

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

**SUPPLEMENTAL DECLARATION OF ELI SABER, PH.D.
IN SUPPORT OF PATENT OWNER'S SUR-REPLY**

TABLE OF CONTENTS

I.	INTRODUCTION	5
II.	BACKGROUND AND QUALIFICATIONS	6
III.	MATERIALS CONSIDERED	6
IV.	LEGAL PRINCIPLES.....	6
V.	LEVEL OF ORDINARY SKILL	6
VI.	CLAIM CONSTRUCTION	7
A.	Element 1/9[e]: “detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”	7
1.	Dr. Bajaj misconstrues “object” as “collection of components.”.....	8
B.	“spatio-temporal model” (element 1/9[a])	31
VII.	PRIORITY DATE OF THE '980 PATENT	37
VIII.	DR. BAJAJ DOES NOT SHOW THAT ANY CHALLENGED CLAIM WOULD HAVE BEEN OBVIOUS	40
A.	The challenged claims are patentable over Zhao (Ground I).....	40
1.	Claims 1/9	40
2.	Claims 4/12	44
3.	Claims 8/16	47
4.	All dependent claims are patentable over Zhao.....	54
B.	The challenged claims are patentable over Steffens in view of Zhao (Ground II).	55
1.	Claims 1/9	55

2.	Claims 4/12	57
3.	Claims 8/16	58
4.	All dependent claims are patentable over Steffens in view of Zhao.	58
C.	The challenged claims are patentable over Ozer in view of Zhao (Ground III).	59
1.	Claims 1/9	59
2.	Claims 4/12	59
3.	Claims 8/16	60
4.	All dependent claims are patentable over Ozer in view of Zhao.....	60
D.	The challenged claims are patentable over TLP (Ground IV).	61
1.	Alleged Public Accessibility of TLP	61
2.	Claims 1/9	64
3.	Claims 3/11	68
4.	Claims 8/16	68
5.	All dependent claims are patentable over TLP.....	72

EXHIBIT LIST

Exhibit	Description
2001	Declaration of Jaime F. Cardenas-Navia in Support of Motion for Admission Pro Hac Vice
2002	Zou et al., Object Detection in 20 Years - Survey
2003	Email dated December 29, 2023
2004	Declaration of Eli Saber in Support of Patent Owner's Response
2005	Curriculum Vitae of Eli Saber, Ph.D.
2006	Deposition of Dr. Chandrajit Bajaj
2007	Supplemental Declaration of Eli Saber in Support of Patent Owner's Sur-Reply
2008	2024-07-11 Deposition of Dr. Chandrajit Bajaj
2009	2004-02-13 Email
2010	2003-11-17 Email
2011	2024-07-18 Deposition of Dr. Sylvia Hall-Ellis

I. INTRODUCTION

1. I have been retained as an independent expert consultant on behalf of VL Collective IP LLC (“Patent Owner”) in this *Inter Partes* Review proceeding (“IPR”) before the Patent Trial and Appeal Board (“Board” or PTAB”) of the United States Patent and Trademark Office (“PTO”) regarding U.S. Patent No. 7,436,980 (“the ’980 Patent”) (EX1001).¹ I am being compensated for the time I spend on this matter at my usual hourly rate, in addition to expenses. My compensation in this matter is in no way contingent on the nature of my findings, the presentation of my findings in testimony, or the outcome of this or any other proceeding. I have no other interest in this proceeding.

2. In particular, I have been asked to consider, and provide my opinions regarding, arguments and opinions set forth in Petitioners’ Reply (Paper 32) and Dr. Bajaj’s supplemental declaration (EX1052). This declaration is based on the information currently available to me. To the extent that additional information becomes available, I reserve the right to continue my investigation and study. My opinions are set forth below.

¹ In this Declaration, I reference documents by exhibit numbers that, as I understand, are attached to exhibits in this *Inter Partes* Review of the ’980 Patent.

3. I make this declaration as a supplement to my initial declaration (EX2004), in response to arguments and opinions set forth in Petitioners' Reply (Paper 32) and Dr. Bajaj's supplemental declaration (EX1052).

II. BACKGROUND AND QUALIFICATIONS

4. My background and qualifications are summarized in my previous declaration. EX2004, §II.

III. MATERIALS CONSIDERED

5. In forming the opinions in this Declaration, in addition to the documents that I reviewed in preparing my initial declaration (EX1052), I reviewed Petitioners' Reply (Paper 32), Dr. Bajaj's supplemental declaration (EX1052), and other documents I address herein.

IV. LEGAL PRINCIPLES

6. I set forth my understanding of certain legal principles in my initial declaration. EX2004, §IV.

V. LEVEL OF ORDINARY SKILL

7. I set forth my understanding of the level of ordinary skill in my initial declaration. EX2004, §V.

VI. CLAIM CONSTRUCTION

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

8. I understand that Patent Owner sets forth the following claim constructions for element 1/9[e]:

<u>Term</u>	<u>Patent Owner’s Construction</u>	<u>Petitioners’ Construction²</u>
“detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object”	“detecting the object in any image based on comparing the probabilities for each image of the plurality of images to a threshold for detecting the object”	“detecting the collection of components in any image upon comparing the collection of component probabilities for each image to a threshold for detecting the collection of components”

9. I understand that Petitioners and Dr. Bajaj dispute the following: (1) whether “object” should be rewritten as “collection of components”; (2) whether “the probabilities” are for the object in the plurality of images or for the collection of components in each image; and (3) whether “upon” means “based on.”

² I understand that Petitioners and Dr. Bajaj do not formally propose this construction but apply it in their invalidity analysis.

1. Dr. Bajaj misconstrues “object” as “collection of components.”

10. Dr. Bajaj treats “object” and “collection of components,” which are two distinct claim terms, as synonymous and interchangeable. Specifically, for limitation 1[e], Dr. Bajaj argues that “comparing the probabilities for each image” should be understood as “comparing the (collection of component) probabilities for each image.” EX1052, ¶13. Dr. Bajaj bases his interpretation his assertion that objects are just a collection of components. EX2008, 90:7-11 (“The object is the sum of its parts.”), 92:19-93:1 (“The object is the collection of components.”), 18:2-19:5-21, 21:8-22:4, 24:5-12, 86:18-87:7.

11. Bajaj replaces “object” with “collection of components” every time the claims refer to object “probabilities.” EX1052, ¶¶12-13, 16-17. He confirmed this interpretation of the claims at his deposition.

5	Q.	And when you said just now
6		"comparing the probabilities to a threshold,"
7		you're saying your interpretation of the claim is
8		that the probabilities are probabilities for
9		components, right?
10	A.	Comparing the probabilities for the
11		collection of components because that's what
12		constitutes the object in each image.

EX2008, 24:5-12

1 of components. And then determining a
2 probability that the object -- you can always put
3 parentheses again -- of collection of components
4 because the object is always a collection of
5 components. And then detecting that object which
6 is again a collection of components is upon
7 comparing the probabilities for each image to a
8 threshold for detecting the object.

EX2008, 19:1-8

2 Q. Okay. Referring specifically to the
3 Element 1[e]/9[e], the words "collection of
4 component" are not literally present in Element
5 1[e]/9[e], right, as you've written it?

6 A. They are implicitly, and that's why
7 I put that -- put the words -- "collection of
8 components" is part of the entire plane.

9 Q. Yeah, that's -- that's what I want
10 to -- to clarify here. You said that they're --
11 that they're there implicitly and so -- you said
12 that they're there implicitly, but you agree that
13 they're not there expressly or literally?

14 A. There's nothing to agree about your
15 statement. The claims are as they are. It's
16 very clear the claims starting from the
17 beginning; and, you know, we can go back to the
18 very beginning of the claim.

19 But the words I repeated were in
20 1[c]. It says, "measuring the object as a
21 collection of components in each image."

22 So objects are always a collection

1 of components. And then determining a
2 probability that the object -- you can always put
3 parentheses again -- of collection of components
4 because the object is always a collection of
5 components. And then detecting that object which
6 is again a collection of components is upon
7 comparing the probabilities for each image to a
8 threshold for detecting the object.

9 Q. Okay. So going back to Paragraph 13
10 of your declaration, at the end of the first
11 sentence, you refer to a threshold.

12 Do you see that?

13 A. Yes.

14 Q. Okay. And what is that threshold
15 for?

16 A. When the probabilities of each
17 component in the collection of components in an
18 image is beyond a certain threshold, that value,
19 then your likelihood that that object is present
20 in the image is high. And that's what thresholds
21 are used for.

EX2008, 18:2-19:21

8	What I'm asking is: Is the
9	threshold in Paragraph 13 that you're talking
10	about, is that the threshold that you compare the
11	component probability to?
12	A. I'm reading the claims as they are
13	written, and as a person of ordinary skill in the
14	art would interpret it.
15	The object is the collection of
16	components in each image. You determine a
17	probability that the object is in each image and
18	detect the object in any image upon comparing the
19	probabilities. Since the object is a collection
20	of components, there is a probability for each of
21	the components.
22	And the joint probability, as you

1	learn later on, as far as that claim is
2	concerned, it's comparing the probabilities of
3	each of the collection of components in that
4	image to a threshold.

