

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TVISION INSIGHTS, INC.,

Petitioner,

v.

THE NIELSEN COMPANY (US), LLC,

Patent Owner.

Case No. IPR2023-01014

U.S. Patent No. 11,470,243

**PETITIONER'S REPLY TO
PATENT OWNER'S RESPONSE**

TABLE OF CONTENTS

	<u>Page</u>
I. GROUND 1	4
A. “reduc[ing] a resolution.”	4
1. Undisputed teachings of Tian satisfy the limitation at issue.....	4
2. Tian teaches performing the act of reducing a resolution of an image.....	5
3. PO’s argument distinguishing Tian based on unsupported limitations should be rejected.	8
B. Disputed limitations of claims 5, 6, 12, 14, and 19.....	10
1. Undisputed disclosures show that Ground 1 teaches the limitations at issue.....	10
2. PO’s mischaracterization of claim scope should be rejected.	12
C. A POSITA would have found it obvious to combine Lu-577 and Tian.....	14
1. Analyzing reduced-resolution images saves processing resources.....	15
a. The resource savings outweigh any purported cost.....	16
b. PO’s purported redundancy does not exist.	19
c. Tian’s accuracy findings are applicable to audience measurement.	19
2. Applying Tian’s pose estimation teachings would have achieved the explained benefits.	20
a. Pose estimation improves the performance of face recognition.	21
b. Pose estimation is not redundant.	22
c. Pose estimation facilitates determining if a person is watching television.	22
3. Modifying Lu-577 based on Tian’s face tracking teachings would have achieved the explained benefits.	23

D.	The Petition’s use of incorporated teachings of Lert and Lu-000 does not combine different embodiments.	24
II.	GROUND 3	26
A.	The combination of Ground 3 teaches the disputed limitations.	26
B.	PO’s mischaracterization of Petitioner’s combination of Nielsen-372 and Steinberg should be rejected.	26
C.	PO’s mischaracterization does not undermine the obviousness showing of Ground 3 regardless.....	27
D.	PO’s attacks on the motivation to combine with Tian fail for the same reasons as explained for Ground 1.....	29
III.	ANALOGOUS ART	29

PO's arguments that the combinations of Ground 1 and Ground 3 do not teach the disputed limitations mischaracterize the prior art and the claim requirements. And PO's attacks on the motivation to combine Lu-577 or Nielsen-372 and Steinberg with Tian are contradicted by both experts and based on importing unrecited limitations into the claims. PO's bodily incorporation argument against the combination of Nielsen-372 and Steinberg fails because it is based on a misunderstanding of the combination.

I. GROUND 1

A. “reduc[ing] a resolution.”

1. Undisputed teachings of Tian satisfy the limitation at issue.

The Institution Decision found “no dispute that Tian does disclose a method of reducing the resolution of an image[.]” D.I., 21. The POR largely rehashes its POPR arguments and still does not dispute that (1) Tian teaches down-sampling, (2) down-sampling is a method of reducing image resolution, and (3) resolution-reduction techniques were widely known by 2011 (*see also* Ex-1003, ¶15; Ex-1033, 31:6-23). Notably, PO acknowledges that Tian teaches “resolution reduction,” despite contending that it is performed in “pre-processing.” POR, 10-11. None of the claims at issue limit resolution reduction to a certain time or stage

of processing.¹ Therefore, even under its own characterization, PO cannot deny Tian’s explicit teaching of reducing image resolutions. PO’s assertion that Tian’s “experimental setup” or “structure” does not include a step of reducing a resolution (POR, 3, 5-6, 9), even if accepted,² does not negate the obviousness of doing so in light of Tian’s teachings.

PO argues that the Board misapprehends the requirements of the claims. POR, 4-7. This argument is misplaced because the Board found that Tian teaches obtaining a reduced-resolution image by down-sampling, rather than directly from a database. D.I., 22 (“Tian plainly teaches obtaining reduced-resolution images and teaches down-sampling as shown in the Petition.”), 21.

2. Tian teaches performing the act of reducing a resolution of an image.

PO contends that “Tian does not *create* the down-sampled images” (POR, 6; *see also* POR, 3-4, 7-10) based on Dr. Saber’s opinion that Tian directly obtains pre-processed images from the Cohn-Kanade database, which includes images of different resolutions (Ex-2001, ¶¶46-49; Ex-2020, ¶¶4-7).

¹ In *Chewy, Inc. v. Int’l Bus. Machines Corp.*, the Federal Circuit determined that a “storing” limitation requires pre-fetching an object but not retrieving an object on demand. 94 F.4th 1354, 1358-62 (Fed. Cir. 2024). *Chewy* is not applicable because PO’s exclusion of pre-processing has no basis in the claims.

² This assertion is wrong. *Infra* §I.A.2.

