

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.; INSTAGRAM, INC.; WHATSAPP LLC; META
PLATFORMS TECHNOLOGIES, LLC; AND GIPHY, INC.

Petitioners

v.

VL COLLECTIVE IP LLC

Patent Owner

Case IPR2023-00924
U.S. Patent 7,436,980

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	BACKGROUND	3
A.	Overview of the '980 Patent.....	3
B.	The Claims of the '980 Patent.....	8
C.	Overview of Zhao (EX1006).....	9
D.	Overview of Steffens (EX1007).....	11
E.	Overview of Ozer (EX1008).....	13
F.	Overview of TLP (EX1009).....	14
III.	PRIORITY DATE OF THE '980 PATENT	16
IV.	LEVEL OF ORDINARY SKILL	16
V.	RELATED MATTERS	16
VI.	DISCRETIONARY DENIAL IS WARRANTED.....	16
A.	Petitioners' abusive citation practices warrant discretionary denial under § 314(a).....	16
VII.	PETITIONERS HAVE NOT DEMONSTRATED A REASONABLE LIKELIHOOD OF PREVAILING ON ANY CHALLENGED CLAIM.....	21
A.	The Petition fails to identify, "with particularity," arguments or evidence sufficient to establish a reasonable likelihood that TLP or Zhao was publicly accessible, disseminated, or otherwise qualifies as prior art.	21
1.	The Petition identifies insufficient evidence to meet Petitioners' burden.	21
2.	Petitioners should not be permitted a reply or to submit supplemental information to address prior art or printed publication status.....	29

B.	The Petition fails to demonstrate that Zhao (Ground I) discloses or renders obvious every limitation of claims 1 and 9.....	31
1.	1/9[a]: “providing a spatio-temporal model for an object to be detected”.....	31
2.	1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”.....	41
C.	The Petition fails to demonstrate that Steffens in view of Zhao (Ground II) renders obvious every limitation of claims 1 and 9.....	45
1.	1/9[a]: “providing a spatio-temporal model for an object to be detected”.....	45
2.	1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”.....	52
D.	The Petition fails to demonstrate that Ozer in view of Zhao (Ground III) discloses or render obvious every limitation of claims 1 and 9.....	56
1.	1/9[a]: “providing a spatio-temporal model for an object to be detected”.....	56
2.	1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”.....	61
E.	The Petition fails to demonstrate that TLP (Ground IV) discloses or render obvious every element of claims 1 and 9.....	63
1.	1/9[a]: “providing a spatio-temporal model for an object to be detected”.....	63
2.	1/9[e] “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”.....	65

F.	The Petition fails to demonstrate a reasonable likelihood that dependent claims 2-8 and 10-16 will be proven unpatentable.....	68
VIII.	CONCLUSION.....	68

PATENT OWNER'S EXHIBIT LIST

Exhibit	Description
2001	Declaration of Jaime F. Cardenas-Navia in Support of Motion for Admission <i>Pro Hac Vice</i>
2002	Zou et al., Object Detection in 20 Years: A Survey (May 16, 2019).

VL Collective IP LLC (“Patent Owner”) submits this Preliminary Response to the Petition of Meta Platforms, Inc.; Instagram, Inc.; WhatsApp LLC; Meta Platforms Technologies, LLC; and Giphy, Inc. (“Petitioners”) seeking *inter partes* review (“IPR”) of U.S. Patent No. 7,436,980 (“the ’980 patent”).

I. INTRODUCTION

Institution should be denied for several reasons. To begin with, the Board should exercise its discretion under 35 U.S.C. § 314(a) to deny the Petition in view of Petitioners’ incorporations by reference and abusive citation practices for circumventing the wordcount limitations on petitions. Between Petitioners’ explicit incorporations by reference and repeated practice of supporting conclusory assertions with little more than bare citations to supporting declarations, the Petition incorporates by reference thousands of additional words and the entirety of Petitioners’ substantive positions (less the conclusory assertions in the Petition itself) regarding the prior art status of two references (TLP and Zhao) relied on in all four asserted Grounds. Whether the Board deems Petitioners’ abusive citation practice a violation of the rules worth denying the Petition, or instead a substantive defect that results in Petitioners’ failure to satisfy its burden of proof for institution, the Board should deny the Petition.

Second, Petitioners fail to demonstrate a reasonable likelihood that their TLP and Zhao references constitute prior art printed publications under the law. For both

references, Petitioners fail to identify, with sufficient particularity, arguments and supporting evidence of public accessibility *in the Petition*. Petitioners' attempt to improperly incorporate arguments and evidence from other documents through open ended and unexplained citations are improper attempts to circumvent the wordcount limits and should not be condoned. Further, even if the Board were to fully consider Petitioners' improperly incorporated testimony and evidence, Petitioners' arguments regarding the TLP reference are legally insufficient to show it constitutes prior art.

Third, the Petition fails to demonstrate a reasonable likelihood of proving the challenged claims are obvious. Petitioners' four asserted Grounds of unpatentability include two single-reference Grounds (I and IV) and two combination Grounds (II and III). Petitioners fail to show that the asserted reference in each single-reference Ground (Zhao and TLP, respectively) discloses or renders obvious every element of independent claims 1 and 9. Petitioners' combination Grounds fail because (a) they rely on purported disclosures of Zhao that do not disclose every element of claims 1 and 9, (b) the additional references fail to disclose claim elements for which they are relied upon, and (c) Petitioners' motivation to combine and reasonable expectation of success arguments are legally deficient.

II. BACKGROUND

A. Overview of the '980 Patent

The '980 Patent is concerned with object detection. Object detection—determining and tracking what is within images and videos—is a vital part of numerous web-based technologies, including social media and communications platforms. At base, “[t]he objective of object detection is to develop computational models and techniques that provide one of the most basic pieces of information needed by computer vision applications: What objects are where?” EX2002.

Object detection presents fundamental challenges. What may seem simple for a human is a complex exercise for a computer, which in the past required great computational resources. It is difficult to program a computer to know a cat from a dog or a car from a boat, and, most importantly, to do so in a way that doesn’t require the programmer to manually identify all possible objects. Object detection is made even more challenging by partially covered objects, rotated objects, scaled objects (such as a small house in the background of a photo), etc.

To address these problems, the inventors of the '980 patent came up with a method to detect and track an object within a video by measuring the object as a collection of components in each image and, using a spatio-temporal model, determining the probability that the object is in any given image that makes up the video. *See* EX1003, 3 (“We represent objects as spatio-temporal two-layer graphical

models, where each node corresponds to an object or component of an object at a given time.”). This approach was completely unheard of in the field of object detection, but after the invention of the ’980 Patent, its method became widely adopted. While it was relatively common to “represent objects as collections of features with distinctive appearance, spatial extent, and position,” *see* EX1001, 1:15-21, the ’980 patent’s solution to the object detection problem was unknown at the time of the invention, and even, but for the ’980 patent, for years after. This demonstrates the unique problems inherent in object detection and highlights the inventive nature of the ’980 patent’s solutions.

Critical to the insight of the ’980 patent is the use of a spatio-temporal model for an object that is to be detected. As the ’980 patent explains, “[t]he spatio-temporal model is a graphical model comprising nodes corresponding to each of the collection of components and to the object.” EX1001, 1: 41-42. The patent’s spatio-temporal model thus comprises nodes representing two types of things (1) an object, and (2) the object’s components. The ’980 patent describes this as a “two-layer graphical model” as having a “course, object [layer], [which] corresponds to an entire appearance model of the object” and a “fine, component [layer], [which] includes a set of loosely connected parts.” EX1001, 2:55-59, 4:20-23. The object layer is in turn “connected to all constituent components.” EX1001, 4:23-24. For example, in the case of a pedestrian (an object), the components might be the head,

the left arm, the right arm, and the legs. EX1001, 4:18-33. Thus, the component layer of the spatio-temporal model would be made up of nodes corresponding to those components, while the object layer is the entire object—the pedestrian. EX1001, 2:55-59, 4:18-33, Figs. 2A and 3 (annotated below).

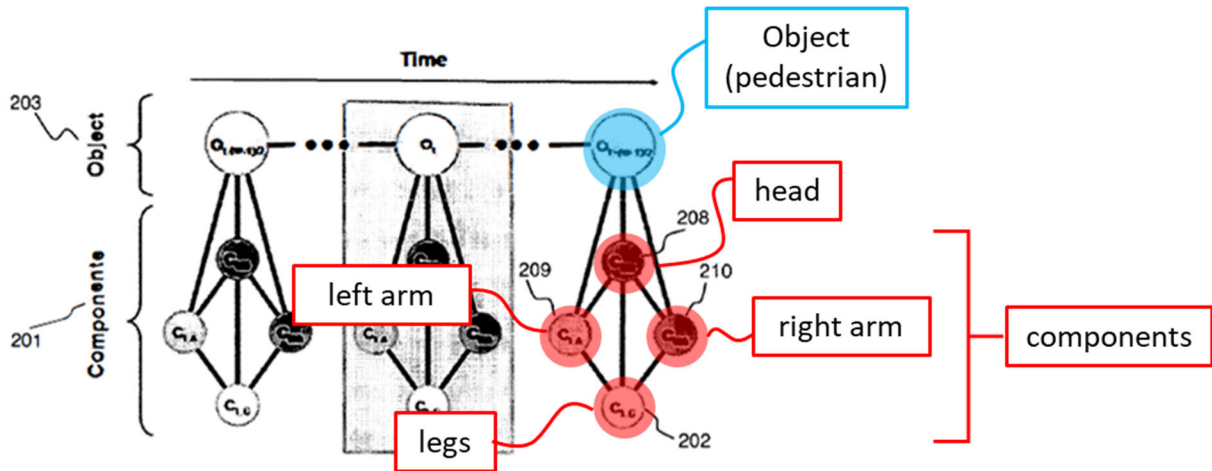
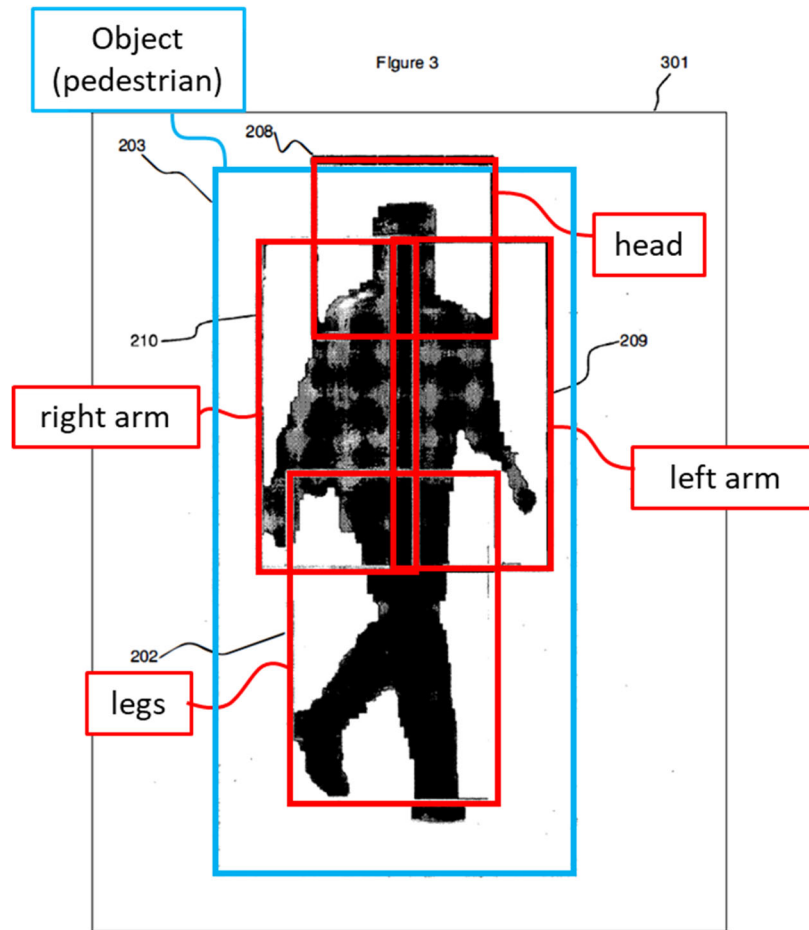


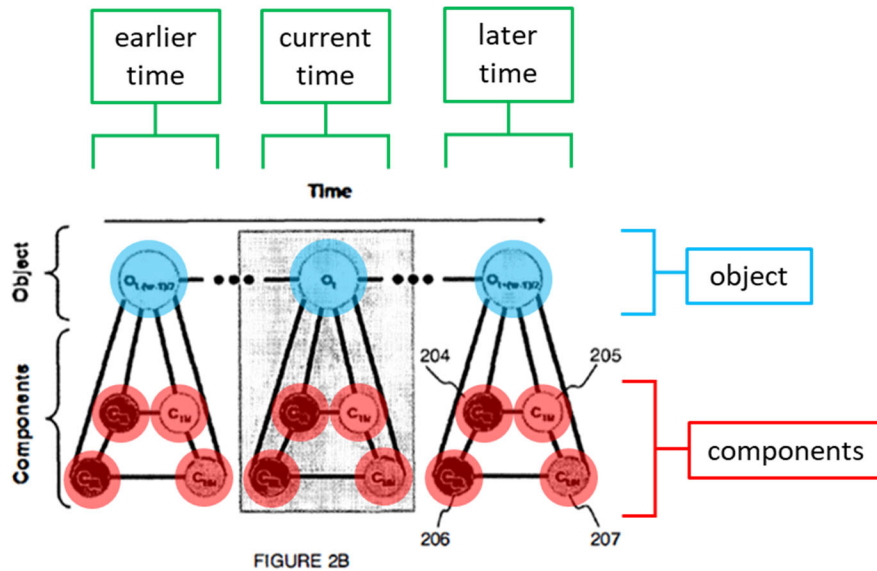
FIGURE 2A



The spatio-temporal model of the '980 patent models an object over both space and time. *See* EX1001, 1:22-24, 3:47-49. As the patent explains, “[e]ach node in the graph represents either the object or a component of the object *at time t*” and may have an associated state vector of the object (X_t^O) or component ($X_t^{C_n}$), respectively, defining its “real-valued position” *at time t*.¹ EX1001, 3:47-53 (emphasis added). For a single frame, “objects are represented using a two-layer

¹ Where time t goes from 1-T and component number n goes from 1-N. *See* EX1001, 3:47-4:7.

spatial graphical model” (as explained above), and “[t]o integrate temporal constraints” resulting in a spatio-temporal model, “the spatial graphical model[] [is] extended over time to an arbitrary length temporal window.” EX1001, 4:34-337, Fig. 2B (annotated below).