EX2008, 21:8-22:4

12. The '980 patent describes new techniques for improving the detection of “objects.” Objects in the '980 patent are modeled, as are their “collection of components.” Dr. Bajaj errs in his interpretation of the claims by collapsing these two related but distinct entities (“objects” and “collection of components”) into a

VL Collective Exhibit 2007

IPR2023-00924

Meta Platforms, Inc. v. VL Collective IP LLC

Page 12 of 72

single, interchangeable concept. His view conflicts with the intrinsic evidence and is not how the POSITA would understand the terms.

13. The claims recite steps “for *object* detection” where “a spatio-temporal model for an *object* to be detected” is provided, and each step expressly refers to “the *object*” detected using the model. *See* EX1001, claims 1 and 9. Specifically, the claims recite:

1/9[a]: providing a **spatio-temporal model for an object to be detected**;

1/9[b]: providing a video comprising a **plurality of images including the object**;

1/9[c]: **measuring the object** as a collection of components in each image;

1/9[d]: **determining a probability that the object** is in each image; and

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

See EX1001, claims 1 and 9. In each of the bolded portions above, the POSITA would understand the claims to be referring to the “object” as a distinct entity from the “collection of components, which is separately recited (and only addressed in) element 1/9[c]. Given the claim’s express reference to both an “object” *and* a “collection of components,” the POSITA would have recognized them as distinct elements and therefore would have expected the patentee to have referred to the

“collection of components” (rather than “object”) wherever that was intended. The POSITA would have also understood that by requiring the object to be measured “as a collection of components,” the patentee intended for there to be a distinction between the “object” and the “collection of components,” otherwise the clause “as a collection of components” would be superfluous—for example, the claim could have just said “measuring the object in each image.” Accordingly, the POSITA would have understood that “object” means “object”—not “collection of components”—wherever it is stated.

14. Dr. Bajaj incorrectly interprets “object” in elements 1/9[d] and 1/9[e] as “collection of components.” He does so based on element 1/9[c]’s step of “measuring the object as a collection of components in each image.” *See* EX1052 ¶¶13-14; EX2008 at 17:6-19:8.

1 A. Yes.

2 Q. And do you see that the words
3 "collection of component" are in parentheses as
4 well?

5 A. Yes, I wrote this.

6 Q. Okay. And so then, you would agree
7 that the words in parentheses there are not
8 literally present in Element 1[e]/9[e], right?

9 A. The words "collection of components"
10 is present in 1[e]/9[e], and the probabilities
11 are the likelihood of the presence of an object
12 in an image. And that's what the sentence refers
13 to.

14 Q. Okay. Let's take a look at
15 Paragraph 12, which is just above Paragraph 13 on
16 Page 8 of your declaration.

17 You -- you include here the text of
18 Element 1[e]/9[e]. And it says, "detecting the
19 object in any image upon comparing the
20 probabilities for each image to a threshold for
21 detecting the object."

22 Do you see that?

1 A. I see the elements, yes.

2 Q. Okay. Referring specifically to the
3 Element 1[e]/9[e], the words "collection of
4 component" are not literally present in Element
5 1[e]/9[e], right, as you've written it?

6 A. They are implicitly, and that's why
7 I put that -- put the words -- "collection of
8 components" is part of the entire plane.

9 Q. Yeah, that's -- that's what I want
10 to -- to clarify here. You said that they're --
11 that they're there implicitly and so -- you said
12 that they're there implicitly, but you agree that
13 they're not there expressly or literally?

14 A. There's nothing to agree about your
15 statement. The claims are as they are. It's
16 very clear the claims starting from the
17 beginning; and, you know, we can go back to the
18 very beginning of the claim.

19 But the words I repeated were in
20 1[c]. It says, "measuring the object as a
21 collection of components in each image."

22 So objects are always a collection

1 of components. And then determining a
2 probability that the object -- you can always put
3 parentheses again -- of collection of components
4 because the object is always a collection of
5 components. And then detecting that object which
6 is again a collection of components is upon
7 comparing the probabilities for each image to a
8 threshold for detecting the object.

EX2008 at 17:6-19:8

That is, he interprets “the object” in those elements to refer to the “collection of components” in 1/9[c] rather than the “object” recited in the previous elements. *Id.* Indeed, Dr. Bajaj’s claim construction analysis in his declaration paraphrases the claims in order to use the words “collection of components” or “component collection” when the claims themselves do not. EX1052 ¶¶13-14 (referring to “(collection of component) probabilities” and “component collection model”). Dr. Bajaj confirmed at his deposition that under his interpretation, every instance of the word “object” in the claims can be replaced with “collection of components.” EX2008 at 18:9-19:8. (“And then determining a probability that the object -- you can always put parentheses again -- of collection of components because the object is always a collection of components.”).

9 Q. Yeah, that's -- that's what I want
10 to -- to clarify here. You said that they're --
11 that they're there implicitly and so -- you said
12 that they're there implicitly, but you agree that
13 they're not there expressly or literally?

14 A. There's nothing to agree about your
15 statement. The claims are as they are. It's
16 very clear the claims starting from the
17 beginning; and, you know, we can go back to the
18 very beginning of the claim.

19 But the words I repeated were in
20 1[c]. It says, "measuring the object as a
21 collection of components in each image."

22 So objects are always a collection

1 of components. And then determining a
2 probability that the object -- you can always put
3 parentheses again -- of collection of components
4 because the object is always a collection of
5 components. And then detecting that object which
6 is again a collection of components is upon
7 comparing the probabilities for each image to a
8 threshold for detecting the object.

EX2008 at 18:9-19:8

VL Collective Exhibit 2007
IPR2023-00924

Meta Platforms, Inc. v. VL Collective IP LLC
Page 18 of 72

Accordingly, under Dr. Bajaj's claim construction, "object" is equated or made synonymous with "collection of components" such that every time the inventors meant to distinguish between "object" and "collection of components," the distinction is removed. I disagree that the POSITA would interpret the claims that way.

15. The POSITA would also understand that for the object to be *measured* as a collection of components, the object would have an identity and characteristics distinct from its constituent components. This understanding, and the distinction between the object and its components, is further confirmed by the patent's specification. For example, the patent teaches that "[m]odeling objects by an arrangement of image-based, and possibly overlapping, components facilitates detection of complex articulated objects as well as helps in handling partial object occlusions or local illumination changes." EX1001 at 2:50-54. This teaches that objects are modeled as distinct entities from their components. *Id.* This distinction is evident in Figures 2A and 2B, which shows that "object detection" is formulated as an inference in a model that distinctly models the object (e.g., as an object layer or object node) in addition to its components (e.g., as a component layer or component nodes). *Id.* at 2:55-59. The patent also teaches that the object is modeled distinctly from its components at *each* point in time, as "[e]ach node in the graphical

model corresponds to a position and a scale of the component or the object as a whole in an image at a given time instance.” *Id.* at 2:60-63. The patent further teaches that the object is modeled as a probabilistic spatio-temporal model in which the object is modeled distinctly from its components, for example, as shown by the joint probability distribution for the spatio-temporal model described at 3:47-4:7, which expressly includes distinct states of the object (“*O*”) ($X_0^O \dots X_T^O$) and each component (“*C*”) ($X_0^{C_0}, X_0^{C_1} \dots X_0^{C_N} \dots X_T^{C_0}, X_T^{C_1} \dots X_T^{C_N}$) at each point in time. *Id.* The joint probability distribution for the spatio-temporal model also includes multiple probabilistic constraints (potential functions) that specifically distinguish between the object and components: $\psi_{ij}(X_i^O, X_j^O)$ for object-to-object compatibility; $\psi_{ik}(X_i^O, X_i^{C_k})$ for object-to-component compatibility, and $\psi_{kl}(X_i^{C_k}, X_i^{C_l})$ for component-to-component compatibility. *Id.* It also expressly accounts for distinct local evidence object state ($\phi_i(X_i^O)$) and local evidence of component state ($\phi_i(X_i^{C_k})$). *Id.* at 3:47-4:7. The patent teaches that the spatio-temporal object model is built using a spatial model that includes both a component layer of “loosely connected parts” and a distinct object layer that “corresponds to an entire appearance model of the object.” *Id.* at 4:18-26. In this way, “*objects are modeled using ... components.*” *Id.* at 4:26-27 (emphasis added). “To integrate temporal constraints,” which are probabilistic temporal constraints that make the object model spatio-

temporal, the spatial model is “extended over time to an arbitrary length temporal window.” *Id.* at 4:34-37. “The resulting spatio-temporal graphical models are shown in Figs. 2A and 2B.” *Id.* Figures 2A and 2B depict the “object as a whole” 203 (also referred to as “entire appearance model” and “ O_t ”) as a distinct entity from the collection of “components” 202 in the spatio-temporal model. EX1001, 4:18-48. This distinctness has a purpose: it “allows the inference process to reason explicitly about the object as a whole” and assume “conditional independence of components” for purposes of detecting the *object*. EX1001, 2:55-59, 4:37-42. Even when “the whole appearance of an object is too complicated to model as a whole ... and can only be modeled in terms of components” (*e.g.*, because the object is too complicated to train a likelihood proposal for its entire appearance), “a uniform likelihood can be assumed over the object state space” for the object layer. *Id.* at 6:13-22; EX1051 at 71:12-75:20. In other words, the spatio-temporal object model still includes a distinct object layer/object node (and a distinct “object state space” with its own distinct “likelihood”), but the probabilistic distribution of the object over its components is assumed to be uniform (rather based on trained relationships). *Id.* As the POSITA would understand, the object is referred to as an “entire appearance model” because the object, although it includes components, is mathematically represented in the spatio-temporal model as a distinct entity with