This argument is contradicted by Tian’s “Database” sub-section, which discusses the “DFAT subset of Cohn-Kanade expression database” used by Tian’s experiment. Ex-1005, 000005. Tian explains that “[t]he database contains 704 image sequences from 97 subjects” and “[t]he image sequences ... were digitized into 640x480 pixel arrays[,]” for which “[t]he size of head region is about 280x380 pixels.” *Id.* Tian obtains images of the original resolution (i.e., 640x480) from the Cohn-Kanade database and creates reduced-resolution images by down-sampling the original images. *E.g.*, Ex-1005, 000002 (“The lower resolution images are down-sampled from the originals.”), Abstract.

Tian teaches face regions of five different resolutions: 288x384 (Original), 144x192, 72x96, 36x48, and 18x24. Ex-1005, 000002 (Table 1). The 288x384 face region (i.e., “about 280x380 pixels”) is directly cropped from the original images. Ex-1005, 000005. Head regions of the other resolutions are created using down-sampling. Ex-1005, 000002, Abstract.

Tian’s teaching of down-sampling images is consistent with statements quoted in the POR. In the statement “[a]n example in five different resolutions from Cohn-Kanade database” (Ex-1005, 000002 (quoted at POR, 8)), “in five different resolutions” and “from Cohn-Kanade database” both modify “example.” This demonstrates down-sampling an example from the Cohn-Kanade database into five different resolutions. In the statement “five different resolutions of the

head region are studied ... based on Cohn-Kanade AU-Coded Face Expression Image Database” (*id.*), the term “based on” confirms that the images of different resolutions are created based on original images in the database.

Dr. Saber’s opinion is flawed because he misunderstood the Cohn-Kanade database. First, Dr. Saber admitted that Tian’s “Database” subsection described the resolution of the image sequences that Tian uses. Ex-1033, 110:11-113:18. He conceded that only one resolution is mentioned in this subsection, and could not explain why it does not mention other resolutions had the database included them. *Id.*; Ex-2020, ¶54.

Second, Tian states that “[m]ore details about the [Cohn-Kanade] database can be found at paper [16].” Ex-1005, 000005. This paper confirms that the images sequences in the database “were digitized into either 640x490 or 640x480 pixel arrays” without mentioning any reduced resolutions and explains that “image data [in the database] may be ... down-sampled to examine the effects of reduced resolution ...” Ex-1031, 000006.

Third, PO’s argument should be rejected because Dr. Doermann obtained a copy of the Cohn-Kanade database and confirmed that it does not include reduced-resolution images. Ex-1034, ¶¶2-4. A POSITA reading Tian would have known that the Cohn-Kanade database, admittedly well-known in the field (Ex-1033, 113:19-114:13), does not include reduced-resolution images and that Tian teaches

creating its own. In contrast, Dr. Saber had no support or knowledge for his theory regarding the Cohn-Kanade database. Ex-1033, 116:5-20, 119:19-120:19.

PO's argument that Tian does not teach "circuitry, instructions or a device" (POR, 3, 9-10) attacks a strawman because the Petition relies on other references for these conventional hardware or software features. Pet., 13-16, 32, 47-49, 64-65; Ex-1033, 31:24-32:12.

3. PO's argument distinguishing Tian based on unsupported limitations should be rejected.

PO's arguments that Tian's teachings would not result in a usable audience measurement system (POR, 4, 10-13) are incorrect and directed to unrecited limitations. *Phillips v. AWH Corp.*, 415 F.3d 1303, 1323 (Fed. Cir. 2005) (warning against confining claims to specific embodiments).

First, PO argues that Tian relies on pre-processing images and would not be able to activate the illumination source for succeeding frames. POR, 11. As explained above, Tian does not rely on pre-processed images. *Supra* §I.A.2. More importantly, the feature of activating an illumination source is only recited in claims 24-25, which are not at issue in this IPR.³ Nor is activating an illumination source part of either ground of the Petition.

³ PO's continuing reliance on the illumination feature confirms Petitioner's explanation that none of the challenged claims recite the '243 patent's purported inventive idea. Pet. 2-3.

Any purported necessity of activating illumination based on image analysis is contradicted by PO's expert, who admitted that cameras had the ability to activate flashlight with each photo or activate flashlight based on detection of ambient light by 2011. Ex-1033, 136:14-24. Dr. Saber also admitted that well-lit images can be taken with natural light or artificial light in an audience measurement environment, with or without a camera's illumination source. Ex-1033, 135:10-136:13.

Second, PO argues that Tian's teachings would not result in a usable system because "down-sampling does not increase the contrast of images." POR, 11-12. Again, none of the challenged claims require some level of contrast for image analysis or recite contrast improvement as a purpose of resolution reduction.

PO argues that "applying *Tian*'s teachings ... would result in low-resolution, low-contrast images not suitable for analysis." POR, 12; Ex-2020, ¶16, ¶18. Tian proves PO wrong by reporting face detection and pose estimation being performed on down-sampled images with high accuracy. Ex-1005, 000005; Ex. 1003, ¶68. Dr. Saber admitted that he had no reason to challenge Tian's credibility. Ex-1033, 123:1-10.

