

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 IPR2023-01014
Patent No. 11,470,243

**PATENT OWNER'S RESPONSE TO PETITION
FOR *INTER PARTES* REVIEW**

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	GROUND 1: THE COMBINATION OF LU-577 AND TIAN DOES NOT RENDER CLAIMS 4-6, 8, 11-14, AND 18-20 OBVIOUS	2
A.	PETITIONER’S PROPOSED COMBINATION FAILS TO TEACH THE CLAIM LIMITATIONS.....	2
1.	<i>Tian Does Not Teach the Claimed “Reduc[ing] a Resolution”</i>	<i>3</i>
2.	<i>Petitioner’s References, Alone or in Combination, Do Not Teach the Claimed Two-Frame Process</i>	<i>13</i>
B.	THE PETITION FAILS TO ESTABLISH A SUFFICIENT MOTIVATION TO COMBINE THE REFERENCES	17
1.	<i>The Combination Does Not Achieve Petitioner’s Proposed Benefit, and Any Benefit Would Be Outweighed by Drawbacks Petitioner Fails to Address.....</i>	<i>19</i>
2.	<i>The Petition Fails to Provide Any Motivation to Combine the Teachings of Lu-000 and Lert with Lu-577</i>	<i>41</i>
III.	GROUND 3: NIELSEN-372 IN VIEW OF STEINBERG AND IN FURTHER VIEW OF TIAN DOES NOT RENDER OBVIOUS THE CHALLENGED CLAIMS	44
A.	OVERVIEW OF THE CITED REFERENCES AND PETITIONER’S PROPOSED COMBINATION	44
1.	<i>Petitioner’s Proposed Modification of Nielsen-372 to incorporate the teachings of Steinberg.....</i>	<i>44</i>
2.	<i>The Institution Decision Misapprehends the Combination Proposed in the Petition</i>	<i>46</i>
B.	PETITIONER’S PROPOSED COMBINATION FAILS TO TEACH THE CLAIM LIMITATIONS.....	49
1.	<i>Tian Does Not Teach Instructions, a Device, or a Step for “Reduc[ing] a Resolution”</i>	<i>49</i>
2.	<i>Petitioner’s References, Alone or in Combination, Do Not Teach Performing Head Location Determination in a First Image and Face Recognition on the Determined Location in a Second Image.....</i>	<i>49</i>
C.	THE PETITION FAILS TO ESTABLISH A SUFFICIENT MOTIVATION TO COMBINE THE REFERENCES	50

1.	<i>A POSITA Would Not Have Found It Obvious and Been Motivated to Modify Nielsen-372 with the Teachings of Steinberg</i>	51
2.	<i>Modifying Nielsen-372/Steinberg with Tian’s Teachings Does Not Achieve Petitioner’s Proposed Benefit, and Any Benefit Would Be Outweighed by Drawbacks Petitioner Fails to Address</i>	57
IV.	NEITHER TIAN NOR STEINBERG IS ANALOGOUS ART TO THE ’243 PATENT	59
A.	NEITHER THE ’243 PATENT NOR PETITIONER’S ART IN THE FIELD OF “MEDIA ANALYSIS”	60
B.	PETITIONER’S ART IS NOT REASONABLY PERTINENT TO THE PROBLEM BEING SOLVED BY THE ’243 PATENT	62
V.	CONCLUSION	64

TABLE OF AUTHORITIES

Cases:

<i>10X Genomics, Inc. v. Bio-Rad Laboratories, Inc.</i> , IPR2018-00489, Paper 13, (PTAB June 2018).....	30-31
<i>Abiomed, Inc., v. Maquet Cardiovascular, LLC</i> , IPR2017-01026, Paper 8, (PTAB Sept. 20, 2017)	41-42
<i>Airbus S.A.S. v. Firepass Corp.</i> , 941 F.3d 1374, (Fed. Cir. 2019)	63
<i>Arctic Cat Inc. v. Bombardier Recreational Prods. Inc.</i> , 876 F.3d 1350, (Fed. Cir. 2017)	20, 25
<i>Chemours Company FC, LLC v. Daikin Industries, Ltd.</i> , No. 20-1289, Slip Op. (Fed. Cir. Jul. 22, 2021)	44
<i>Cook Group Incorporated v. Boston Scientific Scimed, Inc.</i> , IPR2017-00440, Paper 7	19, 47
<i>IBM v. Chewy, Chewy, Inc. v. Int’l Business Machines Corp.</i> , 94 F.4th 1354, (Fed. Cir. Mar. 5, 2024)	1
<i>In Re Keller</i> , 642 F.2d 413, CCPA 1981).....	45
<i>In re Gordon</i> , 733 F.2d 900, (Fed. Cir. 1984).....	54
<i>In re Stepan</i> , 868 F.3d 1342, (Fed. Cir. 2017).....	41
<i>Intel Corp. v. Qualcomm Inc.</i> , 21 F.4th 784.....	25
<i>Netflix v. DivX</i> , No. 22-1138, (Fed. Cir. Sept. 11, 2023)	60, 63
<i>Oticon Medical AB v. Cochlear Limited</i> , IPR2019-00975, Paper 52, (PTAB Oct. 2019)	20-21, 36, 37, 40
<i>RJ Reynolds Vapor Co. v. Fontem Holdings I B.V.</i> , IPR2016-01692, Paper 45, (PTAB March 2018).....	31
<i>SAS Inst. Inc. v. Iancu</i> , 138 S. Ct. 1348, (2018).....	48
<i>SK Hynix Inc. et al., v. Netlist, Inc.</i> , IPR2020-01042, Paper 14, (PTAB December 2020).....	45
<i>TCT Mobile (US), Inc. v. Fundamental Innovations Systems Int’l LLC</i> , IPR2021-00597, Paper 8, (PTAB Aug. 25, 2021).....	54, 56

<i>Victoria’s Secret Stores LLC v. Andra Group, LP</i> , IPR2020-00853, Paper 12, (PTAB Oct. 22, 2020)	54
--	----

<i>Virtek Vision International ULC v. Assembly Guidance Systems, Inc. d/b/a Aligned Vision</i> , Nos. 2022-1998, (Fed. Cir. Mar. 27, 2024).....	2
---	---

Regulations:

37 C.F.R. § 42.8.....	66
37 C.F.R. § 42.24.....	66

I. INTRODUCTION

Patent Owner The Nielsen Company (US), LLC (“Patent Owner”) requests the Board find all challenged claims not unpatentable because Petitioner TVision Insights, Inc. (“Petitioner”) fails to show the unpatentability of any challenged claims 4–6, 8, 11–14, and 18–20 of U.S. Patent No. 11,470,243 (“the ’243 Patent”).

Petitioner’s mapping ignores the plain language of the claims and the requirements of the claims as a whole and fails to provide a viable motivation to combine. First, Petitioner ignores that the claims require affirmatively reducing the resolution of images as part of the claimed audience measurement system and method. Despite this plain requirement, Petitioner maps this limitation to a reference that relies on pre-reduced images of various resolutions that are stored in a separate database. There is no explanation of how this teaches the claim limitation. In fact, a recent Federal Circuit decision highlights the importance of this distinction. *IBM v. Chewy, Chewy, Inc. v. Int’l Business Machines Corp.*, 94 F.4th 1354, 1351-62 (Fed. Cir. Mar. 5, 2024) (distinguishing pre-processing versus processing on demand).

Second, the claims require a two-frame process for facial recognition versus the one-frame process used by the prior art. In the patent’s two-frame process, the first frame identifies a head location, and the second frame uses the identified head location from the first frame to perform facial recognition in the second frame.

The one-frame process in *Tian* does not use the face location from the first frame for facial recognition in the second frame.

Third, the motivations to combine are utterly lacking and contrary to the strict requirements of Federal Circuit law. *See, e.g., Virtek Vision International ULC v. Assembly Guidance Systems, Inc. d/b/a Aligned Vision*, Nos. 2022-1998, -2022 (Fed. Cir. Mar. 27, 2024) (“It does not suffice to meet the motivation to combine requirement to recognize that two alternative arrangements ... were both known in the art.”)

Petitioner fails to show the unpatentability of any of the challenged claims, and the Board should find all challenged claims not unpatentable.

II. GROUND 1: THE COMBINATION OF LU-577 AND TIAN DOES NOT RENDER CLAIMS 4-6, 8, 11-14, AND 18-20 OBVIOUS

A. Petitioner’s Proposed Combination Fails to Teach the Claim Limitations

Tian’s system does not teach affirmatively reducing a resolution, as required by the challenged claims, but instead relies on images from a database that have been pre-processed (at some previous time by some unspecified party) to reduce their resolution. *Tian* also fails to teach performing face recognition on a *different* image from that in which the face was detected and its orientation determined, as required by Challenged Claims 5-6, 8, 12-14.

1. *Tian Does Not Teach the Claimed “Reduc[ing] a Resolution”*

Challenged Claims 4, 11, and 18 recite similar limitations for processing circuitry, a device, or a method step to “reduce a resolution of a first image,”¹:

4. The audience measurement system of claim 1, wherein the processor circuitry is to:

reduce a resolution of a first image of the one or more of the images of the media exposure environment to obtain a reduced-resolution image, and
determine the orientation of the head with respect to the camera based on the reduced-resolution image.

11. The audience measurement system of claim 9, wherein the audience measurement device is to:

reduce a resolution of the first image of the media exposure environment to obtain a reduced-resolution image, and
determine the orientation of the head with respect to the camera based on the reduced-resolution image.

18. The method of claim 16, further comprising:

reducing a resolution of the first image of the media exposure environment to obtain a reduced-resolution image, and
wherein the determining of the orientation of the head includes determining the orientation of the head with respect to the camera using the reduced-resolution image.

'243 Patent, 15:56-58, 16:49-51, 17:41-43. *Tian* does not teach actually reducing the resolution of the images, much less circuitry, instructions, or a device to do so. Instead, *Tian* discloses a pre-existing “public database” of images that “are down-

¹ Claim language is underlined and *italicized* for ease of reference.

sampled” versions of the original images. *Dec.* 46, 48. *Tian* fails to disclose the audience measurement system of claims 4 or 11 or the method of claim 18, which teach a specific process of reducing the resolution of images as part of the overall system or method. The specification describes how the image sensor of the device reducing the resolution enhances the contrast and allows a low-light frame to trigger illumination for a succeeding frame—thus enabling the advantage of the invention. ’243 Patent, 11:5-36. In other words, the specification never describes the use of pre-processed images. Instead, the specification *only* describes using circuitry *in the audience measurement system* to reduce the resolution of images as part of a specific system or method. *See id.* at FIG. 3.

a) The Institution Decision (“ID”) misapprehends the requirements of the claims

Claim 4 includes an affirmative limitation of processor circuitry to execute ... machine-readable instructions to ... reduce a resolution of a first image. *See* limitations [1a] and [4a]. Claim 11 affirmatively recites that the audience measurement device is to ... reduce a resolution of the first image[.] *See* limitations [11p] and [11a]. Claim 18 recites a method that includes an affirmative step of reducing a resolution of the first image in the method. *See* limitation [18a].² Despite

² For ease of references, these limitations are collectively referred to as “the resolution reduction limitations.”

this, the Institution Decision determines that “Tian plainly teaches obtaining reduced-resolution images[.]” *Institution Decision*, 22. This conclusion ignores the specific requirements of the claims as a whole.