respect to the components. EX1001, 3:47-448. Indeed, the patent teaches that non-parametric belief propagation (BP) can be used for inferring both “the state of the object *and* its components” using the spatio-temporal object model. *Id.* at 6:23-25 (emphasis added). For example, the patent teaches that “information from all components is integrated in the object layer node” using non-parametric BP to represent the distinct object entity (which is not a component). *Id.* at 6:41-49. In this way the distinct object entity can be detected based on the distinct state of the object itself and the state of its components. *Id.* at 7:23-26. The distinction between the object and its components is also reflected in Figure 5 and the specification’s corresponding description, which further shows the patent’s consistent distinction between “object” and “components.” *Id.* at 7:45-55; Fig. 5. Indeed, as summarized in the patent, the patent exemplifies its framework using examples of objects (vehicles and pedestrians) for which state information of “the objects *and* their components” can be reliably inferred. *Id.* at 8:9-17 (emphasis added). From the context of the claim language itself, and in view of the rest of the patent’s specification (as discussed above), the POSITA would understand that, in the claims, the “object” is distinct from the “collection of components.” Accordingly, the POSITA would understand that the patent teaches determining whether a distinct

probability that the *object* is present is greater than a threshold for detecting the *object*: “An object is considered present if $h_k(I) \geq \frac{1}{2} \sum_{k=1}^K \alpha_k$.” *Id.* at 5:58-6:13.

16. Dr. Bajaj’s statement that object probability is determined “as a function of” component probabilities supports this distinction and undermines Dr. Bajaj’s opinion to the contrary. EX1052, ¶¶16-17. Dr. Bajaj essentially suggests that object probability (an output) is a function of component probabilities (inputs). But this shows that objects and components are not the same—otherwise there would be no need for a function to get from component probabilities to the object probability. His opinion confirms that there is a distinct object probability separate and apart from any component probabilities. Moreover, the specification does not disclose comparing each *component* probability to the threshold for detecting the *object*, as Dr. Bajaj’s construction requires. *See* EX1052 ¶13 (“These claim 1 and 9 limitations clearly specify ... comparing the (collection of component) probabilities for each image to a threshold.”); EX2008 at 19:14-21 (“When the probabilities of each component in the collection of components in an image is beyond a certain threshold, that value, then your likelihood that the object is present in the image is high. And that’s what thresholds are used for.”).

14	Q.	Okay. And what is that threshold
15		for?
16	A.	When the probabilities of each
17		component in the collection of components in an
18		image is beyond a certain threshold, that value,
19		then your likelihood that that object is present
20		in the image is high. And that's what thresholds
21		are used for.

EX2008 at 19:14-21

The threshold in the claims is “for detecting the object,” and the specification does not disclose comparing each *component* probability to the threshold for detecting the *object* – it always discloses comparing the object probability to the threshold. EX1001 at 5:59-67, 7:23-44. Comparing the *component* probabilities to an *object* threshold would also make little sense. For example, if the object were a person, then comparing the probability that something is a left arm to the threshold associated with the person would be nonsensical.

17. Because the POSITA would understand “object” and “collection of components” to be distinct entities, claim 1/9 would not be reasonably or logically understood such that “object” means “collection of components” or vice versa. It is true that an object can include a collection of components, but as the POSITA would

understand, “object” and “collection of components” are nevertheless distinct concepts and entities. For example, when “object” means the same thing as “collection of components” (as under Dr. Bajaj’s construction), element 1/9[c] would mean “measuring the [collection of components] as a collection of components,” which is redundant and adds no further meaning to the claim. Furthermore, as a matter of ordinary meaning and common sense, the POSITA would understand that measuring the object as a collection of components necessarily contemplates a difference between the object and its components (e.g., a person may have a left arm, but a left arm is not a person).

18. Under Dr. Bajaj’s construction, which equates “object” with “collection of components,” the claims would exclude the embodiments of the ’980 specification where $w > 1$. Specifically, Dr. Bajaj’s construction results in the object detection analysis of limitation 1/9[e] involving just a *single* image ($w=1$), and never a “plurality of images” ($w > 1$). That is contrary to the plain meaning of the claims. Limitation 1/9[c] indicates that there are “a plurality of images” (i.e., two or more), while 1/9[d] requires “determining a probability that the object is in each image.” This means that there is *a probability* (singular) associated with *the object* for *each image* (of the plurality). The single *probability* for *each image* results in two or more *probabilities* for the *plurality of images*. The reference to “the probabilities

for each image” in limitation 1/9[e] must therefore be referring to the *probabilities* for the *plurality of images*, as there are no other *probabilities* (plural). Under Dr. Bajaj’s construction, however, the object detection step of 1/9[e] is only ever based on the probabilities for a *single* image—the one being analyzed—which is the “w=1” scenario. Under Bajaj’s construction, the claims *only* cover this scenario, at the exclusion of the “w>1” scenarios, which the patent teaches are preferred over the w=1 scenario. See EX1001, 7:20-22, 7:56-8:3.

19. Furthermore, equating “object” with “collection of components” (under Dr. Bajaj’s construction) confuses whether 1/9[e]’s “probabilities for each image” refer to the same or different “probability that the object is in each image” in 1/9[d]. This is because element 1/9[d] introduces only “*a* probability” (singular) for the object in each image, not *multiple* probabilities for the collection of components in each image, and the singular/plural mismatch creates confusion. To make sense, Dr. Bajaj’s construction would require rewriting 1/9[d] to say “probabilities that the object is in each image.” It also creates inconsistencies between claims 1/9 and their dependent claims 3/11. Claims 3/11 requires “nodes” for both “the object” and “the collection of components,” thus they require that the “object” and the “collection of components” be distinct and not the same, which is inconsistent with Dr. Bajaj’s interpretation of claims 1/9 (where they are not

distinct). Similarly, claims 4/12 expressly recite detecting an object based on an analysis that includes the object probabilities for “a prior image and a later image relative to the current image.” EX1001, claims 1, 4, 9, 12. This is the “w=3” (or greater) scenario. Yet because claims 1/9 only encompass the w=1 scenario under Dr. Bajaj’s construction, claims 4/12 are inconsistent with claims 1/9.

20. The POSITA would understand the claims to be unambiguous at w=1 is excluded. And that makes sense because the w=1 scenario is not part of the inventive “spatio-*temporal*” model at the core of the invention, as discussed in section VI.B below. EX1001, 7:6-22 (distinguishing “single frame component based fusion” from spatio-temporal model). Dr. Bajaj agrees that w=1 does not have a temporal element. *See* EX1052, ¶¶28-29 (w=1 does not fuse, combine, or reference probabilities from other image frames); EX2008, 51:14-52:21 (w=1 “is not using any information from a previous frame or a future frame), 56:1-59:5 (with w=1 “you cannot look at the previous or the future frame”). It is instead a single-frame simple fusion model like those the patent seeks to improve over. EX1001, 7:6-22 (distinguishing “single frame component based fusion” from spatio-temporal model), 7:56-8:3 (distinguishing “simple fusion” from spatio-temporal model); EX1052, ¶29 (w=1 uses probabilities within a single image). That the patent discusses the non-preferred w=1 scenario does not suggest that w=1 is part of the

claims. Instead, the fact that the $w=1$ implementation was discussed in the specification and contrasted from $w>1$ indicates to the POSITA that inventors wanted to distinguish $w=1$ as a disfavored embodiment. I therefore disagree with Dr. Bajaj's view that that the '980 patent does not contrast $w=1$ with the invention because "[i]f $w=1$ was not intended to be included, then the specification could have indicated that w is an odd integer > 1 or w is an odd integer ≥ 3 ." EX1052, ¶¶31-32. The patent did so by stating that "with $w=1$, the algorithm resembles single frame component-based fusion" and describing this as a scenario in which performing inferences over time with the spatio-temporal model "is not an option due to the lengthy off-line processing needed." EX1001, 7:6-22. The " $w=1$ " simple fusion scheme is also incapable of achieving the patent's function and purpose, as shown by its use as the basis for comparison with the "3-frame spatio-temporal object model." EX1001, 7:56-8:3.

21. I understand that Petitioners argue that Figure 5 allegedly supports a broader reading of the claims. I previously explained why Petitioners' reasoning is erroneous. EX2004, ¶68. The language in 505 does not match the specification's description of 505, which is fully consistent with the claim language, suggesting that 505 contains a typographical error (i.e., which could be fixed by changing "image" to "images"). *Compare* Figure 5 (step 505) *with* 7:53-55 and 8:40-43.