Third, PO argues that down-sampling without antialiasing does not produce an image ready for analysis and, if antialiasing is performed, it may hinder facial recognition. POR. 12-13. The '243 patent does not claim or otherwise mention

antialiasing (Ex-1033, 95:18-96:15) and discloses a resolution-reduction method (e.g., pixel binning) without a separate antialiasing step (Ex-1033, 68:10-14). *Cf.* Ex-2018, 44:16-45:1. To the extent PO is contending antialiasing is an implied necessary step, Tian's experimental results (Ex-1033, 123:1-10) show otherwise by achieving high accuracy (Ex-1005, 000005; Ex. 1003, ¶68) while not performing antialiasing (Ex-2020, ¶20). Ex-1034, ¶9. Additionally, Dr. Saber admits that the claims encompass capturing a sequence of images (e.g., as frames of a video) before analyzing any of them. Ex-1033, 138:10-139:22. Therefore, antialiasing is not required and, even if performed, would not delay capturing the next image.

B. Disputed limitations of claims 5, 6, 12, 14, and 19.⁴

1. Undisputed disclosures show that Ground 1 teaches the limitations at issue.

The disputed limitations all require performing face recognition on a region/portion of a second image *corresponding to* a location/region of the head in a first image, whose resolution may be reduced. POR, 14-15; Ex-1001, 15:61-16:9, 16:54-61, 17:48-57.

Taking claim 5 as an example, the Petition establishes the following facts:

⁴ While PO's arguments for Ground 1 focus on claims 5, 6, 12, and 19, PO's similar arguments for Ground 3 additionally cover claim 14. POR, 49-50. Since PO raises no argument specifically for claim 14, it is addressed here.

- Lu-577 (incorporating Lu-000) teaches performing face recognition on a region of an image corresponding to a face/head location. *E.g.*, Pet. 28; Ex-1003, ¶¶125-127; Ex-1008, Abstract, 2:20-38, 6:49-51, 7:25-27, 7:39-8:6.⁵
- Tian teaches determining a face/head location in a first image and using this location for subsequent images unless a change is detected. *E.g.*, Pet., 28; Ex-1003, ¶129; Ex-1005, 000001, 000003.
- Face/head location often does not change across a sequence of images taken in a short period of time. *E.g.*, Pet., 9; Ex-1003, ¶64.
- Tian teaches performing face detection on a reduced-resolution image. *E.g.*, Pet. 10-11; Ex-1003, ¶¶68-69; Ex-1005, Abstract, 000002, 000005.

Based on the above, the Petition explains that it would have been obvious to determine a person's face/head location in a first image, whose resolution may be reduced, and perform face recognition on regions of this location in a second image. Pet. 28, 76; Ex-1003, ¶130, ¶¶318-319.

PO does not dispute the aforementioned facts and fails to explain why, when (1) face recognition is performed on a region of a second image corresponding to

⁵ For Ground 3, Steinberg teaches performing face recognition on a region of an image corresponding to a face/head location. *E.g.*, Pet. 76; Ex-1003, ¶¶316-317; Ex-1007, ¶0077, ¶0092, ¶0096, ¶0124, ¶¶0126-0127, ¶0130, FIGs. 1(a)-(c).

the head location and (2) the head location in the second image is the same as that in the first image, the analyzed region in the second image would somehow not correspond to the head location/region in the first image. POR, 15-17. PO cannot reasonably dispute that the limitations at issue are obvious under Ground 1.

Unwired Planet, LLC v. Google Inc., 841 F.3d 995, 1002 (Fed. Cir. 2016)

(“[C]ombinations of prior art that sometimes meet the claim elements are sufficient to show obviousness.”).

2. PO’s mischaracterization of claim scope should be rejected.

Rather than address the claim language directly or propose a claim construction, PO invents the concept of a “two-frame process.” POR, 13. The “two-frame process” concept serves as a disguise for an implicit negative limitation: that the determination of the head location cannot be performed on the second image. PO relies on this negative limitation because, rather than contend there is anything that Tian does not perform, PO faults Tian for confirming the head location in subsequent images. For example, PO argues that Tian does not teach the “two-frame process” because “*Tian* updates (“tracks”) the location of the head/face for each frame to perform face analysis for that frame.” POR, 14-16.

This negative limitation has no basis in the claims. Independent claims 1, 9, and 16 all recite “comprising,” indicating the inclusive nature of the claims. Ex-1001, 15:15, 16:21, 17:16. The disputed limitations all recite the term

“corresponding to,” which is consistent with confirming that it is proper to analyze a region in the second image corresponding to the head location in the first image. Ex-1001, 15:61-16:9, 16:54-61, 17:1-7, 17:48-57; Ex-1003, ¶62. In fact, Dr. Saber disparages PO’s negative limitation, which requires blindly using a head location of a first image for a second image without confirmation, by stating that “a person could move their head between image captures in a sequence, and thereby render the previous location useless in recognizing the face in the current image.” Ex-2020, ¶73.

