

IPR2024-00963

Patent Owner's Response

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

VUSIONGROUP SA,

Petitioner

v.

HANSHOW TECHNOLOGY CO., LTD.,

Patent Owner

Case No.: IPR2024-00963

Patent No. 10,701,321

PATENT OWNER'S RESPONSE

CONTENTS

I. INTRODUCTION	1
II. BACKGROUND.....	3
A. Overview of the '321 Patent	3
B. Prosecution History of the '321 Patent	6
III. ASSERTED PRIOR ART REFERENCES.....	7
A. Laganiere363 (Ex-1005)	7
B. Laganiere817 (Ex-1006)	8
C. Laganiere461 (Ex-1007)	10
D. Bedros (Ex-1008).....	11
E. Chen (Ex-1009).....	13
IV. LEVEL OF SKILL IN THE ART.....	14
V. CLAIM CONSTRUCTION.....	14
VI. PETITIONERS HAVE NOT ESTABLISHED BY A PREPONDERANCE OF THE EVIDENCE ANY CLAIM IS UNPATENTABLE	16
A. Ground 1A and 1B: Petitioner fails to show the Laganiere363, Laganiere817, Laganiere461 (and/or Chen) combination renders any challenged claim obvious	16
1. Claims 1 and 13 (and therefore all challenged claims)	16
B. Ground 2: Petitioner fails to show Bedros renders any challenged claim obvious	33
1. Claims 1 and 13 (and therefore all challenged claims)	33
VII. CONCLUSION	49

EXHIBIT LIST

Exhibit Number	Description
Ex-2001	Amended Docket Control Order
Ex-2002	Fahy Declaration in Support of Motion for Pro Hac Vice Admission
Ex-2003	Deposition transcript of Dr. Henry Houh, dated February 25, 2025
Ex-2004	Declaration of Mingshao Zhang, Ph.D., in Support of Patent Owner's Response
Ex-2005	Curriculum Vitae of Mingshao Zhang, Ph.D.

LISTING OF CHALLENGED CLAIMS¹

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 For convenience purposes only, Patent Owner adopts for this Response the numbering of claim limitations advanced by Petitioner in the Petition (Paper 2). Patent Owner does not concede the appropriateness of such grouping and numbering of claim limitations.

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	
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	
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 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 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 fails to prove that any of the challenged claims are invalid. U.S. Patent No. 10,701,321 (the “’321 Patent”) is directed to an innovative surveillance system with local and remote video analytics systems. The claimed system and method for distributed analysis of image data in a video surveillance system separates the object detection and object analysis systems across a network. The object detection system ranks the image data’s quality and transmits the ranking, through metadata, with pre-processed image data to the video analytics system. The object analysis system reads the ranking and decides whether or not to analyze the data. The inventive distributed system reduces the required network bandwidth, lowers the cost of surveillance cameras, promotes autonomy and flexibility in the central processing, and still ensures the quality of analysis.

Petitioner raises three grounds of unpatentability in its Petition. All of these grounds, however, rely on distinguishable prior art that Petitioner has failed to show renders the challenged claims unpatentable.

With respect to Ground 1A, Petitioner does not stand a reasonable likelihood of prevailing primarily because: (1) Petitioner misunderstands Laganriere817 at such a fundamental level that, even when combined with Laganriere363 and Laganriere461, the proposed combination fails to render obvious “ranked by the object detection system” and “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,” as recited by claims 1 and 13; and (2) the Petition fails to establish a legally sufficient motivation to combine Laganiere363, Laganiere817 and Laganiere461 in the convoluted disassembly-and-reassembly way proposed by Petitioner to arrive at the invention of claims 1 and 13.

For Ground 2, Petitioner does not stand a reasonable likelihood of prevailing because Bedros fails to render obvious “the video analytics system receives the image data from multiple object detection systems after being transmitted over the network” and “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,” as recited by claims 1 and 13.

Both Ground 1A and Ground 2 fail to render obvious claims 1 and 13. Because Petitioner's arguments for dependent claims 2-12 and 14-21 rely on the unproven assumption that base claims 1 and 13 are invalid, Petitioner has also failed to show that the challenged dependent claims are unpatentable under Grounds 1A, 1B, and 2. Ex-2004, ¶¶ 70 and 123. It is axiomatic that, if the prior art does not demonstrate that an independent claim is unpatentable, it cannot be used to demonstrate that any corresponding dependent claim is also unpatentable. *See Hartness Int'l Inc. v. Simplimatic Eng'g Co.*, 819 F.2d 1100, 1108 (Fed. Cir. 1987).

II. BACKGROUND

A. Overview of the '321 Patent

Video surveillance is increasingly deployed within public and private institutions and typically includes common components such as (i) surveillance cameras that capture image data that is typically then sent over one or more data networks, (ii) video analytics systems that analyze the image data to detect and determine temporal and spatial events, and (iii) control systems that control the surveillance systems and coordinate the operation of the components. *Id.*, 1:6-7, 16-26. The methods employed by video analytics systems for detecting and analyzing classes of objects in image data have become more accurate and more important in video surveillance systems, enabling sophisticated functions including object detection to determine the presence of an object or a type of object, or even higher level functions include object analysis, such as object identification. *Id.*, 1:27-36; Ex-2004, ¶26.

In prior art video surveillance systems, the surveillance cameras typically send their image data over the networks to video analytics systems located within the network for further processing. *Id.*, 1:54-59. Advances in video technology have resulted in an increase of the resolution of the image data that the surveillance cameras capture and transmit. *Id.*, 1:59-63; Ex-2004, ¶27. As a result, these prior art video surveillance systems can increase the volume of data traffic sent over the

networks. *Id.*, 1:46-48. This increased data traffic volume can limit the available bandwidth of the networks, which can cause performance problems for components of the video surveillance systems and other nodes that communicate over the networks. *Id.*, 1:48-53. Further, the video analysis methods employed by video analytics systems for analyzing the image data are typically memory processor intensive tasks. *Id.*, 1:65-2:2. This typically increases the cost of the video analytics systems, which correspondingly increases the cost of the overall video surveillance systems. *Id.*, 2:2-5. Further, analysis methods are constantly being improved and so the systems must be periodically updated, adding to the cost of the system. *Id.*, 2:5-7; Ex-2004, ¶28.

To resolve the problems of bandwidth and cost, the '321 Patent discloses an innovative video surveillance system that includes separate object detection and object analysis systems for processing and analyzing the image data. The object detection system detects objects of interest and classes of objects of interest within the image data. *Id.*, 2:11-12. The object analysis system then analyzes the image data including the detected objects of interest, in one example, to recognize (e.g. identify) the detected objects of interest. *Id.*, 2:12-15; Ex-2004, ¶29. This division can be used to lower the volume of image data transmitted over the intervening networks by locating the object detection system closer to the cameras and then sending only the “interesting” portions of the image data to the more remote object analysis system.

Id., 2:15-20; Ex-2004, ¶30. Moreover, the object detection system can rank the detected objects of interest and save the rankings in the metadata of the image data, and the object analysis system of the video analytics system can read the rankings within the metadata and decide whether or not to analyze the image data, so as to reduce the memory and processor requirements for both the object detection system and the video analytics system, thus reducing costs. *Id.*, ¶¶31-32.

An embodiment of the video surveillance system of the '321 Patent is shown in FIG. 5 of the '321 Patent, included below:

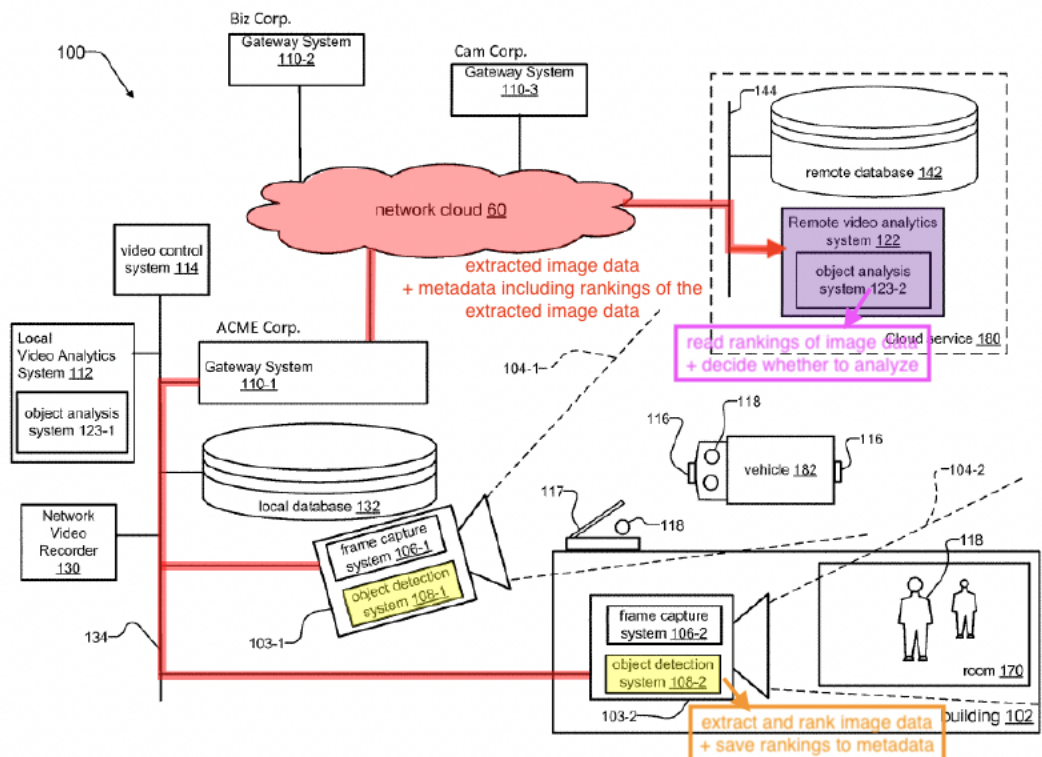


FIG. 1

Paper 2, Ex-1001, FIG. 1 with annotations (further annotated)

Unlike the prior art, the '321 Patent discloses a distributed video surveillance system that separates the object detection system (yellow) and object analytics system (purple) to achieve network (red) bandwidth savings, and allows the object analytics system to selectively process image data based on rankings generated by the object detection system and saved in the metadata of the image data. Ex-2004, ¶¶33-34. This approach reduces bandwidth usage and lowers system costs by minimizing data transmission and allowing for more efficient processing, while maintaining accurate object recognition and analysis.