22. Dr. Bajaj also misinterprets the prosecution history. His assertion that the examiner “concluded that Cham disclosed the probabilistic framework of the ’980 Patent claims” is untrue and ignores the examiner’s findings. EX1052, ¶20. In the Notice of Allowance, the Examiner began his explanation of why the claims of the ’980 patent are allowable over the prior by stating that the features of limitations 1/9[a], 1/9[d], and 1/9[e] “are not taught or suggested by the art of record,” which includes Cham. EX1002, 162. Accordingly, the examiner determined that Cham does not disclose the features of 1/9[a], 1/9[d], and 1/9[e]. *Id.* at 162-163. Nothing in the Examiner’s subsequent discussion of Cham is inconsistent with that finding. In describing what Cham discloses, the Examiner used the claim language for element 1/9[b] and 1/9[c] only—not 1/9[a], 1/9[d], or 1/9[e]. The Examiner determined that Cham discloses “a kinematic model” (not the spatio-temporal model of 1/9[a]) and “performing feature search operations for object features by comparing a total probability with a predetermined threshold” (not the steps of 1/9[d] and 1/9[e]). The Notice of Allowance therefore indicates twice that Cham does not disclose the features of elements 1/9[a], 1/9[d], and 1/9[e]. The examiner’s findings make sense because element 1/9[e] uses the spatio-temporal model of element 1/9[a]. Dr. Bajaj agrees that the detecting step in element 1/9[e] uses the spatio-temporal model of element 1/9[a], which he calls a “precursor” for performing the

detection in 1/9[e]. EX1052, ¶14; EX2008 at 32:3-10 (“Q. So why is providing a spatio-temporal model for an object to be detected a precursor to performing the Steps [c], [d], and [e]? A. Obviously, what am I going to detect? I’m going to detect an object, so you have to provide me what type of object are you trying to detect.”); 32:21-33:2 (“Q. Okay. So you need the spatio-temporal model to do the other measuring, determining, and detecting steps? A. You need the video as well.”).

3	Q.	So why is providing a
4		spatio-temporal model for an object to be
5		detected a precursor to performing the Steps [c],
6		[d], and [e]?
7	A.	Obviously. What am I going to
8		detect? I'm going to detect an object, so you
9		have to provide me what type of object are you
10		trying to detect.

EX2008 at 32:3-10

21	Q.	Okay. So you need the
22		spatio-temporal model to do the other measuring,

Page 33		
1		determining, and detecting steps?
2	A.	You need the video as well.

And, as I've explained, the POSITA would understand that element 1/9[e] indeed requires and uses the spatio-temporal model of element 1/9[a]. EX1051, 59:14-63:11; EX2004, ¶¶70-71. It is therefore no surprise that the Examiner found that Cham does not disclose the spatio-temporal model of 1/9[a] or the detecting step of 1/9[e]. EX1002 at 3. The Examiner's findings support my interpretation of the claims, not Dr. Bajaj's.

23. Dr. Bajaj also misinterprets Cham. Cham's "total probability" is not an object probability; it is a probability of matching a feature in a region. Cham "registers a plurality of object models in at least one image, one *feature* per model at a time." EX1053 at 3:2-3 (emphasis added). Cham matches *features* by searching for them in *regions* that have a match probability above a threshold. *Id.* at 3:16-4:3. The "total probability" is "formed by summing the probabilities associated with each of the found search *regions*." *Id.* (emphasis added). Accordingly, Cham's "total probability" is not "the total probability that an object is in an image frame." EX1052, ¶22.

B. "spatio-temporal model" (element 1/9[a])

24. I understand that Patent Owner sets forth the following claim constructions for spatio-temporal model:

<u>Term</u>	<u>Patent Owner's Construction</u>	<u>Petitioners' Construction³</u>
"spatio-temporal model"	spatial model where an object is detected in an image based on the probabilities that the object is in preceding and/or subsequent images	spatial model where an image's analysis takes into account information from preceding and/or subsequent images

25. The inventors repeatedly and unambiguously defined "spatio-temporal model" in the '980 patent. Starting with the claims, the claims expressly state that the function of the "spatio-temporal model" is "for an object to be detected." EX1001, claim 1. The POSITA would understand that function to be a property of the "spatio-temporal model" (not just an indication of how it can be used), as the claims and specification teach that object detection is an essential and fundamental characteristic of the claimed invention. EX1001, EX1001, 1:28-33, 1:37-39, 1:48-52, 2:7-15, 2:43-54; 3:11-18, 7:6-8:8.

26. The specification also repeatedly uses the term "spatio-temporal model" consistent with PO's construction. The specification refers to "spatio-temporal model" (or slight variations, *e.g.*, "spatio-temporal graphical model," "spatio-temporal directed graphical model," "spatio-temporal object model," etc.) at

³ I understand that Petitioners do not formally propose this construction but apply it in their invalidity analysis.

least 14 times. EX1001, 1:28-2:18, 3:47-4:48, 7:45-8:8. The patent repeatedly indicates that the function of the spatio-temporal model is to detect an object, and to do so based on the probabilities that object is in earlier or later images. EX1001, 1:28-33, 1:37-39, 1:48-52, 2:7-15, 2:59-60, 3:47-4:8, 4:34-52, 5:58-67, 7:45-55, 7:56-8:8, Figures 2A-2B.

27. The spatio-temporal model is a spatial model that is “extended over time” or “replicat[ed] ... along the temporal domain ... and then connect[ed] ... across time.” EX1001, 4:34-45, Figs. 2A-2B; EX1003, 7. The extension or connection over time is probabilistic, *i.e.*, based on the probabilities that the object is in preceding and/or subsequent images. The patent repeatedly explains that a spatio-temporal model has probabilistic connectivity, *e.g.*, “edges,” “constraints,” or “compatibilities” connecting the object in multiple frames. EX1001, 2:59-60, 3:47-4:8; 4:49-52; Figs. 2A and 2B; EX1003, 4-7 Figs. 2(a)-2(d). The use of object probabilities from multiple frames in the spatio-temporal model is further shown in the joint probability distribution for the spatio-temporal model described at 3:47-4:7, which expressly includes the object state at multiple points in time ($X_0^O \dots X_T^O$) (*i.e.*, in each image) and probabilistic constraints for the object state across images $\psi_{ij}(X_i^O, X_j^O)$ (*i.e.*, over time). EX1001, 3:47-4:7.

(*id.*, 7:6-22) to be achieved. The patent demonstrates achievement of these things using a spatio-temporal model, in comparison to a single-frame (non-spatio-temporal) model that had “poor” performance. EX1001, 7:56-8:3.

29. Dr. Bajaj’s construction also incorrectly interprets the “temporal” aspect of “spatio-temporal model” as being satisfied if the model is able “to change configuration in time.” EX1052, ¶14; EX2008 at 35:16-20 (“Q. Okay. So then the ability for the model to change configuration in time, is that what provides the temporal element of it being a spatio-temporal model? A. That’s correct, obviously.”).

16	Q.	Okay. So then the ability for the
17		model to change configuration in time, is that
18		what provides the temporal element of it being a
19		spatio-temporal model?
20	A.	That's correct. Obviously.

EX2008 at 35:16-20

The POSITA would understand that his construction is not consistent with how the term “spatio-temporal” is used in the ’980 patent. For example, the POSITA would understand that a purely spatial model can be used to model something (*e.g.*, a person or a car) at multiple discrete points in time. And if the configuration of the thing being modeled changes in time (*e.g.*, if it moves), the spatial model’s configuration

VL Collective Exhibit 2007

IPR2023-00924

Meta Platforms, Inc. v. VL Collective IP LLC

Page 35 of 72

likewise changes in time. For example, if you have a set of images containing a person, you can use a spatial model to model the person in the first image, then in the second, then in the third, etc. But, as the POSITA would understand, the fact that you can use the spatial model in each image on an image-by-image basis does not make the spatial model a spatio-*temporal* model as that term is used in the '980 patent—it's just a spatial model used on a set of images, one at a time. The POSITA would understand, based on the context of the specification and intrinsic evidence discussed above, that the *temporal* element of the spatio-temporal model is the extension of the model using probabilistic temporal connectivity for the object over time (*i.e.*, across images). EX1001, 2:49-50, 2:59-60, 3:47-4:7, 4:34-48, 4:49-52, 7:6-22, 7:56-8:3, Figs. 2A and 2B; EX1003, 4-7, Figs. 2(a)-2(d). Under Dr. Bajaj's construction of "spatio-*temporal* model," probabilistic temporal connectivity is not required. EX2008 at 56:19-57:7 ("Q. So you said 'spatial connectivity conditioned way.' Does it not require temporal connectivity conditions as well? A. Not when you're looking at a single frame.").

19 Q. So you said "spatial connectivity
20 conditioned way."
21 Does it not require temporal
22 connectivity conditions as well?

Page 57

1 A. Not when you're looking at a single
2 frame. You're looking at detecting an object in
3 that frame. You're saying is the component that
4 you're saying is likely to be the head, is it
5 close to the torso or maybe that component is
6 really a hand, because they have logs in the
7 image.

EX2008 at 56:19-57:7

Dr. Bajaj's construction of spatio-temporal model therefore does not align with the POSITA's understanding of the patent's description of what a *spatio-temporal* model is within the meaning of the '980 patent.