When PO’s negative limitation is not imported into the challenged claims, PO’s arguments fail. For example, PO’s characterization that “Tian must accurately determine the location of the head/face in the current frame...” (POR, 16) is not precluded by the claims. Nor can PO reasonably dispute that face tracking involves analysis of an earlier frame in addition to a current frame—e.g., scanning a localized region in a second image around a location where a face is detected in a first image. Ex-1003, ¶64 (citing Ex-1010, ¶0003). PO’s characterization that “Tian updates the facial feature locations for each frame ... after the initial head location is determined” based on Dr. Doermann’s testimony is, again, not precluded. POR, 17. There is no dispute that when the head location does not change in two images (which often occurs), any updating will determine a

region in the second image corresponding to the head location in the first image.

E.g., Ex-1003, ¶135.

C. A POSITA would have found it obvious to combine Lu-577 and Tian.

In the DI, the Board determined that “Petitioner’s showing as to motivation to combine the relevant teachings of Lu-577 and Tian is well-supported, reasonable and persuasive” (D.I., 25), enumerating a list of Petitioner’s showings in support (D.I., 22-25). The POR addresses some showings but neglects others (*see* POR, 17-40), including:

- Lu-577 explicitly invites employing “known head location and face recognition software.” D.I., 23 (citing Pet., 8; Ex. 1003, ¶59; Ex. 1004, ¶39).
- Motivation to apply Tian’s teaching of taking a sequence of images. D.I., 23 (citing Pet., 9; Ex-1003, ¶61; Ex-1010, ¶3; Ex-1022, 7).
- The combination involves an application of known techniques (e.g., Tian’s facial analysis techniques) to a known system (e.g., Lu-577’s audience measurement system) and method ready for improvement. D.I., 24 (citing Pet., 11, Ex. 1003, ¶71).
- The combination simply uses Tian’s teachings in the manner taught by Tian to implement functionalities in Lu-577. D.I., 25 (citing Pet., 11; Ex-1003, ¶72).

These showings provide independent and unrebutted reasons for combining Lu-577 and Tian.

PO claims that “[i]ntentionally reducing the resolution of images prior to facial analysis is novel,” but cannot sustain this claim without relying on an ambiguous concept of “intentionally” and a feature of “increasing the contrast,” neither of which is claimed. POR, 17-18; Pet. 3. The challenged claims merely require reducing image resolution before performing facial analysis, and there is nothing novel about this because it is exactly what Tian does.⁶ PO’s attack on Tian for being an academic reference and bald assertion that “no contemporary image processing system would have taught or suggested internationally reducing the resolution” (POR, 17) are absurd and incorrect because, for example, Exhibit 1014 teaches intentionally reducing the resolution.⁷ Ex-1014, 2:43-53; Ex-1015, ¶0022; Ex-1003, ¶69; Pet. 10. As explained below, PO fails to rebut concrete motivations for modifying Lu-577 based on Tian’s teachings.

1. Analyzing reduced-resolution images saves processing resources.

⁶ No novelty is added by Dr. Doermann’s willingness to try what PO’s counsel himself called “a magical system,” where higher resolution images can be analyzed “with perfect accuracy and without incurring any time or computational resources,” because such a system does not exist. Ex-2018, 71:1:9; POR, 18.

⁷ PO’s arguments distinguishing Exhibit 1014 (POR, 26) lack merit. Exhibit 1014 teaches saving processing resources for face detection, which includes the step of deriving integral images, by processing reduced-resolution images. Ex-1034, ¶8; Ex-1003, ¶69; Ex-1014, 1:23-58, 2:29-53, Fig. 2.

The Petition establishes the obviousness of performing face detection or pose estimation on reduced-resolution images because doing so saves processing resources without reducing accuracy. Pet., 10-11 (citing Ex-1003, ¶69). PO fails to rebut this showing.

a. The resource savings outweigh any purported cost.

PO does not dispute that analyzing reduced-resolution images saves processing resources but contends that the saving is outweighed by cost for creating the reduced-resolution images. POR, 21-24. This argument is contradicted by PO's own expert. When questioned, Dr. Saber admitted that face detection is computationally complex, including steps such as data dimension reduction, pre-processing, feature extraction, weighing, evaluation, and comparison. Ex-1033, 75:13-91:4; Ex-2034, 14-15. The feature extraction step, by itself, includes computing convolution multiple times (e.g., horizontal, vertical) for each pixel. Ex-1033, 85:14-88:21. Head orientation determination also requires feature extraction. Ex-1033, 91:10-92:8.

In contrast, down-sampling may simply involve "pick[ing] a pixel out of every four[.]" Ex-1033, 33:15-34:1, 35:18-38:13; *see also* Ex-1003, ¶15. Dr. Saber conceded that face detection and head orientation determination are each computationally more complex than resolution-reduction methods. Ex-1033, 98:24-102:2.

