with the image data*”). *Houh*, ¶222.

j. 1.5.3

Bedros teaches 1.5.3. Optional processing layer 220 in server 306 applies quality measure scores and/or prioritization results (“*ranked...to create a ranking*”) to facial images 116 and stores the quality measure scores and prioritization results in a record of image data structure 160 (“*ranking is saved in metadata with the image data*”). §VII.C.2.i; *Houh*, ¶¶223-224.

Multiple servers 306 submit images to server 308 that “executes the facial recognition software.” §VII.C.2.f; *Bedros*, 8:46-49. Specifically, after processing in optional processing layer 220 is complete, facial images 116 are submitted to the facial image queue (*id.*, 7:19-21), where queue processing “includes submitting 242

images from the queue to the facial recognition program and receiving 244 facial recognition data in response” (*id.*, 7:26-30). Bedros discloses “submitting 230 the facial image 116 to the queue may include **removing** and reordering facial images 116 within the queue awaiting submission to the facial recognition program.” *Bedros*, 7:21-25. Thus, certain facial images 116 may be removed from the queue of images awaiting processing by the facial recognition software as subsequent images arrive from any of servers 306. POSITAs would have understood this sort of queue processing—reordering the queue as new images arrive from different places—would have taken place at server 308 (“*video analytics system*”), where all quality scored images would be stored together for processing with the same structure. *Houh*, ¶225. POSITAs would have thus understood the facial image queue resides and is managed on server 308, where facial recognition takes place, because ranked images are received from multiple servers 306. *Id.*

If PO argues, or the Board finds, Bedros does not teach processing the facial image queue at server 308, POSITAs would have been motivated to do so at least because (1) all facial images are ultimately processed at server 308 and (2) servers 306 would not know what ranking other servers 306 assign to the images they send. *Houh*, ¶226. Processing the queue at server 308 would have been more efficient and reduced bandwidth by relieving servers 306 from reading the queue to reorder or remove facial images from it. *Id.* Performing queue processing on image data at the

server receiving information was well-known in the art. *Houh*, ¶227-229 (citing Laganier⁴⁶¹, ¶¶24, 25; *Gali*, ¶¶53-54, 94-95, Fig. 1; *Deever* '965, 14:45-15:35). POSITAs would have been further motivated to separate and delineate tasks performed at servers 306 and servers 308 to streamline operation and to reduce transmission bandwidth. *Houh*, ¶¶230-232 (citing *Yang*, ¶¶14, 20).

POSITAs would have had a reasonable expectation of success of performing queue processing at server 308 because such processing represents nothing more than the combination of known elements used according to their ordinary and expected operations. *Houh*, ¶233. This modification was well within the skill of POSITAs and would have required nothing more than the application of known, conventional techniques (utilizing more than one computer to send image data to a centralized analysis system). *KSR*, 550 U.S. at 416; *Houh*, ¶233.

Furthermore, to reorder or remove images from the queue, POSITAs would have recognized queue processing must first read the quality scores of the images in the queue. *Houh*, ¶234 (citing *Deever* '965, 14:49-60). POSITAs would have also understood the removed facial images 116 would not be further processed using the facial recognition program (“*whether or not to analyze the image data*”). *Houh*, ¶234. This selection of facial images 116 for removal is “based on processing outcomes within the optional processing layer 220.” *Bedros*, 6:49-55, 7:24-25. Therefore, analysis of facial images 116 (“*the image data*”) is determined by the

processing outcomes (e.g., quality measure score and/or prioritization processing) (“*ranking*”). *Houh*, ¶234. POSITAs would have further understood the methods executed on server 308 to perform reading the metadata and reordering or deleting image data was a subset of the overall methods performed by server 308. *Id.* As such, POSITAs would have recognized the subset of methods as the claimed “*object analysis system*.” *Id.*

Accordingly, Bedros teaches queue processing and/or recursion module (“*object analysis system*”) of server 308 (“*video analytics system*”) prioritizing facial images within the queue of images (“*decides whether or not to analyze the image data*”) based on a prioritization rule (“*reads the ranking within the metadata*”). *Houh*, ¶235.