B. Prosecution History of the '321 Patent

The '321 Patent issued from U.S. Application 14/634,428, filed on February 27, 2015. Ex-1001, [cover]. During prosecution, the applicant responded to various office actions, successfully establishing the patentability of the '321 Patent over prior art identified by the examiner. In particular, the applicant successfully distinguished the claims from prior art that carried out extensive video analytics on the source end of the video surveillance system. *E.g.* Ex-1002, 210-211, 268-269. Following several rounds of amendment, the patent issued on June 30, 2020. Ex-1001, [Cover].

III. ASSERTED PRIOR ART REFERENCES

A. Laganriere363 (Ex-1005)

Laganriere363 (U.S. Patent App. Pub. No. 2012/0195363) teaches “a method for performing video analytics includes receiving at a source end video data” relating to an event of interest. Ex-1005, [Abstract]. “Using video analytics, pre-processing of the video data is performed at the source end to reduce the bandwidth requirement for transmitting the video data to below a bandwidth limit of a Wide Area Network (WAN) over which the video data is to be transmitted.” *Id.* The pre-processed video data is transmitted to a central server via the WAN where other video analytics processing is performed. *Id.* Based on a result of the other video analytics processing, a signal is generated for performing a predetermined action, in response to an occurrence of the event of interest at the source end. *Id.* Optionally, the bandwidth is reduced by transmitting only one or a few frames having a suitable quality relative to the other frames. *Id.*, [0041]; Ex-2004, ¶36.

A simplified flow diagram of Laganriere363's system is shown in its FIG. 4. *Id.*, [0038]. As shown in step 408, all the actions in the later video analytics system are predetermined in nature. Ex-2004, ¶¶37-38.

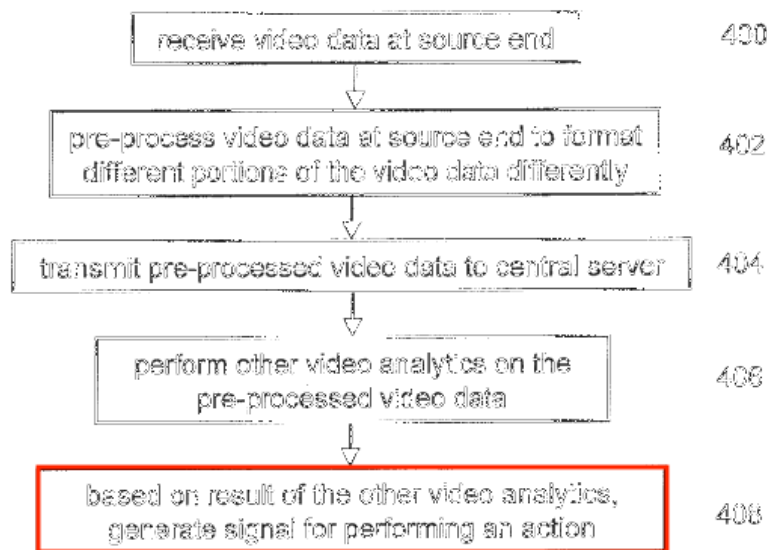


Fig. 4

Ex-1005, FIG. 4 (annotated)

B. Laganieri817 (Ex-1006)

Laganieri817 (U.S. Patent No. 8,139,817) teaches a system and method for constructing face image logs. Ex-1006, [Abstract]. Because face image logs capture a high rate of face images, any human operator or automated biometric face recognition systems trying to process the face image log in real-time may be overwhelmed. *Id.*, 1:32-37. Large face image logs also result in wasteful processing. *Id.*, 1:37-39. Accordingly, there is a need for quality appraisal of face images for the purpose of selecting high quality faces from video sequences. *Id.*, 1:53-55; Ex-2004, ¶¶39-40.

As shown in figure 1, Laganier817's system comprises a face detection module for determining face images of one or more subjects from candidate regions within each frame of the video sequence; a quality analysis module for receiving face images from the face detection module and determining a quality score based a plurality of criteria of the face images; and a storage means for storing the analyzed face images in a face image log with an associated quality score, wherein the face image log is periodically purged based upon the score of the acquired images. *Id.*, 2:29-39; Ex-2004, ¶41.

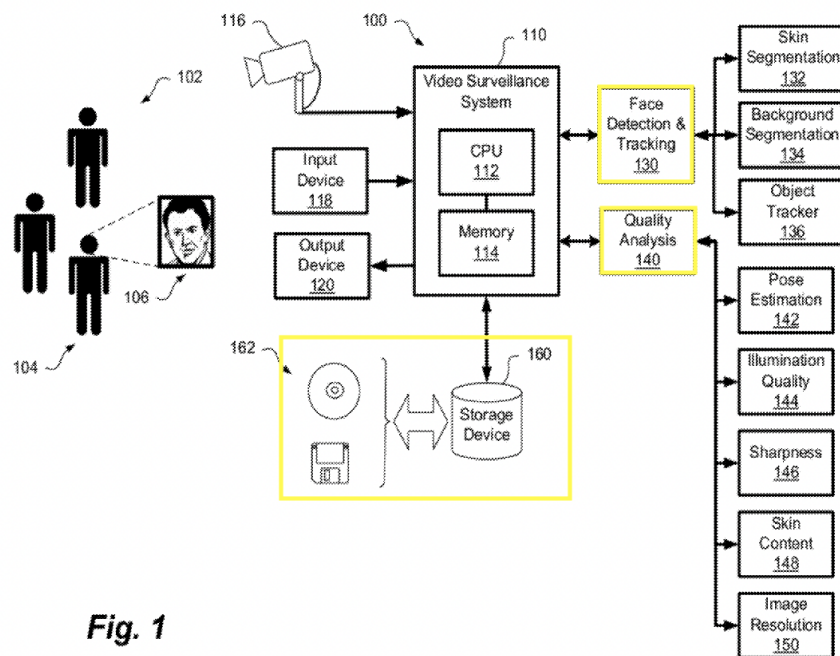


Fig. 1

Ex-1006, FIG. 1 (annotated)

Laganier817's system is designed to acquire high quality face images for each person but store only one (or very few) of them per person. For example, figure

7 of Laganier817 shows a purging process in step 722 that “[p]urge[s] face history based upon quality.” Ex-1006, FIG. 7. Thus, the resulting face image log after processing is likely intended for security personnel to review these images manually, rather than process them using a video analytics system. Ex-2004, ¶42.

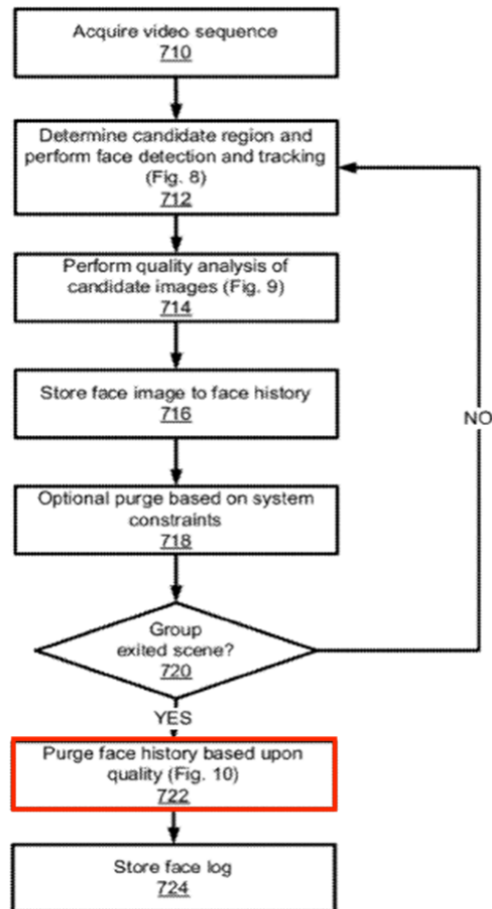


Fig. 7

Ex-1006, FIG. 7 (annotated)

C. Laganier461 (Ex-1007)

Laganier461 (U.S. Patent App. Pub. No. 2011/0113461) teaches a method and system 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. Ex-1007, [0002]. The selective deletion is performed under the control of a central server based on the results of video analytics processing of the video data. *Id.*; Ex-2004, ¶¶43-44.