Dr. Saber admitted that performing facial analysis on fewer pixels significantly saves processing resources. Ex-1033, 83:7-18, 88:22-89:1. Under Dr. Saber's examples, the savings of performing facial analysis on a reduced-resolution image with 1/4 of the pixels (i.e., eliminating 25 multiplications and 24 additions on each of 3/4 of the pixels) undoubtedly outweigh any processing needed to reduce the resolution (e.g., picking 1/4 of the pixels with no calculation). Ex-1033, 87:14-22, 33:15-34:1.

PO focuses on two special scenarios, antialiasing and interpolation, in an attempt to show cost for reducing image resolution. Both scenarios are untethered from the claims and grounds.

First, PO argues that “additional processing resources [for antialiasing] can outweigh any reduction in processing resources[.]” POR, 22-23. However, antialiasing is neither required by the '243 patent nor necessary for resolution reduction. *Supra* §I.A.3. None of the Petition's grounds require antialiasing.

Second, PO argues that resolution reduction by interpolation may “require analyzing all pixels of an image.” POR, 22-23 (citing Ex-2016; Ex-2020, ¶40, ¶44). Again, the '243 patent does not even mention interpolation, and Tian's down-sampling admittedly includes computationally more efficient methods such as

subsampling. Ex-2020, ¶43; Ex-1033, 33:15-34:1, 35:18-38:13.⁸ Regardless, Dr. Saber admits that interpolation is computationally simpler than face analysis. Ex-1033, 96:23-98:3, 100:3-14, 101:18-102:2.

PO's allegation that Petitioner has not met its burden to weigh advantages against disadvantages (POR, 20-21) overlooks the record. Dr. Doermann explains that resolution-reduction techniques included "simply subsampling the image and dropping pixels." Ex-1003, ¶15; Pet., 2. He next explains that facial analysis is "resource-consuming" and that "[t]he processing resources and time can be reduced if the resolution of the image is reduced." Ex-1003, ¶69 (citing Ex-1014, 2:29-53; Ex-1015, ¶0022); Pet. 10-11. Here, unlike *Oticon*, the advantage of performing facial analysis on reduced-resolution images easily outweigh PO's purported disadvantage. Ex-1034, ¶¶5-8. Ultimately, even if there is tradeoff between resource savings and additional costs, such a tradeoff does not obviate the motivation to combine Lu-577 and Tian because "[i]t's not necessary to show that a combination is the best option, only that it be a suitable option." *Intel Corp. v. Qualcomm Inc.*, 21 F.4th 784, 800 (Fed. Cir. 2021) (internal quotation mark omitted); *Medichem, S.A. v. Rolabo, S.L.*, 437 F.3d 1157, 1165 (Fed. Cir. 2006) ("A given course of action often has simultaneous advantages and disadvantages, and this does not necessarily obviate motivation to combine.") (citation omitted).

⁸ Tian's down-sampling does not require interpolation. Ex-1034, ¶9.

b. PO's purported redundancy does not exist.

PO's argument that the advantage of resolution reduction is redundant because head location and orientation are known from head tracking (POR, 24-25) misunderstands the combination and the claims. The Petition explains that the combination of Lu-577 and Tian teaches "reduc[ing] the resolution of a first image of the one or more of the images and perform[ing] face detection or head pose estimation on the reduced-resolution image." *E.g.*, Pet. 26-27; Ex-1003, ¶120. The combination further teaches, after determining the face/head location in a reduced-resolution image, performing face tracking in subsequent images (e.g., a second image). *E.g.*, Pet. 28; Ex-1003, ¶¶129-130. Because face detection and head pose estimation are performed on a reduced-resolution image and face tracking is performed on different, subsequent, images, PO's purported redundancy does not exist.

c. Tian's accuracy findings are applicable to audience measurement.

PO contends that Tian's findings are not applicable to audience measurement, for which a face region has a lower resolution. POR, 27-28. To begin with, PO's assumption that "the camera ... captures an image of the entire media exposure environment" (POR, 27) is contradicted by PO's expert, who testified that the camera just needs to capture part of the environment (Ex-1033, 153:22-154:9).

PO's contention relies on Dr. Saber's analysis of two examples. POR, 27-28; Ex-2020, ¶¶55-59. The first example, Figure 1 of the '243 patent, does not illustrate an image taken by a camera in a media exposure environment, but instead includes people's back and the camera (110) itself. Ex-2020, ¶56; Ex-1001, Fig. 1, 5:9-28. The second example is an image from an online shopping webpage for a sofa, where individuals are not even watching television. Ex-2020, ¶57; Ex-2037. Dr. Saber's testimony based on inapplicable examples carries no weight.