3. Independent Claim 13

Bedros teaches claim 13 for substantially the same reasons as claim 1 (§VII.C.2) because the claims are substantively identical, and Bedros further discloses steps of a method “for analyzing video in a video surveillance system.” *Bedros*, Abstract; compare 1.pre-1.5.3, with 13.pre-13.5.3; *Houh*, ¶236.

Claim 13.pre is substantially similar to 1.pre, except it recites a “method” instead of a “system.” *Houh*, ¶237. As discussed in §VII.C.2.a, Bedros describes the recited method because it discloses steps to perform the claimed video analysis of claim 13. *Bedros*, Figs. 1-2; *Houh*, ¶237.

The table below correlates the reasons why Bedros teaches the limitations of claim 1 with limitations 13.1-13.5.3. *Houh*, ¶238.

Claim 13		
Limitation	Substantially Similar Limitation(s)	Reference
13.1	1.1	§VII.C.2.b
13.2.1	1.2	§VII.C.2.c
13.2.2	1.3.2	§VII.C.2.e
13.3	1.3.1, 1.4.2	§§VII.C.2.d, VII.C.2.g
13.4	1.4.1	§VII.C.2.f
13.5.1	1.5.1	§VII.C.2.h
13.5.2	1.5.2	§VII.C.2.i
13.5.3	1.5.3	§VII.C.2.j

4. Claims 2, 14

Claims 2 and 14 depend from claims 1 and 13, respectively, which as discussed above, Bedros teaches. *Houh*, ¶239. Bedros further teaches claim 2’s additional limitations. *Id.*

Bedros extracts facial image 116 (“*cropped images*”) from first image 110 (“*image data*”). *Bedros*, 3:42-46; *Houh*, ¶240. Person 102’s face is “*the detected objects of interest*” and method 150 extracts “a face image 116 from the first image 110” such that “face image 116 includes the face of the person 102.” §VII.C.2.c; *Bedros*, 3:42-46. Facial image 116 is a cropped image created from first image 110 because it is created by identifying box 114 containing a person’s face and altering the boundaries of first image 110 based on the coordinates of box 114 such that areas outside of box 114 is excluded. *Id.*, 3:28-55, Fig. 1; EX1016, 44; *Houh*, ¶240. The

extraction process is performed by server 306 (“*object detection system*”), which sends facial images 116 (“*cropped images*”) to server 308 (“*video analytics system*”). §VII.C.2.c; *Bedros*, 8:46-49; *Houh*, ¶240.

Claim 14 is substantially similar to claim 2, except it recites a “method” instead of a “system.” *Houh*, ¶241. For the reasons discussed above for claim 2, *Bedros* teaches claim 14. *Id.*

5. Claims 3, 15

Claims 3 and 15 depend from claims 1 and 13, respectively, which as discussed above, *Bedros* teaches. *Houh*, ¶242. *Bedros* further teaches claim 3’s additional limitation. *Id.*

POSITAs would have understood “the image data include video data” to mean the image data is included in video data as frames of video data. *Houh*, ¶243. *Bedros*’s server 306 (“*object detection system*”) extracts cropped portions of frames of video data (“*portions of the frames*”) containing a person’s face (“*detected objects of interest*”) and sends those frames to server 308 (“*video analytics system*”) for subsequent processing. §§VII.C.2.b-VII.C.2.c; *Houh*, ¶243. POSITAs reading claim 3 in light of claim 1 would have understood the recited “provides **frames from the video data** that include the detected objects of interest” as referring back to claim 1’s “extracted portions of the frames.” *Houh*, ¶243. Accordingly, if “*frames*” is interpreted as “*portions of frames*” (which it should be), *Bedros* teaches claim 3.

Claim 15 is substantially similar to claim 3, except it recites a “method” instead of a “system.” *Houh*, ¶244. For the reasons discussed above for claim 3, Bedros teaches claim 15. *Id.*

6. Claims 4¹¹, 6, 16

Claims 4, 6, and 16 depend from claims 1 and 13, respectively, which as discussed above, Bedros teaches. *Houh*, ¶245. Bedros further teaches claim 4’s additional limitation. *Id.*

Bedros teaches a subset of methods (“*object analysis system*”) in server 308 operating on image data in the image queue (“*objects of interest detected by the object detection system*”). §VII.C.2.j; *Houh*, ¶246.