The inventors of the '980 patent recognized the value in their approach to object modeling. As the patent states, “[h]aving a two-layer graphical model allows the inference process to reason explicitly about the object as a whole . . . and reduce the complexity of the graphical model by allowing the assumption of the conditional independence of components . . . over time.” EX1001, 4:37-41. The inventors recognized this as an improvement over prior modeling techniques—including the techniques in Petitioners’ asserted TLP reference (EX1009)—in which images and videos were analyzed as a single layer. *See* EX1003, 6 (“Unlike the model presented in [TLP] having a second-layer also allows us to assume conditional independence

of components over time given the whole appearance, hence reducing the complexity of the overall graphical model.”). In addition to speeding up object detection, the patented method “facilitates detection of complex articulated objects as well as helps in handling partial object occlusions [e.g., partially covered items] or local illumination changes [e.g., changes in light].” EX1001, 2:51-54.

B. The Claims of the '980 Patent

Representative claim 1 of the '980 patent is reproduced below. Key concepts discussed above and relevant to this Preliminary Response are italicized.

1. A computer implemented method for object detection comprising:

providing a spatio-temporal model for an object to be detected;

providing a video comprising a plurality of images including the object;

measuring the object as a collection of components in each image;

determining a probability that the object is in each image; and

detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object.

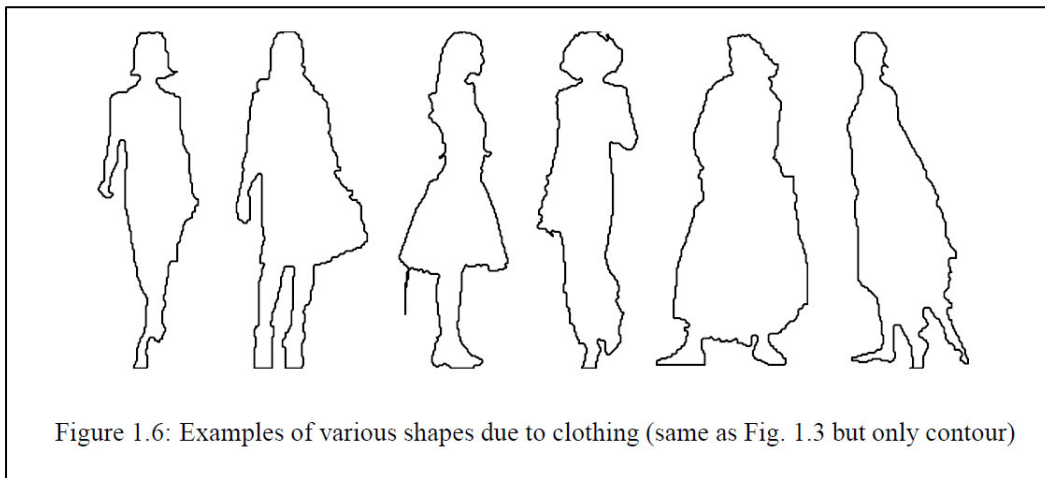
The '980 patent claims novel methods and systems that relate to new and inventive techniques for detecting objects in video images using the two-layered

spatio-temporal model described in the '980 patent, as described above. Claims 1 and 9 are the independent claims. Claims 2-8 depend from claim 1, and claims 10-16 depend from claim 9.

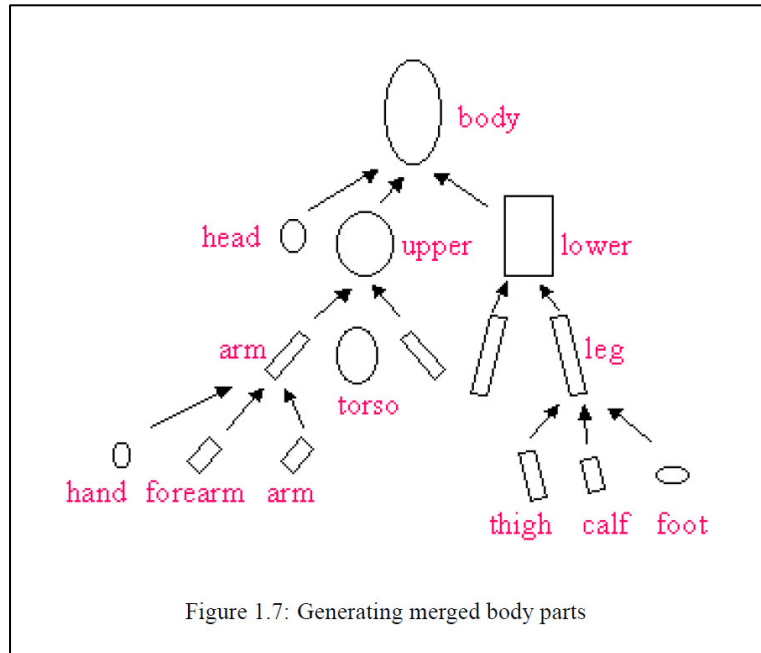
The Petition addresses independent claims 1 and 9 together, making the same arguments as to both claims. Accordingly, Patent Owner treats claim 1 as representative of claims 1 and 9 for purposes of this Preliminary Response.

C. Overview of Zhao (EX1006)

Zhao states that “[a]lthough humans can easily detect other humans and their body part locations from a single image, these problems are inherently difficult for a computer.” EX1006, 20. Zhao teaches a “recognition scheme [that] is [] based on the shapes of body parts and the relationships between them.” *Id.* To that end, Zhao addresses “how to decompose a silhouette into parts, and how to represent the shapes and the relationships between the parts.” *Id.*



In Zhao's approach, "body parts are organized in a coarse-to-fine hierarchy and body part matching is performed in a coarse-to-fine manner." *Id.* 23, 50.² Zhao generates a "human model" by constructing a "representation of the shapes of the body parts and their size and spatial relationships." *Id.* 23.



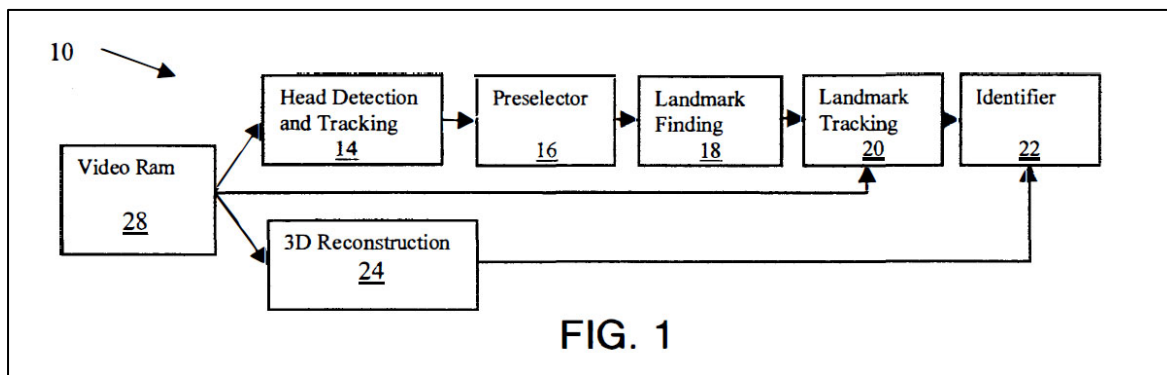
In Zhao's model, "a body part is parameterized with a vector (a, l, x, y, θ) , where $a = w/l$ is the aspect ratio that captures the general shape of a ribbon, (x, y) are the coordinates of the origin in the coordinate frame of the parent part, and θ is the intersection angle between the major axes of this part and its parent part." *Id.* 43. Zhao further explains that "[a]ssuming that the human body model consists of m body parts that are parameterized with vectors $f_1, f_2, \dots, f_m$, where $f_i = (a_i; l_i; x_i; y_i;$

² Zhao mentions "fine-to-course" on pages 23 and 93, which appear to be typographical errors.

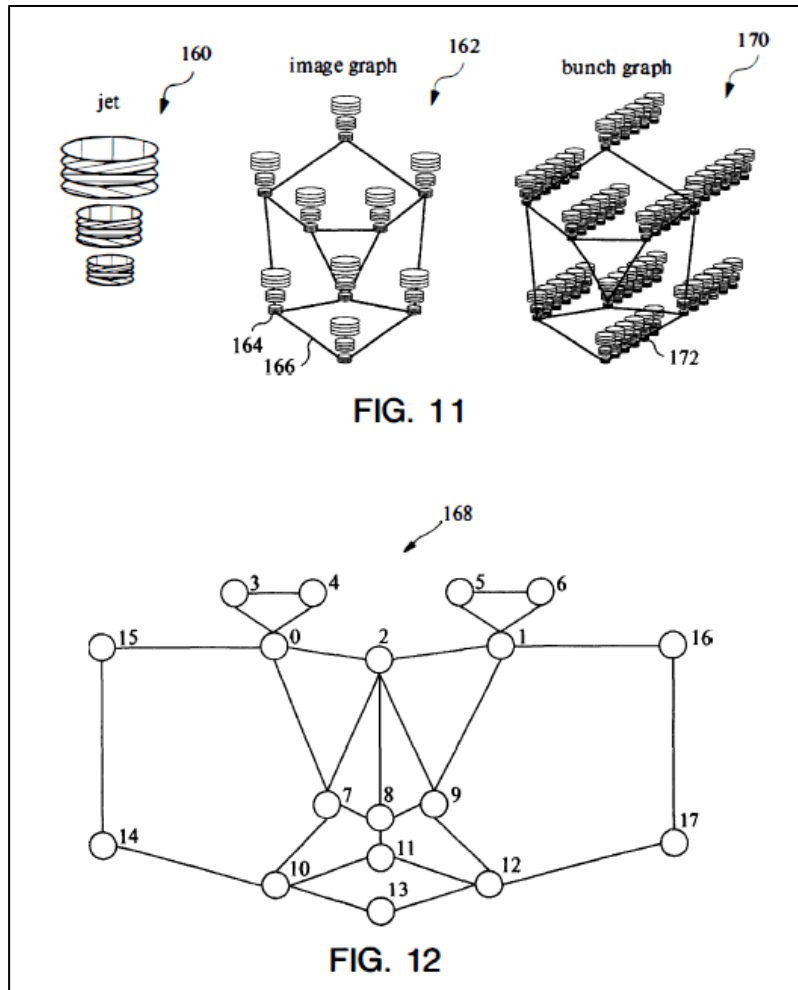
θ_i). Then, the human body model is parameterized with four model matrices: the aspect ratio vector $A = \{a_1, \dots, a_m\}$ the length ratio matrix $S = \{s_{ij}\}, i, j = 1, \dots, m$, where $s_{ij} = l_i/l_j$, the relative position vector $X = (x_1, y_1, \dots, x_m, y_m)$, and the orientation or posture vector $\Theta = \{\theta_1, \dots, \theta_m\}$.”

D. Overview of Steffens (EX1007)

Steffens relates to “vision-based object detection and tracking, and more particularly, to systems for detecting objects in video images, such as human faces, and tracking and identifying the objects in real time.” EX1007, 1:14-17; *see id.* 3:36-37 (“The object may be, for example, a head having particular facial characteristics.”). With reference to Figure 1 (reproduced below), Steffens’ process includes “receiv[ing] a person’s image from a video source” as “a stream of digital video image frames” and employing separate processes for “head detection and tracking process 14”; using a “landmark finding process 18 for detecting the individual facial features”; and using a “landmark tracking process 20” to “track [] the landmarks,” among other things. EX1007. 3:47-4:61.



According to Steffens, “landmark locations on an image correspond[] to readily identifiable features,” i.e., facial features of a head. *Id.* 3:34-46. “The facial landmarks and features of the head may be located using an elastic graph matching technique.” *Id.* 7:51-53. The graph matching technique involves “determining the similarity between [an] image graph and [a] model graph.” *Id.* 7:51-67. An “image graph” is “used to describe the aspects of an object (in this context, a face),” in which “nodes 164” in the image graph “refer to points on the object and are labeled by jets 160” and “define the graph topology.” *Id.* 8:56-67. “A model graph 168,” on the other hand, includes “nodes” corresponding to facial features (e.g., right eye pupil, left eye pupil, top of the nose, etc.). *Id.* 9:1-21. The model graph may be a “bunch graph,” in which “a whole bunch of jets 172 (a bunch jet) are attached to each node.” *Id.* 9:22-45, Figs. 11-12.



“When matching a bunch graph to an image, each jet extracted from the image is compared to all jets in the corresponding bunch attached to the bunch graph and the best-matching one is selected.” *Id.*

E. Overview of Ozer (EX1008)

Ozer states that it relates to a “method and apparatus to detect the presence of articulated objects, e.g., human body, and rigid objects and to identify their activities.” EX1008, 1:15-21. Ozer says that it seeks to provide a “system” that can “detect non-rigid (e.g., human body) and rigid object parts and recognize their

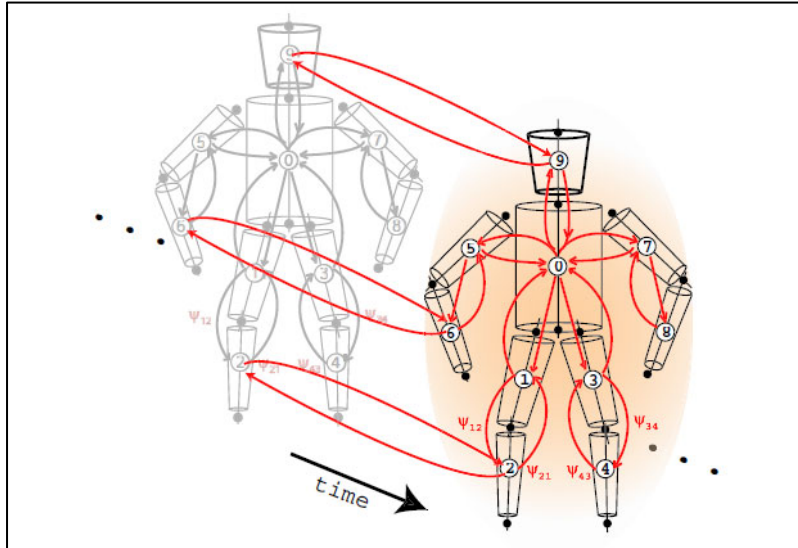
activities.” *Id.* 3:55-57. According to Ozer, its technique uses “two levels, namely low and high levels,” where “[t]he low-level part performs object detection and extracts parameters for the abstract graph representation of the image being processed” and “[t]he high level part uses dynamic programming to determine the activities of the object parts, and uses a distance classifier to detect specific activities.” *Id.* 3:55-64.