VII. PRIORITY DATE OF THE '980 PATENT

30. Dr. Bajaj does not address the priority date of the '980 patent in his supplemental declaration. *See* EX1052. Thus, I understand that he does not dispute my opinions demonstrating that the claims of the '980 patent have written

description support in the provisional application (EX1003). I understand that Petitioners nevertheless dispute whether elements 1/9[d] and 1/9[e] (and only those elements) have written description support in the provisional application. *See* Reply at 5-7.

31. Petitioners' assertion that my written description support opinions are "conclusory" is untrue. Reply, 6. I provided a detailed analysis that identified specific parts of the provisional application that indicate to the POSITA that the inventors were in possession of the claimed invention. EX2004 ¶¶35-36.

32. Petitioners' argument that the provisional does not include the '980 patent's Figure 5 flowchart or the exact text of element 1/9[e] ignores the evidence that I cited. Reply, 6. I am informed that written description support does not require the exact claim language to appear in the provisional. Based on that understanding, I identified portions of the provisional application that indicate to the POSITA that the inventors possessed the claimed invention. For example, as I explained, the provisional describes an "inference algorithm" that uses a "probabilistic image likelihood model $\phi_i(X_i)$," which is "the likelihood of observing the [] image conditioned on the state X_i of node i ." EX1003, 9. It further states that "[t]he output of the AdaBoost classifier is a confidence $h_k(I)$ " that is converted to a likelihood function, *i.e.*, a probability, where "[i]t is customary to consider object present if

$h_k(I) \geq \frac{1}{2} \sum_{k=1}^K \alpha_k.$ ” EX1003, 10. Based on these disclosures (and the others identified in my initial declaration), the POSITA would understand that that the inventors were in possession of element 1/9[d].

33. As also explained in my initial declaration, the provisional application further states that “[o]bject detection and tracking are formulated as inference in a two-layer graphical model,” where “[i]nferring the state of the object and its components in our framework is defined as estimating a belief in a graphical model” using “a form of non-parametric belief propagation.” EX1003, 4, 8. “The approach is a generalization of particle filtering” where “the conditional distribution used in standard particle filtering is replaced by product of incoming messages,” where “we integrate information from all components in the second-layer node,” (*i.e.*, in the object) and send messages between “the set of neighbors of node i” (as well as between frames) based on “the local evidence (or likelihood) associated with the node i” ($\phi_i(X_i)$) and “the potential designating the compatibility between the states of node i and j” (which can be in different frames). EX1003, 8-9. In other words, the object is detected based on object probabilities from multiple frames. Furthermore, a “w-frame windowed smoothing algorithm” can be used for object detection, “where w is an odd integer ≥ 1 (see Figs. 2c and 2d),” to achieve “temporal integration” and “temporal consistency” between frames.” *Id.* Based on these

disclosures (and the others identified in my initial declaration), the POSITA would understand that the inventors were in possession of element 1/9[e].

VIII. DR. BAJAJ DOES NOT SHOW THAT ANY CHALLENGED CLAIM WOULD HAVE BEEN OBVIOUS

A. The challenged claims are patentable over Zhao (Ground I).

1. Claims 1/9

a. Elements 1/9[e]

34. Dr. Bajaj's supplemental declaration does not address (and so I understand does not dispute) my opinions regarding Ground I and the Zhao reference. I understand that Petitioners rely on Dr. Bajaj's new claim construction arguments to argue that Zhao disclose element 1/9[e]. Accordingly, as I understand it, the dispute boils down to two issues: (1) whether Patent Owner's or Petitioners' claim constructions are correct, and (2) whether Zhao's RCR algorithm detects an object in a given image based on the probability that the object is in another image using a spatio-temporal model.

35. As to the first issue, I explained how the POSITA would interpret element 1/9[e] in my initial declaration (EX2004), and I supplemented that explanation herein in response to Dr. Bajaj's opinions. I also explained herein how the POSITA would interpret "spatio-temporal model" in response to Dr. Bajaj's

opinions. In doing so, I explained how Dr. Bajaj’s constructions are inconsistent with how the POSITA would interpret the claims in light of the intrinsic evidence.

36. As to the second issue, I previously explained that Zhao’s RCR process operates on a single image and does not detect an object (a human) in the image under consideration based on the probability that the object (the human) is in another image. EX2004 ¶¶107-112. I also explained that Zhao’s body model has no temporal probabilistic connectivity. EX2004 ¶¶111, 126. Petitioners’ new arguments fail to show otherwise.

37. Petitioners’ argument that “Zhao is directed to much more than a ‘simple fusion scheme’” is incorrect. Reply, 8. Zhao’s RCR algorithm does not fuse or combine *object* probabilities from *multiple* frames. The probabilistic analysis is confined to the *single frame* under consideration, as I explained in my previous declaration. EX2004, ¶¶81-85, 107-112. Zhao’s body model also does not combine object probabilities from multiple frames—the elements of the *body* model have no temporal components at all, as indicated by the lack of temporal indicia. See EX1006 at 43-45; EX2004, ¶¶75-78, 111, 126.

part is that the location of the origin becomes invariant to the body part's orientation. In summary, 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.

Assuming that the human body model consists of m body parts that are parameterized with vectors $f_1, f_2, \dots, f_m$, where $f_i = (a_i, l_i, x_i, y_i, \theta_i)$. 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 $\Theta = \{\theta_1, \dots, \theta_m\}$. Obviously, A and S are TRS-invariant.

EX1006 at 43

independent of each other. Based on the above assumptions, the probability distributions of the three model matrices A, S, U can be estimated separately:

$$A \sim \mathcal{N}(\bar{A}, \Sigma_A) \quad (2.4)$$

$$S \sim \mathcal{N}(\bar{S}, \Sigma_S) \quad (2.5)$$

$$U \sim \mathcal{N}(\bar{U}, \Sigma_U). \quad (2.6)$$

Assuming that the variations of the aspect ratios of the body parts are statistically independent², then

$$N(\bar{A}, \Sigma_A) = \prod_i N(\bar{a}_i; \sigma_{a_i}^2) \quad (2.7)$$

Assuming that the variations of the ratios of the lengths between the body parts and a reference body part f_i are independent³, then

$$N(\bar{S}_i, \Sigma_{S_i}) = \prod_{j \neq i} N(\bar{s}_{ji}; \sigma_{s_{ji}}^2) \quad (2.8)$$

where $S_i = \{s_{ji}\}$, $s_{ji} = l_j/l_i$, $j = 1, \dots, m$, $j \neq i$. Thus, S_i is used to constrain the relative lengths of the body parts with respect to the length l_i of a reference part f_i .

EX1006 at 43-44.

The POSITA would understand that because Zhao's body model and RCR algorithm do not include temporal probabilistic connectivity, Zhao does not disclose a spatio-temporal model. Petitioners' argument that it does rests on Dr. Bajaj's incorrect construction of "spatio-temporal model," which encompasses a purely spatial model with no temporal connectivity (not spatio-temporal) that is used on each single frame independently. Reply, 8 (citing EX1052, ¶¶6-35). As I explained above in Section VI, Dr. Bajaj's construction is incorrect.

38. Petitioners' argument that the POSITA would understand Zhao's use of motion information to generate a "prediction from previous frames (constant angular velocity assumed)" as indicating that the RCR algorithm contour updating and prediction procedure would "contain" a comparison of the object probability in the previous frame to a threshold is incorrect. Reply, 10-11. As explained above and in my first declaration, the probabilistic analysis in Zhao's RCR algorithm operates on a single frame at a time. And while the RCR process can use some limited information from previous frames, Zhao indicates that such information is "motion information." EX1006, 91-92. Such motion information does not include

probabilities and creates no temporal probabilistic connectivity. The POSITA would therefore understand that Zhao's "motion information," *e.g.*, "how the left knee angle changes with time" (*i.e.*, angular velocity) does not include object or component probabilities from other frames. *Id.* Nor did I "acknowledge" that Zhao teaches probabilistic temporal connectivity, as required by the spatio-temporal model in element 1/9[a] and element 1/9[e], during my deposition. Reply, 11. To the contrary, during my deposition I repeatedly explained that Zhao does not teach the use of probabilistic temporal connectivity and that motion information does not include probabilities from other image frames. EX1051 at 180:3-194:2; *see also generally id.* at 167-251. Accordingly, Zhao's RCR algorithm contour updating and prediction process does not "contain" probabilities from other image frames.

2. Claims 4/12

39. Petitioners claim that Patent Owner ignores "Zhao's disclosure" of using measurements from "adjacent frames," where "adjacent frames" (according to Petitioners) "necessarily includes previous and later frames." Reply, 11. Yet Petitioners cite no "disclosure" in Zhao (and I am aware of no such disclosure) regarding "adjacent frames" in Zhao. *Id.*; Pet. 29; EX1004, ¶107; EX1052 ¶¶36-39. Petitioners make a new argument in their Reply based on Zhao's discussion of "adjacent features," but Petitioners misinterpret Zhao. The "adjacent features" are

“features” within model (analogous to a model’s parts or components) used to perform “feature matching” in the image frame under consideration—there is no discussion of adjacent frames or looking to adjacent frames, let alone previous *and* later frames. EX1006 at 90.