Patent Owner submits that the claims do not recite a method step of “obtain[ing] a reduced-resolution image[.]” Rather, the claims recite instructions, a device, and an affirmative step to “reduce a resolution” of a first image—i.e., to actually perform the processing needed to reduce the resolution of an image. “[O]btain[ing] a reduced-resolution image” is not “reduc[ing] a resolution,” because it does not perform the claimed processing. *Tian*’s experimental setup describes no such instructions, device, or step. Indeed, as shown below, *Tian* shows in FIG. 1 the “basic structure” (i.e., the steps) of the algorithms tested, and no “reduce the resolution” step is included:

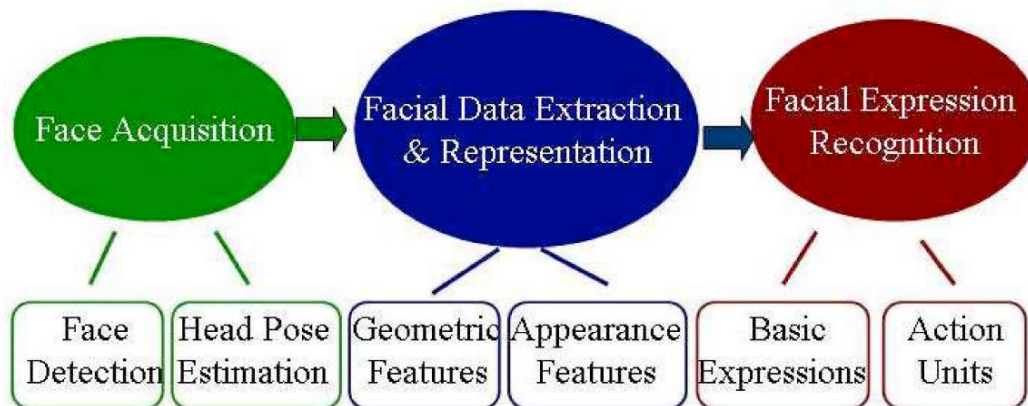


Figure 1: Basic Structure of Facial Expression Analysis Systems.

Tian, FIG. 1.

The Institution Decision also states that *Tian* “teaches down-sampling as shown in the Petition.” *ID*, 22. But *Tian* teaches *obtaining* images that “are down-sampled” by some undetermined process; down-sampling is not a part of the “Basic Structure” of *Tian*’s “Facial Expression Analysis Systems.” *Tian*, FIG.1 (*supra*). *Tian* does not perform such down-sampling or disclose any computer-readable instructions, device, or method step for performing the down-sampling. *Tian*, *generally*.

Petitioner’s mapping mischaracterizes the teachings of *Tian*. *Petition*, 26 (“*Tian creates* lower-resolution images by down-sampling from original images.”) (*emphasis added*). Petitioner’s characterization of *Tian*’s language is wrong, as *Tian* does not *create* the down-sampled images; it just accesses existing down-sampled images. *Dec* 48 (Dr. Saber opining that a POSITA would not understand that *Tian* teaches a facial expression analysis system that creates lower-resolution images or performs any resolution reduction step.).

The distinction between “obtaining reduced-resolution images” and the resolution-reduction limitations of the claims is important. These limitations are part of the claims and must be afforded patentable weight. *See, e.g., Chewy*, 94 F.4th at 1351-62 (Fed. Cir. Mar. 5, 2024) (affirming summary judgement of non-

infringement based on distinguishing pre-processed content versus content processed on demand).

b) Tian’s system does not include a process, circuitry, or a device for reducing the resolution of an image

Tian is directed to “explor[ing] the effects of different image resolutions for each step of facial expression analysis.” *Tian*, Abstract. *Tian* uses images derived from a “widely used public database” at five different resolutions. *Id.* To clarify that these images are identical but for resolution, *Tian* states that “lower resolution images are down-sampled from the originals.” *Id.*, 000002. This is the basis for Petitioner’s contention that *Tian* teaches the resolution-reduction limitations. *Petition*, 26. Petitioner mischaracterizes *Tian*’s teachings in support of the claim that “Tian **creates** lower-resolution images by down-sampling from original images.” *Id.* However, a POSITA would understand Tian’s language to mean that the lower-resolution images are down-sampled *versions* of the original images. *Dec.* 46-50.

Petitioner’s carefully selected quotes do not establish that *Tian* teaches actively creating the reduced-resolution images. *Id.* Rather, Tian relies on a single set of **pre-processed images** to test a variety of algorithms at each resolution. *See, e.g., Tian*, Table 3 (summarizing the results for a variety of image resolutions); *see also Supp. Dec.*, 4-7 (Dr. Saber opining that a POSITA would read *Tian* as using the pre-processed images from the database rather than repeatedly reducing the resolution of each image for each algorithm).

The caption for Table 1 begins by characterizing the images as “[a]n example in **five different resolutions from Cohn-Kanade database.**” *Tian*, 000002. *Tian* goes on to state, “five different resolutions of the head region are studied ... **based on** Cohn-Kanade AU-Coded Face Expression Image Database[.]” *Id.* Neither Petitioner nor their declarant makes any attempt to reconcile *Tian*’s statement that the “five different resolutions” are “from [the] Cohn-Kanade database” with the actual requirements of the resolution-reduction limitations. *Petition*, 26; Ex. 1003, ¶ 119. *Tian* is not performing the down-sampling. *Dec.*, ¶¶ 47-49. If *Tian* were actively reducing the resolution in the *Tian* process (as required by claims 4, 11, and 18), then there would be no reason for *Tian* to describe its process as using the five different resolutions “based on” the Cohn-Kanade database. *Dec.* ¶ 49. As described above, the claims require “reducing a resolution[.]” not simply obtaining reduced-resolution images.

Analyzing the facial expressions at whatever resolution was available from the Cohn-Kanade database would also be consistent with *Tian*’s actual purpose, which is to investigate expression analysis in “real environments[.]” *Tian*, Abstract; *Dec.*, ¶ 50. *Tian* does not teach reducing the resolution of images as a part of performing facial analysis; rather, *Tian* investigates pre-processed low-resolution images. *Tian*, 000002 (“Most systems attempt to recognize facial expressions from data collected in a highly controlled laboratory situation with very high resolution

faces. **However, in real applications, the face region is often in lower resolutions.**") (internal parentheticals omitted). In other words, a POSITA would read *Tian* as teaching that expressions are analyzed in images at whatever resolutions are available from the database. *Dec.*, ¶ 50. There would be no reason to actively reduce the resolution of an image further for the purpose of performing expression analysis. *Dec.*, ¶¶ 50-51.

As noted above, *Tian* provides a high-level outline of the structure (*i.e.*, the steps) of the Facial Expression Analysis Systems being evaluated. *Tian*, FIG. 1. Notably, this structure fails to include any step of “reduce a resolution[.]” *See also Supp Dec.*, 8-10 (Dr. Saber opining that a step of reducing the resolution would not fall into any of the depicted steps or sub-steps, and a POSITA would not be motivated to include a resolution-reduction process within any of these steps). *Tian* views low-resolution images as a natural occurrence, undesirable but unavoidable. *Supp. Dec.*, 11. In light of the plain view of the claims, *Tian* does not disclose the resolution-reduction limitations.

Finally, as noted above, claims 4 and 11 require “circuitry [] to [] reduce a resolution” and an “audience measurement device [] to [] reduce a resolution[.]” Petitioner’s declarant admits that *Tian* includes no circuitry. *Ex. 2018, Doermann Declaration*, 27:14-21 (“‘Circuitry’ does not appear in *Tian*.”). Similarly, Petitioner’s declarant admits that *Tian* includes no disclosure of computer-readable

instructions, and any such instructions must be inferred. *Id.*, 30:19-21 (“[T]he algorithms here [in Tian] are an abstraction of the implementation that **surely would have been done** in order to get the results.”). As such, regardless of how *Tian*’s images are reduced, it is undisputed that *Tian* includes no disclosure of the claimed hardware that performs the resolution reduction process.

c) Applying *Tian*’s teachings would not result in a useable “audience measurement system” as claimed

The resolution-reduction limitations represent a pertinent step in the claimed system and method and are important because the claimed step of reducing a resolution occurs in the context of an audience measurement system, which imposes unique operating constraints for image processing. For example, the specification explains that reducing the resolution achieves the advantage of improved contrast in low-light images, common in a media exposure environment. ’243 *Patent*, 10:58-63 (unilluminated images are too dark and low-contrast for analysis); 11:24-29 (reducing the resolution can increase the amount of light per pixel, thereby enabling analysis of low-light images). The resolution-reduction process can be performed quickly enough to be processed for use in the next frame (the “two-frame process”) so as to minimize the use of the illumination source, which is distracting to the audience. *Id.* at FIG. 3; 11:5-36; 9:18-31; 10:1-13; 4:40-51.

Tian does not teach this process, and *Tian*’s teachings of pre-processed reduced-resolution images would not result in a usable audience measurement

system and method. *Supp. Dec.*, 12-14. First, *Tian* relies on pre-processing for resolution reduction. *Tian*, 000002 (characterizing Table 1 as “[a]n example in five different resolutions from Cohn-Kanade database.”) Using a pre-processed image to determine head orientation does not make sense because the head orientation is used to determine whether to activate the illumination source. *'243 Patent*, 13:41-51; *see also unchallenged claims* 24-25. By relying on images captured and pre-processed before the analysis begins, a system relying on *Tian*’s teachings would not be able to activate the illumination source in order to perform analysis of succeeding frames,³ because the succeeding frames would have been captured before the analysis of the first frame could trigger illumination. Because no frames would be illuminated, the system would never be able to perform facial analysis of the unilluminated images. *'243 Patent*, 10:44-54 (images suitable for facial recognition are those that are high-resolution and well-lit). Such a system or method would thus also be unable to identify audience members and would be of no use as an audience measurement system or method. *Supp. Dec.*, 15.

Second, relying on *Tian*’s teachings would fail to result in a useable audience measurement system or method for another reason as well: images captured in a media-exposure environment are typically low-contrast, and down-sampling does

³ As recited, for example, in claims 5, 6, 12, 14, and 19.

not increase the contrast of images. '243 *Patent*, 3:44-50; *Supp. Dec.*, 16-17; *see also Ex. 2018, Doermann Deposition Transcript* 63:21-64:6 (Dr. Doermann agreeing that *Tian* would assume that it would be beneficial to compare images that are identical but for resolution). Because the reduced-resolution images in the Cohn-Kanade database are produced by down-sampling, applying *Tian*'s teachings (rather than a suitable technique) would result in low-resolution, low-contrast images not suitable for analysis. '243 *Patent*, 10:58-67 (when the illuminator is not lit, even high-resolution image data is likely to be dark and low-contrast and thus not suitable for analysis); *see also Supp. Dec.* 18. The other techniques identified by Petitioner, subsampling and interpolation, also fail to increase the contrast of images and thus share the same fatal flaw. *Petition*, 2; *Supp Dec.*, 19. Because the resulting low-contrast images would not be suitable for analysis, such a system would be unable to identify frames suitable for facial recognition. Again, such a system or method would be unable to identify audience members and would not be functional as an audience measurement system or method.

Third, *Tian*'s teachings of down-sampling would not result in a useable audience measurement system or method because down-sampling alone does not produce an image ready for analysis. Dr. Saber discusses how down-sampling without an appropriate antialiasing step can lead to aliasing, which in turn could interfere with head detection and render the system inoperable. *Supp. Dec.*, 20.

Adding the necessary antialiasing step, however, also adds processing requirements that may extend processing time unacceptably. *See also* § (B.1.a)(1) *infra* (discussing the processing requirements of antialiasing). For example, if a complex antialiasing algorithm is used, the scene in the media exposure environment may have changed before the time the head detection step can trigger facial recognition, frustrating the claimed two-frame process because the head has moved in the interim. *Supp. Dec.*, 21-24. Unlike the '243 Patent, Neither *Tian* nor Petitioner begins to address the complexities a POSITA would face when actually implementing down-sampling in an audience measurement system.

Because (1) *Tian* does not teach the resolution-reduction limitations, and (2) the only technique mentioned by *Tian* for producing reduced-resolution images would not result in a usable audience measurement system, the Board should find these claims not unpatentable.

2. *Petitioner's References, Alone or in Combination, Do Not Teach the Claimed Two-Frame Process*

Claims 5, 6, 12, and 19 require a two-frame process for facial recognition.

Broadly:

1. The location of a head is identified in a first frame; and
2. The identified head location from the first frame is used to perform facial recognition in a second frame.

This is different from the prior art, including Tian, which identifies the face location and performs facial analysis in the same frame. For every subsequent frame, Tian updates the face location and does not use the face location from the first frame. Simply put, Tian uses a single-frame process for facial analysis. *Supp. Dec.*, 26.

The key claim limitations reciting this two-frame process are reproduced below:

4. The audience measurement system of claim 1, wherein the processor circuitry is to:

reduce a resolution of a first image of the one or more of the images of the media exposure environment to obtain a reduced-resolution image, and determine the orientation of the head with respect to the camera based on the reduced-resolution image.