Bedros’s “facial recognition program is a program that operates to receive a digital image including a face and to **respond with an identity of an individual matching the face from the received image.**” *Bedros*, 7:63-66. Facial image 116 contains person 102. §VII.C.2.c. The facial recognition program determines person 102’s identity by “**compar[ing] facial features from the received image** with a facial database” which “typically includes a database of **known facial feature landmarks of a known subject faces.**” *Id.*, 7:66-8:2. Accordingly, Bedros teaches or suggests identifying the person in the image (“*identifying objects of interest*”).

¹¹ See note 5.

Houh, ¶247. Examples of facial feature landmarks include characteristics of components of a face, such as “eyes, nose, cheekbones, jaw, facial scars.” *Bedros*, 8:3-5. The comparison process utilizes several image analysis algorithms. *Id.*, 8:5-12. Therefore, Bedros teaches “*comparing the image data to reference images*” and “*applying image processing functions*.” *Houh*, ¶247.

Claims 6 and 16¹² are substantially similar to claim 4, except claim 16 recites a “method” instead of a “system” and both claims 6 and 16 recite the “video analytics system” performs the recited function of “*comparing the image data to reference images*.” *Houh*, ¶248. If claim 16 requires performance of additional functions, Bedros further teaches “*applying image processing functions*.” *Id.* As explained above, Bedros’ subset of methods is part of server 308 (“*video analytics system*”). Accordingly, POSITAs would have understood server 308 performed the methods. *Id.* For the reasons discussed above for claim 4, Bedros teaches claims 6 and 16. *Id.*

7. Claims 5, 18

Claims 5 and 18 depend from claims 1 and 13, respectively, which as discussed above, Bedros teaches. *Houh*, ¶249. Bedros further teaches claim 5’s additional limitation. *Id.*

¹² See note 6. If the Board finds claim 16 does not have a typo, Bedros still teaches it. *Houh*, ¶247.

Bedros discloses server 306 (“*object detection system*”) and at least one camera 104 (“*one or more surveillance cameras*”). §§VII.C.2.b, VII.C.2.c; *Houh*, ¶250. Camera 104 is equipped with “suitable optics to capture images of scene as needed based on requirements for the particular embodiment.” *Bedros*, 2:47-50. Adjustments to the optics may be made “**based on commands determined by a control system** based on determined environmental changes within the assigned scene 100.” *Id.*, 2:54-60. Server 306 can be computer 410 that “executes image capturing and processing software.” *Bedros*, 8:20-21, 8:54-55. Computer 410 includes input 416 that is **directly connected to a camera**. *Id.*, 9:14-17. Therefore, computer 410 (i.e., server 306) is a control system that controls the image capture of camera 104 and adjustments to camera optics. *Houh*, ¶250. POSITAs would have recognized camera 104 requires a control system and thus with a direct connection between camera 104 and server 306 would have considered them both to be a part of the same camera system. *Id.*

If PO argues, or the Board finds, Bedros does not disclose this limitation, Bedros renders it obvious. *Houh*, ¶251. POSITAs would have recognized camera 104 requires a control system and would have been motivated to integrate server 306 into camera 104 because server 306 issues image-capturing commands to camera 104. *Houh*, ¶251. It was also well-known to POSITAs at the relevant time to integrate camera controlling hardware and software with image processing

components into surveillance cameras. *Id.* For example, IP-based cameras contain built-in video analytics that recognize significant features within a series of frames and transmit only important events. EX1017, 25-28; *Houh*, ¶251. It was also well-known that transmitting high-resolution images lead to increased network traffic, and utilizing built-in video analytics in cameras provide the benefit of reducing network strain. *Houh*, ¶251. POSITAs would have also been motivated to integrate server 306 into camera 104 because this implementation would have required nothing more than the application of known, conventional techniques. *Id.*; *KSR*, 550 U.S. at 416; *ZUP, LLC*, 896 F.3d at 1371-73. Furthermore, POSITAs would have had a reasonable expectation of success because smart IP cameras were well-known at the time. *Houh*, ¶251.