40. Dr. Bajaj’s opinions that “it would have been obvious to extrapolate from later frames” do not “clarify” his prior opinions—they are new opinions based on new evidence and obviousness rationale. EX1052 ¶¶36-39 (newly citing EX1015 at 7, EX1011 at 3 and Fig. 5). As I previously explained, Dr. Bajaj’s first declaration provided no explanation as to how the references he discussed disclose or render obvious the specific limitations of claim 4/12. EX2004, ¶119. In response, Dr. Bajaj discusses new portions of evidence and attempts to explain his reasoning for the first time. EX1052 ¶¶36-39 (newly citing EX1015 at 7, EX1011 at 3 and Fig. 5). Dr. Bajaj’s new opinions and evidence do not show that claim 4/12 was disclosed or obvious.

41. EX1011 addresses a technique for “tracking a non-rigid deforming object” to purportedly improve over prior techniques in which “it is not possible to extrapolate *forward* in time.” EX1011, 1-2 (emphasis added). The technique is based on modal analysis (using “vibration modes”), and it generates “‘trained’ vibration modes” from a set of training data that do not assume constant elasticity of

the object. EX1011, 1-6, 9. The training data includes observations of nodal (point) displacements for the object at multiple points in time. *Id.* at 2-3. It's use of training data "assume[s] an observed set of matched, aligned shape vectors consisting of nodal (or point) positions" observed at multiple points in time. *Id.* This technique for tuning vibration modes using training data does not teach or suggest determining the probability that an object is in an image, nor doing so by detecting the object in a current image using measurements of the object from prior and later images.

42. His new evidence from EX1015 is a sentence that mentions "forward-backward smoothing" in an article co-authored by one of the inventors of the '980 patent. The forward-backward smoothing mentioned in EX1015 relates to extending a model to track components (*i.e.*, follow detected components across frames), but the component detection is still performed within a single frame on a frame-by-frame basis ("static detection"). EX1015 at 2, 4, 6-7. The authors proposed a "loose-limbed" body model where "each limb is treated quasi-independently with soft constraints between the position and orientation of adjacent parts," but indicated that detection techniques were "beyond the scope of this paper." *Id.* at 4. This loose-limbed body model does not teach or suggest determining the probability that an object is in an image or doing so by detecting the object in a current image using measurements of the object from prior and later images.

3. Claims 8/16

43. As I previously explained, claim 8/16's joint probability distribution includes elements for a spatio-temporal model that do not exist in Zhao's framework (thus Zhao does not disclose it), and Dr. Bajaj fails to show why it would have been obvious to modify Zhao to include them. EX2004, ¶¶122-127. Dr. Bajaj does not dispute my opinions that Zhao, Steffens, and Ozer do not disclose or render obvious claims 8/16, as he focuses entirely on the TLP reference. EX1052, ¶¶65-85.

44. Dr. Bajaj argues that lacking certain specific parts of the claimed distribution “does not affect the *formula* of the joint probability distribution.” EX1052, ¶66 (emphasis added). And he argues that “the same joint probabilistic distribution *formulas*” as claims 8/16 are allegedly disclosed in alleged prior art. *Id.*, ¶¶67-68 (emphasis added). In other words, Dr. Bajaj construes claim 8/16 as only requiring a generic joint probability distribution *formula*—regardless of the distribution's actual parts. However, Dr. Bajaj cites no intrinsic evidence that the POSITA would interpret claims 8/16 that way. The POSITA would not distill claim 8/16 to a generic “formula” by removing terms and notations that give specific meaning to the formula shown in the claims. To the contrary, as I explained in my previous declaration, the POSITA would give meaning to each term in claim 8/16's joint probability distribution, which includes terms specifically for the state of the

object X^O (indicated with superscript “ O ”), the state of each of the plurality of components X^C (with superscript “ C_0 ” through “ C_N ”), local evidence ϕ for the object state X^O , local evidence ϕ for each component state X^{C_k} , and three potential functions ψ_{ij} , ψ_{ik} , and ψ_{kl} , which represent object-to-object temporal compatibility, object-to-component spatial compatibility, and component-to-component spatial compatibility, respectively. EX2004, ¶¶122-127. The formula of claim 8/16 includes these terms, and Dr. Bajaj does not show that any prior art discloses that formula or renders it obvious. Dr. Bajaj attempts to argue that his cited references teach all of terms of claims 8/16 based the idea that the “object” includes the “components.” EX2008 at 86:18-87:7 (“But the object is a collection of components.”), 90:7-11 (“The object is the sum of its parts.”), 92:19-93:1 (“The object is the collection of components.”).

18 Q. So in claim 8, the size and phi in
19 claim 8 distinguish between object state and
20 component state, right? Isn't that what the
21 superscript "O" and "C" mean? Those correspond
22 to whether it's phi of something corresponding to

Page 87

1 the object or the component?

2 A. Yeah. But the object is a
3 collection of components. And that's why I tried
4 to explain it with a figure on Page 48, is to
5 detect an object which is a collection of
6 components as also a graphical model and also a
7 spatio-temporal model. Same.

EX2008 at 86:18-87:7

7 Q. Okay. And so since we're talking
8 about a potential function in the superscript
9 "O," we're just talking about the object, not the
10 ones that are superscript "C," right?

11 A. The object is the sum of its parts.

EX2008 at 90:7-11

19	Q.	So the difference between the object
20		state XO and the component state XC, those are
21		different states, right?
22	A.	Not at any point in time. The
Page 93		
1		object is the collection of components.

EX2008 at 92:19-93:1

But this interpretation reads out any distinction between the “object” terms and the “component” terms, which renders the formula in claim 8/16 mathematically non-sensical. As discussed in Section VI above, the POSITA would appreciate and respect the patent’s distinction between the “object” and the “components” in the claims, not ignore it.

45. Despite their obvious differences to the POSITA, Dr. Bajaj asserted during his deposition that the formulas in his cited references are “exactly the same” as claim 8/16, which, as the POSITA would know, is mathematically impossible. EX2008 at 86:2-6.

2	Q.	So you think the formula shown in
3		claim 8 is exactly the same as the others that
4		you show in your declaration in Paragraphs 67 and
5		68?
6	A.	Exactly.

EX2008 at 86:2-6

Dr. Bajaj eventually admitted that the formulas in his cited references would have to be rewritten in order to be like that of claim 8/16. EX2008 at 94:7-96:3.

7 Q. So in claim 8 of the '980 patent,
8 that formulation separates out object from
9 component by using a superscript O and C, right?

10 A. Just a small rewriting of that
11 formula, which is already there. It's not very
12 obvious. It's just rewriting the same formula in
13 a slightly expanded fashion.

14 Q. I see. So you would have to rewrite
15 the formula to do it like --

16 A. No, no. It's --

17 Q. -- like the one --

18 A. I'm not rewriting it. It's what the
19 patent owners rewrote by looking at what was
20 there in the prior art and saying, Oh, let's just
21 rewrite it at this nomenclature.

22 But it captures the same totality.

1 It is the same form, one over Z times pairwise
2 times singular potential functions.

3 Q. Okay. So what's shown in claim 8 is
4 written differently than what's shown in
5 Paragraph 67 and 68 of your declaration, right?

6 A. Marginally different; it's just
7 expanded slightly. But it's very obvious to a
8 person of ordinary skill in the art that the form
9 and structure of the formula is identical. It
10 captures the same joint probability distribution
11 over spatio-temporal object.

12 Q. So you're saying the form and
13 structure is identical, but you have to rewrite
14 it. I don't understand.

15 A. I'm not saying you have to rewrite
16 it. You've rewritten it in a slightly different
17 way. That different way is you put superscripts
18 for the time and subscripts for the spatial
19 indexing.

20 And there it's combined, but they
21 are all over all edges. And the reference in the
22 figure I showed you, and if you look at other

1 figures in these other arts, which I did not
2 repeat, it's absolutely the same calculation of
3 belief propagation.

94:7-96:3

But the POSITA would understand that the formulas in the cited references cannot simply be rewritten or expanded to yield claim 8/16 because they do not include representations of each piece of information in claim 8/16, including those I identified above.

46. Regardless, Dr. Bajaj does not argue that Zhao discloses or renders obvious claim 8/16 even under his new construction. EX1052, ¶¶65-85. For the reasons explained above and in my initial declaration, it does not.

4. All dependent claims are patentable over Zhao.

47. I understand that dependent claims include every limitation of the prior claims that they depend on. Accordingly, it is my opinion that claims 2-8 and 10-16 are patentable over Zhao for the same reasons as claims 1 and 9.

B. The challenged claims are patentable over Steffens in view of Zhao (Ground II).

1. Claims 1/9

a. Elements 1/9[d]

48. Dr. Bajaj admits that Steffens’s similarity measure “is not explicitly stated as calculating a ‘probability’ that an object is in an image.” EX1004, ¶144. To the extent Dr. Bajaj was suggesting that Steffens’ approach is implicitly probabilistic, he never explained how Steffens supports that argument. Dr. Bajaj now attempts to argue for the first time that Steffens’ discussion of “similarity” is actually a discussion of “probability.” EX1052, ¶42. As the POSITA would understand, however, it is not. Dr. Bajaj attempts to show compatibility between Steffens’ graph matching similarity measure and Zhao’s Bayesian (probabilistic) similarity measure by identifying attenuated parallels between them, *see* EX1052 ¶¶43-51, 54-55, but those parallels do not change the fact that Steffens’ approach is deterministic and not probabilistic and therefore fundamentally different from and incompatible with Zhao’s. The key error in Dr. Bajaj’s analysis is that he incorrectly characterizes “the normalized scalar product of the amplitude vector” of two jets in Steffens as a probability. EX1052, ¶¶52-53. As the POSITA would know from basic linear algebra, however, normalizing the absolutes of two vectors does not yield a probability. *See* EX1007 at 8:56-67.