5. The audience measurement system of claim 4, wherein the processor circuitry is to determine the audience identification information by: (i) generating a facial signature from a region of a second image of the one or more of the images corresponding to a location of the head in the reduced-resolution image, and (ii) comparing the generated facial signature to a database of facial signatures.

6. The audience measurement system of claim 4, wherein the processor circuitry is to:

identify a region corresponding to the head within the first image from which the reduced-resolution image was obtained;

analyze a portion corresponding to the identified region of a second image; and

based on the analysis of the region of the second image, determine that the head matches the known person.

12. The audience measurement system of claim 11, wherein the audience measurement device is to:

identify a region corresponding to the head within the first image from which the reduced-resolution image was obtained; and

analyze a portion of a second image corresponding to the identified region to determine whether the head matches the known person.

19. The method of claim 18, further comprising:

identifying a region corresponding to the head within the first image from which the reduced-resolution image was obtained;

analyzing a portion of the second image corresponding to the identified region; and

based on the analyzing of the region of the second image, determining whether the head matches the known person.

Petitioner maps *Tian*'s teachings of "determining a face/head location in a first image and using this location for subsequent images unless a change is detected" to the claimed two-frame process. *Petition*, 28, 30, 37, 42, 77, 83, 87. Yet, these portions of *Tian* highlight that *Tian* is a single-frame, not two-frame, process.

Tian updates ("tracks") the location of the head/face for each frame to perform face analysis for that frame. *Tian*, 000001 ("a face detector to detect a face in each frame or just detect face in the first frame and then track the face in the remainder of the video sequence."); *Supp. Dec.*, 27. Regardless of whether *Tian* detects the head/face location in each frame or tracks the movement of the head/face in the current frame based on prior frames, *Tian* is locating and conducting facial analysis of a single frame (the "single-frame process"). *Supp. Dec.*, 27. This is in contrast to the claimed invention that determines a head/face location in the first frame and then

uses that location in the second frame to perform facial recognition (the “two-frame process”).

Tian’s tracking of the head and face between frames shows that the location of the head/face *for the current frame* is used for facial analysis of a given frame rather than the location of the head/face for a *previous* frame. *Supp Dec.* 28-29. *Tian* further explains that images must be “well aligned” (*i.e.*, tracked accurately) for facial analysis. *Tian*, 000006; *see also Supp Dec.* 29-30. Simply put, *Tian* must accurately determine the location of the head/face in the current frame to perform facial analysis of that same frame. *Tian* does not teach the claimed method of using the location of the head/face in a first frame to perform facial analysis of later frames.

Despite the obvious difference between the claimed two-step process and *Tian*’s one-step process, Petitioner’s declarant glosses over this difference and claims “[v]ideo cameras were able to take multiple images in a relatively short period of time, in which a person’s head or face is unlikely to move significantly.” *Ex. 1003*, ¶ 64 (citing *Ex-1010*, ¶0003) (“Once a face is detected, its location is recorded and a localized region around that face is scanned by a face detector in the next frame. Thus, once a face is initially detected, it can be accurately tracked from frame to frame without a need to run a face detector across the entire image”). This ignores the fact that *Tian* is determining the location of the head and face in each frame, regardless of the level of movement. *Supp. Dec.*, 31. When questioned, Dr.

Doermann admitted that *Tian* updates the facial feature locations for each frame in the remainder of the video sequence after the initial head location is determined. *Ex. 2018*, Doermann Deposition Transcript, 36:20-37:1, 39:18-21 (Dr. Doermann agreeing that Tian updates the facial feature locations for each successive frame based on the previous frame). 1. Thus, Petitioner's declarant admits that the face locations are updated for every frame. As such, a combination relying on *Tian's* teachings would not teach the claimed method of using the location of the head/face in a first frame to perform facial analysis of later frames.

Because *Tian* does not teach the claimed two-frame process, Petitioner has failed to meet its burden. As such, the Board should find claims 5, 6, 12, and 19 not unpatentable for at least this reason as well.

B. The Petition Fails to Establish a Sufficient Motivation to Combine the References

The Petition relies on *Tian*, an academic reference exploring the effects of low resolution, to provide teachings of reducing a resolution and thereby modify Lu-577/Lu-000/Lert, an existing audience measurement system. This is needed because no contemporary image processing system would have taught or suggested intentionally reducing the resolution. Petitioner argues that disclosures of the '243 Patent relate to "selectively activating an illumination source and processing images taken without illumination to enhance the lighting while reducing the resolution and analyzing them for certain purposes." *Petition*, 3. The "certain purposes" Petitioner

glosses over include increasing the contrast to enable image analysis in unilluminated media exposure environments. '243 Patent, 10:58-11:4. Intentionally reducing the resolution of images prior to facial analysis is novel. This is why the *Tian* mapping fails. Tian suggests that conventional approaches rely on performing “facial analysis” on “very high resolution faces[.]” *Tian*’s investigation shows the highest accuracy for both face detection and head-pose estimation with full-resolution images. *Tian*, 000002; Table 3; *Supp Dec.* 32-33.

Indeed, Petitioner’s declarant agreed that a system that could perform image analysis on images upscaled to higher resolution with perfect accuracy and without incurring any time or computational resources would “improve [his] results” and “in general, [he] would try it. Of course.” Ex. 2018, *Doermann Deposition Transcript*, 71:1-9. Though Dr. Doermann couched such a system as “fiction” and “magic” (*id.* at 69:15-21, 70:8-10), such an approach is simple in the context of Petitioner’s proposed combination: simply do not intentionally reduce the resolution of images prior to performing image analysis. As such, a POSITA would understand that the best results would be obtained with the highest resolution images and would not seek to intentionally seek to reduce resolution.

Because intentionally reducing the resolution of images prior to facial analysis runs counter to conventional wisdom and *Tian*’s results, Petitioner can find no non-hindsight motivation to modify the system of *Lu-577/Lu-000/Lert* to do so. This

highlights the impermissible hindsight of Petitioner’s combination. Petitioner takes a functional system and adds in processing steps that are detrimental and unnecessary, which is improper under the law. *Cook Group Incorporated v. Boston Scientific Scimed, Inc.*, IPR2017-00440, Paper 7, 25-26 (using hindsight to employ a “redundant and unnecessary” modification does not provide a “sufficient rationale” for obviousness).

1. The Combination Does Not Achieve Petitioner’s Proposed Benefit, and Any Benefit Would Be Outweighed by Drawbacks Petitioner Fails to Address

For the combination of *Lu-577/Lu-000/Lert* with *Tian*, Petitioner alleges numerous benefits:

- Performing head detection and head orientation determination on reduced-resolution images requires fewer processing resources (claims 4, 11, and 18). *Petition*, 10.
- Determining whether an image is suitable for face recognition would improve the “performance” of face recognition (claims 4, 11, and 18). *Id.*
- Head tracking within a video sequence would save further processing resources by reducing the number of times face detection needs to be applied (claims 5, 6, 12, 14, and 19). *Id.* at 9.

As set forth below, each of these modifications involves complications and drawbacks that Petitioner fails to weigh properly against the purported benefits, and none of the alleged benefits themselves stand up to scrutiny.

a) Petitioner’s Proposed Modification to Process Reduced-Resolution Images Would Not Reduce the Net Processing Resources Required

Petitioner proposes performing head orientation determination on a reduced-resolution image in order to allegedly “reduce the time and resource consumption of tasks such as face detection and pose estimation, without substantially reducing the accuracy of these tasks.” *Petition*, 10 (citing Ex. 1003, ¶¶69-70). However, Petitioner’s modification *would not* save time and processing resources and *would* reduce the accuracy of face detection and pose estimation. Petitioner’s failure to consider the drawbacks of the proposed modification is fatal to their proposed combination.

(1) Petitioners Fail to Consider the Additional Processing Resources Required to Reduce the Resolution

When combining prior art references, “[e]vidence suggesting reasons to combine cannot be viewed in a vacuum apart from evidence suggesting reasons not to combine.” *Arctic Cat Inc. v. Bombardier Recreational Prods. Inc.*, 876 F.3d 1350, 1363 (Fed. Cir. 2017). It is Petitioner’s burden to provide “adequate evidentiary support for how a person of ordinary skill would weigh any advantages or disadvantages” for the proposed combination. *Oticon Medical AB et al. v. Cochlear*

Limited, IPR2019-00975, Paper 52, 36-41 (PTAB Oct. 14, 2020) (holding Petitioner’s failure to weigh the disadvantages of a proposed modification against the benefits showed that the modification was based on impermissible hindsight); *see also id.* at 38) (“[T]he burden of proof, which rests with Petitioner, [is] to prove that the disadvantages of [a modification] would not be significant.”).

Here, Petitioner claims a reduction in processing resources required to analyze reduced-resolution images but fails to weigh (or even consider) the *additional* processing resources that would be required to produce the reduced-resolution images in the first place. Neither Petitioner nor Petitioner’s declarant addresses whether the increase in resources needed to reduce the resolution of the images would outweigh any decrease in resources needed to process the lower-resolution images. *See Petition* at 10-11 (claiming the benefit of “reduce[d] the time and resource consumption” for reducing image resolution but failing to address any time and resource consumption associated with the resolution reduction process). Because Petitioner has not met their burden (or, indeed, made any effort) to show that the benefits of their proposed modification outweigh the associated disadvantages, Petitioner has not provided persuasive evidence to show predictable results. *Oticon Medical* at 39 (Petitioner’s failure to provide evidence that disadvantages would not be significant was a failure to show predictable results). As such, Petitioner’s obviousness theories are fatally flawed.

Here, though Petitioner has not addressed these disadvantages—much less met their burden of showing that they are not significant—the proposed modification includes disadvantages more than sufficient to outweigh any potential reduction in processing power. For example, down-sampling, as proposed by Petitioner, is known to introduce image artifacts unless filtering techniques such as antialiasing are also performed. *Supp. Dec.* 20-21 (citing Ex. 2028). These imaging artifacts can interfere with image analysis such as head orientation determination. *Id.* As Dr. Saber opines, reducing the resolution of an image without introducing artifacts that can interfere with subsequent processing is a multi-step process that involves processing each pixel of the image multiple times. *Supp. Dec.*, 37-39, 41-43. The additional processing resources required can outweigh any reduction in processing resources from performing head detection and head orientation determination on the reduced-resolution image. *Id.* at 39.

Petitioner also overstates the processing required for face detection and head orientation determination. In describing the processing resources required to perform face detection, Petitioner’s declarant states that “[s]ome approaches of face detection may require analyzing all pixels of an image to search for a face.” Ex. 1003, ¶ 69. However, this analysis ignores the fact that resolution reduction itself, even excluding the extra processing needed to avoid artifacts, may also “require analyzing all pixels of an image.” *Supp Dec.*, 40, 44. Indeed, a prior patent for which Dr.

Doermann is named as an inventor characterizes the same resolution reduction method proposed by Dr. Doermann as “slow” because it requires multiple floating-point calculations for each pixel. *Ex. 2016*, U.S. Patent Application Publication No. 2007/0257934A1 to Doermann, et al., [0061]; [0065]; *see also Ex. 2018, Doermann Deposition Transcript*, 40:23-41:10 (Dr. Doermann admitting that forms of interpolation used in image scaling typically require calculations for each pixel).

Furthermore, the resolution reduction process commonly includes an anti-aliasing filter, such as a low-pass filter, to avoid aliasing artifacts. *Supp. Dec.*, 41. In fact, Dr. Doermann’s own publication history supports the proposition that “[i]n digital video, a low pass filter is often applied to filter a high-resolution image before subsampling in order to eliminate aliasing.” *Ex. 2015, Text Extraction, Enhancement and OCR in Digital Video*, 7; *see also Ex. 2018, Doermann Deposition Transcript*, 59:14-60:4 (Dr. Doermann that *Ex. 2015* represents his contemporaneous opinion “in general”). Typical anti-aliasing algorithms rely on convolution, which requires processing each pixel of the original multiple times. *Supp. Dec.*, 41-43 (citing the computational complexity of convolution algorithms). This results in the Petitioner’s proposed combination requiring *more* processing resources for typical audience measurement scenarios than simply performing all steps on the full-resolution images. *Id.* Accordingly, a POSITA would not be motivated to modify the system

of Lu-577/Lu-000/Lert to perform facial analysis on reduced-resolution images in order to save processing resources.