Claim 18 is substantially similar to claim 5, except it recites a “method” instead of a “system.” *Houh*, ¶252. For the reasons discussed above for claim 5, Bedros teaches claim 18. *Id.*

8. Claims 7, 17

Claims 7 and 17 depend from claims 1 and 13, respectively, which as discussed above, Bedros teaches. *Houh*, ¶253. Bedros further teaches claim 7’s additional limitation. *Id.*

Bedros discloses identifying person 102 (“*identified objects of interest*”). Bedros explains the facial recognition program (operating on server 308) receives

facial image 116 and “respond[s] with an identity of an individual matching the face from the received image.” §VII.C.6; *Bedros*, 7:63-66. That information is stored in “image data structure 160 [that] may be a data structure held in...a database table record” as “IDENTITY.” *Bedros*, 3:58-4:3, 7:45-49. Data about the image stored in image data structure 160 is metadata. §VII.C.2.i; *Houh*, ¶254.

Claim 17 is substantially similar to claim 7, except it recites a “method” instead of a “system.” *Houh*, ¶255. For the reasons discussed above for claim 7, *Bedros* teaches claim 17. *Id.*

9. Claims 8, 19

Claims 8 and 19 depend from claims 1 and 13, respectively, which as discussed above, *Bedros* teaches. *Houh*, ¶256. *Bedros* further teaches claim 8’s additional limitation. *Id.*

Bedros’s server 308 is a “*video analytics system*.” Server 306 is connected to “a computer network” through communication connection 420 of computer 410 and transmits processed image data from camera 302. §VII.C.2.d; *Bedros*, 9:9-17. In particular, *Bedros* explains server 306 may include a “router...or other common network node.” *Bedros*, 9:20-22; *Houh*, ¶257. Furthermore, images from camera 302 are sent to server 306 before being sent to server 308. *Bedros*, 8:32-49. Indeed, *Bedros* further explains computer 410 (e.g., server 306) can be directly connected to the camera and to a network via connection 420. *Bedros*, 9:14-17; *Houh*, ¶257.

POSITAs would have further recognized Bedros's server 306 as a "*gateway system*" because routers are commonly recognized as gateways or nodes. '321*Pat.*, 3:26-27; *Houh*, ¶257 (citing *Tanenbaum*, 405). Furthermore, "[t]he communication connection may include...**the Internet**, or other networks." *Bedros*, 9:22-25. Internet is a type of "public network." *Id.*, 8:30-31. Server 306 ("*gateway system*") provides access to server 308 ("*video analytics system*") and both are connected to the Internet ("*public network*"). *Houh*, ¶257.

Claim 19 is substantially similar to claim 8, except it recites a "method" instead of a "system." *Houh*, ¶258. For the reasons discussed above for claim 8, Bedros teaches claim 19. *Id.*

10. Claims 9, 20

Claims 9 and 20 depend from claims 8 and 19, respectively, which as discussed above, Bedros teaches. *Houh*, ¶259. Bedros further teaches claim 9's additional limitation. *Id.* Bedros teaches server 306 ("*object detection system*" and "*gateway system*"). §§VII.C.2.c, VII.C.9; *Houh*, ¶259.

Claim 20 is substantially similar to claim 9, except it recites a "method" instead of a "system." *Id.*, ¶260. For the reasons discussed above for claim 9, Bedros teaches claim 20. *Id.*

11. Claims 10, 21

Claims 10 and 21 depend from claims 1 and 13, respectively, which as discussed above, Bedros teaches. *Houh*, ¶261. Bedros further teaches claim 10's additional limitation. *Id.*

Server 306 (“*gateway system*”) provides access to servers 308 (“*video analytics system*”) and both are connected to the Internet (“*public network*”). §VII.C.9; *Houh*, ¶262. Furthermore, POSITAs would have recognized camera 104/302 and server 306 as part of the same camera system or would have been motivated to integrate server 306 into camera 104. §VII.C.7; *Houh*, ¶262. Indeed, network-connectable cameras were well-known in the art at the relevant time, and it was well within the skill of POSITAs to incorporate object detection into the camera. *Houh*, ¶262. Accordingly, POSITAs would have understood, or found it obvious, camera 104 (“*one or more surveillance cameras*”) includes server 306 (“*gateway system*”) that is connected to servers 308 (“*video analytics system*”) through the Internet (“*public network*”). *Id.*

Claim 21 is substantially similar to claim 10, except it recites a “method” instead of a “system.” *Id.*, ¶263. For the reasons discussed above for claim 10, Bedros teaches claim 21. *Id.*