A simplified flow diagram of Laganieri461's system is shown in its FIG. 4. *Id.*, [0025]; Ex-2004, ¶45.

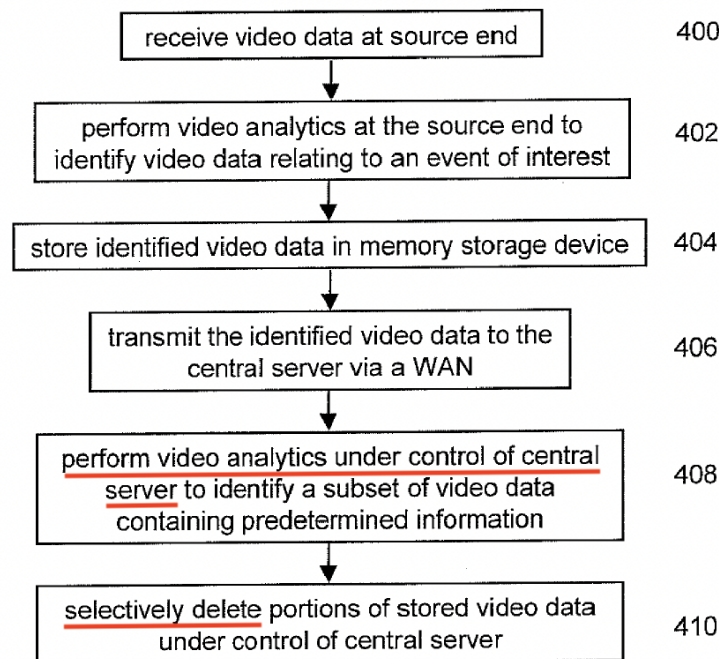


Fig. 4

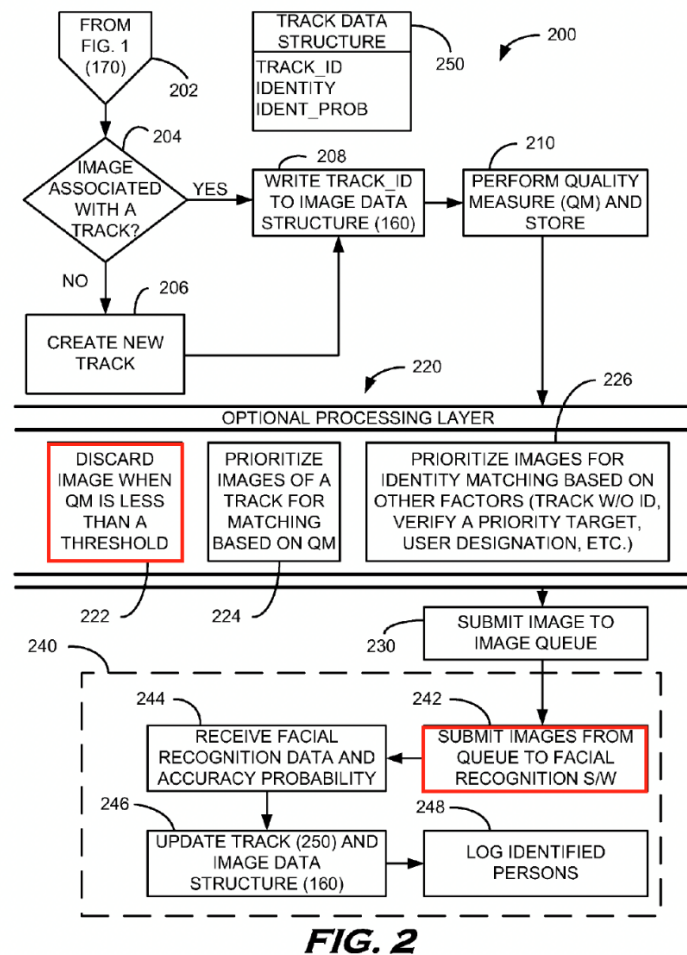
Ex-1007, FIG. 4 (annotated)

D. **Bedros (Ex-1008)**

Bedros (U.S. Patent No. 8,634,601) teaches systems, methods and software that utilizes a single high-resolution camera to capture images, locate individuals,

and provide images to a facial recognition process. Ex-1008, 1:58-62. Bedros's system includes one particular server 306 that receives digital images or a stream of digital video frames from a camera 302 over the network 304, extracts facial images 116, and optionally submits each facial image 116 to one or more servers 308 that execute facial recognition software. *Id.*, 8:20-49; Ex-2004, ¶46.

Some embodiments of Bedros's system also include prioritization and filtering that choose which captured images from a stream of images to process and when to process them when there are many images to be processed. *Id.*, 1:62-65. A method of filtering and prioritizing facial images for submission to the facial recognition system executed by the one particular server 306 is shown in its FIG. 2. *Id.*, 4:26-29, 8:20-22; Ex-2004, ¶47. As the annotation below illustrates, step 222, part of an optional processing layer 220, discards an image when the quality measure of the image is less than a threshold, *several steps before* step 242 submits images from queue to facial recognition software (and therefore *several steps before* the facial recognition software receives images). Ex-2004, ¶48.



Ex-1008, FIG. 2 (annotated)

As illustrated above, the Bedros's system follows a fixed sequence of predetermined steps, with specific decision points that cannot be arbitrarily altered during operation. Ex-2004, ¶49.

E. Chen (Ex-1009)

Chen (U.S. Patent App. Pub. No. 2014/0044348) is only used in Petitioner's ground 1B for dependent claims 4, 6, 16. Paper 2, 2. Chen teaches method, software and computer for assessing the quality of an image, for example, quality of the image

of a face that indicates the suitability of the image for use in face recognition. Ex-1009, [0002]; Ex-2004, ¶50.

IV. LEVEL OF SKILL IN THE ART

For purposes of this IPR only, Patent Owner applies Petitioner's definition of a POSITA. Paper 2, 6.

V. CLAIM CONSTRUCTION

Petitioner proposed that no formal claim constructions are necessary. Paper 2, 6. Thus, Petitioner and Patent Owner are in agreement that no construction is necessary and that the words in the claims should be given their plain and ordinary meaning. Ex-2004, ¶52; Paper 2, 6²; Ex-1003, ¶51 (Petitioner's expert claims to

² Despite Petitioner's assertion that no construction is necessary, it then halfheartedly proposed that several terms be construed as means-plus-function terms. Paper 2, 6-7; Ex-1003, ¶52. But Petitioner concedes that there is a rebuttable presumption that §112(f) does not apply to these terms and Petitioner bears the burden of going forward with evidence to rebut this presumption by a preponderance of the evidence. Paper 2, 7; *see also Apex Inc. v. Raritan Comput., Inc.*, 325 F.3d 1364, 1372 (Fed. Cir. 2003). Petitioner conspicuously failed to meet this burden, presenting solely a hypothetical argument "if the Board finds the presumption is overcome for any of the 'system' limitations." Paper 2, 7; Ex-1003, ¶52; Ex-2004,

“have given all the claim terms of the Challenged Claims their plain and ordinary meaning”).

¶53. Petitioner's ill-argued and ill-defined “constructions” cannot prevail, and the Board correctly rejected them. Paper 8, 10-11. In addition, Petitioner argues that the three “system” terms “lack[] sufficient disclosure of sufficient algorithms to perform the claimed functions.” Paper 2, 7-8; Ex-1003, ¶52; *see also* Ex-2003, 25:16-21 (“I did not see any algorithms that perform the claim functions”), 48:3-49:1. But if so, then the Board should find that no challenged claim is unpatentable because under Petitioner's proposal, the Board cannot compare corresponding structure to structure in the prior art. *See Samsung Elecs. Am., Inc. v. Prisia Eng'g Corp.*, 948 F.3d 1342, 1353 (Fed. Cir. 2020).

VI. PETITIONERS HAVE NOT ESTABLISHED BY A PREPONDERANCE OF THE EVIDENCE ANY CLAIM IS UNPATENTABLE

A. Ground 1A and 1B: Petitioner fails to show the Laganiere363, Laganiere817, Laganiere461 (and/or Chen) combination renders any challenged claim obvious

1. Claims 1 and 13 (and therefore all challenged claims)

a. Laganiere363, Laganiere817 and/or Laganiere461 fail to disclose or render obvious “ranked by the object detection system” ([1.5.1], [13.5.1]) and “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” ([1.5.3], [13.5.3])

Claim 1 recites an “object detection system” that ranks the detected image data corresponding to objects of interest according to different factors and saves the ranking in metadata with the image data, and separately, an “object analysis system” that reads the rankings within the metadata and decides whether or not to analyze the image data. Ex-2004, ¶¶54-55. Because the “object detection system” and the “object analysis system” are recited separately, they represent two separate and distinct elements of the claimed invention. *See Becton, Dickinson & Co. v. Tyco Healthcare Grp., LP*, 616 F.3d 1249, 1254 (Fed. Cir. 2010). This is supported by the plain language of claims, which requires the object detection system and the video analytics system (containing object analysis system) to be separated by, and transmit data over, a network. This separation represents an important technical distinction that saves the bandwidth of the intervening networks by sending only the portions

of the image data that includes the detected objects of interest and reduces the memory and processor requirements for both the object detection system and the video analytics systems, thus lowering system costs and allowing for more efficient processing while maintaining accurate object recognition and analysis. Ex-2004, ¶56.