According to Ozer, “[o]bject detection is achieved by matching the relational graphs of objects with the reference model.” *Id.* 4:24-34. Then, says Ozer, “[a]fter the detection of the object parts, the system is ready to recognize the activities of each object part and the overall activity of the object.” *Id.* 4:35-37. Ozer states that “the system can identify each object part separately (in this case hands 100, head 101, and torso 102) after comparing the object attributes with the model database via a graph-matching algorithm.” *Id.* 5:66-6:3. Ozer also states that the system can “be set up to ... recognize a wide variety of activities, such as gestures ranging from pointing a gun to waving arms, leaving unattended objects, entering a prohibited area, tailgating at security doors, spending too much or too little time in an area, etc.” *Id.* 6:4-27.

F. Overview of TLP (EX1009)

TLP is a paper co-authored by Leonid Sigal (one of the inventors of the '980 patent). EX1009, 1. According to TLP, it describes a “method for tracking human

bodies” using “a loose-limbed body model . . . in which limbs are connected via learned probabilistic constraints.” *Id.* 1. “Each node in the graphical model represents . . . a limb (Figure 1).” *Id.*



TLP teaches that “[m]ost current techniques model the body as a kinematic tree,” and that “[w]hen such algorithms lose track (as they always do), the dimensionality of the state space makes it difficult to recover.” *Id.* 1. However, according to TLP, “[w]hile the full body pose is hard to recover directly, the location and pose of individual limbs is much easier to compute.” *Id.* To that end, TLP teaches that “the body is represented by a graphical model in which each graph node corresponds to a body part (upper leg, torso, etc.).” *Id.* 2; *see also id.* at 5 (describing 30 nodes for 10 “body parts” at 3 points in time).

III. PRIORITY DATE OF THE '980 PATENT

Petitioners state that “[t]he earliest potential effective filing date of the '980 Patent’s claims is May 27, 2004.” Pet. 5. The '980 patent, which issued from U.S. Application No. 11/135,210, filed on May 23, 2005, claims priority to U.S. Provisional Application No. 60/574,799, filed on May 27, 2004 and incorporates its content by reference. EX1001 1:4-6. Petitioners do not dispute that any of the '980 patent’s claims are entitled to the priority date of May 27, 2004. Pet. 5.

IV. LEVEL OF ORDINARY SKILL

Petitioners set forth a level of ordinary skill. Pet. 11. Patent Owner reserves the right to dispute Petitioners’ definition if this IPR is instituted. For present purposes, this Preliminary Response establishes that institution should be denied even under Petitioners’ preferred definition of the level of ordinary skill.

V. RELATED MATTERS

Patent Owner filed suit against Petitioners in the District of Delaware on May 24, 2022, alleging infringement of the '980 patent. *VideoLabs, Inc., et al. v. Meta Platforms, Inc. et al.*, No. 1-22-cv-00680 (D. Del.).

VI. DISCRETIONARY DENIAL IS WARRANTED

A. Petitioners’ abusive citation practices warrant discretionary denial under § 314(a).

Incorporation by reference, whether express or through broad and unexplained citations to supporting declarations, is prohibited to prevent petitioners

from engaging in self-help to circumvent the wordcount limit on petitions. *Cisco Systems, Inc. v. C-Cation Techs., LLC*, IPR2014-00454, Paper 12 at 7-10 (PTAB Aug. 29, 2014) (informative). Petitioners' Certificate of Compliance states that the Petition includes 13,998 words—two words under the limit. Pet. 76; 37 C.F.R. § 42.24. However, the Petition expressly, and through the use of broad and unexplained citations to supporting declarations, incorporates by reference thousands of additional words, including argument, thereby circumventing the wordcount limit for petitions.

Petitioners expressly “incorporated by reference” an entire section of their expert’s declaration, amounting to over approximately 2,000 words. Pet. 5 (“Petitioners’ challenge is based on the following prior art, for which a detailed overview is provided in the Declaration of Dr. Bajaj, which is incorporated by reference. EX1004, Section VIII (Overview of the Prior Art).”). Worse, this section includes arguments and characterizations that go beyond providing a mere “overview.” For example, Dr. Bajaj argues that Zhao and TLP were allegedly publicly available and disseminated, which is not an “overview” of the reference, but instead seeks to demonstrate that they constitute prior art under the legal standard. *See* EX1004, ¶¶ 48, 65-67. Dr. Bajaj also argues that the references disclose claim elements, such as the “spatio-temporal model,” and “temporal models” beyond the scope of their disclosure. *See, e.g., id.* ¶¶ 51, 64.

The Petition also incorporates by reference (through the use of broad, unexplained citation) a second section of Dr. Bajaj’s declaration amounting to another approximately 1,400 words. Pet. 11 (“An overview of the ’980 patent, including a summary of the technology, claims, and prosecution history is provided in the Declaration of Dr. Bajaj. EX1004, Section VI.”). But this section includes arguments attempting to demonstrate that claimed subject matter (including the claimed “spatio-temporal model”) were “well known” in the art and to a POSITA before the ’980 patent’s priority date (*see* EX1004, ¶¶ 34-36), linking claim terms to specification disclosures (*see id.*, ¶¶ 38-42), and interpreting the prosecution history of the patent (*see id.*, ¶¶ 43-46).

The Petition similarly incorporates by reference (through broad, unexplained citation) portions of the Bajaj, Benner, and Mageski declarations that provide the substantive argument, explanation, and supporting evidence regarding the prior art status of Petitioners’ TLP and Zhao references. Pet. 5 (citing EX1004 (Bajaj), EX1013 (Benner), and EX1014 (Mageski)); *see* § VII.A.1, *infra*. Petitioners incorporate by reference their entire substantive argument regarding prior art status from these supporting declarations.

The Board has authority to deny a Petition for violations of its rules. *See Intervet Inc. v. Boehringer Ingelheim Vetmedica, Inc.*, IPR2018-01788, Paper 9, at 8-9 (P.T.A.B. Apr. 6, 2019) (“Here, we discern a violation of [37 C.F.R.

§ 42.22(a)(2) and § 42.6(a)(3)] that pervades every ground of unpatentability raised in the Petition, and warrants denial of review on that basis alone.”). Petitioner’s improper incorporations by reference here pervade every ground of the Petition, at least because each ground is based on Zhao (alone or in combination with another reference) or TLP, for which Petitioners incorporate by reference the entirety of their substantive argument and supporting evidence regarding prior art status. Pet. 6. Petitioners’ incorporations by reference also sweep in substantive arguments from the Bajaj declaration pertinent to each asserted Ground.

Alternatively, the Board can decline to consider testimony and evidence improperly incorporated by reference. *See, e.g.*, Patent Trial and Appeal Board Consolidated Trial Practice Guide, November 2019, at 35-36; *Cisco*, IPR2014-00454, Paper 12 at 9-10. This is so even when the consequence of non-consideration is that the petitioner’s argument or ground fails, resulting in non-institution. *See Clim-A-Tech Industries, Inc. v. William A. Ebert*, IPR2017-01863, Paper 13 at 27-28 (P.T.A.B. Feb. 12, 2018) (“Even if Petitioner’s expert declaration were to provide such particularity, arguments and information that are not presented and developed in the Petition, but instead are incorporated by reference, are not entitled to consideration, as it is improper to incorporate by reference arguments from one document into another document.”); *Icon Health & Fitness, Inc. v. Nautilus, Inc.*, IPR2017-00494, Paper 11, at 4 (P.T.A.B. Oct. 27, 2017) (“There is no question that

Mr. Ganaja’s testimony in these paragraphs addresses the issues that we found lacking in the Petition. The problem for Petitioner, however, is that the Petition itself did not include these arguments Citing to these paragraphs of Mr. Ganaja’s declaration without providing any explanation of the significance thereof in the Petition itself is insufficient. *See* 37 C.F.R. § 42.6(a)(3).”); *Unified Patents Inc. v. William Grecia*, IPR2016-00789, at 15-16 (P.T.A.B. Sept. 9, 2016) (“Such reliance on Mr. Cherukuri’s declaration testimony is an improper incorporation by reference of arguments asserted in a declaration, which we will not consider here. *See* 37 C.F.R. § 42.6(a)(3) Petitioner’s asserted ground fails for this reason.”); *Valeo North America*, IPR2014-01206, Paper 13, at 14; *Conopco, Inc. v. The Proctor & Gamble Co.*, IPR2013-00510, Paper 9, at 8-9 (P.T.A.B. Feb. 12, 2014).

Whether the Board considers Petitioners’ flagrant and unabashed incorporations by reference of multiple thousands of words a sufficiently egregious violation of its rules to warrant denial of the Petition, or instead declines to consider the improperly incorporated testimony and evidence and finds the Petition insufficient as a result, the outcome should be the same—the Petition should be denied. Either way, the basis for denial boils down to a violation of the Board’s rules, thus the Board can (and should) discretionarily deny the Petition under its § 314(a) authority.

**VII. PETITIONERS HAVE NOT DEMONSTRATED A REASONABLE
LIKELIHOOD OF PREVAILING ON ANY CHALLENGED CLAIM**

A. The Petition fails to identify, “with particularity,” arguments or evidence sufficient to establish a reasonable likelihood that TLP or Zhao was publicly accessible, disseminated, or otherwise qualifies as prior art.

**1. The Petition identifies insufficient evidence to meet
Petitioners’ burden.**

To demonstrate a reasonable likelihood of success, the Petition must “identify *with particularity* the grounds for institution and evidence supporting such grounds,” including “the prior art relied upon *and evidence that it qualifies as such.*” *Hulu, LLC, v. Sound View Innovations, LLC*, IPR2018-01039, Paper 29 at 12–13 (PTAB Dec. 20, 2019) (Precedential) (emphasis added). “This standard is far more than is required in typical notice pleading, which only require that a party make a plausible claim.” *Id.* “Thus, at the institution stage, the petition must identify, *with particularity*, evidence sufficient to establish a reasonable likelihood that the reference was publicly accessible before the critical date of the challenged patent and therefore that there is a reasonable likelihood that it qualifies as a printed publication.” *Id.* (emphasis added).

“Whether a reference qualifies as a printed publication under § 102 is a legal conclusion based on underlying fact findings,” including as to “public accessibility.” *Acceleration Bay, LLC v. Activision Blizzard, Inc.*, 908 F.3d 765, 772 (Fed. Cir. 2018) (citing cases). “Public accessibility” has been called “the touchstone in

determining whether a reference constitutes a ‘printed publication.’” *Id.* (quoting *Jazz Pharms., Inc. v. Amneal Pharms., LLC*, 895 F.3d 1347, 1356 (Fed. Cir. 2018) (cleaned up)). “A reference is considered publicly accessible if it was disseminated or otherwise made available to the extent that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, can locate it.” *Id.* (cleaned up). Petitioners bear the burden to prove that the asserted references qualify as printed publications. *Id.*

TLP. Petitioners fail to show that TLP was either publicly disseminated or publicly available at the relevant time. TLP itself bears no indicia of publication or dissemination. *See* EX1009. Petitioners identify no relevant dates or indicia of dissemination on TLP. *See* Pet. 12. Petitioners rely solely on purported evidence extrinsic to TLP, none of which is sufficient to satisfy the legal test.

Beyond its conclusory assertion of public accessibility, Petitioners cite two paragraphs of the Bajaj declaration and EX1010 without providing any explanation or discussion of them or the relevance of their contents. Pet. 6. Petitioners do not, for example, explain how anything in the Bajaj declaration or EX1010 demonstrates when or how TLP allegedly was “widely disseminated” or “publicly available.” *Id.* Conclusory assertions followed by bare citations to expert declarations amount to improper incorporation by reference. *Cisco*, IPR2014-00454, Paper 12 at 9 (“This practice of citing [a] [d]eclaration to support conclusory statements that are not

otherwise supported in the Petition ... amounts to incorporation by reference.”); *Valeo North America, Inc. v. Magna Electronics, Inc.*, IPR2014-01206, Paper 13, at 14 (P.T.A.B. Dec. 23, 2014) (“The Petition simply cites the lengthy explanation in the Declaration and does not otherwise provide an argument or explanation. That is insufficient.”). The Board should thus ignore Dr. Bajaj’s testimony. *See, e.g.*, Patent Trial and Appeal Board Consolidated Trial Practice Guide, November 2019, at 35-36; *Cisco*, IPR2014-00454, Paper 12 at 9-10.

A panel of the Board previously found that a conclusory public availability assertion coupled with a citation to a supporting declaration was insufficient to meet a petitioner’s burden in the context of a post-grant review. *See, e.g., Dynaenergetics US, Inc. v. Geodynamics, Inc.*, PGR2018-00065, Paper 8 at 14–19 (P.T.A.B. Nov. 15, 2018). Although the institution standard for post-grant review³ is different than that for IPR, the panel’s assessment in *Dynaenergetics* is apt: “We view Petitioner’s naked assertion that [the reference] is publicly available, and blanket citation to paragraphs 1 and 12 of the [supporting] declaration, as an invitation to be an archeologist of the record in search of pertinent facts to satisfy the criteria for public accessibility,” which “however, is not [the Board’s] role.” *Id.* at 18.