(2) Petitioner's Modifications Are Redundant

Elsewhere, Petitioner proposes further modifying *Lu-577* to incorporate *Tian's* teachings of “determining a face/head location in a first image and using this location for subsequent images[.]” *Petition*, 9. However, Tian’s face/head tracking process updates the location of facial features in each frame based on the previous frame, so the head location and the head orientation are known. *Ex. 1005, Tian*, 000001; *see also Ex. 2018, Doermann Deposition Transcript*, 39:15-21 (Dr. Doermann agreeing that Tian updated the facial feature locations for each successive frame). Given that the head location and the head orientation are known from face tracking, reducing the resolution has computational costs (from reducing the resolution) without any benefit (because head detection and head orientation determination would be unnecessary), thus negating any benefit from reducing the resolution. *Cf. Ex. 1003*, ¶ 69 (“face detection may require analyzing all pixels of an image to search for a face”) *with id.* at ¶ 64 (face tracking avoids the need to run a face detector across the entire image). Thus, by rendering the purported advantages of Petitioner’s resolution-reduction modification redundant, this head-tracking modification further shifts the net effects of Petitioner’s proposed combination towards the “disadvantages” side of the scale. While a POSITA is not required to

choose the best option from all of those available in order to render a claim obvious, no motivation to modify a reference exists when the motivation *not* to combine outweighs any motivation to combine. *Cf. Intel Corp. v. Qualcomm Inc.*, 21 F.4th 784, 800 (Fed. Cir. Dec. 7, 2017) (“It’s not necessary to show that a combination is the best option, only that it be a suitable option.”) (internal quotes omitted) *with Arctic Cat* 876 F.3d at 1363 (“Evidence suggesting reasons to combine cannot be viewed in a vacuum apart from evidence suggesting reasons not to combine.”). Here, the advantages, slight to begin with, are made redundant by the other modifications in Petitioner’s proposed combination, while the disadvantages are significant. *Supp. Dec.*, 45-48.

(3) Petitioner’s Analysis of Accuracy is Factually Flawed

The only potential drawback of Petitioner’s proposed modification addressed by the Petition is the effect on face detection and pose estimation accuracy of performing these tasks on reduced-resolution images. *Petition*, 10. Petitioner contends that reducing the resolution prior to performing the step of head detection and head orientation determination can be done “without substantially reducing the accuracy of these tasks.” *Id.* In support of this claim, Petitioner cites to *Tian’s* experimental findings and, in passing, Exs. 1013-1015. *Id.* However, none of this material would lead a POSITA to the conclusions that Petitioner wishes to draw. *Supp. Dec.*, 49-59.

First, none of Exs. 1013-1015 relate to reducing the resolution of an image, and none support the proposition that face detection and head orientation determination can be performed on a reduced-resolution image without loss of accuracy:

- Ex. 1013 describes the processing requirements of a particular view-based face detection algorithm that the reference characterizes as “not the appropriate choice for real-time applications[,]” such as that described by the ’243 Patent. *Ex. 1003*, 000001; *see also Supp Dec.*, 50.
- Ex. 1014, as cited by Petitioner, characterizes the processing requirements of an “integral image[,]” a heavily processed image where “[e]ach element of the integral image is calculated as the sum of intensities of all points above and to the left of the point in the image.” As such, Ex. 1014 is also irrelevant to the processing requirements of the unprocessed images of the ’243 Patent. *Ex 1014*, 2:43-53; 1:27-29; *see also Supp Dec.* 51.
- Ex. 1015 describes reducing the resolution of the *saturation* of an image (*i.e.*, making the colors more uniform so as to be more easily able to pick out the uniform colors of a face) without changing the image resolution itself. Ex. 1015, ¶ [0022] (“By **reducing the resolution of the saturation** of the image, computational requirements are

substantially reduced and a relatively simple method may be used.”)

(**emphasis** added); *see also Supp. Dec. 52* (Dr. Saber opining that Ex.

1015 describes reducing the saturation resolution of individual pixels).

In short, Petitioner’s only evidence allegedly supporting the claim that performing facial analysis on reduced-resolution images can be done “without substantially reducing the accuracy of these tasks” is *Tian*.

However, Petitioner also fails to consider the applicability of *Tian*’s findings in the context of the claimed audience measurement system. As discussed below, *Tian*’s field of endeavor is not audience measurement. § IV, *infra*. This is important because *Tian* begins with full-face images, where the face region takes up a significant portion of the original image. *Supp. Dec.*, 53-54 (Dr. Saber computing that, based on the full resolution for the original image and the full resolution of the face region, the face makes up approximately 36% of the pixels of original image).

By contrast, the camera in an audience measurement system captures an image of the entire media exposure environment. *See, e.g.*, ’243 Patent, claim 9 (reciting a system operable to “collect a first image of the media exposure environment with a camera”); *see also id.* at FIG. 1 (depicting the field of view for a typical media exposure environment). As Dr. Saber points out, an image captured of such a media exposure environment will begin with the face of an audience member taking up significantly less of the image. *Dec. 55-59* (Dr. Saber calculating the size of the head

region for a viewer in a media exposure environment takes up less than 1% of the original image and is close to Tian's minimum viable size, even before resolution reduction). Because the starting resolution of the head portion of the image is significantly smaller than the starting resolution in Tian's experiments, *Tian's* results cannot be blindly applied to conclude that reducing the resolution does not affect the ability to detect heads or determine head orientation.⁴ *Supp. Dec.*, 58-59. Yet Petitioner and Petitioner's declarant ignore these issues completely. As such, Petitioner's limited analysis of this proposed modification is incomplete and cannot be said to weigh the alleged benefit against even this drawback accurately.

b) Petitioner's Proposed Modification to Perform Head Orientation Determination Is Flawed

Petitioner motivates the modification of *Lu-577/Lu-000/Lert* to perform head orientation determination by suggesting that doing so would "determine whether an image is suitable for face recognition" (thereby "improv[ing] the performance of

⁴ Patent Owner emphasizes that this analysis does not rely on reading any particular resolution requirements or image framing requirement from the Specification of the '243 Patent into the claims; rather, it demonstrates that the realities of the field of media exposure analysis would not lead a POSITA to apply Tian's findings in this field of endeavor as Petitioner contends; *Tian's* results are simply inapplicable in the field of audience measurement.

face recognition”) and enable the system to determine “whether a person is watching the television[.]” However, Petitioner misapprehends what “performance of face recognition” means in the context of audience measurement.

(1) Petitioner’s Modification Would Reduce Face Recognition Performance By Reducing Face Recognition Opportunities

Petitioner’s proposed combination relies on modifying *Lu-577/Lu-000/Lert* (and, specifically, modifying the face detector of *Lu-000*) with *Tian’s* teachings of head orientation determination to identify candidate images before performing facial recognition in order to increase the chances of a successful face recognition of a given image. *Petition*, 10. Petitioner thus identifies, as the benefit of the modification, improved chances of facial recognition. However, Petitioner mistakenly focuses on the chances of success for facial recognition of a particular image over the chance of success for recognizing a member of the viewing audience over all images. For example, performing facial recognition on 10 consecutive frames, each with a 25% chance of success, is significantly more likely to identify an audience member than performing face recognition on a single frame with a 50% chance of success. *Supp. Dec.*, 61. Indeed, Petitioner’s declarant agreed that “in general, analyzing more images would typically improve the results.” *Ex. 2018, Doermann Deposition Transcript*, 78:15-16. Because of this, Petitioner’s proposed modification would not improve the performance of facial recognition but would make it worse, as set forth below.

A POSITA would understand that, in the context of audience measurement, many images of a potential audience member will be captured. *Supp. Dec.* 62-63. In such a context, the chance of recognizing an audience member (and analyzing their media exposure) is maximized by performing face recognition over **all** of the candidate images of the media exposure environment where an audience member is present. *Id.* Indeed, this is the approach taken by *Lu-000*: reference images are captured for a variety of head poses so that audience members can be recognized regardless of the head orientation of an audience image in a particular frame. *Lu-000*, 2:32-36 (“The extracted pattern image signature is compared with **each of the stored pattern image signatures** to identify a particular one of the predetermined audience members.”). Such continuous measurement is particularly desirable in the audience measurement context, as it allows more granular measurement of the precise times an audience member is in the media exposure environment. *Supp. Dec.*, 64-65. Thus, Petitioner’s modification would be detrimental in an audience measurement system.

(2) Petitioner’s Modification Would Be Redundant As Applied to *Lu-000*’s Face Recognition Approach

Because *Lu-000* can recognize images without requiring a particular orientation, the *Lu-577/Lu-000/Lert* system would not benefit from the proposed modification of filtering images to those with a particular orientation. *Dec.* ¶ 60. As a result, there would be no motivation for this modification. *10X Genomics, Inc. v.*

Bio-Rad Laboratories, Inc., IPR2018-00489, Paper 13, 15-16, 20-21 (PTAB June 28, 2018) (denying institution where the Petitioner did not demonstrate a reasonable likelihood of success because the primary reference “already” achieves the benefit asserted by Petitioner to motivate the combination). The Board should find Petitioner’s combination “not sufficient to render” the claim limitation obvious because Petitioner does not “persuasively explain why” a POSITA would have sought to modify the system of *Lu-577/Lu-000/Lert* when *Lu-000* already provides the purported improvement. *RJ Reynolds Vapor Co. v. Fontem Holdings I B.V.*, IPR2016-01692, Paper 45, 24-26 (PTAB March 2018).

Lu-000’s brute force method has no need to determine the head orientation and filter for suitable images because the system already accounts for various head poses when recognizing audience members. *Lu-000*, 8:36-45 (“An individual audience member face image may be learned multiple times (R) with R possible different extracted signatures resulting, depending on any changed facial expressions or various profiles of the audience member.”), 7:31-35 (“Multiple face images are learned for each of the audience members P by sequentially processing a series of video images of the video camera 28 by the image signal processing blocks of Fig. 6...”). *Lu-000*’s system extracts feature signatures from digital representations generated from video images and compares these signatures to **all stored signatures** to **determine a match**. *Lu-000*, 4:64-5:2, 7:41-55. Multiple signatures for each

person are stored to account for “various profiles of the audience member.” *Lu-000*, 8:36-45.

Because *Lu-000* stores various profiles of the audience members and performs facial recognition against the stored profiles, *Lu-000* already performs facial recognition for a variety of head orientations. *Dec.*, ¶ 58. As Dr. Saber explains, because the *Lu-000* system compares each captured signature to **all stored signatures**, employing head orientation for the purpose of “improved accuracy of facial recognition” is unnecessary and redundant. *Dec.*, ¶ 58. And, because various poses are stored, a variety of poses in the feature-extracted image creates a match with one of the stored signatures. *Dec.* ¶¶ 59-60. Thus, in the modified *Lu-577*, the digital representations generated from the video image already serve as “suitable” images for face recognition **without needing to determine a head pose**. *Id.*

In fact, a POSITA would be dissuaded from making such a modification, because doing so would add unnecessary additional steps to *Lu-000*’s recognition system. *Dec.* ¶¶ 61-65; *In re Schweickert*, 676 F. App’x 988, 995 (Fed. Cir. 2017) (reversing Board’s obviousness determination where proposed modification is “unnecessary” and would add unwanted “complexity to the system”). The additional unnecessary steps would be a wasted processing step, thereby incurring unnecessary processing time. Dr. Saber opines that *Lu-000*’s learning mode and recognition mode would require additional processing to filter images to identify frontal poses for

recognition. *Dec.* ¶¶ 61-65. Including *Tian*'s head pose classification would add multiple steps to *Lu-000*'s learning mode and recognition mode of the system to determine the orientation and detect whether the head orientation is frontal. *Id.*

(3) Petitioner's Modification Would Not Allow Determining When a Person Is Watching Television

Petitioner provides a secondary motivation to modify *Lu-577/Lu-000/Lert* with the head orientation determination teachings of *Tian*: "to determine whether a person is watching the television." *Petition*, 10. In support of this alleged motivation, Petitioners turn to Dr. Doermann, who cites *Weinblatt* (Ex. 1012) in support of Petitioner's proposed motivation. Ex. 1003, ¶ 67. In particular, Dr. Doermann cites portions of *Weinblatt* discussing gaze tracking (*i.e.*, what the subject is "looking at") to determine if the subject is watching the TV program:

For example, the facial recognition unit 11 may use the **measurements of the angle of a face, if the eyes of the face are open or closed, and/or if the face of the person is 'looking at' the camera** or at the broadcast program to determine if the person is watching the broadcast program.