Petitioner concedes that Laganriere363 does not disclose either of these, but then attempts to equate these distinct limitations to a single system for face log creation in a separate reference Laganriere817. Paper 2, 23. Petitioner's arguments are either conclusory, self-defeating, or guided by hindsight. Ex-2004, ¶57.

Petitioner maps the “object detection system” of claim 1 to a face image quality scoring in Laganriere817's face log creation and the “object analysis system” of claim 1 to purging lower quality images, again in Laganriere817's face log creation. Paper 2, 23-25, 28. However, Petitioner's own evidence and arguments demonstrate that the alleged “quality scoring” and “purging lower quality images” are not implemented by separate and distinct elements so they cannot teach or suggest the separate “object detection system” and “object analysis system” of the claims. Ex-2004, ¶58.

Figure 1 of Laganriere817 shows that the system for face tracking and face log creation includes one video surveillance system 100 associated with video camera 116. The video surveillance system 100 comprises “a central processing unit for executing software (CPU) for performing video analysis.” *Id.*, 3:55-66, 4:3-4. By

executing code stored in the memory 114, the face detection and tracking module 130 analyzes the video frame and provides face images to the quality analysis module. *Id.*, 4: 21-26. “The quality analysis module 140 analyzes the face image to assign a quality score to the image and to determine if it is of suitable quality for storage in the face log.” *Id.*, 4: 28-30. Thus, the same quality analysis module creates a quality score for a face image and determines whether to store it in the face log. Ex-2004, ¶59.

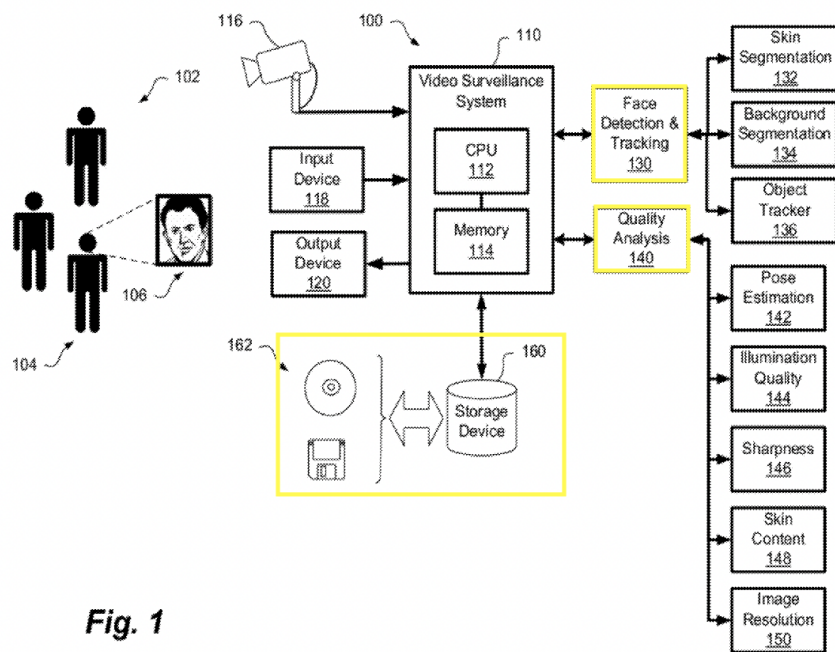


Fig. 1

Ex-1006, FIG. 1 (annotated)

Petitioner then argues that “Laganieri817 discloses purging lower quality images from the face log once a face (or license plate) has been recognized...[b]y 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.” Paper 2, 28. But nowhere does Lagnaniere817 disclose a central server separate from its video surveillance system, let alone a central server purging lower quality images from the face log based on their quality rankings (created by the video surveillance system). Moreover, the cited materials fail to teach or suggest that the alleged “quality scoring” and “purging lower quality images” are implemented by separate entities or separate hardware. *Id.* (citing Ex-1006, 10:50-54). In fact, Lagnaniere817 explicitly discloses the opposite, that “the purging of the face log may be performed by the quality analysis module 140,” that is, the same quality analysis module within the video surveillance system creates a quality score for a face image and performs the purging of lower quality images based on the quality score. Ex-1006, 9:42-45; Ex-2004, ¶60. Likewise, a POSITA would understand that the linear process described in steps in figure 7 of Lagnaniere817 is intended for execution in a single locality, not two or more distributed systems separated by a network. *Id.*, ¶61.

Further, modifying Lagnaniere817 in the way proposed by Petitioner – saving lower quality images together with quality scores in the face log for future processing – would require the face recognition system outside of Lagnaniere817's system to receive quality scores in addition to images below a certain quality threshold, only for it to read the quality scores and discard the lower quality images. This is contrary

to Lagnaniere817's explicit teaching that "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," as Petitioner recognizes. Paper 2, 24 (citing Ex-1006, 7:48-51, 9:61-63); Ex-2004, ¶62. This is also contrary to the purpose of Lagnaniere817's face log construction, that is, to reduce capacity requirements and operator intervention or judgment and *avoid overwhelming face recognition systems*. Ex-1006, 2:26-63, 3:4-6. A POSITA would not reconstruct Lagnaniere817's face log creation system only to *increase* capacity requirements and operator intervention or judgment and overwhelm *face recognition systems* with quality scores in addition to the excessive numbers of lower quality images, "unnecessarily consum[ing] [its] processing or storage resources." *Id.*, 2:51-55. In addition, a POSITA would understand that the proposed modification would allow low-quality images to arrive at face recognition systems which may or may not discard them, thus compromising the reliability of the face recognition results. Ex-2004, ¶¶63-64.

Recognizing that the Lagnaniere363-Lagniere817 combination, individually or in combination, does not teach a "central server" that purges images from the face log, Petitioner resorts to Lagnaniere461. Paper 2, 28. However, Lagnaniere461 has nothing to do with quality rankings or face logs, instead, it describes a system similar to Lagnaniere363. *Id.*, 29. Thus, Lagnaniere461, just like Lagnaniere363, does not

disclose the central server purges images “from the face log based on their quality rankings,” nor either of the limitations [1.5.1]/[13.5.1] and [1.5.3]/[13.5.3]. Ex-2004, ¶¶65-66.

Instead, the selective deletion of portions of video data in Lagniere461 is “based on a result of *video analytics processing* of the video data.” Ex-1007, [0010]. As shown in FIG. 4, the central server has to perform video analytics on all the transmitted video data before selectively deleting portions of stored video data. *Id.*, [0025], FIG. 4. Lagniere461 further discloses the *central server* performing *video analytics* “to identify a single frame of data in a sequence of frames that provides an image of the face of an individual that is of sufficient quality to enable identification of the individual.” *Id.*, [0028]. This is the opposite of the central server “read[ing] the rankings within the metadata” created by a separate object detection system before “decid[ing] whether or not to analyze the image data.” Ex-2004, ¶67.

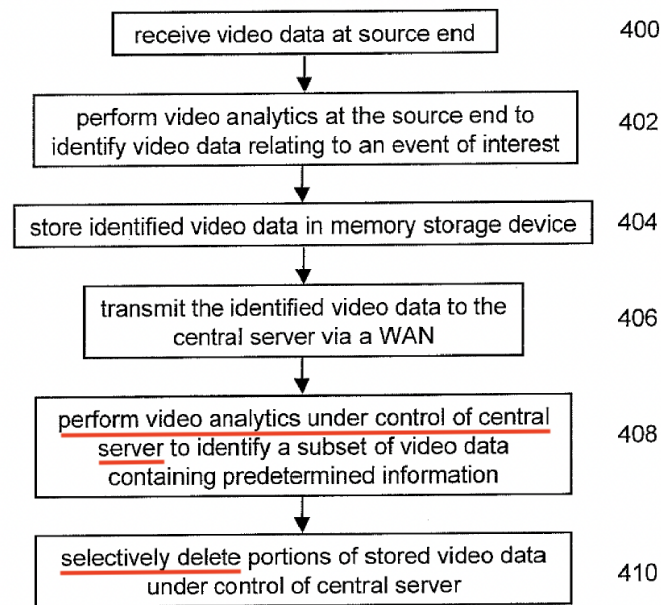


Fig. 4

Ex-1007, FIG. 4 (annotated)

Even as to Lagniere817's purported disclosure of "quality scoring" and "purging lower quality images," Petitioner could not explain how the alleged procedures are implemented by separate and distinct elements. Thus, Petitioner fails to establish that Lagniere817 discloses separate and distinct "object detection systems" and "object analysis systems" respectively implementing the limitations [1.5.1]/[13.5.1] and [1.5.3]/[13.5.3] as the claims require. Ex-2004, ¶68.

A POSITA would not view those alleged steps for face log creation to be separate in Lagniere817's system. In fact, because Laganriere817 is directed to a method and system for "generating concise and complete face logs from a video sequence" that "contain only the best images available for each individual observed"

(Ex-1006, 2:20-22, 2:29-31, [abstract]), a POSITA would understand that Laganriere817 contemplates an integral end-to-end processing from a video sequence to a video log, not two separate and distinct systems. Ex-2004, ¶69.