³ For a PGR, the standard for institution is whether it is “more likely than not” that the petitioner would prevail at trial. 35 U.S.C. § 324(a).

Conclusory assertions coupled with citations to supporting declarations is also insufficient for IPRs. *See, e.g., Metall Zug Ag v. Carl Zeiss Meditec AG*, IPR2020-00916, Paper 8 at 21-22 (P.T.A.B. Nov. 4, 2020) (“Such open-ended incorporation . . . fails to satisfy Petitioner’s burden of showing with particularity why the patent it challenges is unpatentable.”); *Floodbarrier, Inc. v. ILC Dover, LP*, IPR2019-01543, Paper 12 at 15-16 (P.T.A.B. Mar. 11, 2020) (incorporation by references “fails to satisfy the particularity requirement under 35 U.S.C. § 312(a)(3)); *Apple Inc. v. Contentguard Holdings, Inc.*, IPR2015-00442, Paper 9, at 9-10 (P.T.A.B. Jul. 13, 2015) (“[T]he Petition fails...to provide a sufficiently detailed explanation of the significance of the citations to these references and the Sherman Declaration—as required under 35 U.S.C. § 312(a)(3), 37 C.F.R. § 42.22(a)(2), and 37 C.F.R. § 42.104(b)(4) – (5).”); *ZTE (USA), Inc. v. Huawei Devices USA, Inc.*, IPR2015-00028, Paper 12, at 5-7 (P.T.A.B. Mar. 20, 2015) (“We agree with Patent Owner that ‘standing on its own, the Petition lacks any ‘detailed explanation of the significance of the evidence,’ in violation of 37 C.F.R. § 42.22(a)(2).’ Thus, we conclude that the Petition fails to meet the statutory requirement of 35 U.S.C. § 314(a) because the information presented in the Petition does not show that there is a reasonable likelihood that Petitioner would prevail with respect to at least one of the challenged claims.”); *DIRECTV, LLC v. Qurio Holdings, Inc.*, IPR2015-02007, Paper 6, at 14-15 (P.T.A.B. Apr. 4, 2016).

Even if the Board were to consider the Bajaj declaration, Dr. Bajaj’s testimony fails to show that TLP was disseminated or publicly available. Dr. Bajaj vaguely asserts that papers for the 2004 CVPR are done “in advance of the actual conference” and were due by “November 19, 2003.” EX1004, ¶66. Dr. Bajaj cites EX1010 in support of the alleged paper and manuscript submission dates but never explains what EX1010 is, its origin or source, its date, or whether it shows accurate information. *Id.* Critically, Dr. Bajaj asserts, without any factual support, that “[a]fter submission, the papers are widely disseminated and made publicly available to academics for peer review,” and that “TLP was widely disseminated and made publicly available through the submission process, and no later than April 9, 2004.” EX1004, ¶66. However, Dr. Bajaj’s unsupported assertion that TLP was “widely disseminated” and made “publicly available” through submission for peer review, if accepted, is insufficient to meet the legal standard.

The Federal Circuit has found that accessibility of a paper to a “peer-review committee” in connection with submitting the paper to a conference “suggest[s] an *absence* of actual public accessibility” and is insufficient to demonstrate public accessibility. *SRI Int’l, Inc. v. Internet Sec. Sys.*, 511 F.3d 1186, 1197 (Fed. Cir. 2008) (emphasis added). In *SRI*, the inventors drafted a paper called the “Live Traffic paper” and submitted it to the Symposium on Network and Distributed System Security (“SNDSS”) by placing it on an FTP server for retrieval. *Id.* at 1190.

The court found that “the pre-publication Live Traffic paper, though on the FTP server, was not catalogued or indexed in a meaningful way and not intended for dissemination *to the public*.” *Id.* at 1197 (emphasis added) (citing cases). The court found that “the peer-review feature also suggests no intent to publicize,” and that “without additional evidence as to the details of the [paper’s] accessibility,” the evidence was insufficient to demonstrate public accessibility. *Id.* at 1197-98. The court reiterated that the submission “does not evidence that the Live Traffic paper was accessible to anyone other than the peer-review committee, thus further suggesting an absence of actual public accessibility.” *Id.* at 1197; *see Constant v. Advanced Micro-Devices, Inc.*, 848 F.2d 1560, 1569 (Fed. Cir. 1988) (“[D]issemination and public accessibility are the keys to the legal determination whether a prior art reference was ‘published.’”).

Petitioners purported evidence of TLP’s pre-publication submission to CVPR for peer review in this case is similarly deficient. As in *SRI*, notwithstanding Dr. Bajaj’s assertion that TLP would have been subject to peer review, Petitioners and Dr. Bajaj fail to identify or analyze any evidence of TLP’s *public* accessibility or dissemination—the “keys” to determining whether TLP legally constitutes a printed publication. *See SRI*, 511 F.3d at 1197-98; *Constant*, 848 F.2d at 1569; *Acceleration Bay*, 908 F.3d at 772. Indeed, Petitioners and Dr. Bajaj offer no evidence that TLP was “catalogued or indexed in a meaningful way [or] intended for dissemination to

the *public*.” *Id.* (emphasis added). Dr. Bajaj’s conclusory assertion that TLP was “widely disseminated” and “publicly available to academics *for peer review*,” which lacks any explanation of how TLP was purportedly disseminated or accessible to the *public*, is therefore insufficient to show that TLP constitutes prior art. *Id.* (emphasis added).

Petitioners also assert that TLP “was disclosed during prosecution” and “is cited as prior art on the face of the ’980 Patent and its provisional application.” Pet. 6. This too is insufficient to show public accessibility of TLP. In *SRI*, the “Live Traffic paper” was cited in an information disclosure statement of the patents-in-suit during prosecution. *SRI*, 511 F.3d at 1190. The paper was nevertheless found not to qualify as prior art. *See id.* at 1194-98. Moreover, “[i]t is well settled ... that a patentee’s inclusion of a reference in an information disclosure statement does not constitute an admission that the reference is prior art.” *Stamps.com Inc. v. Endicia, Inc.*, 437 F. App’x 897 (Fed. Cir. 2011) (citing 37 C.F.R. § 1.97(h); *ResQNet.com, Inc. v. Lansa, Inc.*, 594 F.3d 860, 866 (Fed. Cir. 2010); *Abbott Labs. v. Baxter Pharm. Prods., Inc.*, 334 F.3d 1274, 1279 (Fed. Cir. 2003)). Here, the applicant’s IDS listing TLP includes the date of “2004” (no month or day), which does not demonstrate public availability before the ’980 patent’s undisputed priority date of May 27, 2004. *See* EX1002, 40. Petitioners have therefore failed to demonstrate TLP constitutes prior art.

Zhao. Although Petitioners assert that Zhao “is dated July 26, 2001,” their conclusory assertion that it “was publicly available no later than September 23, 2002” is not supported by particular evidence or argument. Petitioners’ cite to EX1013 for the similarly conclusory assertion that “the Zhao Dissertation would have been accessible to the public at the Engineering & Science Library by September 23, 2002.” *Id.* EX1013 is a declaration by Jessica Benner, but Petitioners do not identify or discuss any particular evidence from the Benner declaration that supports their assertions. Pet. 5. This is another improper incorporation by reference and fails to meet Petitioners’ burden of demonstrating public accessibility for purposes of institution. *See Cisco Systems*, IPR2014-00454, Paper 12 at 9; *Dynaenergetics*, PGR2018-00065, Paper 8 at 14–19; *Clim-A-Tech*, IPR2017-01863, Paper 13 at 27–28.

Petitioners’ additional assertion that “Zhao was also stored in the ProQuest repository and publicly available at least as early as June 25, 2002” is similarly deficient, as it is followed by only an unexplained citation to EX1014 (declaration of Carl Mageski)—another improper incorporation by reference—and fails to identify or discuss any particular evidence from the Mageski declaration that supports Petitioners’ assertions. *See Cisco Systems*, IPR2014-00454, Paper 12 at 9; *Dynaenergetics*, PGR2018-00065, Paper 8 at 14–19; *Clim-A-Tech*, IPR2017-01863, Paper 13 at 27–28.

2. Petitioners should not be permitted a reply or to submit supplemental information to address prior art or printed publication status.

“The rules of *inter partes* review do not automatically authorize a petitioner to file a reply to a patent owner preliminary response; instead, a petitioner seeking to file a reply prior to institution must make a showing of “good cause.” *See* 37 C.F.R. § 42.108(c) (2019).” *Victoria’s Secret Stores, LLC v. Andra Group, LP*, IPR2020-00853, Paper 11, at 2 (P.T.A.B. Aug. 13, 2020) (finding no good cause). “When determining whether a petitioner has shown good cause for a reply to a preliminary response, the Board may assess whether the argument at issue in the preliminary response was reasonably foreseeable.” *Id.* at 3-4 (citing cases). To the extent appropriate and permitted, “submission of supplemental information or evidence occurs after institution of trial,” not before. *See RayVio Corp. v. Nitride Semiconductors Co., Ltd. et al*, IPR2018-01139, Paper 14, at 8-9 (P.T.A.B. Oct. 31, 2018) (citing 37 C.F.R. §§ 42.64(b)(2), 42.123). In any event, “[t]he provision for submitting supplemental information under Rule 42.123(a) ‘is not intended to provide a petitioner an advantageous “wait-and-see” opportunity to use a patent owner’s preliminary response ... in order to refine or bolster petitioner’s position.’” *Artesian Home Products et al v. Gutterglove, Inc.*, IPR2018-00031, Paper 24, at 3 (P.T.A.B. May 31, 2018) (citing cases).

No good cause exists to permit Petitioners a reply to this Preliminary Response to address or cure issues of public accessibility and prior art status. That Patent Owner would challenge the sufficiency of Petitioners' evidence of public accessibility is entirely foreseeable, which demonstrates a lack of good cause. *Victoria's Secret Stores*, IPR2020-00853, Paper 11, at 7; *BMW of N. Am. LLC v. Intellectual Ventures*, IPR2017-01558, Paper 8 at 2–3 (PTAB Oct. 25, 2017). Petitioners should not be permitted to bolster their Petition through a “wait-and-see” reply when it was their burden to provide their arguments and evidence in the Petition, *Artesian Home Products*, IPR2018-00031, Paper 24, at 3, especially in light of Petitioners' significant violations of the word count limits through incorporation by reference, *see* §§ VI.A and VII.A.1, *supra*, and because the § 315(b) bar date has passed, *Mylan Pharmaceuticals Inc. v. Boehringer Ingelheim Pharmaceuticals Inc.*, IPR2016-01565, Paper 15, at 2-3 (P.T.A.B. Dec. 7, 2016); *Teva Pharmaceuticals USA, Inc., v. Indivior UK Limited*, IPR2016-00280, Paper 21, at 4 (P.T.A.B. May 27, 2016).

The Board is capable of assessing Patent Owner's arguments—Petitioners need not be heard again on an issue for which they bear the burden. *See, e.g., Unified Patents, LLC f/k/a Unified Patents Inc. v. Velos Media, LLC*, IPR2019-01663 (PTAB) (denying request to file a reply to preliminary response to address case law on public accessibility); *RayVio Corp. v. Nitride Semiconductors Co., Ltd.*,

IPR2018-01139 (PTAB) (denying request for reply to “direct [the Board] to portions of the record that support its position, and, separately to provide supplemental information to ‘clarify’ that [a reference] is a publication”); *Franklin Electric Co., Inc. et al. v. Liberty Pumps, Inc.*, IPR2017-00112 (PTAB) (finding no good cause for a reply); *Xactware Solutions, Inc. et al. v. Pictometry International Corp.*, IPR2016-00593 (PTAB) (“To the extent that Petitioner seeks to identify new citations or bolster arguments in the Petition, we are not required to allow Petitioner to do so.”).

B. The Petition fails to demonstrate that Zhao (Ground I) discloses or renders obvious every limitation of claims 1 and 9.

1. 1/9[a]: “providing a spatio-temporal model for an object to be detected”

Petitioners fail to demonstrate that Zhao discloses or renders obvious “providing a *spatio-temporal model* for an object to be detected.” (emphasis added). Petitioners assert that “Zhao uses an object model that takes the form of a ‘human model.’” Pet. 13-14. Petitioners argue that this “human object model” of Zhao discloses or renders obvious the claimed “spatio-temporal model” because it “includes spatial and temporal elements.” Pet. 14. Petitioners’ argument fails for multiple reasons.

a. Petitioners fail to show that Zhao discloses or renders obvious a “spatio-temporal model.”

Zhao’s human model does not disclose or render obvious a “spatio-temporal model.” Petitioners assert that Zhao’s model includes a “spatial” element in the form of “model matrices.” Pet. 14-15. According to Petitioners, Zhao’s model uses these matrices to encode “spatial relationships between body parts.” Pet. 15. Zhao’s model matrices are those of its “TRS-Invariant Body Model.” *See* EX1006, 42-43 (section titled “TRS-Invariant Body Model” describing Zhao’s four model matrices), 23 (“In this dissertation, I develop a TRS-invariant representation of the shapes of the body parts and their size and *spatial relationships*.”), 69 (“A TRS-invariant probabilistic model is designed to encode the shapes of the body parts and the context information — the size and *spatial relationships* between body parts.”).

Zhao’s “spatial” human body model, however, is not spatio-*temporal*. Zhao’s model includes no temporal, variables, labels, or other elements that model the human body (or its parts) over time. All of Zhao’s body part modeling variables— w (part width), l (part length), a (aspect ratio = w/l), x (part’s original coordinate position along x direction), y (part’s original coordinate position in y direction), and θ (intersection angle between major axes of a part and its parent part)—encode spatial or kinematic information; and all of the variable subscripts— i and j —differentiate among discrete body parts in the model. EX1006, 42-43. In other

words, Zhao’s model is only a spatial model—it has no temporal dimension and does not model the human body (or its parts) over time. Unlike the spatio-temporal model of the ’980 patent, which is “extended over time”—thus it may include nodes that “represent[] either the object or a component of the object *at time t*” over a “temporal window” and account for the “temporal compatibility of object state between frames *i* and *j*” (EX1001, 3:47-50, 4:34-47)—Zhao does not disclose a spatio-temporal model that models the human body (or its part) over any length of time.