12. Claim 11

Claim 11 depends from claim 1, which as discussed above, Bedros teaches. *Houh*, ¶264. Bedros further teaches claim 11’s additional limitation. *Id.*

Server 308 (“*video analytics system*”) receives facial image 116 (“*image data*”) from multiple servers 306 (“*object detection system*”) over network 304 (“*transmitted over the network*”). §VII.C.2.f. Network 304 can be “**an internal network 304 of an entity** implementing the system 300.” *Bedros*, 8:24-27. POSITAs would have understood the “internal network 304 of an entity” teaches an “*enterprise network*” because it connects devices within an entity. *Houh*, ¶265.

13. Claim 12

Claim 12 depends from claim 1, which as discussed above, Bedros teaches. *Houh*, ¶266. Bedros further teaches claim 12’s additional limitation. *Id.*

The ’321 patent discloses “**network cloud 60 is preferably a public/service provider network such as the Internet.**”¹³ ’321Pat., 5:13-14. Server 308 receives facial image 116 (“*image data*”) from processing units 402 of multiple servers 306 (“*object detection system*”) over the Internet (“*public network*”). §VII.C.9; *Bedros*, 8:27-31. POSITAs would have understood the Internet includes a cloud network.

¹³ Although the ’321 patent uses the phrase “network cloud,” POSITAs would have understood that to also refer to a cloud network. *Houh*, ¶266.

Houh, ¶267. Moreover, at the relevant time, accessing servers “via the cloud” was well-known and POSITAs would have understood because server 308 provides a service (facial recognition), it could be implemented as a service via the cloud. *Id.*

VIII. DISCRETIONARY DENIAL WOULD BE INAPPROPRIATE

Petitioners expect PO will not raise *Fintiv*. Although trial is currently scheduled for October 20, 2025, slightly before the expected FWD (December 2025), that trial date conflicts with case dispositive deadlines in a related matter between the parties in the same court. EX1020, 1. Petitioners have proposed to move (and PO did not oppose) the trial date to January 12, 2026, after the expected FWD deadline in this IPR. *Id.* Further, the *Fintiv* factors do not support discretionary denial here, regardless of the trial date. Interim Procedure for Discretionary Denials (June 21, 2022).¹⁴ In particular, *Fintiv* factor 2 is neutral because the current trial date is within a month of the projected FWD, and as noted above, the parties expect the trial date to be delayed. *Fintiv* factor 3 weighs in favor of institution because it is early in the district court proceeding, with little investment by the court or parties. The parties have not yet exchanged terms for claim construction, and the court has not issued any substantive orders. Under *Fintiv* factor 4, this petition challenges

¹⁴ https://www.uspto.gov/sites/default/files/documents/interim_proc_discretionary_denials_aia_parallel_district_court_litigation_memo_20220621_.pdf

non-overlapping claims of the '321 patent. Finally, given the strong (or compelling) merits of the petition, *Fintiv* factor six weighs heavily in favor of institution (or precludes discretionary denial altogether).

The Board should not exercise its discretion under §325(d) to deny institution. *Becton, Dickinson & Co. v. B. Braun Melsungen AG*, IPR2017-01586, Paper 8, 17–18 (Dec. 15, 2017) (precedential); *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6, 9–11 (Feb. 13, 2020) (precedential). This Petition’s grounds rely on five prior art references. Of those, only Chen was before the Office during prosecution. Although Chen was before the office, the Examiner did not consider Chen in combination with any of the references on which this Petition relies. Thus, Section 325(d) is inapplicable to the grounds raised in this Petition.

IX. CONCLUSION

Petitioner respectfully requests the Board institute *inter partes* review trial and cancel the challenged claims as unpatentable.

X. Mandatory Notices

A. Real Party in Interest

Petitioner identifies itself as real party in interest. Out of an abundance of caution, Petitioner discloses VusionGroup GmbH and Captana GmbH as potential real parties in interest.

B. Related Matters

To the best of Petitioner's knowledge, the '321 patent has been involved in one district court litigation, *Hanshow Technology Co., Ltd. v. SES-imagotag SA, et al.*, Case No. 2:23-cv-00617 (E.D. Tex.). EX1018.