Petitioner therefore fails to meet its burden to show that the claimed invention was rendered obvious under Ground 1A and 1B. Ex-2004, ¶70. *Workspot, Inc. v. Citrix Systems, Inc.*, IPR2019-01002, Paper 39, 46–50 (PTAB Nov. 17, 2020) (rejecting Petitioner's challenge because Petitioner mapped two claim elements (a “content server” and “web server”) to the same component of a prior art reference without proving they were distinct); *see also Comcast Cable Communs., LLC v. Promptu Sys. Corp.*, 838 Fed. Appx. 551, 553 (Fed. Cir. 2021) (affirming Board's decision that “speech recognition system” and “wireline node” are distinct elements and should not be mapped to a single element); 37 C.F.R. § 42.104 (“The petition must specify where each element of the claim is found in the prior art patents or printed publications relied upon.”).

b. A POSITA would not have been motivated to combine Laganriere363, Laganriere817 and/or Laganriere461 in the way proposed by the Petitioner

Petitioner apparently argues that a POSITA would implement Laganriere817's quality scoring (and only that) in the first video analytics (Paper 2, 27) while also implementing Laganriere817's purging of lower quality images (and only that) in a central server (Paper 2, 28-30). But Petitioner fails to articulate a reason why a

POSITA would engage in such a convoluted disassembly-and-reassembly of disclosures. Ex-2004, ¶71.

i. A POSITA would not have been motivated to combine Laganriere363 and Laganriere817 in the way proposed by the Petitioner.

Petitioner acknowledges that Laganriere363 “reduces bandwidth by [the first video analytics] selecting particular frames ‘with a suitable license plate image’ and transmitting only the ‘one or few frames having a suitable quality relative to the other frames.’” Paper 2, 23 (citing Ex-1005, [0041]). Petitioner’s expert also confirmed that this process is performed by the first video analytics on the video source side of Laganriere363’s system, before transmission of data to the central server. Ex-2003, 66:22-71:10; Ex-2004, ¶72. Petitioner then argues that “POSITAs would have understood the face log in Laganriere817 was a suitable method of accomplishing Laganriere363’s goals of reducing bandwidth by *transmitting only high-quality object images for further processing*.” Paper 2, 25. By this rationale, if Laganriere817 were to be combined with Laganriere363, POSITAs would implement *both* Laganriere817’s quality scoring *and* Laganriere817’s purging lower quality images in Laganriere363’s first video analytics before transmission to the central server for further processing, arriving at a completely different system from the claimed invention. Ex-2004, ¶73.

Laganiere363 discloses a first video analytics for detecting the event of interest in a portion of the video data and selecting and transmitting frames having a suitable quality, which is analogous to Laganiere817's face detection and tracking, quality scoring and purging lower quality images for face log creation. Ex-1005, [0040]; Ex-1006, 4:20-30. Separately, Laganiere363 discloses a central server for performing second video analytics and determining information relating to the event of interest, which is analogous to subsequent facial recognition, outside Laganiere817's method and system. Ex-1005, [0040]; Ex-1006, 3:4-7; Ex-2004, ¶74.

Even if POSITA could implement Laganiere817's face detection and tracking and Laganiere817's quality scoring *in Laganiere363's first video analytics* according to Laganiere817's flow, such combination would not render obvious the claimed invention because it fails to disclose the architectural division of two separate and distinct "object detection system" and "object analysis system" across the network. Ex-2004, ¶75.

Meanwhile, a POSITA will not have a reasonable expectation of success in implementing Laganiere817's quality scoring (and only that) in Laganiere363's first video analytics while also implement Laganiere817's purging lower quality images in Laganiere363's central server because it requires artificially dividing the organic whole of Laganiere817's disclosures, thus complicating the implementation. Ex-2004, ¶76. Moreover, this implementation cannot be expected to achieve the alleged

goal of bandwidth reduction at all. Paper 2, 25. A POSITA would know that if Laganriere817's purging of lower quality images was implemented in Laganriere363's central server, the lower quality images would have to be transmitted to the central server over the network, occupying precious bandwidth, only to be discarded later. This disassembly-and-reassembly of disclosures would negate the advantages of reduced bandwidth and thus cannot stand. A POSITA would further understand that sending the lower quality images to the central server (which may decide to analyze them) may compromise the reliability of the video analytics outcome in the central server. Ex-2004, ¶77; see *McGinley v. Franklin Sports, Inc.*, 262 F.3d 1339, 1354 (Fed. Cir. 2001) ("If references taken in combination would produce a 'seemingly inoperative device,' we have held that such references teach away from the combination and thus cannot serve as predicates for a prima facie case of obviousness.").

Petitioner then argues that "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." Paper 2, 28. Yet, a motivation for the central server to "start with the highest quality images" is not a motivation for the central server to determine whether or not to analyze the image data. A POSITA would understand that, if the central server is to "start with the highest quality images", then it cannot discard the lowest quality images, because it has to keep

them in storage until it has finished processing the highest quality images that it started with, so as to process them after the highest quality images. Ex-2004, ¶¶78-79. So even if Petitioner establishes this motivation (it did not), this cannot render obvious the claimed “deci[sion] whether to analyze the image data.”

Indeed, nowhere in the references is it taught that the central server has an advantage, or even the possibility, of “start[ing] with the highest quality images.” On the contrary, Laganier363 explicitly discloses that the second video analytics either analyze only one or a few frames selected by the first video analytics or analyze all the frames stored local to the source and discard later. Ex-1005, [0041]. Further, Laganier363 expresses an objective of “speeding *real-time* security response.” Ex-1005, [0004] (emphasis added). A server aiming at providing a real-time response in a video surveillance system cannot arbitrarily change the processing order to “start with the highest quality images,” for the simple reason that the server will not know when or whether higher-quality images than the current one will arrive, and cannot afford to postpone the processing of the current image indefinitely to wait for higher-quality images. Ex-2004, ¶80. Petitioner’s expert did not explain how the so-called combination teaches “starting with the highest quality images,” instead resorting to the conclusory assertion that a POSITA would find it obvious to start with the highest quality images. Ex-2003, 83:11-85:9. This is no more than a hindsight explanation for the proposed unnatural disassembly-and-reassembly of the

prior art's teachings. *See In re NTP, Inc.*, 654 F.3d 1279, 1299 (Fed. Cir. 2011) (“Care must be taken to avoid hindsight reconstruction by using the patent in suit as a guide through the maze of prior art references, combining the right references in the right way so as to achieve the result of the claims in suit.”) (quoting *Grain Processing Corp. v. American-Maize Prods. Co.*, 840 F.2d 902, 907 (Fed. Cir. 1988)); *see also Ruiz v. A.B. Chance Co.*, 357 F.3d 1270, 1275 (Fed. Cir. 2004). (“This form of hindsight reasoning, using the invention as a roadmap to find its prior art components, would discount the value of combining various existing features or principles in a new way to achieve a new result – often the very definition of invention.”)

ii. A POSITA would not have been motivated to arrive at the claimed invention in view of Laganier461's teachings.

Third, Petitioner argues that “[i]f PO argues the Laganier363-Laganier817 combination does not teach the central server purges image from the face log,” “POSITAs would find it obvious” based on Laganier461's teachings. Paper 2, 28-29. Apparently, Petitioner thinks Laganier461's post-video-analytics selective erasing can be equated to Laganier817's immediate-post-facial-detection purging, but these functions by nature must be performed at opposite ends of video analytics. Laganier461's selective erasing is based on the results of and after video analytics, while Laganier817's purging largely relies on simple criteria that would be

immediately available after facial detection and would not rely on any further video analytics. Ex-2004, ¶81. In fact, Petitioner's expert recognized that Laganiere817's purging occurs intentionally at the end of facial detection and without consideration of when facial recognition takes place. Ex-2003, 89:9-91:22. However, Petitioner apparently argues that they can somehow be merged into one and ordered the way Petitioner wishes them to be (not post video analytics as in Laganiere461, nor immediately post facial detection as in Laganiere817, but after the frames have been unnecessarily transmitted over the network, occupying bandwidth, and still before video analytics is performed). A POSITA would not disregard the different structures of Laganiere461 and Laganiere817 and create a chimera out of their parts, one that results in the opposite of the purpose of *either* of them. Ex-2004, ¶81.

Unsurprisingly, Petitioner also failed to provide an actual motivation for this never-made-clear "combination." Petitioner argues that "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." Paper 2, 28-29. But Laganiere461 discloses selectively erasing video data *after* the central server has performed video analytics on the video data. Ex-1007, [0002], [0010], FIGS. 3-4; Ex-2004, ¶82. Petitioner's expert tries to confuse the issue by arguing that the purging can also happen after the first video analytics on the video source end, but could not refute FIG. 4's explicit disclosure of purging in the central server after

both video analytics. Ex-2003, 92:1-96:10. Thus a POSITA would understand that, contrary to Petitioner's unsubstantiated claim (Paper 2, 30), Laganier⁴⁶¹'s purging will *not* reduce the amount of the data the central server must process. Ex-2004, ¶82. This "motivation" is based not on any teaching of the references but purely on hindsight.