Petitioners assert that Zhao’s use of “motion information obtained from previous frames . . . to predict the orientation . . . and to solve the ambiguity” of overlapping limbs demonstrates that Zhao’s model “incorporates temporal elements for an object to be detected” and therefore discloses or renders obvious the claimed “spatio-temporal model for an object to be detected.” *See* Pet. 15-17. But neither Petitioners nor Zhao explains what the “motion information” is or how it is allegedly incorporated into Zhao’s model. *See* Pet. 15-17; EX1006, 91-94. As is evident from Zhao’s modeling matrices discussed above, Zhao’s spatial model encodes only spatial or kinematic information and does not model a human body (or body part(s)) over time. Accordingly, even if Zhao’s spatial model incorporates the unspecified “motion information” Petitioners identify to “to resolve the ambiguity with limbs’ orientations” in space (i.e., to disambiguate left limbs from right limbs in space that “tend to be confused in a side-view”), Petitioners fail to show how the use of such

“motion information” in this way discloses anything more than the use of just another spatial or kinematic input into Zhao’s explicitly spatial model. *Id.*; see EX1002, 163 (distinguishing a “kinematic model (spatial model)” from the spatio-temporal model of the ’980 patent).

Petitioners alternatively assert that Zhao “at least rendered [the spatio-temporal model] obvious,” but fail to support that assertion. See Pet. 15, 17. Petitioners assert that “it would have been *known to a POSITA* to incorporate temporal elements into an object model . . . by incorporating information measured from multiple image frames,” citing examples of ““motion-based’ recognition,” “temporal propagation,” and “temporal extrapolation” purportedly described in non-asserted references. Pet. 15-16 (emphasis added). But Petitioners do not explain how any of those techniques purportedly disclose the claimed “spatio-temporal model” under their own apparent interpretation (i.e., “incorporate[ing] temporal elements into an object model”), let alone *why* a POSITA would modify Zhao to adopt those techniques instead of Zhao’s explicitly spatial model. *Arendi S.A.R.L. v. Apple Inc.*, 832 F.3d 1355, 1361 (Fed. Cir. 2016), *cert. denied*, 137 S. Ct. 1329 (2017) (“[A] patent can be obvious in light of a single prior art reference if it would have been *obvious to modify that reference* to arrive at the patented invention.” (emphasis added)).

In fact, Zhao teaches away from the use of “motion-based” detection techniques, stating that “there are several limitations with these schemes,” including: (1) “the human’s feet or legs should be visible in order to extract cyclic features,” (2) “the recognition procedure requires a sequence of images, which delays the identification until several frames later and increases the processing time,” and (3) “the procedure cannot detect stationary humans and humans performing unconstrained and complex movement such as wondering [*sic*] around, turning, jumping, etc.” EX1006, 82. Zhao thus discourages the use of “motion-based” detection techniques that require a sequence of images, while leading the reader toward its TRS-invariant model, which, as explained above, is not spatio-temporal. *See AstraZeneca AB v. Mylan Pharm. Inc.*, 19 F.4th 1325, 1337 (Fed. Cir. 2021) (“[A] prior art reference is said to teach away from the claimed invention if a skilled artisan upon reading the reference, would be discouraged from following the path set out in the reference, or would be led in a direction divergent from the path that was taken in the claim.” (cleaned up)). There is no evidence to suggest (and Petitioners identify none) that Zhao’s use of “motion information” for disambiguating already-detected limbs teaches incorporating “motion based” methods of “human detection,” which Zhao teaches away from in favor of its TRS-invariant technique. *See* EX1006, 82, 91-92. As explained above, the Petition (and Zhao) do not explain what “motion information” is or how it is incorporated into

Zhao's TRS-invariant model. Zhao's teaching away from the use of "motion-based" detection thus undercuts Petitioners vague and shallow obviousness argument based on what a POSITA would have purportedly "known" about motion-based recognition. *Id.* ("[A] reference that properly teaches away can preclude a determination that the reference renders a claim obvious." (cleaned up)).

Even if Petitioners had shown that a POSITA knew of the claimed spatio-temporal model, "a patent composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art." *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007). Instead, there must have been "an apparent reason to combine the known elements in the fashion claimed by the patent at issue." *Id.* at 417-18. Thus, even though Petitioners' obviousness argument is based on a single reference (i.e., Zhao), Petitioners were still required to show that "it would have been obvious to modify that reference to arrive at the patented invention." *Arendi S.A.R.L.*, 832 F.3d at 1361 (Fed. Cir. 2016); *SIBIA Neurosciences, Inc. v. Cadus Pharm. Corp.*, 225 F.3d 1349, 1356 (Fed. Cir. 2000).

Petitioners' assertion that "it would have been known to a POSITA to incorporate temporal elements into an object model for an object to be detected by incorporating information measured from multiple image frames" at most suggests that a POSITA *could* have modified Zhao to arrive at the claimed invention but does

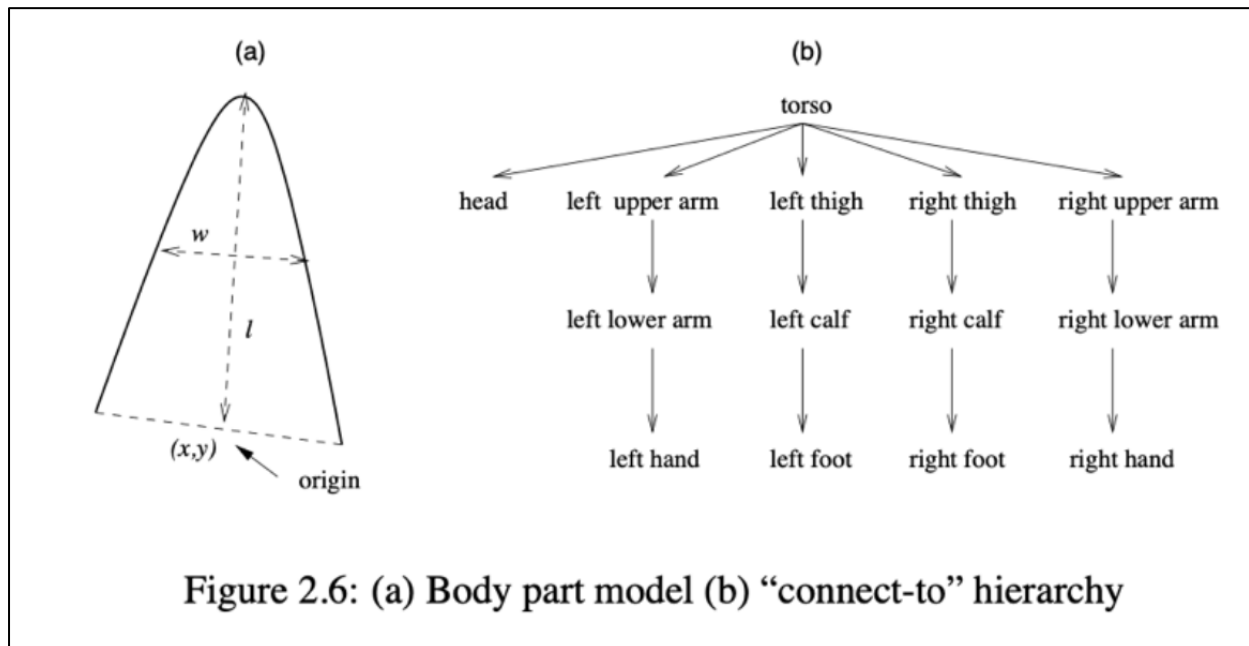
not explain why a POSITA *would have been motivated* to do so, especially when Petitioners seem to assert that Zhao alone discloses it. *See* Pet. 15-21. *See Belden Inc. v. Berk-Tek LLC*, 805 F.3d 1064, 1073 (Fed. Cir. 2015) (“[O]bviousness concerns whether a skilled artisan not only *could have made* but *would have been motivated to make* the combinations or modifications of prior art to arrive at the claimed invention.”); *InTouch Techs., Inc. v. VGo Communs., Inc.*, 751 F.3d 1327, 1352 (Fed. Cir. 2014). Dr. Bajaj’s testimony similarly misses the mark, suggesting at most that a POSITA knew how to “incorporate temporal elements into an object model” and that spatio-temporal models are “not new,” but failing to explain why a POSITA would have been motivated to modify Zhao to arrive at the claimed invention, particularly in light of Zhao’s teaching away of the use of motion-based techniques. *See* EX1004, ¶¶ 78-79. Arguments and testimony so lacking amount to hindsight reconstruction of the claimed invention and are insufficient to support a conclusion of obviousness. *See TQ Delta, LLC v. Cisco Sys.*, 942 F.3d 1352, 1359-63 (Fed. Cir. 2019) (analyzing cases and reversing Board’s obviousness finding).

b. Petitioners’ arguments under their alternative interpretation of “spatio-temporal model” also fail.

Petitioners present additional arguments under an alternative interpretation of “spatio-temporal model”—“a graphical model comprising nodes corresponding to each of the collection of components and to the object”—that they neither advocate

for nor support with a claim construction analysis or intrinsic evidence support. *See* Pet. 12, 18-21. Nevertheless, Petitioners’ arguments fail under their alternative construction.

Petitioners first argue that that Zhao’s “ribbons” shown in Fig. 2.6(a) “are graphical representations of body parts and are ‘nodes’ corresponding to each of the collection of components and the object.” Pet. 45-46.



EX1006, 42

Petitioners fail to identify what alleged “nodes” in Zhao correspond to the “components” and what node corresponds to the “object” as claimed. Petitioners quote Zhao as stating that “[e]ach human body model consists of six *main body parts* (the head, the torso, two arms and two legs) and twelve *subparts* of the four limbs,” and assert that Zhao’s model is “comprised of ‘ribbons’ *for these body parts*.” Pet.

18 (emphasis added). (quoting EX1006, 42). After enumerating the “main body parts” and “subparts,” Zhao also states that “[t]he body *parts* are modeled with ribbons.” EX1006, 42 (emphasis added). Thus, both Zhao and Petitioners treat Zhao’s “main body parts” and “subparts” as a collective class of body “parts” modeled with “ribbons.” This is confirmed by Petitioners’ assertion that “[t]he ribbons are defined as graphical representations of a human’s body *parts*” without differentiating between “main body parts” and “subparts.” Pet. 19 (emphasis added).

But despite identifying a single class of ribbons for Zhao’s collective body “parts,” Petitioners inexplicably conclude without citation or explanation that the body part ribbons “act as ‘nodes’ corresponding to each of the collection of *components* and to the *object*.” Pet. 19. That is, Petitioners identify one class of ribbons—i.e., for body “parts”—but conclude without explanation that they disclose two classes of nodes—i.e., “object” and “component.” Petitioners fail to identify which alleged “nodes” in Zhao (if any) are “object” nodes and which are “component” nodes. Petitioners therefore fail to show that Zhao’s body “part” ribbons disclose a “spatio-temporal model” under its alternative construction.

Petitioners next argue that Zhao’s “merged body parts are also modeled with ribbons,” and that Zhao’s model “includes ‘levels between the whole body [*i.e.*, the object] and the major body parts to include the merged body parts [*i.e.*, the collection of components].” Pet. 19-20. Thus, Petitioners also argue that “the whole body”

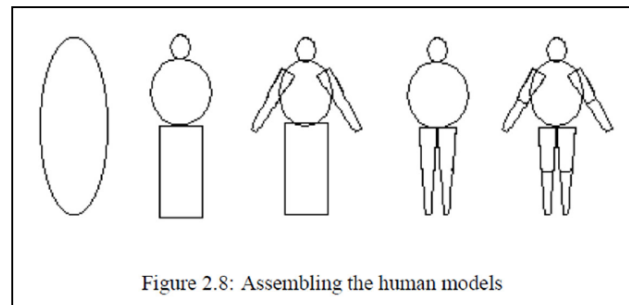
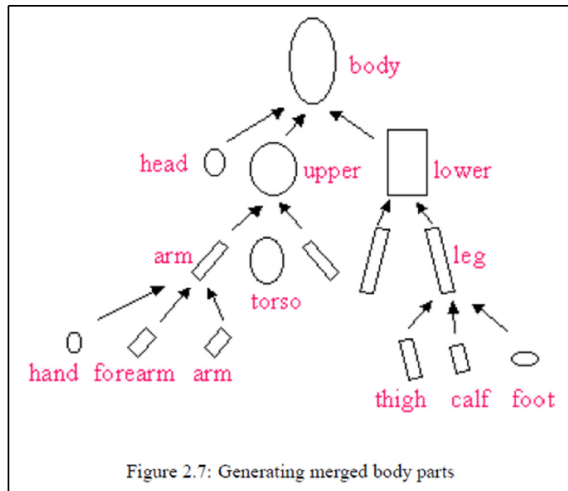
corresponds to the “object” and that “the merged body parts” correspond to “the collection of components.” *Id.* But then Petitioners switch to mapping the “object” to Zhao’s “merged parts,” and the “collection of components” to Zhao’s “component body parts.” Pet. 21. Accordingly, as summarized in the table below, Petitioners’ “merged body parts” argument apparently asserts that the “whole body” or other lower level “merged body parts” constitute a “node” corresponding to the “object.”

ribbon	“object”	“collection of components”
“whole body” (merged)	X	
“merged body parts”	X	X
“component body parts”		X

Table 1: Petitioners mappings of Zhao’s “ribbons” to nodes corresponding to the “object” and the “collection of components.”

Both arguments fail, however, because Petitioners fail to explain how a given human model in Zhao includes “ribbons” (the alleged “nodes”) for both a merged part (the alleged “object”) *and* for its component parts (the alleged “collection of components”). Indeed, Petitioners’ alternative interpretation of “spatio-temporal model” is “a graphical model comprising nodes corresponding to each of the collection of components *and* to the object.” *See* Pet. 12 (emphasis added). But, as shown in Zhao’s Figs. 2.7 and 2.8 cited by Petitioners, merged body parts (i.e., “body,” “upper,” “lower,” “arm,” and “leg”) are represented in the assembled human

model by only one ribbon (alleged “node”) corresponding to the merged part—ribbons for the merged part’s component parts are not *also* present.