Ex. 1003, ¶ 67 (quoting *Weinblatt*, [0039]). Here, it is not even necessary to go beyond the quoted material to realize that Petitioner's proposed modification would **not** result in a system capable of determining television-watching activity for an audience member, as Petitioner contends. This is the case because *Tian*'s system does not determine any of the criteria *Weinblatt* uses to determine if the person is watching the broadcast program.

Tian performs head pose estimation to determine whether a subject's head pose is suitable for facial expression analysis. *Dec.* ¶ 91. *Weinblatt*, on its face, describes three options for gaze tracking: (1) measurements of the angle of a face; (2) if the eyes are open or closed; and (3) whether the face of the person is “looking at” the camera or television. Given a head pose as classified by *Tian*, it would not be possible to determine **any** of this information. *Dec.* ¶¶ 89-94.

In particular, *Tian* does not measure a face angle or determine whether the eyes are closed; as discussed above, *Tian* merely classifies whether a face image is suitable for expression analysis. *Tian*'s classifications do not break down neatly into “looking at the camera” or “not looking at the camera” as would be required for *Weinblatt*'s gaze tracking. *Dec.* ¶¶ 91-94. As such, the claimed benefit would not actually be achievable in Petitioner's proposed combination. *Dec.* ¶ 94.

Furthermore, and contrary to Dr. Doermann's unsupported conclusions, it would not be beneficial to only recognize audience members who are actively looking at the television. *Supp. Dec.*, 68-69. For example, *Lu-577* gives, in motivating the invention, “changes in methods of measuring the reception of television or **radio programming**.” *Lu-577*, ¶ [0007]. A POSITA would recognize that an audience member does not stop being exposed to media simply because they are not looking at the television. *Supp. Dec.*, 69. Indeed, by Dr. Doermann's logic, advertisers would be equally served by totally silent television ads, since any

audience member not actively looking at the television is not exposed to their advertisement. *Id.* However, in practice, a POSITA would appreciate that it would be desirable to include in audience measurements all people in the room who can be recognized, even if they look away from the television. *Id.*

c) Petitioner's Proposed Modification to Rely on Head Tracking Would Not Reduce the Net Reduction of Processing Resources Required and Would Decrease the Performance of Face Recognition

Petitioner motivates modifying *Lu-577/Lu-000/Lert* to include head tracking within a video sequence by stating that doing so would save further processing resources by reducing the number of times face detection needs to be applied. *Petition*, 9.; *see also* Ex. 1003, ¶ 63-64. However, much like their other proposed motivations, Petitioner fails to consider the drawbacks of the additional processing resources required for their proposed modification is fatal to their proposed combination.

Petitioner's proposed combination fails because Petitioner completely fails to consider the drawbacks of the modification and weigh them against the claimed benefit. And these drawbacks are significant:

- First, Petitioner claims the benefit of reduced processing resources resulting from performing head detection fewer times because the head is tracked between video frames, without accounting for the **increase** in processing costs associated with performing the head tracking.

- Second, Petitioner fails to account for the need to perform head detection regardless, in order to identify new audience members entering the frame.
- Third, Petitioner entirely fails to weigh the reduction in facial recognition performance associated with using the head location from the first image to perform face recognition in the second image.

Because Petitioner has not met their burden to show that the benefits of their proposed modification outweigh the associated disadvantages, Petitioner has not provided persuasive evidence to show predictable results. *Oticon Medical* at 39.

First, Petitioner does not even attempt a cursory accounting for the **increase** in processing resources needed to perform the head tracking against a potential decrease resulting from performing head detection fewer times. In fact, as described below, the sole reference cited by Petitioner's declarant in support of this modification immediately goes on to emphasize the need to continue to perform face detection even after a face is detected. Ex. 1010, ¶ [0003]; *see also* Ex. 2018, *Doermann Deposition Transcript*, 75:19-24 (Dr. Doermann agreeing that, based on Ex. 1010, it would still be necessary and desirable to perform face detection over the other portions of the image). Nonetheless, assuming *arguendo* that head detection could be replaced by head or face tracking, thereby saving the resources associated with head detection, a POSITA would appreciate that performing face tracking itself

requires processing resources. *Supp. Dec.*, 70-71; *see also Ex. 2018, Doermann Deposition Transcript*, 35:5-11 (Dr. Doermann agreeing that head tracking in a frame would require processing the frame based on a previous frame in order to update the location of the head in the current frame); 39:15-21 (Dr. Doermann agreeing that Tian's face tracking process updates the facial locations for each successive frame based on the previous frame). Petitioner has provided no evidence of how large or small the processing resources required by face tracking would be or whether they would be greater or larger than the processing resources required by performing head detection. As such, Petitioner has failed to meet "the burden of proof, which rests with Petitioner, *i.e.*, to prove that the disadvantages [of the proposed modification] would not be significant." *Oticon Medical*, 38.

Second, Petitioner ignores the teaching from their own background reference that teaches that it is both desirable and necessary to perform head detection across each image even when head tracking is used:

Once a face is detected, its location is recorded and a localized region around that face is scanned by a face detector in the next frame. Thus, once a face is initially detected, it can be accurately tracked from frame to frame without a need to run a face detector across the entire image. The "located" face is said to be "locked" by the face tracker. **Note that it is still desirable to scan the entire image or at least selected portions of the image with a face detector as a background task in order to locate new faces entering the field of view of the camera.**

Ex. 1010, ¶ [0003] (**emphasis added**). In other words, a POSITA would understand that in Petitioner's proposed combination, **no savings in processing resources could be realized** because performing head detection in each new frame would still be necessary. *Supp. Dec.*, 72. Once again, Petitioner's claimed benefits fail to stand up to scrutiny. Though Petitioner has failed to meet its burden of showing evidence of how a POSITA would have weighed the benefits of this modification against the drawbacks, the lack of benefit (owing to the need to perform head detection regardless) would suggest that a POSITA would find that drawbacks (in terms of increasing the processing costs of performing head tracking) outweighed any benefit, and the POSITA would not be motivated to modify *Lu-577-Lu-000/Lert* as Petitioners propose. *Id.*

Third, a POSITA would understand that performing facial recognition in a region of an image where the head **was** is necessarily going to be less accurate than performing facial recognition in a region of an image where the head **is now**. *Supp. Dec.*, 73-74. As Dr. Saber opines, facial analysis accuracy is improved by starting with an up-to-date identification of the head region and is decreased by an out-of-date identification of the head region. *Supp. Dec.*, 73 (citing *Tian's* need for "well-aligned" heads for facial analysis). Once again, Petitioner fails to weigh this drawback (decreased accuracy) against the claimed benefit (decreased processing resources from fewer head detections). Indeed, in proposing other modifications,

Petitioner makes precisely the opposite tradeoff. *Petition*, 9-10 (performing head detection and head orientation determination on reduced-resolution images but performing face recognition on full-quality images because “it was known that the performance of face recognition may be adversely affected by the reduction of image resolution.”).

In justifying Petitioner’s modification to rely on out-of-date head location information, Dr. Doermann discusses how this approach is “applicable when a sequence of images is taken for a person by a video camera” because “a person’s head or face is unlikely to move significantly” and “the region corresponding to the person’s head or face likely remains the same across the sequence of images.” Ex. 1003, ¶ 64. However, this analysis ignores the fact that the background references he relies on all rely on head tracking so that the head “can be accurately tracked from frame to frame.” *Id.*

For example, *Tian*, which Petitioner relies on for their mapping, expressly states that face tracking should be used to update the face location for a video sequence and provides the possibility that “head finding and head tracking” may **also** be needed “to handle large head motion.” *Tian*, 000001; *see also Supp Dec.*, 75 (Dr. Saber opining that Tian’s data set represents the best case for reusing a head position and that a real-world scenario would require updating the head position before performing face detection.). *Tian* further states that even when the “face image is

continually in the frontal or near frontal view in an image sequence,” the location of facial features must be updated prior to further processing by “searching for the dark pixels around their positions in the last frame.” *Tian*, 000003; *Supp. Dec.*, 76. Ex. 1010, also relied upon by Dr. Doermann, suggests that a previously located face must be relocated in each frame by a process in which “a localized region around that face is scanned by a face detector in the next frame.” Ex. 1010, ¶ [0003]. 1010, ¶ [0003]; *Supp. Dec.*, 74. In other words, Petitioner cites **no art** suggesting that acceptable (much less un-degraded) facial recognition performance can be achieved by using un-updated face locations from the previous frame.

Once more, Petitioner’s failure to weigh these disadvantages of the proposed modification against the claims benefit is fatal to their proposed combination. *Oticon Medical*, 36-41. Indeed, it is doubtful Petitioner could ever meet this burden, given their contradictory rationales for the different modifications to *Lu-577-Lu-000/Lert*. Cf. *Petition 10-11* (arguing that a POSITA would spend computing resources to perform facial recognition on full-resolution images in order to avoid adversely affecting facial recognition performance) with *Petition 9* (arguing that a POSITA would choose not to update the face location before performing facial recognition in order to save processing resources). As such, claims 5-6, 8, 12-14, and 19-20 should be found not unpatentable for at least this reason.

2. *The Petition Fails to Provide Any Motivation to Combine the Teachings of Lu-000 and Lert with Lu-577*

Although Ground 1 is *Lu-577* combined with *Tian*, Petitioner heavily relies on *Lu-577*'s "incorporation" of *Lu-000* and *Lert* for teaching several limitations of the challenged claims.⁵ *Petition*, 12-15, 18-21, 23-25. *Lu-000*'s and *Lert*'s disclosures provide, at most, additional embodiments that could be incorporated into *Lu-577*'s audience measurement system. Patent Owner submits that combining features from a reference that is incorporated by reference is akin to combining two embodiments of a single reference. Both Federal Circuit case law and this Board's case law require a motivation to combine two (or more) embodiments of a reference. *Abiomed, Inc., v. Maquet Cardiovascular, LLC*, IPR2017-01026, Paper 8, 7-8, 10-11 (PTAB Sept. 20, 2017); *In re Stepan*, 868 F.3d 1342, 1346 n.1 (Fed. Cir. 2017) (Board stating an obviousness argument based on "combining multiple embodiments from a single reference" must include a motivation and a reasonable expectation of success).

By analogy, Petitioner must show why a POSITA would have combined the features from *Lu-000* and *Lert* with *Lu-577*, especially considering Petitioner is

⁵ Patent Owner cites the limitations in claim 1 mapping *Lu-000* and *Lert*. Petitioner references the same arguments in the other challenged independent claims. Thus, Patent Owner's arguments apply for all independent claims.

further proposing modifying *Lu-000*. *Abiomed, Inc.*, Paper 8 at 10-11. (Board finding Petitioner “failed to establish a reasonable likelihood of prevailing” because of Petitioner’s “failure to explain sufficiently and support” a combination of two alternative embodiments in a single reference).

Petitioner’s mapping applies the teachings of *Lu-000* and *Lert* throughout Ground 1. *Petition*, 3. For example, Petitioner asserts *Lu-000*’s capturing of images and detection of heads within the images teaches the claimed step of detecting a head recited in the independent claims. *Petition*, 18-21. Petitioner also maps *Lu-000*’s teaching of comparing stored signatures with extracted feature signatures to the “audience identification information” limitations recited in the independent claims. *Petition*, 23-25. Petitioner applies *Lert*’s audio processor circuitry and *Lu-000*’s various disclosed processing circuitries for teaching the “processor circuitry to execute the machines readable instructions” of claim 1. *Petition*, 15-16. Thus, teachings from all three references must be used for Petitioner’s Ground 1 to render the challenged claims “obvious,” even though Petitioner omits both *Lu-000* and *Lert* from Ground 1. *Petition*, 3. Through this omission, Petitioner attempts to improperly benefit from the teachings of three references, without performing the required obviousness analysis. Because Petitioner did not provide **any** motivation to combine the features of *Lu-000* and *Lert* with *Lu-577*, Petitioner has not met its burden.