Instead, Laganier⁴⁶¹ teaches the central server performing post-video-analytics selective deletion "for reducing the amount of space that is required to store video data." Ex-1007, [0002]. Specifically, Laganier⁴⁶¹'s central server performs video analytics "to identify a single frame of data in sequence of frames that provides an image of the face of an individual that is of sufficient quality" and storage of the single frame with high quality for each individual "reduces the storage requirement significantly and facilitates identification by human users." *Id.*, [0028]; Ex-2004, ¶83. Again, POSITAs, given Laganier⁴⁶¹'s explicit teaching of reducing storage requirements, would be dissuaded from, rather than motivated to, transmitting unnecessarily the lower quality images (accompanied by their metadata) to the central server, taking even more storage space, only for the central server to discard them later. *See Lab'y Corp. of Am. Holdings v. Ravgen, Inc.*, Nos. 2023-1342, 2023-1345, at *10 (Fed. Cir. Jan. 6, 2025) (requiring the Board "to consider evidence that 'suggests reasons that a skilled artisan would be discouraged from pursuing such a combination.'") (citing *Arctic Cat Inc. v. Bombardier Recreational Prods. Inc.*, 876

F.3d 1350, 1363 (Fed. Cir. 2017)). Ex-2004, ¶84. Even if a POSITA were to implement *both* Laganriere817's quality scoring *and* Laganriere817's purging lower quality images in Laganriere461's central server, that is, the central server performs video analytics to calculate a quality score to identify frames of sufficient quality, and purging lower quality images, it still fails to teach or suggest the architectural division in the claimed invention. *Id.*, ¶85.

Overall, Petitioner simply failed to show how and why a POSITA would look to combine the prior arts in its proposed manner. *See KSR International Co. v. Teleflex Inc.*, 550 U.S. 398, 401 (2007) ("Although common sense directs caution as to a patent application claiming as innovation the combination of two known devices according to their established functions, it can be important to identify a reason that would have prompted a person of ordinary skill in the art to combine the elements *as the new invention does.*"). Even if a POSITA could implement Laganriere817's quality scoring and face log purging (two steps of an integral method) together in Laganriere363's first video analytics to reduce bandwidth or implement Laganriere817's quality scoring and face log purging (two steps of an integral method) together in Laganriere461's central server to reduce the amount of storage space, such combination still fails to teach or suggest an "object detection system" that ranks the detected image data corresponding to objects of interest according to different factors and saves the ranking in metadata with the image data, and separately, an

“object analysis system” that reads the rankings within the metadata and decides whether or not to analyze the image data.

Petitioner fails at its attempt at a combination that satisfies limitation [1.5.1]/[13.5.1] by implementing Laganriere817's quality scoring in Laganriere363's first video analytics, and also limitation [1.5.3]/[13.5.3] by implementing Laganriere817's purging lower quality images (and/or something from Laganriere461) in Laganriere363's central server. It does not, and indeed cannot, provide a sufficient motivation for a POSITA to combine the teachings of the prior arts in such a way that destroys the explicitly sought advantage of reduced bandwidth or storage requirement. Ex-2004, ¶86. Petitioner's whole proposition can only be based on hindsight, and the Board should reject it.

iii. Petitioner has not established a sufficient motivation to combine the prior arts to achieve the claimed invention.

The Board preliminarily found that “Laganriere363's disclosure of transmitting one or a few frames having suitable quality relative to other frames does not preclude ranking the images associated with those transmitted frames” and “Laganriere817's teachings can be applied to metadata received at the central server, as Petitioner notes is taught by Laganriere363.” Paper 8, 20-21. However, the *possibility* of a combination alone is different from establishing a sufficient motivation to combine the prior art to achieve the totality of the claimed invention

including “ranked by the object detection system” and “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.” Petitioner has not clearly articulated such a combination, and, as discussed above, to the extent one can be discerned, it fails. *See Ryko Mfg. Co. v. Nu-Star, Inc.*, 950 F.2d 714, 717 (Fed. Cir. 1991) (“When analyzing a patent claim for obviousness, the claim should be considered as a whole, but the differences between the claim and the prior art need to be identified to place the obviousness analysis into proper perspective.”). Accordingly, Grounds 1A and 1B fails to disclose or render obvious the challenged claims.

B. Ground 2: Petitioner fails to show Bedros renders any challenged claim obvious

1. Claims 1 and 13 (and therefore all challenged claims)

a. Bedros fails to disclose or render obvious “the video analytics system receives the image data from multiple object detection systems after being transmitted over the network” ([1.4.1], [13.4])

Petitioner argues that Bedros teaches this limitation because “Bedros discloses server 308 (“*video analytics system*”) receives facial image 116 (“*image data*”) over a network from server 306 (“*object detection system*”).” Paper 2, 61. But Petitioner fails to identify a disclosure of “multiple object detection systems” in this mapping. Ex-2004, ¶87. Bedros explicitly discloses that “*one particular* server 306” – as compared to “one or more servers 308” – executes image capturing and processing

software. Ex-1008, 8:20-22. Petitioner argues that Bedros discloses the computing device depicted in Figure 4 may be server 306 and the computing device can be a “computer 410.” Paper 2, 61 (citing Ex-1008, 8:55-56, 9:9-17). It is unclear how this shows “multiple” servers 306, especially since FIG. 4 shows a single computer 410. Such oblique reasoning forces the Board and Patent Owner to guess at its mappings. Ex-2004, ¶88.

Petitioner then argues that “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.” Paper 2, 61-62. But nowhere does Bedros disclose a need for multiple cameras monitoring multiple locations. Ex-2004, ¶89. Petitioner first cites the “individuals and their locations” language out of context in an attempt to create such a need. Paper 2, 62. Yet the context makes it clear that this language is about locating individuals in images from *a* monitored environment. Ex-1008, 1:6-10. The “locations” here are locations of individuals within an image (from one camera), not multiple locations requiring coverage by multiple cameras. Ex-2004, ¶90. In fact, the cited materials shows that Bedros *disparages* the multiple-camera solution because of “deficiencies identified such as the added cost of procuring, installing, and maintaining multiple cameras.” Ex-1008, 1:45-46. Bedros further disparages a *two*-camera solution as presenting problems of latency and missed opportunity. Ex-1008, 1:46-57; Ex-2004,

¶¶91-92. Thus, the disclosure of Bedros explicitly teaches away from utilizing multiple object detection systems. *See In re Gurley*, 27 F.3d 551, 553 (Fed.Cir.1994) (“A reference may be said to teach away when a person of ordinary skill, upon reading the reference, would be discouraged from following the path set out in the reference or would be led in a direction divergent from the path that was taken by the applicant.”); *see also Crocs, Inc. v. International Trade Com'n*, 598 F.3d 1294, 1308–09 (Fed. Cir. 2010) (finding that claims directed to a foam shoe were not obvious where the prior art criticized the use of foam material for shoe straps); *Allergan, Inc. v. Sandoz Inc.*, 796 F.3d 1293, 1306 (Fed. Cir. 2015) (finding that claims directed to a composition having 200 ppm of BAK were not invalid for obvious when the prior art taught that BAK should be minimized in ophthalmic formulations to avoid safety problems).

Petitioner further cites different embodiments of Bedros as disclosing a motivation 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).” Paper 2, 62 (citing Ex-1008, 5:3-11, 5:26-33). However, in presenting different embodiments, Bedros is disclosing why spatial proximity may vary depending on, partially, *the* (singular) monitored scene 100. Ex-1008, 5:25-26. Bedros discusses hallways and work areas as two examples of scenes that require different spatial proximities. Ex-1008, 5:26-

33; Ex-2004, ¶93. Nowhere does Bedros present these two examples as two scenes that are to be monitored simultaneously. In fact, given Bedros's disparagement of multi-camera systems (Ex-1008, 1:43-57), a POSITA would understand that the different embodiments represents different *uses* of Bedros's method and system (implemented with a *single* camera) in different scenarios by different users. Ex-2004, ¶94. Petitioner failed to cite any other evidence (other than its expert's identically-worded and unsubstantiated declaration) for its counter-factual assertion that a POSITA would have been motivated to utilize "Bedros' multiple computers with cameras." Paper 2, 62.

Petitioner also argues that POSITAs would be motivated "to send multiple different image data from multiple locations to server 308 to cost-effectively utilize the service." Paper 2, 62-63. But Petitioner and its expert offer no evidence to support this assertion. As discussed above, Bedros does not disclose a usage scenario or a need that involves monitoring multiple locations simultaneously. Thus, a POSITA would not be motivated to modify Bedros against the explicit disclosure of "one particular server 306." Ex-1008, 8:20; Ex-2004, ¶95.

Petitioner then argues that a POSITA would have a reasonable expectation of success in implementing multiple computers. Paper 2, 62-63. But this is irrelevant because, as discussed above, Petitioner has not established a motivation for a POSITA to make such a change.

b. Bedros fails to disclose or render obvious “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” ([1.5.3], [13.5.3])

Petitioner argues that Bedros's “queue processing” (Ex-1008, 7:19-30) discloses this limitation. Paper 2, 68-71. Petitioner identifies step 230 (submitting the facial image to the facial image queue, which may include removing and reordering facial images within the queue) of method 200 as performing this “queue processing”, but Bedros clearly distinguishes “queue processing 240” (illustrated with a dashed box in FIG. 2) from “submitting 230.” *Id.*, 69; Ex-1008, 7:19-62, FIG. 2. The “removing and reordering” language describes submitting 230, not queue processing 240. Ex-1008, 7:21-25. Thus, Petitioner's description of Bedros's disclosure is wrong to start with. Ex-2004, ¶96.