In other words, as shown in Zhao’s Fig. 2.8, a human model that includes a ribbon for a merged component (e.g., the whole “body,” which is formed by merging the components shown in Fig. 2.7) does not also include ribbons for the merged component’s constituent parts (and the same is true for other merged parts in other segmented models shown in Fig. 2.8). Accordingly, Petitioners fail to show that Zhao discloses a model that includes both a ribbon for a merged part (alleged “object” node) *and* ribbons for the merged part’s component parts (alleged “component” nodes).

2. 1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”

Petitioners fail to show that Zhao discloses this claim element.

Element 1/9[e] describes a particular step for “detecting the object in any image”—i.e., detecting the object in any singular image. This step of detecting the object in a singular image recites “comparing the *probabilities* for *each image* to a threshold for detecting the object.” Here, “probabilities” is plural and refers to the probabilities for “each image.” And “each image” refers to the “plurality of images including the object,” where “in each image” the object is measured and “a probability that the object is in the image” is determined. In other words, the claims describe providing “a plurality of images” and, in each image, “measuring the object” and “determining a probability that the object is in the image,” i.e., in each image. Accordingly, “detecting the object” in any singular image upon “*comparing the probabilities for each image to a threshold* for detecting the object” requires comparing multiple probabilities to a threshold.

This understanding is consistent with the '980 patent's specification. For example, the specification repeatedly and consistently describes a method that comprises “providing a video comprising *a plurality of images* including the object,” “measuring the object as a collection of components in *each image*” (i.e., of the plurality), “determining a probability that the object is in *each image*” (i.e., of the plurality), and “detecting the object in any image upon comparing the probabilities for *each image* [i.e., of the plurality] to a threshold for detecting the object.” EX1001, 1:28-36, 2:11-18.

The patent teaches that the plurality of images may be the images for which the spatio-temporal model is defined. For example, the patent teaches that “it is possible to perform inference over the spatio-temporal model defined for the entire image sequence.” EX1001, 7:6-10. The patent also teaches, however, that “there are many application[s] for which this is not an option due to the lengthy off-line processing needed.” *Id.* The patent thus teaches that “[a] w-frame windowed smoothing algorithm” may be used “where w is an odd integer ≥ 1 .” *Id.* For example, the patent describes an example using a “3-frame spatio-temporal object model” on a sequence of 55 consecutive frames.” *Id.* 7:56-8:7. The patent further teaches that the temporal window for the spatio-temporal model of an object at time t can encompass earlier and later times. *See id.*, 4:34-48, Figs. 2A-2B.

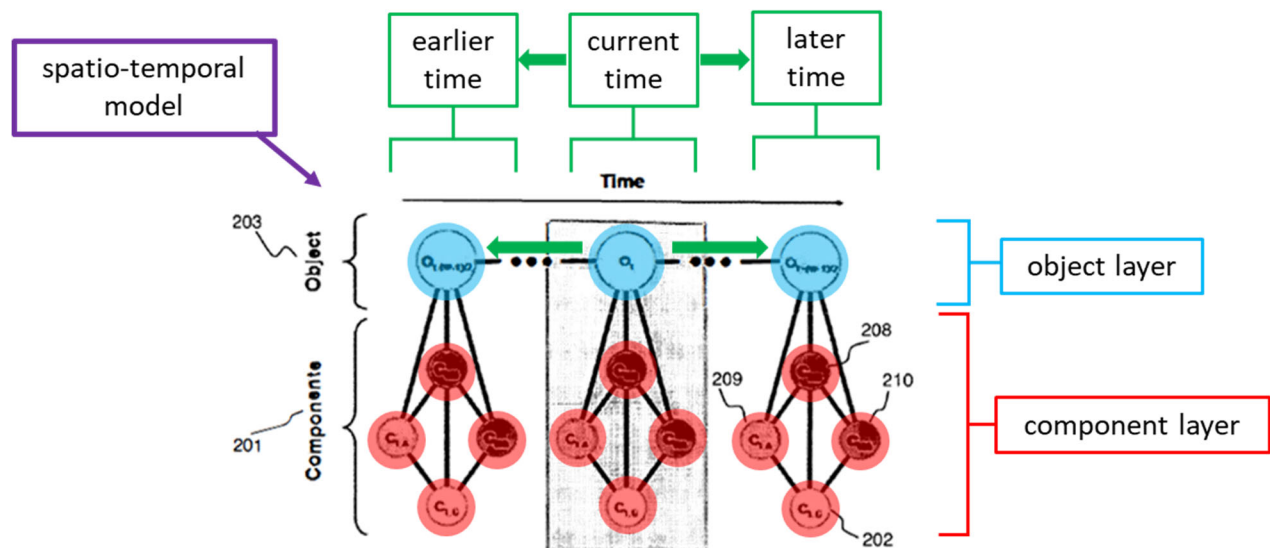


FIGURE 2A

Petitioners fail to show that Zhao discloses the particular step described in element 1/9[e] of “detecting the object *in any image*,” i.e., in a singular image, “upon comparing *the probabilities for each image*,” i.e., for each image in the plurality, “to a threshold for detecting the object.” Petitioners assert that “[a]fter determining the probability that the object is in each image, Zhao compares the probability to a threshold to detect the object (*i.e.*, a human).” Pet. 26. But Petitioners fail to show that Zhao discloses detecting the object in any such image upon comparing “the probabilities for each image” of the plurality to a threshold rather than upon comparing the probability of a single image to a threshold. This fundamental failing of Zhao is further evidenced by its teaching away of the use of “motion-based” detection using “a sequence of images” (which would include a plurality of images). *See* § VII.B.1, *supra*.

Petitioners assert that Zhao discloses a decision rule that “compares the probability of an extracted contour matching a person with a threshold” (i.e., “ $P(\text{person}|C) > \text{threshold}$ ”) and reformulates the rule using Bayes formula (i.e., “ $P(C|H^*, \text{person})P(\text{person}|H^*) > \text{threshold}$ ”). Pet. 27. However, Petitioners fail to show that Zhao discloses “detecting the object *in any image*” i.e., in a singular image, “upon comparing *the probabilities for each image*,” i.e., each image in the plurality,” to a threshold using Zhao’s decision rule. Indeed, Zhao discloses no such thing. At most, Zhao purportedly teaches detecting a human (i.e., the alleged “object”) in a

given image upon comparing the probability that a contour corresponds to a person *for just the given image* to a threshold. *See* EX1006, 64-65; Pet. 26-27. Indeed, Zhao poses the “human detection problem” in the context of evaluating a single contour (i.e., a contour detected in an image): “The human detection problem is: given a detected contour C, determine whether this contour represents the silhouette of a person or not.” EX1006, 64 (emphasis added), *see also id.* 54 (“A part-based similarity measure evaluates the resemblance between a contour and a model based on the best match between their body parts.” (emphasis added)). Petitioners do not cite any portion of Zhao that discloses detecting a human (the alleged “object”) in any singular image upon comparing the probabilities “for each image” in the plurality to a threshold. *See* Pet. 26-27. Zhao rejected the concept of using a plurality of images over time as a disfavored “motion-based” detection scheme. *See* § VII.B.1, *supra.*

C. The Petition fails to demonstrate that Steffens in view of Zhao (Ground II) renders obvious every limitation of claims 1 and 9.

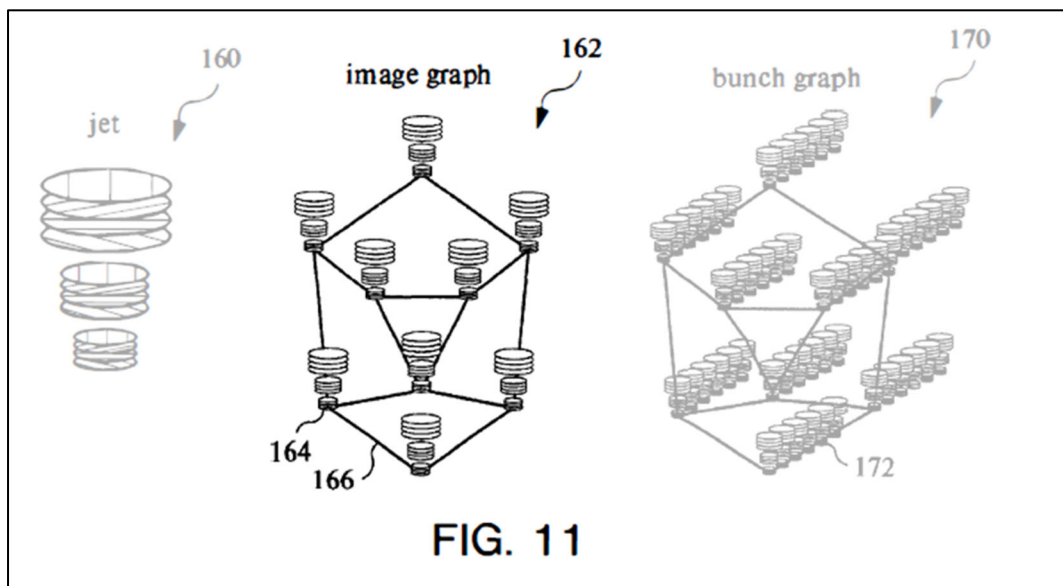
1. 1/9[a]: “providing a spatio-temporal model for an object to be detected”

Petitioners fail to demonstrate a reasonable likelihood that Steffens in view of Zhao discloses or renders obvious “providing a *spatio-temporal model* for an object to be *detected*.” (emphasis added). Petitioners assert that “Steffens provides a model for an object to be detected, which is a ‘spatio-temporal model’” where the “spatio”

element corresponds to Steffens' "labeled image graph" described at 8:58-62 and Figure 11 of Steffens. Pet. 38. Petitioners assert that the "temporal" element corresponds to Steffens' description of "*tracking*" objects and nodes "between images" at 2:6-13 and claims 14, 16, and 17 of Steffens. Pet. 39 (emphasis added). Petitioners argue that these portions of Steffens disclose "the 'spatio' aspect of the spatio-temporal model" and "the use of 'temporal' information, including for example the location of an object and its nodes between image frames." Pet. 39. Petitioners' argument fails for multiple reasons.

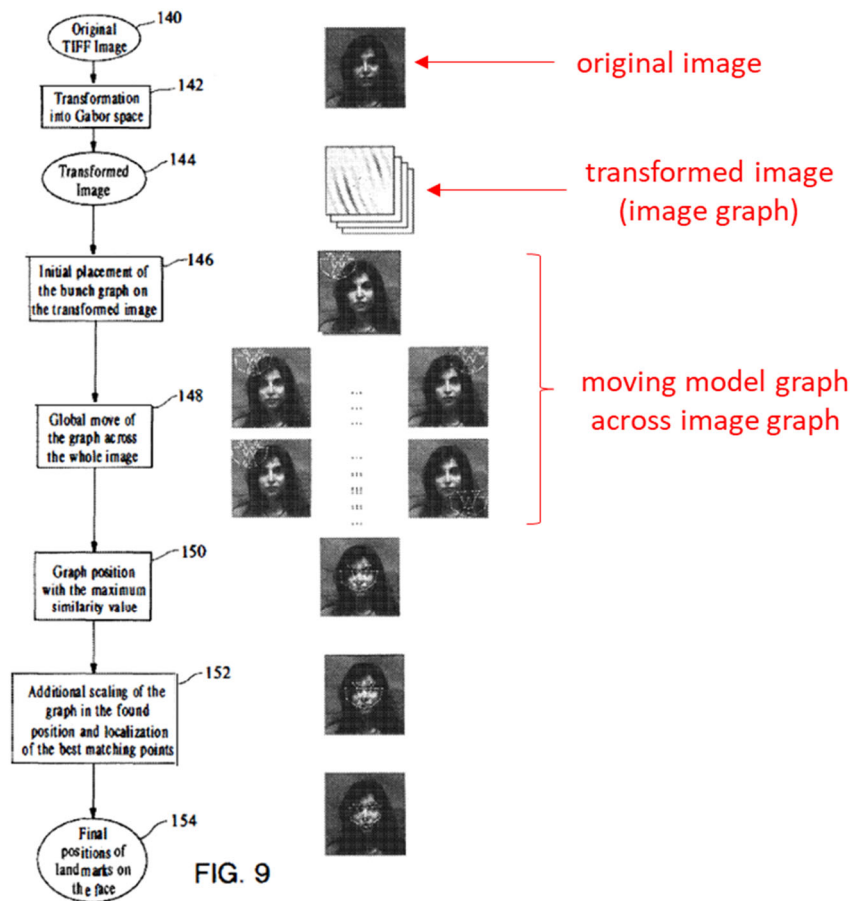
a. Petitioners fail to show that Steffens discloses or renders obvious a "*spatio-temporal model for an object to be detected.*"

Petitioners rely on Steffens' "labeled graph" or "labeled image graph" as containing the alleged "spatio" element of Steffens' model. Pet. 38.



EX1007, Fig. 11 (annotated)

Steffens' labeled image graph is used in Steffens' "graph matching" process for locating "facial landmarks and features of the head" in a given image. EX1007 7:51-53. In this process, "*a captured image . . . is transformed . . . using a wavelet transformation,*" and "a model graph . . . is positioned over *the transformed image* at varying node positions to locate a position of optimum similarity." *Id.* 7:51-63. This is performed "by positioning the model graph in the upper left hand corner of the image, extracting the jets at the nodes, and determining the similarity between *the image graph* and the model graph." *Id.* 7:64-67. Steffens' "image graph" thus corresponds to a single "transformed image." *See id.*; *see also id.* at 8:10-13 ("Note that although the graphs are shown in Fig. 9 with respect to the original image, the model graph movements and matching are actually performed *on the transformed image.*").



Steffens’ “image graph”—which contains the alleged “spatial elements”—is not a spatio-temporal model; it does not model an object over time but instead is a single image transformed into a graph. *See* EX1007, 7:51-8:13. Indeed, when there is “a sequence of images that contain an object of interest,” Steffens “applies graph matching in order to evaluate *each* image,” which, according to the process summarized above, generates a different image graph for each image—i.e., a different alleged “model where ‘spatial’ elements are used.” EX1007, 7:9-18; Pet. 38.