In the Institution Decision, the Board characterizes the argument that “the Petition is deficient because it relies on teachings of Lu-000 and Lert that are incorporated by reference into Lu-577 without providing any additional motivation to combine these teachings” as “legally and factually deficient.” *Institution Decision*, 31-32. Patent Owner acknowledges that this appears to be a question of first impression. However, Patent Owner respectfully submits that, by analogy with the cases cited above, even where the material is incorporated by reference and “is effectively part of the host document as it were explicitly contained therein,” a motivation to combine should be required when the incorporated material is (as here) provided as background material and is not a part of what the inventors viewed as their invention.

Here, Petitioner is **modifying** the invention of Lu-577 with the two secondary references and relying on the teachings of these secondary references in their combination. For example, *Lert* (as characterized by *Lu-577*) is simply “[a]nother approach” among the “many approaches to measuring the usage of electronic entertainment.” *Lu-577*, ¶ [0003]; [0002]. Indeed *Lu-577* **disparages these approaches**, stating that “[c]hanges in the methods of measuring the reception of television or radio programming are required because of a planned change-over from analog to digital broadcasting.” *Lu-577* nowhere suggests that using the specific (outdated) teachings of *Lert* would be desirable. By way of example, a patent could

incorporate a prior art reference addressing the problem solved by the invention and disparaging the prior approach. Indeed, by characterizing the approach of *Lert* as one in which “changes ... are required,” *Lu-577* actually does disparage the approach on which Petitioner relies. If all incorporated material is considered part of an incorporating patent requiring without an obviousness analysis (including a motivation to combine), then a Petitioner could rely on the prior art material together with the invention despite a clear teaching away. Such a result flies in the face of an enormous body of obviousness case law. *See, e.g., Chemours Company FC, LLC v. Daikin Industries, Ltd.*, No. 20-1289, Slip Op. (Fed. Cir. Jul. 22, 2021) (reversing finding of obviousness where the PTAB ignored express disclosure in [a prior art reference] that teaches away from the claimed invention and relied on teachings from other references that were not concerned with the particular problems [the reference] sought to solve).

III. GROUND 3: NIELSEN-372 IN VIEW OF STEINBERG AND IN FURTHER VIEW OF TIAN DOES NOT RENDER OBVIOUS THE CHALLENGED CLAIMS

A. Overview of the Cited References and Petitioner’s Proposed Combination

1. Petitioner’s Proposed Modification of Nielsen-372 to incorporate the teachings of Steinberg

Petitioner’s modification combines *Steinberg*’s “face detection and recognition” into *Nielsen-372*’s metering system. *Petition*, 44. Specifically, Petitioner asserts “Nielsen-372’s system **includes needed hardware (e.g., a**

computer **and a camera**) for implementing **Steinberg’s approach.**” *Id.*; *see also* Ex. 1003, ¶ 200. In particular, Petitioner cites *Nielsen-372*’s ON/OFF device video sensor to teach the camera. *Id.* (citing *Nielsen-372*, [0050]). Dr. Doermann similarly points to the “video sensor” as providing the “**hardware needed to implement Steinberg’s approach.**” Ex. 1003, ¶¶ 200; 199. Thus, Petitioner’s combination modifies *Nielsen-372*’s ON/OFF device to implement *Steinberg*’s facial recognition for identifying “audience members.” *Petition*, 44-45 (asserting “*Nielsen-372* provide[s] the hardware architecture for performing Steinberg’s method”); Ex. 1003, ¶¶ 199-200, 203; *Dec.* ¶ 138.

Patent Owner is mindful of this Board’s precedent of avoiding bodily incorporation of features from a secondary reference into the modification. *In Re Keller*, 642 F.2d 413, 425 (CCPA 1981); *SK Hynix Inc. et al., v. Netlist, Inc.*, IPR2020-01042, Paper 14, 35 (PTAB December 2020). However, here Petitioner and their declarant have made their proposed combination clear: the video sensor of *Nielsen-372* is used as the “hardware architecture” for capturing and processing images using Steinberg’s approach of face detection and recognition. *See, e.g., Petition*, 44; Ex. 1003, ¶ 203 (“*Nielsen-372* provide[s] the hardware architecture for performing Steinberg’s face recognition method.”). As such, Patent Owner is not resorting to bodily incorporation in its arguments regarding a lack of motivation to

combine *Nielsen-372* and *Steinberg*, but rather pointing out issues inherently present in Petitioner's expressly proposed combination.

2. *The Institution Decision Misapprehends the Combination Proposed in the Petition*

In the Institution Decision, the Board determines that Petitioner's proposed combination relies upon the camera of *Steinberg*, rather than that of *Nielsen-372* and, on that basis, rejects Patent Owner's arguments that a POSITA would not be motivated to modify *Nielsen-372* with the teachings of *Steinberg*. *Institution Decision*, 41-43. Patent Owner respectfully submits that this determination represents a combination not advanced by Petitioner. In particular, the Institution Decision cites Petitioner's mapping for claim 1 as representative. *Id.* (citing Petition, 53 "(Steinberg teaches obtaining a sequence of images by a camera (*e.g.*, digital camera, image capture device). Ex-1007, ¶0010.")). On the basis that Petitioner cites *Steinberg* for the claim limitation "obtaining a sequence of images by a camera," the Board concludes that Petitioner's combination relies on the camera of *Steinberg* rather than *Nielsen-372*.

However, Patent Owner submits that this language is, at most, a mapping of the "obtaining" step of *Steinberg's* "process" which (per Petitioner's mapping) is implemented by the "hardware architecture" of *Nielsen-372*. *See, e.g., Petition*, 44; Ex. 1003, ¶ 203 ("Nielsen-372 provide[s] the hardware architecture for performing *Steinberg's* face recognition method."). Consistently with this language, the

combination of the Petition would “obtain[] a sequence of images” (as taught by *Steinberg*) “by a camera” (provided by the hardware architecture of *Nielsen-372*).

As further evidence of this, compare the language used in Petitioner’s mapping of claim element 1d for Ground 1:

“Lu-577’s person identifier identifies “persons watching television programing” based on image recognition. Ex-1004, ¶0039. [...] The images are captured by a camera. Ex-1004, ¶0039 (‘The person identifier 98 may be video camera, an IR camera, or the like.’). [...] Lu-000 teaches obtaining (e.g., capturing) images for audience members in the monitored area using a video camera. Ex-1008, Abstract. [...] To the extent it is argued that Lu-577 does not teach “a sequence of images,” Tian teaches obtaining a sequence of images of human subjects using a camera. Ex-1005, 000005.

Petition at 18-19 (**emphasis** added). In other words, Petitioner uses virtually identical language to that relied upon by the Board for **three separate** references (*Lu-577*, *Lu-000*, and *Tian*) in Ground 1. Following the approach of the Institution Decision, Petitioner’s combination for Ground 1 would include three separate cameras, which would represent significant issues of redundancy for the combination. *Cook Group Incorporated*, Paper 7, 25-26 (a “redundant and unnecessary” modification does not provide a “sufficient rationale” for obviousness). Indeed, in Ground 3, *Petition* again includes virtually identical language for *Tian* as for *Steinberg*: **“Tian further teaches obtaining a sequence of**

images of human subjects using a camera and analyzing the sequence of images to detect a head appearing in one or more of the images” *Petition*, 74. Since this language is used for multiple references in each proposed combination, this language cannot be relied upon to determine which reference is being relied upon to teach a camera.

Rather, Petitioner tells us clearly and unambiguously exactly what Nielsen-372 is used for and what Steinberg is used for when the Petition sets out the proposed combination: Nielsen provides the hardware for implementing Steinberg’s method/approach. *Petition*, 44-45 (“Nielsen-372’s system includes needed hardware (e.g., a computer and a camera) for implementing Steinberg’s approach. [...] A POSITA would have had a reasonable expectation of success given that Nielsen-372 provide the hardware architecture for performing Steinberg’s method and that Steinberg’s techniques were admittedly known.”).

“[I]n an inter partes review the petitioner is master of its complaint[...].” *SAS Inst. Inc. v. Iancu*, 138 S. Ct. 1348, 1355 (2018). “[T]he petitioner’s contentions [...] define the scope of the litigation all the way from institution through to conclusion.” *Id.* at 1357. Here, the Petition is clear: the camera in the proposed combination is that of *Nielsen-372*, not that of *Steinberg*. As set forth below, Petitioner has proposed a modification to *Nielsen-372* that renders it inoperable for its stated purpose and changes its principle of operation. Petitioner’s proposed

combination is fatally flawed, and the Board should not step into Petitioner's shoes to fix it.

B. Petitioner's Proposed Combination Fails to Teach the Claim Limitations

1. *Tian Does Not Teach Instructions, a Device, or a Step for "Reduc[ing] a Resolution"*

Like the mapping for Ground 1, Petitioner's mappings for Ground 3 rely on *Tian* to teach the claim limitation of "reducing the resolution of an image to obtain a reduced-resolution image." *Petition*, 75 (Ground 3 mapping for limitation [4a] citing the mapping in Ground 1.) As set forth above in § II.A.1, *Tian* does not teach or suggest instructions, a device, or a step for "reducing a resolution," and Petitioner's mappings for claims 4, 11, and 18 in Ground 3 fail for the same reason as that in Ground 1.

2. *Petitioner's References, Alone or in Combination, Do Not Teach Performing Head Location Determination in a First Image and Face Recognition on the Determined Location in a Second Image*

Like the mapping for Ground 1, Petitioner's mappings for Ground 3 rely on *Tian* to teach the claim limitations relating to performing head location determination in a first image and face recognition on the determined location in a second image. *Petition*, 77-79 (Ground 3 mappings for these limitations in claim elements [5] and [6b] citing the mapping in Ground 1). As set forth above in § II.A.2, *Tian* does not teach or suggest performing head location determination in a first

image and face recognition on the determined location in a second image, and Petitioner's mappings for claims 5, 6, 12, 14, and 19 in Ground 3 fails for the same reason as that in Ground 1.

C. The Petition Fails to Establish a Sufficient Motivation to Combine the References

Petitioner's Ground 3 (incorporating Petitioner's now-moot Ground 2) relies on the combination of *Nielsen-372*, *Steinberg*, and *Tian*⁶. *Petition*, 71-73. Petitioner's combination is fundamentally flawed. A POSITA would not have found it obvious and been motivated to combine *Nielsen-372* and *Steinberg* as proposed by Petitioner because modifying *Nielsen-372* using the teachings of *Steinberg* as Petitioner proposes would frustrate the purpose of *Nielsen-372* and change its principle of operation. Further, even assuming combining *Nielsen-372* and *Steinberg* as Petitioner proposes, would give a functional audience measurement system, the same problems described above in Ground 1 would arise in modifying that system using the teachings of *Tian*.

⁶ Petitioner provides substantially similar mappings for claims 4, 11, and 18. As such, Patent Owner's arguments for claim 4 also apply to claims 11 and 18. *Compare Petition*, 45-64 with *Petition*, 64-71.