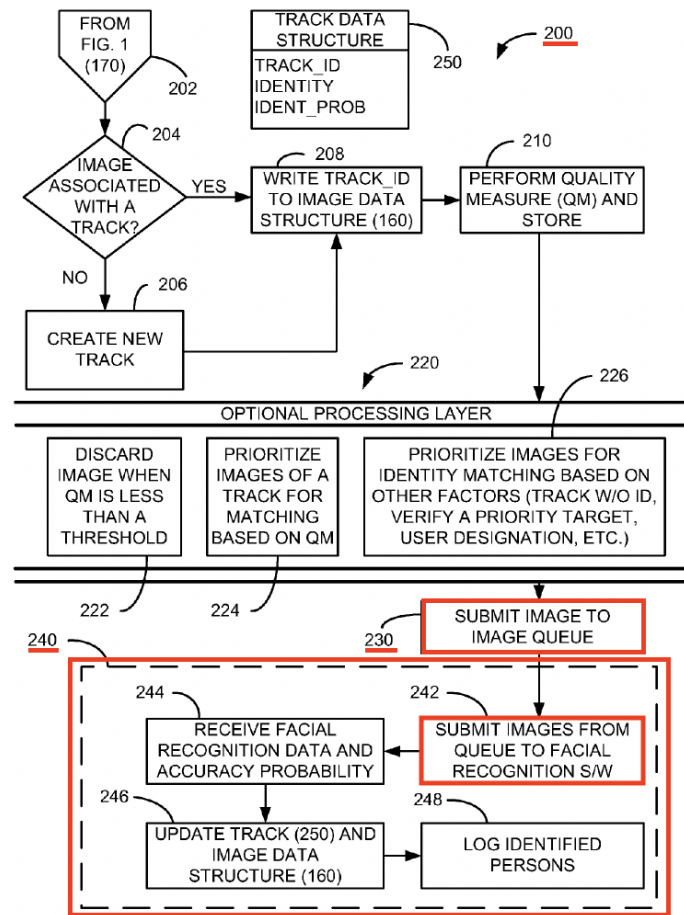


FIG. 2

Ex-1008, FIG. 2 (annotated)

Petitioner's argument has three parts: (1) a queue of images resides and is managed on servers 308; (2) this queue processing involves removing "certain facial images 116. . . from the queue of images" (3) to remove images from the queue, "POSITAs would have recognized queue processing must first read the quality scores of the images in the queue." *Id.* All are wrong because all are the opposite of Bedros's explicit disclosures. Ex-2004, ¶¶97-98.

i. Bedros fails to disclose limitation [1.5.3]/[13.5.3] because the alleged “queue processing” is executed by server 306.

Petitioner first argues that Bedros teaches processing the facial image queue at servers 308. Paper 2, 68-71. However, contrary to Petitioner's mischaracterization, Bedros explicitly discloses that the queue resides on and is managed by the “[o]ne particular server 306” executing software that implements the methods 200, not servers 308. *Compare* Paper 2, 69 with Ex-1008, 8:20-22; Ex-2004, ¶99.

Further, as Petitioner itself cites, Bedros's so-called “queue processing” (submitting 230) is done on “facial images 116 within the queue *awaiting submission to the facial recognition program.*” (Ex-1008, 7:23-24). Petitioner's expert reconfirmed this. Ex-2003, 108:10-14; *see also* Ex-1003, ¶169; Ex-2004, ¶100. A step of “submitting” cannot be performed by the servers 308 executing facial recognition software, who are recipients of the “submission.” Therefore, the “queue processing” can only take place on server 306, otherwise step 230 would not be able to operate on facial images before submission to the facial recognition program. *Id.*, ¶101.

A POSITA, looking at Bedros's explicit disclosures, would understand that none of the images in a queue *awaiting submission to the facial recognition program* have been submitted to the facial recognition program, therefore none of the images in the queue have “arrive[d]” at the facial recognition program. Yet in the very next

sentence, Petitioner simply recharacterizes the queue – without citing any evidence – as a queue “awaiting processing by the facial recognition software as subsequent images arrive from any of servers 306”, that is, a queue of images that has already arrived at the facial recognition programs running on servers 308 (otherwise the language “subsequent images arrive from any of the servers 306” would make no sense). Paper 2, 69; Ex-2004, ¶102. The Board should reject such blatant mischaracterization of prior art disclosure through pure attorney argument. *In re Chapman*, 595 F.3d 1330, 1339 (Fed. Cir. 2010) (remanding the Board’s decision regarding obviousness when the Board based its decision on a misunderstanding of prior art).

Petitioner next argues that “[i]f 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.” Paper 2, 69. Both premises of Petitioner are wrong. Ex-2004, ¶103. First, as discussed above, there cannot be “other servers 306,” which is against Bedros’s explicit disclosure of “one particular server 306”, Ex-1008, 8:20. Even in an imagined scenario involving multiple servers 306 (directly contrary to Bedros’s explicit disclosure), Petitioner has not provided any advantage associated with knowing the rankings from other servers 306. Ex-2004, ¶104. Second, Bedros

teaches one or more servers 308. Ex-1008, 8:23-31. Thus, Bedros contemplates scenarios in which more than one server 308 is present. In this case, no single server 308 will “ultimately process” all facial images. Ex-2004, ¶105. These so-called motivations are nothing but hindsight justifications, unsupported by anything in Bedros's disclosure.

Petitioner's next argument is that “[p]rocessing the queue at server 308 would have been more efficient and reduced bandwidth by relieving servers 306 from reading the queue to [remove and reorder] facial images from it.” Paper 2, 69. This is a false premise. In its effort to map limitation [1.5.1]/[13.5.1], Petitioner acknowledges that optional layer 220 in server 306 applies quality measure scores and/or prioritization results to facial images 116, that is, server 306 would inevitably read and analyze image data to filter and prioritize facial images. Paper 2, 64, 68. If images are not removed by server (not “servers”) 306 and sent to server 308 only to be removed there, this will *increase* bandwidth, not reduce it, and will be *less*, not more efficient. Ex-2004, ¶106.

Petitioner argues that “[p]erforming queue processing on image data at the server receiving information was well-known in the art” by having its expert discussing various references *not* asserted as part of the ground. Paper 2, 69-70. But none of the cited references are related to removing images as a part of “queue processing” (submitting the facial image to the facial image queue) on the server.

Laganiere461 discloses purging data *after* video analytics, not while they are in a queue waiting for processing. Ex-1007, [0024-0025]. Gali discloses sorting the queue (that is, removing and reordering and putting back in the correct order, as a POSITA would understand the term “removing and reordering” in Bedros’s context), not discarding images from the queue. Ex-1021, [0053-0054], [0094-0095], FIG. 1. Deever965 has nothing to do with removing or reordering data items in a queue, as Petitioner’s expert’s own declaration makes clear. Ex-1015, 14:45-15:35, 16:47-17:33; Ex-1003, ¶229; Ex-2004, ¶107. In any case, these references cannot support Petitioner’s repeated false argument that POSITAs would have been motivated to modify Bedros to process the facial image queue at server 308, for the simple reason that Petitioner’s proposed modification will *increase*, not decrease, transmission bandwidth. Ex-2004, ¶108.

Petitioner further argues that “POSITAs would have been further motivated to separate and delineate tasks performed at [server] 306 and servers 308 to streamline operation and to reduce transmission bandwidth.” Paper 2, 70. Petitioner’s expert cites to the reference Yang that is, again, not part of this ground and not relate at all to removing or reordering data items in a queue. See Ex-1003, ¶¶230-232. Yang merely discloses smart cameras detecting objects of interest and removing background, while “the central node or the cloud can perform state of the art high precision algorithms for detection, tracking or recognition to deliver real

time analytics.” *Id.* (citing Ex-1022, [0026], [0020]). The purpose of Yang’s central node is to generate commands and send the commands back to local processor for it to adjust parameters for analyzing the sequence of images. *Id.*, claim 10. Thus, The separation of work between smart cameras and central node or cloud is completely different from the claimed invention. Ex-2004, ¶¶109-110. As with all the other references, Petitioner fails to establish a motivation for a POSITA to separate and delineate Bedros’s integrated method 200 to be performed by server 306 and servers 308 in the particular way Petitioner wishes.

Because Petitioner cannot establish a motivation for modifying Bedros and implementing image-data-discarding in server 308, Petitioner’s argument of reasonable expectation of success (Paper 2, 70) is simply irrelevant.

ii. Bedros fails to disclose limitation [1.5.3]/[13.5.3] because server 308 does not read the rankings within the metadata.

Petitioner further fails to map the “reads the rankings within the metadata” limitation. Petitioner’s whole argument is that, for servers 308 to be able to do the queue processing as it wished, they must be able to read the quality scores of the images. Paper 2, 70. First, as discussed above, servers 308 do not do the “queue processing” (that Petitioner seeks to transplant from step 230 performed by server 306). *See supra* Section B(1)(b)(i); Ex-2004, ¶¶111-112.