To the best of Petitioner's knowledge, the '321 patent has not been challenged in any *inter partes* review prior to this proceeding.

C. Notice of Counsel and Service Information

Pursuant to 37 C.F.R. §§42.8(b)(3), 42.8(b)(4) and 42.10(a), Petitioner appoints the following counsel in the table below:

LEAD COUNSEL	BACK-UP COUNSEL
Jessica Kaiser (Reg. No. 58,937) kaiser-ptab@perkinscoie.com PERKINS COIE LLP 1900 Sixteenth Street, Suite 1400 Denver, CO 80202-5255 Telephone: (303) 291-2300	Matthew J. Moffa (Reg. No. 58,860) moffa-ptab@perkinscoie.com PERKINS COIE LLP 1155 Avenue of the Americas 22nd Floor New York, NY 10036-2711 Phone: 212.262.6900 Patrick McKeever (Reg. No. 66,019) mckeever-ptab@perkinscoie.com PERKINS COIE LLP 11452 EL Camino Real, Suite 3000 San Diego, CA 92130 Tel: (858) 720-5700 John C. Moulder (Reg. No. 70,490) moulder-ptab@perkinscoie.com PERKINS COIE LLP 700 Thirteenth Street, N.W. Suite 800 Washington, DC 20005-3960 Telephone: (202) 654-3337

	Yuanfeng Gao (Reg. No. 79,573) gao-ptab@perkinscoie.com PERKINS COIE LLP 700 Thirteenth Street, N.W. Suite 800 Washington, DC 20005-3960 Telephone: (202) 654-6200 Tyler Knox (Reg. No. 79,948) knox-ptab@perkinscoie.com PERKINS COIE LLP 1900 Sixteenth Street, Suite 1400 Denver, CO 80202 Tel: (303) 291-2300 Grant Kinsel <i>Pro Hac Vice</i> to be submitted kinsel-ptab@perkinscoie.com PERKINS COIE LLP 1201 Third Avenue, Suite 4900 Seattle, WA 98101-3099 Phone: 206.359.8000
--	---

Please address all correspondence to lead and back-up counsel at the addresses shown above. Petitioner consents to electronic service by e-mail.

D. Power of Attorney

A power of attorney is filed herewith according to 37 C.F.R. §42.10(b).

Respectfully submitted,

/ Jessica Kaiser /
Jessica Kaiser
Reg. No. 58,937
Attorney for Petitioner

PERKINS COIE LLP
1900 Sixteenth Street, Suite 1400
Denver, CO 80202-5255

Date: June 3, 2024

CERTIFICATE OF WORD COUNT UNDER 37 CFR § 42.24(D)

Pursuant to 37 C.F.R. §42.24(a), Petitioner hereby certifies that portions of the above-captioned Petition for *Inter Partes* Review of U.S. Patent 10,701,321, in accordance with and reliance on the word count provided by the word-processing system used to prepare this Petition, that the number of words in this paper is 13,997. Pursuant to 37 C.F.R. §42.24(a), this word count is in compliance and excludes the table of contents, table of authorities, mandatory notices under §42.8, certificate of service, certificate of word count, appendix of exhibits, and any claim listing. This word count was prepared using Microsoft Word.

Respectfully submitted,

/ *Jessica Kaiser* /

Jessica Kaiser
Reg. No. 58,937
Attorney for Petitioner

Date: June 3, 2024
PERKINS COIE LLP
1900 Sixteenth Street, Suite 1400
Denver, CO 80202-5255

CERTIFICATE OF SERVICE

The undersigned hereby certifies that true copies of the Petition for *Inter Partes* Review of U.S. Patent No. 10,701,321 and supporting materials (Exhibits and Power of Attorney) were served via overnight delivery on the Patent Owner at the correspondence address of record as listed on Patent Center:

Arch & Lake LLP
203 North LaSalle Street, Suite 2100
Chicago, IL 60601
UNITED STATES

A courtesy copy was also sent via electronic mail to Patent Owner's litigation counsel listed below:

Hao Tan
haotan@archlakelaw.com

Shen Wang
shenwang@archlakelaw.com

Respectfully submitted,

/ *Jessica Kaiser* /

Jessica Kaiser
Reg. No. 58,937
Attorney for Petitioner

Date: June 3, 2024