1. A POSITA Would Not Have Found It Obvious and Been Motivated to Modify Nielsen-372 with the Teachings of Steinberg

As discussed below, Petitioner's combination relies upon the camera of *Nielsen-372* to teach the claim limitations of the '243 Patent related to capturing images of the "media exposure environment" to collect the audience identification information. As such, the camera must be oriented toward the audience. *Dec.*, ¶ 139. However, as designed, *Nielsen-372*'s camera is oriented toward the **television** in order to determine **display activity for the display device**. *See, e.g., Ex. 1006, Nielsen-372*, [0081]. As with other sensors, Nielsen-372's camera is oriented to collect image from the television:

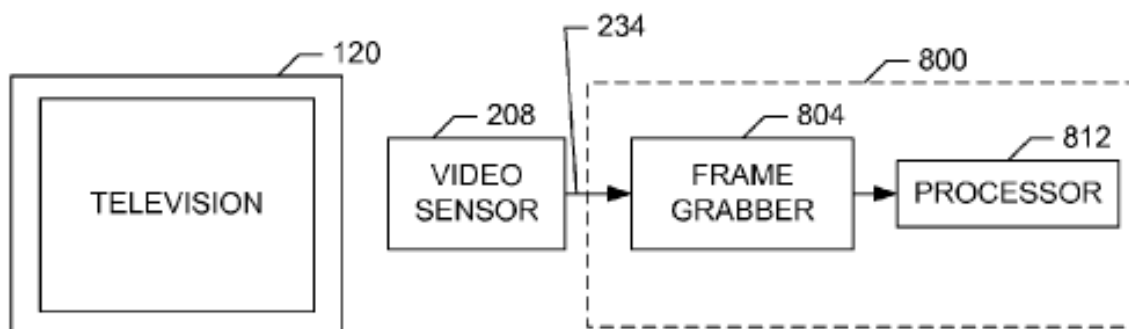


FIG. 8

Nielsen-372, FIG. 8 (depicting video camera oriented towards television).
Petitioner's proposed modification of *Nielsen-372*'s camera to perform facial

recognition of people watching TV would frustrate the purpose of *Nielsen-372*'s metering system by rendering it unable to determine media content information and would change the principle of operation.

a) Petitioner's Proposed Modification Frustrates the Purpose of Nielsen-372's Metering System

Notably, the video sensor of *Nielsen-372* is "positioned to detect the display area of the display device." *Nielsen*, [0050]. Per Dr. Saber, a camera positioned to detect the display area of the display device, *e.g.*, a TV, is not positioned to detect the audience viewing the TV. *Dec.*, ¶ 140. Petitioner's proposed combination would, at best, result in the bizarre situation where the "audience measurement system" identifies the actors present in the media content rather than the audience.

As discussed above, it is the function (and the stated purpose) of *Nielsen-372*'s ON/OFF device generally and the video sensor specifically to determine whether the display showing the viewed content is powered on. Thus, the camera faces the display device, **not the audience watching the displayed content**. *Nielsen-372* [0050] ("The video sensor 208 may be, for example, a camera, a single output analog or digital light sensor, etc., **positioned to detect the display area of the display device**"); *Dec.* ¶¶ 141-142. Furthermore, the video images taken to determine the ON/OFF state of the display device **would not include the audience members**. *Id.*

The *Nielsen-372* camera does not take images of the audience, as the people meter⁷ performs the function of identifying audience members. *Nielsen-372*, [0048]; *Dec.* ¶ 142.

Petitioner’s proposed modification would require the ON/OFF device camera (*i.e.*, the only “camera” of *Nielsen-372*) to face the audience to recognize the faces of the audience members using “Steinberg’s methods.” *Petition*, 44-45; *Dec.* ¶ 143. With the camera facing the audience, the *Nielsen-372* camera would no longer be positioned to take images of the screen to detect when the display is ON or OFF. *Dec.* ¶ 143. As described, detecting when the display device is ON or OFF (to increase the “accuracy” of the generated viewing records) is the purpose of the video sensor in the *Nielsen-372* system. *Nielsen-372*, [0043]; *Dec.* ¶ 144.

Modifying *Nielsen-372* so that the camera faces the audience (to meet the claim limitation of the ’243 Patent) instead of the display (the actual functionality disclosed in *Nielsen-372*) significantly and fundamentally changes how *Nielsen-372*’s system operates and renders it unsatisfactory for its intended purpose, (*i.e.*, generating accurate viewing records by detecting whether the display is on or off).

⁷ The people meter of *Nielsen-372* is a remote-control-like device with buttons for each viewer, allowing individual viewers to register their presence in the media exposure environment voluntarily. *Nielsen-372*, [0048]

Dec., ¶¶ 144-145. Petitioner’s proposed modification would render the ON/OFF device, and consequentially the metering system, “**unsatisfactory for its intended purpose**,” making the proposed modification **not obvious**. *Victoria’s Secret Stores LLC v. Andra Group, LP*, IPR2020-00853, Paper 12, 28-29 (PTAB Oct. 22, 2020) (citing *In re Gordon*, 733 F.2d 900, 902 (Fed. Cir. 1984)).

b) Petitioner’s Proposed Modification Would Change the Principle of Operations of Nielsen-372’s System

Modifying the camera of the ON/OFF device to capture images of the audience and automatically identifying audience members using *Steinberg*’s facial recognition techniques would change the principle operation of the *Nielsen-372* system in two important ways: first, the system of *Nielsen-372* would no longer be able to rely on the video sensor to determine the state of the display, but would be forced to rely on other, less accurate techniques. *Dec.*, ¶¶ 146-149. Second, the system of *Nielsen-372* would be fundamentally transformed from a **voluntary** audience measurement system to an **automatic** audience detection system. *Dec.* ¶¶ 150-152.

When a modification would “change the principles” under which a prior art reference “was designed to operate,” as Petitioner’s proposed modification does, a POSITA would not be motivated to modify the prior art reference in that way. *TCT Mobile (US), Inc. v. Fundamental Innovations Systems Int’l LLC*, IPR2021-00597, Paper 8, 14-15 (PTAB Aug. 25, 2021). A change of the principle of operations

provides “evidence of non-obviousness,” and the Board should reject Petitioner’s proposed modification of *Nielsen-372* for at least this reason. *Id.* at 15, 17.

First, the system of *Nielsen-372* describes a sophisticated combination of audio sensors, video sensors, and emissions sensors making up the ON/OFF device. *Nielsen-372*, FIG. 2, [0133]; *Dec.*, ¶¶ 147-148. Petitioner’s modification would entirely remove one class of sensors from this combination. Dr. Saber opines that the video sensor is an important sensor, as it enables determining when the display device is actually presenting changing content (such as a TV show) to the audience. *Dec.* ¶¶ 147, 149. In other words, removing the video sensor would change the system of *Nielsen-372* from one that can detect when a television is presenting content to an audience into one that cannot distinguish an audience watching television from an audience staring at a blank screen. *Dec.* ¶ 149.

Next, Petitioner’s proposed system would change *Nielsen-372* from a voluntary system into one in which the audience is tracked without their input (and potentially without their knowledge). *Dec.* ¶ 150; *see also* Ex. 1003, ¶ 201. Such a change is fundamental and could significantly impact the willingness of an audience to have *Nielsen-372*’s measurement system in their living room. *Dec.*, ¶ 150. For example, Dr. Saber discusses how audiences might be willing to share the majority of their watching habits using the people meter described in *Nielsen-372* but value the ability not to record some “guilty pleasure” watching. *Dec.* ¶¶ 151-152. Such a

change represents another fundamental change to the principle of operation of the system of *Nielsen-372*, Petitioner's unsupported claims to the contrary notwithstanding. *Cf. Dec.*, ¶¶ 147-153 *with Petition*, 45.

Petitioner asserts that all of the modifications *Nielsen-372* based on *Steinberg* would “not have changed the principle of operation for any of the references.” *Petition*, 45. Petitioner cites Dr. Doermann's opinion to support this argument. *Id.* Dr. Doermann merely parrots the attorney's conclusion, and his testimony should be afforded “little weight.” Ex. 1003, ¶ 203; *Xerox Corp.*, 15. Thus, Petitioner's assertion is little more than attorney argument, which “does not” provide evidence for an understanding of a POSITA. *Flowco Production Solutions*, 17. Ironically, Dr. Doermann notes both of these fundamental changes in passing but performs no analysis and provides no substantial opinions on their effect on *Nielsen-372's* principles of operation. Ex. 1003, ¶ 198 (noting the Nielsen's camera must now “capture one or more images of a television audience”); ¶ 201 (noting that the modified system is now “automatic” and “not affected by the individual's input or lack thereof.”).

For all of the above reasons, Petitioner's proposed combination is fatally flawed, as it would fundamentally change the principle of operation of *Nielsen-372's* metering system in multiple ways. As such, a POSITA would not have found Petitioner's proposed modification to *Nielsen-372* obvious. *TCT Mobile (US), Inc.*,

14-15. Having presented this flawed combination, Petitioner further fails to explain “why” a POSITA would have been motivated to modify *Nielsen-372* “in a way that it would not work for its intended purpose.” *Id.* at 16. Thus, Petitioner fails to establish a *prima facie* showing of obviousness for Ground 3, and Board should reject Petitioner’s proposed modification of *Nielsen-372* for at least this reason. *Endo Pharm., Inc.*, 922 F.3d at 1373.

2. *Modifying Nielsen-372/Steinberg with Tian’s Teachings Does Not Achieve Petitioner’s Proposed Benefit, and Any Benefit Would Be Outweighed by Drawbacks Petitioner Fails to Address*

Petitioner proposes further modifying the system of *Nielsen-372/Steinberg* (introduced in the now-moot Ground 2) to incorporate *Tian*’s teachings in order to render the Challenged Claims obvious for Ground 3. As set forth below, Petitioner has failed to establish that a POSITA would be motivated to modify the *Nielsen-372/Steinberg* system with the teachings of *Tian*, or that such a combination would achieve Petitioner’s proposed benefit.

a) *Petitioner’s Proposed Modification to Process Reduced-Resolution Images Would Not Reduce the Net Reduction of Processing Resources Required*

As in Ground 1, Petitioner’s Ground 3 modification of *Nielsen-372/Steinberg* to incorporate *Tian*’s teachings of processing reduced resolution images is motivated by reduced processing resources and time. The analysis provided is essentially identical between the two Grounds. *Cf. Petition*, 72-73 with *Petition*, 10-11; *cf. Ex.*

1003 ¶ 68-70 with Ex. 1003, ¶¶ 286-288. Petitioner provides no more weighing of the associated disadvantages against the claimed benefits for Ground 3 than for Ground 1. *Id.* In the mapping for element [4a], Petitioner further suggests that Steinberg’s teachings of processing images with different resolutions would motivate a POSITA to incorporate *Tian*’s teachings of actually reducing the resolution. *Petition*, 75-76 (*citing*, Ex. 1003, ¶ 312). However, none of this material remedies Petitioner’s failure to weigh the claimed benefits of processing reduced-resolution images against the additional drawbacks (in the form of the additional processing resources and time required to reduce the resolution of the images).

For the reasons set forth with respect to Ground 1 above in § II.B.1.a), this failure is fatal to the combination of Ground 3, and the Board should find claims 4, 11, and 18 (and by virtue of dependency, the other challenged claims) not unpatentable for at least this reason.

b) Petitioner’s Proposed Modification to Rely on Head Tracking Would Not Reduce the Net Reduction of Processing Resources Required and Would Decrease the Performance of Face Recognition

As in Ground 1, Petitioner’s Ground 3 modification of *Nielsen-372/Steinberg* to incorporate *Tian*’s teachings of head tracking is motivated by reduced processing resources and time. The analysis provided is essentially identical between the two Grounds. *Cf. Petition*, 72 with *Petition*, 9; *cf. Ex. 1003* ¶ 63-64 with *Ex. 1003*, ¶¶ 284-285. Petitioner provides no more weighing of the associated disadvantages

against the claimed benefits for Ground 3 than for Ground 1. *Id.* For the reasons set forth with respect to Ground 1 above in § II.B.1.c), this failure is fatal to the combination of Ground 3, and the Board should find claims 5, 6, 12, 14, and 19 (and by virtue of dependency, claims 8, 13, and 20) not unpatentable for at least this reason.

IV. NEITHER TIAN NOR STEINBERG IS ANALOGOUS ART TO THE '243 PATENT

All of Petitioner's modifications of the primary references across all grounds share a common, fatal flaw: none of Petitioner's secondary references are analogous art. As the Federal Circuit recently affirmed, the test for whether a prior art reference is analogous art is a two-prong one: ““(1) whether the art is from the same field of endeavor, regardless of the problem addressed and, (2) if the reference is not within the field of the inventor's endeavor, whether the reference still is reasonably pertinent to the particular problem with which the inventor is involved.”” *Id.* at *8.

Petitioner alleges that the field of endeavor of the '243 Patent is “media analysis” and that this field of endeavor is shared by the prior art references. *Petition*, 4-6 (identifying the field of endeavor for each of the references). Notably, neither Petitioner nor its declarant makes any showing of the cited references being reasonably pertinent to a problem with which the inventor is involved. However, the field of endeavor of the '243 Patent does not broadly encompass “media analysis,” and the cited secondary references do not share any reasonably scoped field of

endeavor for the '243 Patent. Notably, Petitioner does **not** argue that either *Tian* or *Steinberg* is “reasonably pertinent” to the particular problem being addressed by the inventors of the '243 Patent.