Unlike the spatio-temporal model for object detection of the '980 patent, which is “extended over time”—thus it may include nodes that “represent[] either the object or a component of the object *at time t*” over a “temporal window” and account for the “temporal compatibility of object state between frames *i* and *j*” (EX1001, 3:47-50, 4:34-47)—Steffens’ image graph does not model an object over any length of time. Indeed, Steffens’ image graph reflects spatial relationships between nodes *within* an image frame (not *between* frames). *See* EX1007, 7:51-8:13.

Petitioners assert that Steffens’ “tracking” of objects and nodes “between frames” and “reinitializing a tracked node if the node’s position deviates beyond a predetermined position constraint between image frames” describes “a model for an object to be detected that includes the use of ‘temporal’ information, including, for example, the location of an object and its nodes between image frames.” Pet. 39. In making this assertion, Petitioners rely solely on vague, unspecific characterizations from Steffens’ claims and invention summary section (which tracks the claim language), ignoring the 16 columns (and the vast majority of Steffens’ specification) between them that provide the context for understanding them. *See* Pet. 39 (citing EX1007, 2:6-13 and claims 14, 16, and 17); *see In re Warsaw Orthopedic, Inc.*, 832 F.3d 1327, 1332 (Fed. Cir. 2016) (“When the PTAB examines the scope and content of prior art . . . it must consider the prior art ‘in its entirety, i.e., as a *whole*.’”). The '980 patent explains that “tracking” is a distinct process from “detection” in that

“[t]racking makes use of object detection for initialization and re-initialization during transient failures for occlusions,” whereas “[o]bject detection considers the consistency of the detection over time.” EX1001, 2:47-50. Steffens likewise distinguishes “tracking” from the “finding,” “detecting,” or “locat[ing]” of “landmarks” and facial features. *See* EX1007, 3:54-4:19 (“Once facial landmarks have been found on the facial image, a landmark tracking process 20 may be used to track of the landmarks.”). Steffens performs the landmark finding process using the “image graph,” which, as explained above, is not a spatio-temporal model. EX1007, 7:51-53. After the landmark finding process, Steffens performs landmark “tracking” using a separate “tracking technique.” EX1007, 10:46-61. While the separate tracking process may make use of position information generated from the “landmark finding method,” *id.*, Steffens’ “image graph” used to detect or locate features is nevertheless not a spatio-temporal model (for the reasons explained above), and the landmark “tracking” process does not provide a model “for an object to be *detected*” as recited in the ’980 independent claims.

Accordingly, Steffens does not disclose the claimed “spatio-temporal model.” Petitioners alternatively assert that Steffens renders it obvious. *See* Pet. 38-39. However, Petitioners fail to show that “it would have been obvious to modify [Steffens] to arrive at the patented invention.” *Arendi S.A.R.L.*, 832 F.3d at 1361 (citing *SIBIA Neurosciences, Inc.*, 225 F.3d at 1356).

b. Petitioners’ arguments under their alternative interpretation of “spatio-temporal model” also fail.

Petitioners assert that Steffens’ “bunch graph” and “model graph” disclose a “spatio-temporal model” under Petitioners’ alternative interpretation of “graphical model comprising nodes corresponding to each of the collection of components and to the object.” Pet. 40-42. This argument fails for multiple reasons.

As an initial matter, Petitioners fail to identify which nodes in either type of graph correspond to an “object” and which correspond to the “components” of the object. Petitioners leave Patent Owner and the Board to guess, but it’s not clear.

For the bunch graph, Petitioners assert that “‘nodes’ [] refer to points on the object,” and have “locations at defined position of the head.” Petitioners appear to equate the “object” with the “head,” but Petitioners fail to identify a node for the head, and thus for “the object.” *See* Pet. 40.

For the model graph, Petitioners assert that Steffens’ “coarse-to-fine” approach discloses a model “comprised of ‘nodes’ corresponding to each of the collection of components and to the object,” using the example of representing an eye with a “single node” (coarse) or “more nodes” (fine). Pet. 41-42. However, when the eye is represented as one node, the model lacks either the object node or the component nodes (because there’s only one node). And when it’s represented as multiple nodes, Petitioners identify nodes corresponding to eye *components* (i.e.,

outlining the eye and at the eye pupil) but no additional node that represents the *object* itself (i.e., the eye). *See id.*

2. 1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”

a. Petitioners fail to show that Steffens or Zhao discloses the limitations of element 1[e].

Petitioners do not assert that Steffens discloses this claim element. *See* Pet. 45. Instead, Petitioners rely entirely on their arguments that Zhao discloses this claim element to show that it was allegedly disclosed by the asserted combination of Steffens and Zhao. *See id.* For the reasons explained above, however, Petitioners fail to show that Zhao discloses or renders obvious this claim element. *See* § VII.B.2, *supra*. Accordingly, Petitioners fail to show that the combination of Steffens and Zhao discloses this claim element or that the combination of Steffens and Zhao discloses every element in the independent claims. Petitioners Ground II therefore fails, *Univ. of Strathclyde v. Clear-Vu Lighting LLC*, 17 F.4th 155, 160 (Fed. Cir. 2021) (“An obviousness determination generally requires a finding that ‘all claimed limitations are disclosed in the prior art.’” (quoting *PAR Pharm., Inc. v. TWI Pharms., Inc.*, 773 F.3d 1186, 1194 (Fed. Cir. 2014))), and the Board need not address motivation to combine or reasonable expectation of success, *PAR Pharm.*, 773 F.3d at 1194 (citing *Medichem, S.A. v. Rolabo, S.L.*, 437 F.3d 1157, 1164 (Fed. Cir. 2006) for the proposition that “we consider motivation to combine and

reasonable expectation of success only ‘if all the elements of an invention are found in a combination of prior art references’’).

b. Petitioners fail to sufficiently demonstrate a motivation to combine and reasonable expectation of success of combining Steffens and Zhao for element 1[e].

To the extent the Board considers motivation to combine and reasonable expectation of success, Petitioners fail to satisfy their burden.

It is Petitioners’ “burden to demonstrate both ‘that a skilled artisan would have been motivated to combine the teachings of the prior art references to achieve the claimed invention, and that the skilled artisan would have had a reasonable expectation of success in doing so.’” *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359,1367-68 (Fed. Cir. 2016) (citation omitted). Evidence regarding a motivation to combine is insufficient, however, if it “fails to explain why a person of ordinary skill in the art would have combined elements from specific references *in the way the claimed invention does.*” *ActiveVideo Networks, Inc. v. Verizon Communs., Inc.*, 694 F.3d 1312, 1328 (Fed. Cir. 2012). And “[t]he reasonable-expectation-of-success analysis must be tied to the scope of the claimed invention” and must “prove a reasonable expectation of success in achieving *the specific invention claimed.*” *Teva Pharm. USA, Inc. v. Corcept Therapeutics, Inc.*, 18 F.4th 1377, 1381 (Fed. Cir. 2021).

i. Motivation to Combine

Petitioners fail to explain why it would have been allegedly obvious to combine elements from Steffens and Zhao “*in the way the claimed invention does*” i.e., to specifically include the limitations of element 1[e]. *ActiveVideo Networks*, 694 F.3d at 1328. Petitioners first argument that it would have been obvious to “use a probabilistic framework” for the reasons they provided for claim element 1/9[d] is insufficient because Petitioners’ 1/9[d] arguments (including their section on “Motivation to Combine”) do not address specific prior art elements that allegedly disclose “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object.” *See* Pet. 44-45, § X.B.1 (Motivation to combine); *ActiveVideo Networks*, 694 F.3d at 1328 (finding testimony that is “generic and bears no relation to any specific combination of prior art elements” insufficient).

Petitioners’ next argument that that it would have been obvious “to detect the image by comparing the probabilities for each image to a threshold for detecting the object” because “the use of a ‘threshold’ was widely known” also fails because “a patent composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art.” *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007). Instead, a finding of obviousness “requires ‘identify[ing] a reason that would have prompted a person of ordinary skill

in the relevant field to combine the elements in the way the claimed new invention does.”” *Adapt Pharma Operations Ltd. v. Teva Pharm. USA, Inc.*, 25 F.4th 1354, 1365 (Fed. Cir. 2022) (quoting KSR, 550 U.S. at 418). The Petition does not identify that required reason, *see* Pet. 45, nor does paragraph 147 of the Bajaj declaration (cited by Petitioners), which merely purports to demonstrate that the use of a “threshold” was “known to a POSITA.” EX1003 ¶147.

Petitioners’ third argument is that “a POSITA would have been motivated to look for a probabilistic framework, such as Zhao, to combine with the teaching of Steffens,” citing its “Motivation to Combine” section (§ X.B.1) and its Ground I arguments that Zhao discloses this claim element. This argument is also insufficient because (1) Petitioners’ motivation to combine section (*see* Pet. § X.B.1) does not address specific prior art elements that allegedly disclose “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object,” *ActiveVideo Networks*, 694 F.3d at 1328, and (2) Petitioners’ Ground I arguments do not explain how or why Steffens and Zhao in particular would allegedly be combined to arrive at the claimed invention. *See Metalcraft of Mayville, Inc. v. Toro Co.*, 848 F.3d 1358, 1367 (Fed. Cir. 2017) (“Without any explanation as to how or why the references would be combined to arrive at the claimed invention, we are left with only hindsight bias that KSR warns against).

ii. Reasonable Expectation of Success

Petitioners’ only arguments regarding reasonable expectation of success of combining Steffens and Zhao are in their “Motivation to Combine” section for Steffens and Zhao. Pet. § X.B.1. Nowhere in that section (nor anywhere else), however, do Petitioners tie their reasonable expectation analysis to, or frame it around, the specific requirements of element 1/9[e] of “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object.” *Teva Pharm. USA*, 18 F.4th at 1381 (where a claim required “safe administration of a specific amount of mifepristone,” stating that “the Board was required to frame its reasonable-expectation-of-success analysis around that specific dosage of mifepristone”). Petitioners’ reasonable expectation analysis is thus insufficient. *See id.*

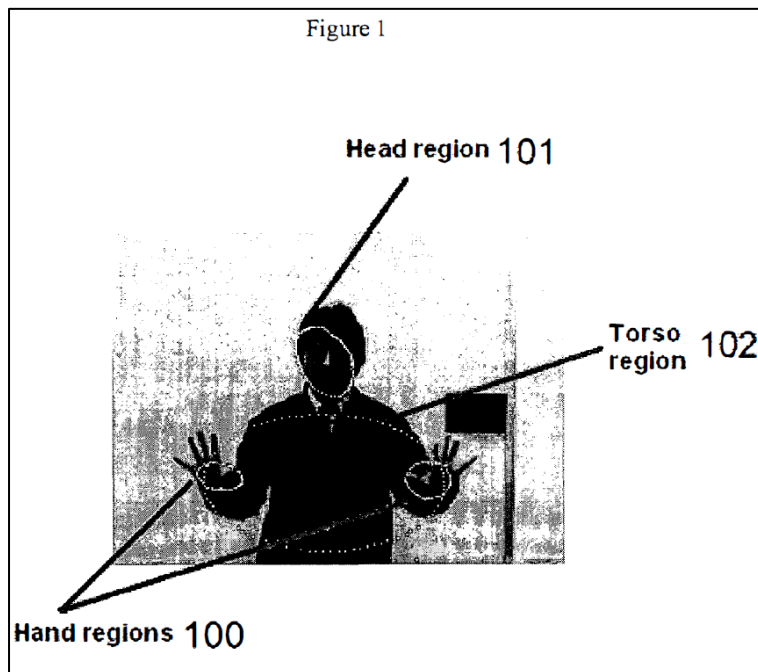
D. The Petition fails to demonstrate that Ozer in view of Zhao (Ground III) discloses or render obvious every limitation of claims 1 and 9.

- 1. 1/9[a]: “providing a spatio-temporal model for an object to be detected”**
 - a. Petitioners fail to show that Ozer discloses or renders obvious a “*spatio-temporal model for an object to be detected.*”**

Petitioners do not identify what specifically is the alleged “spatio-temporal model” in Ozer. *See* Pet. 56-57. Instead, Petitioners separately summarize (a) Ozer’s “graph matching” process and (b) process of “Classifying Over Multiple Frames,” and then assert in conclusory fashion that a POSITA would have found the

claimed “spatio-temporal model” to be “disclosed or at least rendered obvious in view of Ozer.” *Id.* Neither process, however, discloses a “spatio-temporal model” for an “object” to be “detected.”

Petitioners note that Ozer’s graph matching process involves an “object model” and that each “region modeled with ellipses corresponds to a node in the graphical representation of the object of interest.” Pet. 56. Fig. 1 of Ozer (reproduced below), shows Ozer’s “graphical representation of the object of interest.” See EX1008, 5:66-6:3, 7:8-11, Fig. 1.



Ozer indicates that its system builds a separate model “on *each* video frame,” thus at a given point in time and not over time. EX1008, 6:21-24 (“In contrast, this system builds a model of the object of interest on each video frame.”). Ozer’s object

model is thus limited to the state of the object at the time of the corresponding frame and does not disclose a spatio-temporal model. *See id.*

Unlike the spatio-temporal model for objection detection of the '980 patent, which is “extended over time”—thus it may include nodes that “represent[] either the object or a component of the object *at time t*” over a “temporal window” and account for the “temporal compatibility of object state between frames *i* and *j*” (EX1001, 3:47-50, 4:34-47)—Ozer’s object model does not model an object along any length of time. Indeed, Ozer’s object model reflects an object *within* an image frame (not *between* frames). *See* EX1008, 6:21-24.