49. Steffens' wavelet-based approach processes "[a]n original image [] using a Gabor wavelet to generate a convolution result" (*i.e.*, a Gabor-based wavelet), which is a "two-dimensional complex wave field modulated by a Gaussian envelope." EX1007 at 8:14-67. The wavelet has a wave vector restricted by a Gaussian window. *Id.* All wavelet components centered in a single image point are considered as a vector which is called a jet. *Id.* Each component of a jet is the filter response of a Gabor wavelet extracted at a point (x, y) of the image. *Id.* "The normalized dot product of the absolute components of two jets defines the jet similarity." *Id.*

50. As the POSITA would understand, the vectors that Steffens normalizes—which are filter responses of a Gabor wavelet—are not probabilities, and the result of normalizing them does not make them probabilities. *See* EX1007 at 8:14-67. Moreover, Steffens includes no discussion of a probabilistic similarity measure or any mention of probability. The word "probability" does not even appear in Steffens. Steffens' approach is clearly deterministic and has no probabilistic modeling whatsoever. Thus, Steffens does not simply use different terminology than Zhao or the '980 patent as Dr. Bajaj suggests—it employs a fundamentally different mathematical framework for a similarity measure that is incompatible with Zhao's formulation framework.

51. That Steffens and Zhao both relate to detecting and recognizing humans and their body parts does not change the fact that Steffens’ non-probabilistic graph matching approach is fundamentally different from and incompatible with Zhao’s probabilistic approach. Zhao’s teaching away from wavelet-based approaches like Steffens, which I discussed in my first declaration (EX2004, ¶130), underscores the incompatibility of their approaches, notwithstanding the attenuated parallels that Dr. Bajaj relies on.

b. Elements 1/9[e]

52. I understand that Petitioners rely on their arguments regarding claim construction, Ground II element 1/9[d], and Ground I element 1/9[e]. I’ve addressed each of those arguments above and incorporate them here.

2. Claims 4/12

53. Petitioners suggest that tracking “between frames” in Steffens “necessarily involves looking at a prior frame and a later frame relative to a current frame” to detect an object in the current frame. Reply 17-18. Petitioners’ evidence (portions of Steffens and Dr. Bajaj’s supplemental declaration) do not support their inherency argument. As I previously explained, Steffens detection analysis is performed on a single frame, not across or between frames. EX2004, ¶138. I understand that Dr. Bajaj does not dispute those points. Dr. Bajaj and Petitioners

rely on Steffens’ tracking analysis “between” two frames, but Steffens’ tracking analysis does not detect an object in a current frame based on measurements from a later frame *and* a prior frame—it tracks in the forward direction only. *See* EX1007 at 10:64-61. And Dr. Bajaj’s opinions regarding Steffens, which I addressed in Section VII.B.1 above, do not argue such inherency. I also see no support from Petitioners to support their assertion that linear motion prediction “necessarily infers object detection using multiple frames” is also unsupported. The POSITA would understand that an object need not be detected using measurements from prior and later frames just because the object is tracked across frames. For instance, Steffens is an example of one-direction tracking.

3. Claims 8/16

54. I understand that Petitioners rely on their Ground I arguments for Zhao as to this claim in Ground II. I addressed those arguments in Section VII.A above and incorporate them here.

4. All dependent claims are patentable over Steffens in view of Zhao.

55. I understand that dependent claims include every limitation of the prior claims that they depend on. Accordingly, it is my opinion that claims 2-8 and 10-16 are patentable over Steffens in view of Zhao for the same reasons as claims 1 and 9.

C. The challenged claims are patentable over Ozer in view of Zhao (Ground III).

1. Claims 1/9

a. Elements 1/9[a] and 1/9[e]

56. I understand that Petitioners rely on arguments which I’ve addressed above, and I incorporate those discussions here.

2. Claims 4/12

57. Dr. Bajaj’s argument that Ozer’s “activity” classification “encompasses object motion tracking” fails to show that the specific requirements of claim 4/12 are met. EX1052, ¶¶60. The claims require “detecting the object in a current image according to measurements of the object ... determined from a prior image *and* a later image relative to the current image.” EX1001, claim 4/12. But Dr. Bajaj identifies no disclosure in Ozer that teaches or suggests *detecting* the object in a current frame using measurements of the object from both a prior *and* a later image.

58. Instead, Dr. Bajaj argues that Ozer’s motion classification encompasses “object motion tracking” over multiple frames, and Petitioners further argue that the “low level” process (object detection) and “high level” process (activity classification) “work together to detect object parts.” EX1052, ¶¶60-64; Reply 19-20. Their suggestion that Ozer’s detection process uses Ozer’s motion classification process to detect objects (and therefore is based on multiple images) is incorrect.

Dr. Bajaj admits that activity recognition happens “after” the detection process and that Ozer tracks “only objects that fit the model,” *i.e.*, the previously detected objects. EX1052, ¶63. The fact that motion classification over multiple images uses the results of object detection in those images does not mean that object detection spans multiple images. And in Ozer, it does not. As I previously explained, Ozer’s object detection is based on information collected from a single frame, and the detection is made using that information—not information from a previous and a later frame. EX2004, ¶149; EX1008 at 3:65-67, 5:10-20, 6:19-24, 6:52-58.

3. Claims 8/16

59. I understand that Petitioners rely on their Ground I arguments for Zhao as to this claim in Ground III. I addressed those arguments in Section VIII.A above and incorporate them here.

4. All dependent claims are patentable over Ozer in view of Zhao.

60. I understand that dependent claims include every limitation of the prior claims that they depend on. Accordingly, it is my opinion that claims 2-8 and 10-16 are patentable over Ozer in view of Zhao for the same reasons as claims 1 and 9.

D. The challenged claims are patentable over TLP (Ground IV).

1. Alleged Public Accessibility of TLP

61. I understand that Petitioners recently produced a set of emails that they obtained from one of the co-authors of TLP, Michael Isard. Because they were produced only recently on July 10, 2024, I was unable to inspect them or provide opinions related to them in my initial declaration. I inspected some of Mr. Isard's emails that I understand relate to the submission of TLP to the 2004 CVPR conference peer review committee. Based on my review of those emails, it is my opinion that TLP was not made publicly accessible or disseminated outside of the 2004 CVPR peer review committee members responsible for facilitating its review.

62. In an email dated February 13, 2004, to Leonid Sigal (an author of TLP) from two CVPR 2004 Program Chairs (Aaron Bobick and Rama Chellappa), the Program Chairs shared information about the peer review process for 2004 CVPR papers. EX2009. The email indicates that the submitted paper was provided to peer reviewers, who then provided feedback on the paper that were reviewed by the Program Committee Members and Area Chair(s) in charge of review for TLP. It further indicates that the authors' response will be read only by the Area Chair(s) in charge of review for TLP and the Program Chairs. It further indicates that responses may not include any information that might reveal the identity of the paper's authors,

and that “the entire decision process will be conducted double-blind.” It further indicates that if the response contains any identifying information that it will be deleted in its entirety. An email between three of the TLP authors dated November 17, 2003 includes what appears to be a draft of TLP with author-identifying information removed. EX2010.

63. These emails are consistent with my experience as an author of many conference papers as well as a member of several conference committee where I have served in roles such as Technical Program Chair, Area Chair, Special Sessions Chair among others where technical papers are submitted and peer reviewed prior to acceptance. In my experience, papers submitted to technical conferences are managed by the Technical Program Chair(s) charged with facilitating peer review under strict protocols of confidentiality and in some instances “blind” review (*i.e.*, the reviewers do not know whose paper they are reviewing). It is customary and standard practice that papers submitted to a conference for review and potential acceptance are kept strictly confidential during that process, including to protect papers that are not accepted from public disclosure or public association with the authors. Indeed, disclosing a conference submission to the public or to anyone outside of the peer review committee personnel assigned to that submission would be considered unethical and a serious breach of trust that could result in disciplinary

actions, not to mention reputational harm to the violator among professional and academic peers. The blind review process, which ensures that the authors of a submission remain unknown, would also prevent peer reviewers from knowing (or divulging to others) a relationship between an author and a submission. In my experience, which I believe to be representative of academic and professional norms, papers submitted for acceptance to a conference are not made available to the public, not included in a public catalog, public database, or public search tool, and are not publicized or disseminated to the public before they are officially published by the conference (if they are accepted, otherwise, not at all).