Dr. Saber's analysis is also flawed for assuming low original resolutions for images (e.g., 1280x960 and 1200x960). Ex-2020, ¶¶56-57. The assumed resolutions are about 1.2 megapixels and significantly smaller than what cameras in 2011 were capable of capturing. Ex-1032, ¶186; Ex-1033, 158:20-22; Ex-1034, ¶11. Images taken by state-of-the-art cameras would have significantly greater resolution for face regions, allowing successful facial analysis after down-sampling. Ex-1033, 155:4-9; Ex-1034, ¶¶11-13.

Dr. Saber's analysis, even assumed to be correct, cuts against PO's argument. He estimates the head region for an adult to be 75x100 pixels in the second example. Ex-2020, ¶57. This resolution, reduced by a factor of four, becomes 37.5x50 pixels. This is still larger than 36x48, for which Tian reports high-accuracy face detection and pose estimation. Ex-1005, 000005.

2. Applying Tian's pose estimation teachings would have achieved the explained benefits.

PO fails to rebut the obviousness of applying Tian's pose estimation to Lu-577. Pet. 10.

a. Pose estimation improves the performance of face recognition.

PO's argument that selectively performing face recognition would have undermined its performance (POR, 29-30) ignores that certain head orientations are not suitable for face recognition, as both experts agree. Ex-1003, ¶66; Ex-1033, 159:22-161:3. While Lu-000 teaches learning multiple face images for each person (Ex-1008, 7:25-38, 8:36-40), PO does not (and cannot) contend that images of unsuitable head orientations are learned. POR, 30; Ex-1033, 165:22-166:10. It is unquestionably beneficial not to perform face recognition on image with unsuitable head orientations.

PO also ignores that false face recognitions undermine performance. PO's example of face recognition on 10 consecutive frames with a 25% chance of success (POR, 29) ignores that unsuccessful face recognition results in identification of a wrong person (Ex-1033, 163:15-164:7). PO's example would have resulted in more misidentifications (e.g., 75% chance) than correct identifications (e.g., 25% chance), undoubtedly undermining the overall performance of face recognition. Ex-1034, ¶10.

PO argues that Dr. Doermann's testimony somehow supports its argument. POR, 29. However, what Dr. Doermann explained was that analyzing more images

would typically improve face recognition performance because the images can be analyzed to “pick the best one,” which confirms the benefit of selecting images using pose estimation. Ex-2018, 77:18-78:16. In addition, PO’s invented concept of “continuous measurement” (POR, 30) is not claimed by the ’243 patent and irrelevant to the current proceeding.

b. Pose estimation is not redundant.

PO argues that a pose estimation step would have been redundant in light of Lu-000’s “brute force method.” POR, 30-33. As explained above, this argument ignores that (1) there are head orientations not suitable for face recognition and (2) Lu-000 does not store signatures for all possible head orientations. *Supra* §I.C.2.a.

Rather than adding unnecessary processing, the head pose estimation step would prevent wasteful processing for comparing feature signatures extracted from unsuitable images with all stored signatures. Ex-1003, ¶66. This step would also prevent misidentifications resulting from analyzing unsuitable images. *Id.* Thus, pose estimation would have obviated unnecessary processing for the costly “brute force method” and improved the accuracy of face recognition.

c. Pose estimation facilitates determining if a person is watching television.

Contrary to PO’s argument (POR, 33-34), Tian’s pose estimation at least determines (1) “the angle of a face” and (2) “if the face of the person is ‘looking at’ the camera,” two criteria that PO admits to indicate whether a person is watching

television (POR, 33-34 (citing Ex-1012, ¶0039)). In agreement with Dr. Doermann (Ex-1003, ¶67), Dr. Saber admits that Tian’s frontal view “is angled ~0 degrees from the camera” and indicates that that the person “is looking directly toward the camera[.]” Ex-2001, ¶92. Tian groups frontal view and near frontal view together because both are suitable for expression analysis. Ex-1005, 000002-03. This grouping choice does not negate Tian’s ability to determine individuals of the frontal view and the fact that they are looking at the camera, nor does PO dispute the obviousness of doing so. *See generally* Ex-1020.

PO’s argument that advertisers would rather NOT know whether members of the audience look at the television (POR, 34-35) is specious. It is also misplaced because using pose estimation to determine if individuals are watching television does not inhibit the ability to identify those who are not.

3. Modifying Lu-577 based on Tian’s face tracking teachings would have achieved the explained benefits.

None of PO’s purported drawbacks (POR, 35-36) regarding this modification has merit. First, PO’s argument about “increase in processing resources” (POR, 36-37) misses the point because replacing face detection with face tracking reduces the number of pixels analyzed and the processing resources needed to do so. Dr. Doermann explains that face tracking allows a face to be “accurately tracked from frame to frame without a need to run a face detector across the entire image.” Ex-1003, ¶64 (citing Ex-1010, ¶0003); Pet., 9. Dr. Saber

agrees that face tracking saves processing resources, as he contends that this benefit makes the benefit of analyzing reduced-resolution images “redundant.” Ex-2020, ¶¶45-46.