Petitioners also raise Ozer’s “Classifying Over Multiple Frames,” which classifies “*movements*” of the already-classified object parts, where “[t]he *movements* of the object parts are described as a spatio-temporal sequence of feature vectors.” Pet. 56 (quoting EX1008, 7:33-45). Ozer describes this movement classifying process as “tracking” object movement. *See* EX1008, Abstract (“The system tracks the movement of different object parts in order to combine them at a later stage to high-level semantics. For example, the motion pattern of each human body part is compared to the motion pattern of the known activities. The recognized movements of the body parts are combined by a classifier to recognize the overall activity of the human body.”). The '980 patent explains, however, that “tracking” is distinct from “detection” in that “[t]racking makes use of object detection for

initialization and re-initialization during transient failures for occlusions,” whereas “[o]bject detection considers the consistency of the detection over time.” EX1001, 2:47-50. Ozer likewise distinguishes “track[ing]” from the “detect[ing]” and “identify[ing]” of object parts “to decide on the presence of the object of interest.” See EX1008, Abstract, 4:24-36 (“Object detection is achieved by matching the relational graphs of objects with the reference model. ... After the detection of the object parts, the system is ready to recognize the activities of each object part and the overall activity of the object.”). Ozer performs the movement classification or “tracking” process using the “output” of the graph matching or “object detection” process, which, as explained above, does not disclose a spatio-temporal model. EX1008, 7:4-31. After the object detection process, Ozer performs movement “tracking” using a separate process for a *movement* or activity to be “recognized,” classifie[d],” or “determine[d].” EX1008, 4:44-47, 4:54-57, 7:34-46. While the separate tracking process may make use of information generated from the object detection process, *id.*, 7:34-46, Ozer’s “object model” is nevertheless not a spatio-temporal model (for the reasons explained above), and the “tracking” process (a) does not provide a model “for an object to be *detected*,” and, in any event, (b) tracks a “movement” rather than the already-detected *object*.

Accordingly, Petitioners have failed to demonstrate that Ozer discloses the claimed “spatio-temporal model.” Petitioners alternatively assert that Ozer renders

it obvious. *See* Pet. 56-57. However, Petitioners fail to show that “it would have been obvious to modify [Ozer] to arrive at the patented invention.” *Arendi S.A.R.L.*, 832 F.3d at 1361 (citing *SIBIA Neurosciences, Inc.*, 225 F.3d at 1356).

b. Petitioners’ arguments under their alternative interpretation of “spatio-temporal model” also fail.

Petitioners assert that in Ozer’s graphical model, “nodes” for “the classified object parts” disclose nodes corresponding to “each of the collection of components,” and that nodes for “grouped object parts, forming the object” disclose nodes corresponding to “the object.” Pet. 57-58. However, Petitioners fail to show that Ozer discloses a graphical model that includes nodes for **both** the object parts (alleged “components”) **and** the “grouped object parts” (alleged “object”). That is, Petitioners fail to show that when Ozer’s parts are grouped, nodes for the separate parts (i.e., the alleged “components”) in the group are present in the graph along with a node for the group (i.e., the alleged “object”). Petitioners therefore fail to show that Ozer discloses or renders obvious “a graphical model comprising nodes corresponding to each of the collection of components **and** to the object.”

2. **1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”**
 - a. **Petitioners fail to show that Ozer or Zhao discloses the limitations of element 1[e].**

Petitioners do not assert that Ozer discloses this claim element. *See* Pet. 60-61. Instead, Petitioners rely entirely on their arguments that Zhao discloses this claim element to show that it was allegedly disclosed by the asserted combination of Steffens and Zhao. *See id.* For the reasons explained above, however, Petitioners fail to show that Zhao discloses or renders obvious this claim element. *See* § VII.B.2, *supra*. Accordingly, Petitioners fail to show that the combination of Ozer and Zhao discloses this claim element or that the combination of Ozer and Zhao discloses every element in the independent claims. Petitioners Ground III therefore fails, *Univ. of Strathclyde*, 17 F.4th at 160, and the Board need not address motivation to combine or reasonable expectation of success, *PAR Pharm.*, 773 F.3d at 1194.

- b. **Petitioners fail to sufficiently demonstrate a motivation to combine and reasonable likelihood of success of combining Ozer and Zhao for element 1[e].**
 - i. **Motivation to Combine**

To the extent the Board considers motivation to combine and reasonable expectation of success, Petitioners fail to satisfy their burden.

Petitioners fail to explain why it would have been allegedly obvious to combine elements from Ozer and Zhao “*in the way the claimed invention does*” i.e., to specifically include the limitations of element 1[e]. *See ActiveVideo Networks*, 694 F.3d at 1328. Petitioners argue that “a POSITA would have looked to combine the teachings of Ozer with Zhao,” citing its “Motivation to Combine” section (§ X.C.1) and its Ground I arguments that Zhao allegedly discloses this claim element. This argument is insufficient because (1) Petitioners’ motivation to combine section (*see* Pet. § X.C.1) does not address specific prior art elements that allegedly disclose “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object,” *see ActiveVideo Networks*, 694 F.3d at 1328, and (2) Petitioners’ Ground I arguments do not explain how or why Ozer and Zhao in particular would allegedly be combined to arrive at the claimed invention. *See Metalcraft of Mayville*, 848 F.3d at 1367.

ii. Reasonable Expectation of Success

Petitioners’ only arguments regarding reasonable expectation of success of combining Ozer and Zhao are in their “Motivation to Combine” section for Ozer and Zhao. Pet. § X.C.1. Nowhere in that section (nor anywhere else), however, do Petitioners tie their reasonable expectation analysis to, or frame it around, the specific requirements of element 1/9[e] of “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object.”

Teva Pharm. USA, 18 F.4th at 1381 (“the Board was required to frame its reasonable-expectation-of-success analysis around that specific [claim element]”). Petitioners’ reasonable expectation analysis is thus insufficient. *See id.*

E. The Petition fails to demonstrate that TLP (Ground IV) discloses or render obvious every element of claims 1 and 9.

- 1. 1/9[a]: “providing a spatio-temporal model for an object to be detected”**
 - a. Petitioners’ arguments that TLP discloses or renders obvious a “*spatio-temporal model* for an object to be detected” fail to satisfy their burden.**

Petitioners assert that by “*incorporating spatial and temporal elements into a graphical model*,” TLP teaches the claimed “spatio-temporal model.” Pet. 67. Yet, Petitioners pose an alternative interpretation (addressed below) and analyze TLP under that interpretation as well, but without explaining whether or not those two interpretations are consistent with each other or the term’s plain and ordinary meaning to a POSITA in light of the patent’s specification. Petitioners therefore fail to identify how the claims are to be construed for purposes of the Petition’s challenges and thus fail to meet their burden of demonstrating a reasonable likelihood of success. *See* § VI.A., *supra*.

b. Petitioners’ arguments under their alternative interpretation of “spatio-temporal model” also fail.

Petitioners assert that “TLP’s model for an object to be detected is explicitly expressed as ‘a graphical model in which each graph node corresponds to a *body*

part (upper leg, torso, etc.).” Pet. 67 (emphasis added). Petitioners do not, however, identify a node in TLP that corresponds to the body itself or the “human” (i.e., the alleged “object”). This is not surprising because, as Petitioners observed, TLP itself is clear that “each graph node” in TLP “corresponds to a body *part*.” EX1009, 2.

The inventors of the ’980 patent also recognized that TLP’s model lacks a node corresponding to the object. The ’980 patent explains that because the ’980 patent has “a two-layer graphical model” with both a “component layer” node *and* “object layer nodes,” it “allow[s] the assumption of the conditional independence of components ... over time given the overall object appearance.” EX1001, 4:34-44. The inventors recognized this as an improvement over TLP, which does not have an object layer (“second layer”) node in its model. *See* EX1003 (“Unlike the model presented in [2] having a second-layer also allows us to assume conditional independence of components over time given the whole appearance, hence reducing the complexity of the overall graphical model.”).⁴

⁴ Reference to “[2]” in EX1003 is to “L. Sigal, S. Bhatia, S. Roth, M.J. Black and M. Isard. Tracking Loose-limbed People. CVPR, 2004,” Petitioners’ TLP reference. EX.1003, 12. Sigal is a co-inventor of the ’980 patent. The inventors of the ’980 patent thus distinguished their invention from work with which they were familiar.

2. 1/9[e] “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”

Petitioners assert that this element is “disclosed” or “at least rendered obvious” by TLP. Pet. 71. Petitioners fail to show either.

Petitioners do not to show that TLP explicitly discloses this claim element. Petitioners cite no evidence that it does. Indeed, TLP never uses the word “threshold.” Instead, Petitioners assert that (a) TLP “describes the use of an ‘image likelihood model’ and detecting the object is ‘defined as estimating belief in the graphical model,’” and (b) “to infer the belief in a model, it would have been known to a POSITA to compare a determined probability against a threshold.” *Id.* This argument is insufficient for multiple reasons.

First, to the extent Petitioners assert that a POSITA knew of or was aware of the limitations of element 1/9[e], “a patent composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art.” *KSR*, 550 U.S. at 418. Instead, there must have been “an apparent reason to combine the known elements in the fashion claimed by the patent at issue.” *Id.* at 417-18. Thus, even though Petitioners’ obviousness argument is based on a single reference (i.e., TLP), Petitioners were still required to show that “it would have been obvious to modify that reference to arrive at the patented invention.” *Arendi S.A.R.L.*, 832 F.3d at 1361; *SIBIA Neurosciences*,

225 F.3d at 1356. Neither the Petition nor Dr. Bajaj’s declaration supplies such a reason. *See* Pet. 71; EX1004 ¶ 214 (parroting Petitioners’ arguments).

Second, Petitioners’ assertion that “it would have been known to a POSITA to compare a determined probability against a threshold” at most suggests that a POSITA *could* have modified TLP to arrive at the claimed invention, which is insufficient because Petitioners do not explain why a POSITA *would have been motivated* to do so. *See Belden Inc.*, 805 F.3d at 1073; *InTouch Techs., Inc. v. VGo Communs., Inc.*, 751 F.3d 1327, 1352 (Fed. Cir. 2014). Dr. Bajaj’s testimony also misses the mark. *See* Pet. 71; EX1004 ¶ 214 (parroting Petitioners’ arguments). Arguments and testimony so lacking amount to hindsight reconstruction of the claimed invention and are insufficient to support a conclusion of obviousness. *See TQ Delta*, 942 F.3d at 1359-63.

Third, to the extent Petitioners’ argument is interpreted to assert that comparing probabilities against a threshold was inherent to estimating belief, Petitioners and Dr. Bajaj fail to show that this claim element would *necessarily* be present when estimating belief in a graphical model, i.e., that belief propagation necessarily requires “comparing the probabilities for each image to a threshold.” *See* Pet. 71; EX1004 ¶ 214; *Par Pharm., Inc. v. TWi Pharm., Inc.*, 773 F.3d 1186, 1195-96 (Fed. Cir. 2014) (“A party must, therefore, meet a high standard in order to rely on inherency to establish the existence of a claim limitation in the prior art in an

obviousness analysis—the limitation at issue necessarily must be present, or the natural result of the combination of elements explicitly disclosed by the prior art.”). Without such a showing, Petitioners assertion amounts to no more than a suggestion that a POSITA *could* have modified TLP to arrive at the claimed invention, which is insufficient to explain why a POSITA *would have been motivated* to do so. *See Belden Inc.*, 805 F.3d at 1073; *InTouch Techs., Inc. v. VGo Communs., Inc.*, 751 F.3d 1327, 1352 (Fed. Cir. 2014).

Fourth, To the extent Petitioners’ argument is interpreted as asserting it would have been “common sense” or “intuitive” to include this claim element, Petitioners and Dr. Bajaj fail “to identify any actual reason why a skilled artisan would have combined the elements in the claimed manner.” *See* Pet. 71; EX1004 ¶ 214; *In re Van Os*, 844 F.3d 1359, 1361 (Fed. Cir. 2017). Moreover, “common sense” and “ordinary creativity”—whether invoked “to supply a motivation to combine or a missing limitation—cannot be used as a wholesale substitute for reasoned analysis and evidentiary support, especially when dealing with a limitation missing from the prior art references specified.” *DSS Tech. Mgmt. v. Apple Inc.*, 885 F.3d 1367, 1374 (Fed. Cir. 2018).

F. The Petition fails to demonstrate a reasonable likelihood that dependent claims 2-8 and 10-16 will be proven unpatentable.

Claims 2-8 depend from claim 1, and claims 10-16 depend from claim 9. Because Petitioners fail to demonstrate a reasonable likelihood that independent claims 1 and 9 will be proven unpatentable, they fail to demonstrate a reasonable likelihood that their dependent claims 2-8 and 10-16 will be proven unpatentable for the same reasons. *See* §§ VII.A-VII.D, *supra*.

VIII. CONCLUSION

For the reasons stated above, institution of the IPR should be denied.

Dated: September 8, 2023

By: /Christine E. Lehman/
Christine E. Lehman
PTO Reg. No. 38,535
Counsel for Patent Owner

CERTIFICATE OF COMPLIANCE

The undersigned hereby certifies that this Patent Owner's Preliminary Response (the "POPR") complies with the type-volume limits of 37 C.F.R. § 42.24 and consists of 13,975 words, excluding the table of contents, table of authorities, mandatory notices under 37 C.F.R. § 42.8, certificate of service, this certificate of compliance, and table of exhibits. The undersigned has relied on the word count feature of the word-processing system used to prepare this paper (Microsoft Word).

Dated: September 8, 2023

Respectfully submitted,

/Christine E. Lehman/

Christine E. Lehman

PTO Reg. No. 38,535

Counsel for Patent Owner

CERTIFICATE OF SERVICE

The undersigned hereby certifies that a copy of the foregoing **PATENT OWNER'S PRELIMINARY RESPONSE** was served on September 8, 2023, via electronic mail to Meta_VideoLabs@kirkland.com, as agreed to by counsel for Petitioners, upon the following counsel:

W. Todd Baker
todd.baker@kirkland.com
KIRKLAND & ELLIS LLP
1301 Pennsylvania Avenue, N.W.
Washington, D.C. 20004

Ellisen Shelton Turner
ellisen.turner@kirkland.com
KIRKLAND & ELLIS LLP
2049 Century Park East
Los Angeles, CA 90067

Jonathan D. Brit
jonathan.brit@kirkland.com
KIRKLAND & ELLIS LLP
601 Lexington Avenue
New York, NY 10022

Dated: September 8, 2023

/Christine E. Lehman/
Christine E. Lehman