A. Neither the '243 Patent nor Petitioner’s Art in the Field of “Media Analysis”

The term “media analysis” is introduced by Petitioner in an apparent effort to find an umbrella broad enough to cover the '243 Patent and the disparate prior art references. Petitioner’s expert admits that “[t]he phrase [media analysis] does not appear in the ['243] Patent, as I said before, or any of the four specific references that we – that I used.” Ex. 2018, *Doermann Deposition Transcript*, 25:6-8. However, this term does not reflect the content of the '243 Patent, *Tian*, or *Steinberg* and is merely an attempt to obfuscate their disparate fields.

First, the '243 Patent is unambiguous in its general “Field of the Disclosure” statement: “[t]his disclosure relates generally to **audience measurement** and, more particularly, to methods and apparatus to capture images.” '243 *Patent*, 1:36-39. This is supported by the claims: independent claims 1 and 9 recite “an audience measurement system,” and independent claim 16 recites “a method for obtaining exposure data for a media exposure environment.” '243 *Patent*, cols. 1, 9, 16; *see also Netflix v. DivX*, No. 22-1138, *10 (Fed. Cir. Sept. 11, 2023) (“We determine the field of endeavor by reference to explanations of the invention’s subject matter

in the patent application, including the embodiments, function, and structure of the claimed invention.”) (internal quotes omitted).

Next, *Tian*, while not a patent reference, also directly sets out its goal: “[t]he goal was to help understand [the] question: how do we recognize facial expressions in real life.” *Tian*, 000006; *see also Ex. 2018, Doermann Deposition Transcript* 21:13-16, 25:11-20 (Dr. Doermann agreeing that the goal of *Tian* is to understand how we identify facial expressions in real life). To achieve this goal, *Tian* examines images at different resolutions to determine the effects of resolution on the ability to identify facial expressions: “[t]his paper explores the effects of different image resolutions for each step of facial expression analysis.” *Tian*, Abstract. *Nielsen*, by contrast, counts and identifies individual audience members (regardless of their facial expressions) and never addresses facial expressions. *See id.* at 23:4-6 (Dr. Doermann agreeing that the ‘243 Patent never discusses facial expressions).

Finally, *Steinberg*, like the ’243 Patent, is unambiguous in its “Field of the Invention” statement: “[t]he invention relates to digital image processing, particularly to the field of **automatic or semiautomatic grouping and classification of images in a database** or image collection and based on the occurrence of faces in the images and the identification and classification of such faces.” *Steinberg*, [0003]. *Steinberg* goes on to note that “the present invention is presented primarily in the context of collections of consumer digital images which

would be generated by a typical user of a digital camera.” *Steinberg*, [0010]. Thus, *Steinberg*’s field of endeavor is labeling a consumer’s collection of photos with the identities of their friends and family when they appear in those photos. *See, e.g., Steinberg*, [0090], [0098].

As such, neither *Steinberg* nor *Tian* shares the field of endeavor of the ’243 Patent and, as such, a POSITA would not have been motivated to combine them with the respective primary references. *Dec.*, ¶¶ 158-161.

In summary, Petitioner’s characterization of the field of endeavor conflates two entirely distinct fields of endeavor: (1) media **exposure** analysis (*i.e.*, audience measurement); and (2) image processing. *Dec.*, ¶¶ 159-162. As described above, *Tian* is directed to determining the effects of image resolution on facial expression analysis and therefore falls into the category of image processing. *Steinberg* is directed to classifying faces in an image database and is thus likewise directed to image processing. Accordingly, neither of Petitioner’s secondary references is directed to the field of endeavor of the ’243 Patent, that of audience measurement. *Dec.* ¶¶ 159-161.

B. Petitioner’s Art Is Not Reasonably Pertinent to the Problem Being Solved by the ’243 Patent

Although Petitioner does not provide any argument that either *Tian* or *Steinberg* is reasonably pertinent to the problem being solved, Patent Owner addresses the second prong as well. A prior art reference is reasonably pertinent to

the problem solved “only if its subject matter logically would have commended itself to an inventor’s attention in considering his problem.” *Netflix* at *15 (quoting *Airbus S.A.S. v. Firepass Corp.*, 941 F.3d 1374, 1382 (Fed. Cir. 2019)) As described above, the problem being solved by the claimed invention is measuring audiences for media ratings, *i.e.*, counting and identifying viewers of a particular piece of media content. ’243 Patent 1:64-2:3; *Dec.*, ¶ 163.

Tian’s investigation is broadly related to “facial expression analysis.” *Dec.*, ¶ 164. As Dr. Saber opines, a reference investigating the minimum resolution for an image that still allows an expression to be analyzed would not have logically come to the attention of an inventor solving the problem of audience measurement. *Id.*

Similarly, *Steinberg* is broadly related to the classification of images in a database “in the context of collections of consumer digital images which would be generated by a typical user of a digital camera.” *Steinberg*, [0010]. As Dr. Saber opines, a reference related to classifying collections of images from consumer digital cameras would also not have logically come to the attention of an inventor solving the problem of audience measurement. *Dec.*, ¶ 165. As such, neither *Tian* nor *Steinberg* is analogous art under the second prong of the relevant test.

V. CONCLUSION

For the reasons provided herein, Patent Owner submits that Petitioner has not established a reasonable likelihood that any challenged claim is unpatentable, and respectfully requests the Board uphold the validity of all claims.

Respectfully submitted,

/Jennifer C. Bailey/

Jennifer C. Bailey, Reg. No. 52,583
Robin Snader, Reg. No. 66,085
Virginia Brown, Reg. No. 80,538
Megan Redmond (*pro hac vice* filed)
ERISE IP, P.A.
7015 College Boulevard, Suite 700
Overland Park, Kansas 66211
Telephone: (913) 777-5600
Facsimile: (913) 777-5601
Jennifer.Bailey@eriseip.com
Robin.Snader@eriseip.com
Virginia.Brown@eriseip.com
Megan.Redmond@eriseip.com
PTAB@eriseip.com

ATTORNEYS FOR PATENT OWNER

APPENDIX OF EXHIBITS

Exhibit 2001	Declaration of Dr. Eli Saber
Exhibit 2002	<i>Request for Statutory Disclaimer as Filed</i> , October 09, 2023
Exhibit 2003	<i>Complaint for Patent Infringement</i> , C.A. No. 1:22-cv-01345
Exhibit 2004	<i>TVision's Initial Invalidity Contentions Regarding the '243 Patent</i> , Exhibit A1-A4
Exhibit 2005	<i>Notice, Consent, and Reference of a Civil Action to a Magistrate Judge</i> , Nov. 7, 2022
Exhibit 2006	<i>Docket Report for Case No. 1:22-cv-01345</i>
Exhibit 2007	<i>Topia Technology, Inc. v. Egnyte, Inc.</i> , No. 21-cv-01821 (D. Del. Mar. 6, 2023), <i>Hearing Transcript</i> , 5:4-12, 48:16-22
Exhibit 2008	<i>Curriculum Vitae of Dr. Eli Saber</i>
Exhibit 2009	<i>Facial Expression Recognition: A Brief Tutorial Overview</i> , Claude Chibelushi and Fabrice Bourel, Staffordshire University, 2002
Exhibit 2010	<i>Facial Expression Recognition</i> , Maja Pantic, Springer-Verlag Berlin Heidelberg, 2009
Exhibit 2011	<i>Automatic Analysis of Facial Expressions: the State of the Art</i> , Maja Pantic and Leon Rothkrantz, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, No. 12, December 2000
Exhibit 2012	<i>Mechanisms of Face Perception</i> , Doris Y. Tsao and Margaret S. Livingstone, <i>Annu Rev Neurosci.</i> 2008, 31, 411-437
Exhibit 2013	<i>Stipulation and Proposed Order Regarding Claim Construction</i> , October 30, 2023
Exhibit 2014	DSP, Chapter 4 - Oppenheim
Exhibit 2015	Text Extraction, Enhancement and OCR in Digital Video
Exhibit 2016	US Patent Application Publication No. US20070257934A1 to Doermann, et al.
Exhibit 2017	Camera-based analysis of text and documents- a survey
Exhibit 2018	Deposition Transcript of Dr. David Scott Doermann
Exhibit 2019	Declaration of Megan J. Redmond
Exhibit 2020	Supplemental Declaration of Dr. Eli Saber
Exhibit 2021	<i>Machine Analysis of Facial Behavior: Naturalistic and Dynamic Behavior</i> , Maja Pantic, The Royal Society, 2009
Exhibit 2022	<i>Super-Resolution Reconstruction and Its Future Research Direction</i> , Patanavijit, 2009

Exhibit 2023	<i>Analysis and Comparison of Various Image Downsampling and Upsampling Methods</i> , Abdou Youssef, Department of EECS, December 1998
Exhibit 2024	<i>1/2.5 Inch 5Mp CMOS Digital Image Sensor</i> , Aptina Imaging
Exhibit 2025	<i>Aliasing</i> , Bruno Olshausen, October 10, 2000
Exhibit 2026	U.S. Patent No. 7,697,012 to Lee et al. (“ <i>Lee '012</i> ”)
Exhibit 2027	U.S. Patent No. 7,298,866 to Lee (“ <i>Lee '866</i> ”)
Exhibit 2028	U.S. Patent Application Publication No. 2012/0281028 to Orlick et al. (“ <i>Orlick</i> ”)
Exhibit 2029	<i>Quadrature Prefiltering for High Quality Antialiasing</i> , Guenter et al., Georgia Institute of Technology, October 1996
Exhibit 2030	European Patent No. 0694188 to Arbeiter et al. (“ <i>Arbeiter</i> ”)
Exhibit 2031	<i>Filtering by Repeated Integration</i> , Paul Heckbert, ACM, Vol. 20, No. 4, 1986
Exhibit 2032	<i>Using the Convolution Filter Process</i> , Release 9.2, April 14, 2008
Exhibit 2033	<i>Convolution Algorithms</i> , Sidney Burrus, April 10, 2010
Exhibit 2034	<i>Face Recognition Algorithms</i> , Universidad del Pais Vasco, June 16, 2010
Exhibit 2035	<i>Facial Feature Based Head Tracking and Pose Estimation</i> , Jari Hannuksela, Department of Electrical and Information Engineering, University of Oulu, 2003
Exhibit 2036	<i>Is Your TV Too Big for Your Room?</i> , Ann Porter, October 3, 2007
Exhibit 2037	<i>Sectional Sofa With Two Pillows</i> , https://ilovebbcsmlife/product_details/83389857.html
Exhibit 2038	<i>Face Detection and Tracking Techniques: A Review</i> , Intl. Journal of Advanced Research in Computer Science, Vol. 2, No. 5, October 2011

CERTIFICATION OF WORD COUNT

The undersigned certifies pursuant to 37 C.F.R. § 42.24 that the foregoing Patent Owner's Preliminary Response to Petition for *Inter Partes* Review, excluding any table of contents, mandatory notices under 37 C.F.R. § 42.8, certificates of service or word count, or appendix of exhibits, contains 13,823 words according to the word-processing program used to prepare this document (Microsoft Word).

Dated: May 1, 2024

Respectfully submitted,

/Jennifer C. Bailey/

Jennifer C. Bailey, Reg. No. 52,583

ERISE IP, P.A.

7015 College Boulevard, Suite 700

Overland Park, Kansas 66211

Telephone: (913) 777-5600

Facsimile: (913) 777-5601

Jennifer.Bailey@eriseip.com

PTAB@eriseip.com

ATTORNEYS FOR PATENT OWNER

CERTIFICATE OF SERVICE

I hereby certify that a true and correct copy of the foregoing PATENT OWNER'S PRELIMINARY RESPONSE TO PETITION FOR INTER PARTES REVIEW was served on May 1, 2024, by email on the following counsel of record for Petitioner:

Harper Batts (Reg. No. 56,160)
Chris Ponder (Reg. No. 77,167)
Jeffrey Liang (Reg. No. 69,043)

hbatts@sheppardmullin.com
cponder@sheppardmullin.com
jliang@sheppardmullin.com

Dated: May 1, 2024

Respectfully submitted,

/Jennifer C. Bailey/

Jennifer C. Bailey, Reg. No. 52,583
jennifer.bailey@eriseip.com

ATTORNEYS FOR PATENT OWNER