Second, PO asserts that, even with face tracking, face detection is still required for locating new faces entering the field of view. POR, 37-38. This assertion is irrelevant to the current proceeding because all challenged claims concern “a head.” Neither the claims nor the prior-art combinations require locating new faces in any image, let alone an image in which face tracking is performed.

Third, PO argues that face tracking harms the accuracy of face recognition when using out-of-date face locations. POR, 38-40. This argument is untethered from the Petition’s obviousness theory, which states that “Tian teaches determining a face/head location in a first image and using this location for subsequent images *unless a change is detected*.” Pet. 9 (emphasis added). Using a face/head location from a previous image when no change is detected ensures no out-of-date face location is used.

D. The Petition’s use of incorporated teachings of Lert and Lu-000 does not combine different embodiments.

In the DI, the Board correctly determined that “Lu-577’s incorporation by reference of Lu-000 and Lert is not at all akin to combining two embodiments.” I.D. 32. This is because the relied-on teachings of Lert and Lu-000 are both

incorporated to Lu-577's disclosure of its first embodiment, which corresponds to paragraphs 35-43 and Figure 2. Ex-1004, ¶¶0035-43, Fig. 2.

In paragraph 39, Lu-577 teaches “employ[ing] known head location and face recognition software (e.g., as taught by [Lu-000])[.]” Ex-1004, ¶0039. In paragraph 40, Lu-577 explains that “characteristic signatures can be extracted from these representations [of audio portions of television programs] ... in order to identify the television programs[.]” Ex-1004, ¶0040. For this signature-based approach, Lu-577 incorporates Lert by reference. Ex-1004, ¶0003. While PO argues that Lu-577 teaches other approaches (POR, 43; *see* Ex-1004, ¶0002, ¶0004), the Petition does not rely on any other approach or seek to combine it with Lert's signature-based approach. PO's attorney argument that paragraph 7 of Lu-577 disparages Lert (POR, 43) lacks basis because this paragraph is unrelated to Lert's signature-based approach (Ex-1034, ¶14) and does not amount to a teach-away.

Since Lu-577 discloses using Lu-000's and Lert's teachings for its first embodiment, they are “part of what the inventors viewed as their invention” (POR, 43) and are not used to “modify” Lu-577 (*id.*). Because Lu-000 and Lert are properly incorporated by reference into Lu-577 (D.I., 32-33) and are directed to the same embodiment, no motivation to combine is required. *Gen. Elec. Co. v. Raytheon Techs. Corp.*, 983 F.3d 1334, 1352 (Fed. Cir. 2020) (holding that a motivation to combine elements present together in a reference is not required).

II. GROUND 3

A. The combination of Ground 3 teaches the disputed limitations.

PO reiterates its arguments that Tian does not teach the disputed limitations. POR, 49-50. These arguments fail for the same reasons explained above. *Supra* §I.A-B.

B. PO's mischaracterization of Petitioner's combination of Nielsen-372 and Steinberg should be rejected.

PO's attacks the combination of Nielsen-372 and Steinberg (POR, 50-57) by asserting that the combination relies upon the camera of Nielsen-372 (POR, 44-49). The Board correctly pointed out PO's mischaracterization, finding that "[f]or each of [the] limitations in the independent claims that contains the first recitation of 'a camera' ..., Petitioner cites teachings in Steinberg, not Nielsen-372." D.I., 41-43 (citing Pet., 53, 65, 69).

PO insists on its mischaracterization based on the following two sentences in the Petition:

- "Nielsen-372's system includes needed hardware (e.g., a computer and a camera) for implementing Steinberg's approach." POR, 44-45 (citing Pet. 44; Ex-1003, ¶200), 48.
- "Nielsen-372 provide the hardware architecture for performing Steinberg's method..." POR, 45 (citing Pet. 45, Ex-1003, ¶¶199-200, 203.), 48.

In both sentences, Steinberg’s approach/method refers to identifying individuals in the audience using face detection, normalization, and recognition, as opposed to Nielsen-372’s manual-input-based approach. *See, e.g.*, Ex-1003, ¶202 (“Steinberg teaches a different *approach* of identifying individuals by performing face detection, normalization, and recognition ...”); Pet., 44.

Dr. Doermann explains that “modules for performing face detection, normalization, and recognition may be implemented on a computer ..., an image acquisition device ..., or both” (Ex-1003, ¶199) and that it would have been obvious to implement these functions “on either the computer or the camera of Nielsen-372...” (Ex-1003, ¶200) because “Nielsen-372 provide the hardware architecture for performing Steinberg’s face recognition method ...” (Ex-1003, ¶203). Nowhere do the Petition or Dr. Doermann propose physically implementing Steinberg’s camera using Nielsen-372’s camera. PO’s mischaracterization should be rejected.

C. PO’s mischaracterization does not undermine the obviousness showing of Ground 3 regardless.

PO’s attack on the combination of Nielsen-372 and Steinberg essentially asks the Board to find the challenged claims non-obvious for claiming a camera pointing at individuals rather than a television. POR, 50-57. The ’243 patent, as well as Dr. Saber, admits this to be a prior-art feature. Ex-1001, 1:56-66; Ex-1033, 23:10-12; *see* D.I., 43, n.28.