64. Based on the information that I have reviewed, and in light of my personal experience, it is my opinion that TLP had not been made publicly accessible or publicly disseminated as of May 27, 2004 (the priority date of the '980 patent). A draft submission of TLP appears to have been submitted (via a password protected portal) directly to official CVPR personnel that were responsible for facilitating confidential peer review under double-blind conditions. This is consistent with my experience and my understanding of academic/professional norms that are designed and intended to prevent public access to papers submitted for acceptance to a conference.

2. Claims 1/9

a. Elements 1/9[a]

65. As explained in Sections VI.A-B above, the “spatio-temporal model” of 1/9[a] is a “spatial model where an *object* is detected in an image based on the probabilities that the *object* is in preceding and/or subsequent images.” As I previously explained, TLP does not disclose detecting an *object* or determining the probability that an *object* is in each image—instead, it detects components (*i.e.*, limbs), which are distinct from the object. EX2004, ¶¶155-161. Accordingly, TLP does not disclose or render obvious the “spatio-temporal model” of 1/9[a].

b. Elements 1/9[d]

66. Element 1/9[d] requires the determination of an *object* probability. EX1001, claim 1/9. As I explained in Section VI above, in the claims of the '980 patent, the “object” is distinct from the “components.” Accordingly, element 1/9[d] requires the determination of a probability that the *object* (as distinct from the components) is in each image.

67. Petitioners construe the claims to “not require that an object is distinct from its corresponding components,” and on that basis argue that TLP “determines probability that an object is in an image based on the probability that *components* are in the image.” Reply, 7, 22. As explained in Section VI above, however, the

POSITA would not understand the claims that way. Accordingly, I disagree with Petitioners' analysis and conclusions.

68. In any event, Petitioners have not shown that TLP determines a probability that an object is in an image, even one “based on” component probabilities. TLP detects components (*i.e.*, “limbs” or “parts”) based on a probability of component presence (*i.e.*, a “limb likelihood”) and assembles the components into a “lose-limbed body model.” EX1009, 1-4. The component parts are modeled and assembled into a body model, where each node in the model is a component part with a likelihood function. *Id.* But TLP does not disclose a determination of a body (*i.e.*, object) likelihood or probability. That TLP does not disclose an object probability is confirmed by the fact that its model includes no body node or determination of a distinct body likelihood. That is not to say that the claims require a body node—rather, the fact that TLP (which uses a graphical model comprising nodes for limbs) has no body node and does not determine a distinct body likelihood means that TLP does not disclose element 1/9[d], in particular the requirement of an object probability.

69. Dr. Bajaj suggests the '980 patent discloses a scenario where the object is not modeled distinctly from its components, EX1052, ¶76, but he misinterprets the patent. According to the patent, when, even when “the whole appearance of an

object is too complicated to model as a whole ... and can only be modeled in terms of components” (e.g., because the object is too complicated to train a likelihood proposal for its entire appearance), “a uniform likelihood can be assumed over the object state space” for the object layer. *Id.* at 6:13-22; EX1051 at 71:12-75:20. In other words, the spatio-temporal object model still includes a distinct object layer/object node (and a distinct “object state space” with its own distinct “likelihood”), but the probabilistic distribution of the object over its components is assumed to be uniform (rather based on trained relationships). *Id.* This indicates to the POSITA that even in this alternative scenario, the model still represents the object as distinct from components and includes a distinct probability for the object. That is, even when an object cannot be modeled other than in terms of its components, the *spatio-temporal* model of the ’980 patent differs from *traditional* models because it nonetheless models the object as distinct from its components and having its own, distinctly modeled probability. This disclosure reinforces that separately modelling objects and collection of components is necessary and inherent to the claimed spatio-temporal model.

70. Dr. Bajaj’s new argument that the POSITA would have allegedly understood from PAMPAS (EX1019) to select a component in TLP to represent the object (as a second layer of nodes) is incorrect. His argument is based on his

incorrect construction of the claims that does not distinguish the “object” in the claims from the “components.” As the POSITA would understand, treating a component as arbitrarily “representative” of the object does not make that component’s probability a probability of the object. Moreover, PAMPAS does not teach treating a component as the object—it analyzes only components (the “central circle,” “arms,” and “links” are components) and estimates belief for an occluded component (i.e., the central circle node) based on “agreement of the nearby arm nodes” using belief propagation. EX1019, 6-8. It does not, however, select a component to be representative of an object node, as Dr. Bajaj incorrectly suggests.

c. Elements 1/9[e]

71. As I previously explained, TLP does not disclose element 1/9[e] because it does not disclose determining an object probability—only component probabilities. EX2004 ¶¶155-156, 159-161.

72. I understand that Petitioners argue that the claims do not require a difference between object probability and component probability. Reply, 26. In my opinion, Petitioners are again relying on their (and Dr. Bajaj’s) claim constructions to artificially remove any distinction between the “object” and “components” expressly recited in the claims, which, as I explained in Section VI above, is not how the POSITA would understand the claims.

73. I understand that Petitioners also rely (again) on Dr. Bajaj’s opinion that object probability in TLP “can be” determined “based on” component probability. Reply, 26. But, as I explained in Section VIII.D.2.a above with respect to claim 1/9[d], TLP does not determine an object probability—not even one “based on” component probabilities.

3. Claims 3/11

74. Petitioners do not appear to dispute that claim 3/11 requires nodes corresponding to components *and* a node corresponding to the object. Reply, 26-27. But Petitioners and Dr. Bajaj have not identified any nodes in TLP that correspond to the *object* (*i.e.*, the body)—only nodes corresponding to *components* (*i.e.*, body parts). *Id.*; Pet. 67 (“each graph node corresponds to a body part (upper leg, torso, etc.)”), 72; EX1004, ¶203. To the extent Petitioners’ and Dr. Bajaj’s argument is that TLP’s body part nodes (*i.e.*, component nodes) are also nodes that correspond to the object, they are again removing the distinction between “object” and “components” in the claims, which, as I explained in Section VI above, is not how the POSITA would understand the claims.

4. Claims 8/16

75. I understand that Petitioners and Dr. Bajaj assert that “the joint probability distribution equation of claim 8/16 is identical to prior disclosures in the

art.” Reply, 27; EX1052, ¶¶67-68; EX2008 at 86:2-6 (“Q. So you think the formula shown in claim 8 is exactly the same as the others that you show in your declaration in Paragraphs 67 and 68? A. Exactly.”). I disagree.

2	Q.	So you think the formula shown in
3		claim 8 is exactly the same as the others that
4		you show in your declaration in Paragraphs 67 and
5		68?
6	A.	Exactly.

EX2008 at 86:2-6

76. Petitioners and Dr. Bajaj agree that the equation in claims 8/16 includes terms for both the object state (X_i^O) and component states ($X_i^{C_k}$). Pet. 76; EX1004, ¶231; EX1052, ¶82; EX2008 at 86:18-87:7 (confirming that terms with superscript “O” are for the object and with superscript “C” are for components).

18 Q. So in claim 8, the size and phi in
19 claim 8 distinguish between object state and
20 component state, right? Isn't that what the
21 superscript "O" and "C" mean? Those correspond
22 to whether it's phi of something corresponding to

Page 87

1 the object or the component?

2 A. Yeah. But the object is a
3 collection of components. And that's why I tried
4 to explain it with a figure on Page 48, is to
5 detect an object which is a collection of
6 components as also a graphical model and also a
7 spatio-temporal model. Same.

EX2008 at 86:18-87:7

Yet Petitioners and Dr. Bajaj do not identify a joint probability distribution in the alleged prior art disclosures with both an object node state *and* component node states. The distributions that Dr. Bajaj identifies in paragraphs 67 and 68 of his supplemental declaration do not distinguish between an object and components, as shown by their lack of nomenclature or terminology to make such a distinction—as well as by their obvious lack of the actual terms shown in claim 8/16. EX1052, ¶¶67-68 (citing EX1012 and EX1019). Dr. Bajaj and Petitioners do not appear to be

disputing that TLP does not disclose object level nodes, which is what X_i^O corresponds to in claim 8/16.

77. To the extent Dr. Bajaj and Petitioners argue that TLP's body part nodes (*i.e.*, component nodes) are also object nodes, they are again removing any distinction between "object" and "components" in the claims. Dr. Bajaj indicated as such in his deposition by suggesting that his cited references disclose the object-specific terms of claim 8/16 on the basis that "[t]he object is the sum of its parts" and "[t]he object is the collection of components." EX2008 at 90:7-11, 92:19-93:1.

7	Q.	Okay. And so since we're talking
8		about a potential function in the superscript
9		"O," we're just talking about the object, not the
10		ones that are superscript "C," right?
11	A.	The object is the sum of its parts.

EX2008 at 90:7-11

19	Q.	So the difference between the object
20		state X_O and the component state X_C , those are
21		different states, right?
22	A.	Not at any point in time. The

Page 93		
1		object is the collection of components.

78. As I explained in Section VI above, that is contrary to how the POSITA would interpret the claims.

5. All dependent claims are patentable over TLP.

79. I understand that dependent claims include every limitation of the prior claims that they depend on. Accordingly, it is my opinion that claims 2-8 and 10-16 are patentable over TLP for the same reasons as claims 1 and 9.

80. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001, of Title 18 of the United States Code.

Executed on this 24th day of July 2024, in Rochester, NY.



Eli Saber