Second, even if servers 308 would implement some variation of step 230, Bedros never disclosed that step 230 is based on quality scores. Instead, Bedros disclosed that it is based on “processing outcomes within the optional processing layer 220.” Ex-1008, 7:24-25. Quality measures are calculated in step 210, *before* the optional processing layer 220. Ex-1008, 6:4-22, FIG. 2. Therefore, quality scores are *not* outcomes within the optional processing layer 220. Ex-2004, ¶113.

Third, Bedros never disclosed sending anything else other than facial images 116 to servers 308. Bedros disclosed saving quality scores in image data structure 160 (that Petitioner identifies as disclosing “metadata” of [1.5.2], Paper 2, 68) of the facial image 116. Ex-1008, 6:16-19. But Bedros never disclosed sending image data structure 160 to servers 308. Bedros only discloses image 116, not image data structure 160, as being submitted to the queue. Ex-1008, 7:19-25. Thus, the only thing that can be submitted from the queue to facial recognition would be image 116, not image data structure 160. Ex-1008, 7:27-30; Ex-2004, ¶114. In fact, image data structure 160 also contains received identity and accuracy probability data from facial recognition (performed by servers 308), and these fields did not exist when image 116 is submitted to facial recognition. Ex-1008, 7:42-44. This is another reason why Bedros could not have disclosed sending image data structure 160 to servers 308. A POSITA would understand that Bedros's server 306 maintains image data structure 160 through method 200 and does not transmit it to servers 308. Ex-

1008, FIG. 2; Ex-2004, ¶115. Such separate maintenance of meta data (part of image data structure 160) and image (transmitted without being accompanied by other parts of image data structure 160) allows more efficient handling, updating and storage of metadata. Ex-2004, ¶116.

Thus, a POSITA would understand that the quality scores are not accessed by servers 308, and indeed not needed by servers 308 even if servers 308 were to perform a variation of step 230 (which Bedros explicitly allocated, as part of method 200, to server 306, *see* Ex-1008, 8:20-22). Ex-2004, ¶117. Petitioner's argument boils down to that a POSITA would recognize that servers 308 need quality score because the operation that Petitioner wishes to assign to servers 308 needs quality score – a hindsight-driven conclusion.

iii. Bedros fails to disclose limitation [1.5.3]/[13.5.3] because server 308 does not determine whether or not to analyze the image data.

Even if the alleged “queue processing” is artificially carved out of the method 200 to be performed by server 308, Petitioner's argument for the determination step hinges on taking the word “removing” out of its context from Bedros.

Bedros discloses (as part of optional processing layer 220 in server 306) a processing element that discards 222 images when the quality measure of a facial image 116 is less than a threshold value. Ex-1008, 6:38-40. Bedros discloses that “[d]iscarding a facial image 116 may include deleting the facial image 116,

removing a record from the image data structure 160, marking an image data structure 160 record for deletion, *or otherwise removing the facial image 116 from processing within the method 200 and preventing the facial image 116 from being submitted to the facial recognition program.*” Ex-1008, 6:49-55 (emphasis added); Ex-2004, ¶118. Bedros then discloses that “[f]ollowing processing of the optional processing layer 220, the method 200 includes submitting 230 the facial image 116 to the facial image queue. In some embodiments, 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 *based on processing outcomes within the optional processing layer 220.*” Ex-1008, 7:19-25. Thus Bedros explicitly discloses that the “removing and reordering” of step 230 happens *after* the optional processing layer 220 – a fact that Petitioner concedes. See Paper 2, 64 (“Optional processing layer 220 filters and/or prioritizes facial images 116. . . **before** they are being submitted to an image queue at operation 230.”) (emphasis original). But that would mean step 222 has already taken place. A POSITA would thus understand that, when step 230 happens, it is processing facial images 116 which server 306 has *already discarded* by executing the processing element 222 in the optional processing layer. Ex-2004, ¶119.

A POSITA would not understand step 230 as implementing *another* discarding procedure closely following the dedicated discarding processing element

222. Such implementation would not achieve any advantage and would only result in duplicative processing and increased probability of error. Ex-2004, ¶120. Instead, a POSITA would understand the language “removing and reordering” as describing an integral process – removing images from the queue, reordering them based on priority values assigned in step 224 just before 230 (Ex-1008, 6:56-7:3), and then putting the correctly-ordered images back into the queue, so as to achieve the explicitly stated objective that “the facial images 116 that are the most fit for facial recognition processing will be submitted first.” (Ex-1008, 7:1-3) – not “the facial images 116 that are the most fit for facial recognition processing will be submitted *but others will not be submitted at all.*” Ex-2004, ¶121.

Even if a POSITA reads the word “removing” out of the integral “removing and reordering” and thinks step 230 implements some kind of discarding procedure, a POSITA would understand that the “removing and reordering” in step 230 is “based on processing outcomes within the optional processing layer 220.” *Id.*, 7:21-25. Thus, server 308, when solely implementing step 230, would not have independently read the quality measure value, compared it with the quality measure threshold, and determined whether to discard, because such logic has already taken place in the optional processing layer 220. Ex-2004, ¶122.

iv. Bedros fails to disclose limitation [1.5.3]/[13.5.3] even if the same server 306 execute the face recognition software.

The Board preliminarily found that, because Bedros explicitly discloses that the discarding software and the facial recognition software can reside and execute in the same server, in the context of the “systems” recited in claim 1 of the '321 patent, the steps Bedros carries out to perform extraction and prioritizing (ranking) can be characterized as object detection system functions and the steps Bedros carries out to delete and reorder images for submission to the queue for facial recognition process can be characterized as video analysis system functions. Paper 8, 32-34 (citing Ex-1008, 8:22-24). Patent Owner respectfully disagrees because such a mapping conflates the two separate and distinct “objective detection system” and “object analysis system” and overlooks the requirement that the object detection system and the video analytics system (containing object analysis system) be separated by a network and transmit data over it. This mapping would also run afoul of the video analytics system “receiv[ing] the image data from *multiple object detection systems*” and “[being] operated by a different business entity” as recited in limitations [1.4.1] and [1.4.2].³

³ In any case, Petitioner never made this argument and the Board should not rule for Petitioner on a theory it never advanced. *See In re Magnum Oil Tools Int'l, Ltd.*, 829

The Board made its preliminary ruling without having been presented with the full argument against Petitioner's motivation to combine or against Petitioner's mapping of "reads the rankings within the metadata." Patent Owner respectfully requests that the Board reexamine the issue whether Bedros discloses or renders obvious claim limitation [1.5.3]/[13.5.3] on the current, more complete record.

VII. CONCLUSION

For at least the above reasons, Petitioner has not met its burden to show that any claim is unpatentable as obvious over the cited art.

F.3d 1364, 1381 (Fed. Cir. 2016) ("Board must base its decision on arguments that were advanced by a party" and is not "free to adopt arguments on behalf of petitioners that could have been, but were not, raised by the petitioner during an IPR."); 37 C.F.R. § 42.104 ("The petition must specify where each element of the claim is found in the prior art patents or printed publications relied upon.").

Date: March 26, 2025

Respectfully submitted,

/Hao Tan/

Hao Tan, Reg. No. 73,711
Arch & Lake LLP
203 N LaSalle Street, Suite 2100
Chicago, IL 60601
Tel. No. 312.375.9408
Fax No. 312.614.1873
HaoTan@archlakelaw.com

Attorney For Patent Owner

CERTIFICATE OF WORD COUNT

I hereby certify that the foregoing PATENT OWNER'S RESPONSE includes 9,683 words (as calculated by Microsoft Word), excluding those in the table of contents, table of authorities, list of exhibits, list of claims, mandatory notices, signatures, certificate of word count, and certificate of service. Accordingly, this Reply is under the word count limit of 14,000 words.

Date: March 26, 2025

Respectfully submitted,

/ Hao Tan /

Hao Tan, Reg. No. 73,711
Arch & Lake LLP
203 N LaSalle Street, Suite 2100
Chicago, IL 60601
Tel. No. 312.375.9408
Fax No. 312.614.1873
HaoTan@archlakelaw.com

Attorney For Patent Owner

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. § 42.6(e), I hereby certify that, on March 26, 2025, a set of true and correct copies of PATENT OWNER'S RESPONSE and supporting materials has been served via e-mail to the Petitioner by serving the below-listed attorneys of record for Petitioner in the concurrent proceeding:

Jessica Kaiser: kaiser-ptab@perkinscoie.com
Matthew J. Moffa: moffa-ptab@perkinscoie.com
Patrick McKeeever: mckeeever-ptab@perkinscoie.com
John C. Moulder: moulder-ptab@perkinscoie.com
Yuanfeng Gao: gao-ptab@perkinscoie.com
Tyler Knox: knox-ptab@perkinscoie.com
Grant Kinsel: kinsel-ptab@perkinscoie.com

Date: March 26, 2025

Respectfully submitted,

/ Hao Tan /

Hao Tan, Reg. No. 73,711
Arch & Lake LLP
203 N LaSalle Street, Suite 2100
Chicago, IL 60601
Tel. No. 312.375.9408
Fax No. 312.614.1873
HaoTan@archlakelaw.com

Attorney For Patent Owner