First, PO's argument is improperly directed to bodily incorporation, as it focuses on where a specific camera in Nielsen-372's system is pointing at. POR, 47-48, 51-55. This argument also lacks merit because applying Steinberg's teachings to Nielsen-372 does not necessitate turning the Nielsen-372's camera away from the television. Dr. Saber confirmed a POSITA's ability to connect two cameras to a computer to simultaneously capture a television and an audience (Ex-1033, 23:17-24, 24:17-25:9) as well as the availability of cameras with multiple image sensors for different directions or a wide field of view by 2011 (Ex-1033, 24:5-16, 27:8-21). Available technologies would have obviated the need to stop capturing images of the television.

Second, PO contends that not using the video sensor would have frustrated Nielsen-372's purpose and changed its principle of operation. POR, 53-55. However, Nielsen-372 explicitly teaches that the video sensor (e.g., camera) is optional. For example, claim 1 recites "[a] method to determine whether a presentation device is ON or OFF" based on "an audio signal," without requiring video signal. Ex-1006, 17-18 (cl. 1). Thus, not using Nielsen-372's camera to capture images of the television, even if required by the combination, is not inconsistent with the purpose and intended operation of Nielsen-372's system.

Third, PO's claim that "voluntary audience measurement" is a principle of operation of Nielsen-372 (POR, 54-56) lacks support because Nielsen-372 does not

mention the concept of “voluntary audience measurement” or discuss any benefit of such an approach. Nielsen-372 teaches that its people meter is “to capture information about the audience” and “*may* prompt the audience members to indicate that they are present ...” Ex-1006, ¶48 (emphasis added). Dr. Doermann explains that applying Steinberg’s method furthers Nielsen-372’s goal of capturing audience information by reducing interruptions and improving accuracy. Ex-1003, ¶201; Pet., 44. Dr. Saber’s testimony confirms this position by admitting that the purported “voluntary” system interrupts individuals to “identify themselves” and undercounts viewing activities for “certain guilty pleasure shows.” Ex-2001, ¶¶150-151.

D. PO’s attacks on the motivation to combine with Tian fail for the same reasons as explained for Ground 1.

PO’s arguments against the motivation to combine for Ground 3 essentially mimics those for Ground 1. They fail for the same reasons as explained above. *Supra* §I.C.

III. ANALOGOUS ART

Dr. Doermann explains that the field of the ’243 patent is media analysis, including analysis of audio signals and image signals. Ex-1003, ¶46. PO’s proposed field of endeavor, “audience measurement” (POR, 62), is contradicted by Dr. Saber, who accepts Petitioner’s POSITA definition requiring experience in “media analysis.” Ex-2001, ¶¶26-29. When questioned, Dr. Saber further admitted

that he has no experience in audience measurement (Ex-1033, 23:2-9), the “field” of the ’243 patent is “digital image/video processing, computer vision, pattern recognition, machine learning,” (Ex-1033, 19:5-21) and his experience in this field qualifies him as a POSITA and expert (Ex-1033, 13:16-23, 22:15-23). PO’s narrow definition ignores that audience measurement is an application in the field of media analysis. *E.g.*, Ex-1033, 14:25-15:7, 16:8-16, 23:2-9. That the ’243 patent, Tian, and Steinberg does not literally recite “media analysis” (POR, 60) does not negate the fact that all of them are directed to media analysis, which Dr. Saber admits to include analyzing audio data and images. Ex-1033, 21:10-19.

Date: July 24, 2024

Respectfully,

SHEPPARD, MULLIN,
RICHTER & HAMPTON LLP

/Harper Batts/

Harper Batts (Reg. No. 56,160)
Attorney for Petitioner

CERTIFICATE OF COMPLIANCE

Pursuant to 37 C.F.R. § 42.24(d), the undersigned certifies that the foregoing Petitioner's Reply to Patent Owner's Response contains 5,583 words and therefore complies with the 5,600 word type-volume limit specified by 37 C.F.R.

§42.24(c)(1). The word count does not include those portions excepted by 37 C.F.R. §42.24(c) and was calculated by Microsoft Word 365.

Date: July 24, 2024

/Harper Batts/
Harper Batts (Reg. No. 56,160)

CERTIFICATE OF SERVICE

The undersigned hereby certifies that on July 24, 2024, a complete copy of the foregoing Petitioner's Reply to Patent Owner's Response was served via email to all parties to this proceeding at the addresses indicated:

Jennifer C. Bailey
jennifer.bailey@eriseip.com

Megan J. Redmond
megan.redmond@eriseip.com

Robin A. Snader
robin.snader@eriseip.com

PTAB@eriseip.com

Date: July 24, 2024

/Harper Batts/
Harper Batts (Reg. No. 56,160)
Attorney for Petitioner