

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.

U.S. Patent No. 7,436,980

IPR2023-00924

PETITIONERS' REPLY TO PATENT OWNER'S RESPONSE

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	CLAIM CONSTRUCTION	1
	A. 1/9[e] Does Not Require Multi-Image Analysis	1
	B. Claim 1/9's "Spatio-Temporal Model" Is Not a Graphical Model with Nodes Corresponding to Each of Collection of Components and to Object	4
III.	TLP WAS A PUBLICLY AVAILABLE PRINTED PUBLICATION BEFORE THE '980 PATENT'S EARLIEST PRIORITY DATE.....	5
IV.	'980 PATENT IS OBVIOUS	7
	A. Ground I: Zhao Renders the Challenged Claims Unpatentable.....	8
	B. Ground II: Zhao + Steffens Render the Challenged Claims Unpatentable.....	14
	C. Ground III: Zhao + Ozer Render the Challenged Claims Unpatentable.....	18
	D. Ground IV: TLP Renders the Challenged Claims Unpatentable	21
V.	SECONDARY CONSIDERATIONS	27

TABLE OF AUTHORITIES**Page(s)****Cases**

<i>Apple Inc. v. Corephotonics, Ltd.</i> , 81 F.4th 1353 (Fed. Cir. 2023)	2
<i>Liebel-Flarsheim Co. v. Medrad, Inc.</i> , 358 F.3d 898 (Fed. Cir. 2004)	5
<i>Neville v. Found. Constructors, Inc.</i> , 972 F.3d 1350 (Fed. Cir. 2020)	4
<i>Parus Holdings, Inc. v. Google LLC</i> , 70 F.4th 1365 (Fed. Cir. 2023)	5
<i>Phillips v. AWH Corp.</i> , 415 F.3d 1303 (Fed. Cir. 2005)	2, 4, 5
<i>Raytheon Co. v. Roper Corp.</i> , 724 F.2d 951 (Fed. Cir. 1983)	2
<i>Virnetx, Inc. v. Cisco Sys., Inc.</i> , 767 F.3d 1308 (Fed. Cir. 2014)	5
<i>VLSI Tech. LLC v. Intel Corp.</i> , 53 F.4th 646 (Fed. Cir. 2022)	2

TABLE OF EXHIBITS **

Exhibit No.	Description
1001	U.S. Patent No. 7,436,980 ('980 Patent)
1002	File History of U.S. Patent No. 7,436,980
1003	Provisional Application No. 60/574,799
1004	Declaration of Dr. Chandrajit Bajaj
1005	Curriculum Vitae of Dr. Chandrajit Bajaj
1006	Liang Zhao, Dressed Human Modeling, Detection, and Parts Localization, The Robotics Institute, Carnegie Mellon University (2001) ("Zhao")
1007	U.S. Patent No. 6,301,370 ("Steffens")
1008	U.S. Patent No. 7,200,266 ("Ozer")
1009	Leonid Sigal, Sidharth Bhatia, Stefan Roth, Michael J. Black & Michael Isard, <i>Tracking Loose-limbed People</i> , IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2004) ("TLP")
1010	Program Documents for the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2004), http://tab.computer.org/pamitc/archive/cvpr2004/index.html
1011	Adam Baumberg & David Hogg, <i>Generating Spatiotemporal Models from Examples</i> , School of Computer Studies, University of Leeds, http://www.bmva.org/bmvc/1995/bmvc-95-040.pdf ("Baumberg")
1012	Erik B. Sudderth, Alexander T. Ihler, William T. Freeman & Alan S. Willsky, <i>Non-parametric Belief Propagation</i> , IEEE Conference on Computer Vision and Pattern Recognition, (June 2003), https://www.ics.uci.edu/~ihler/papers/tr2551.pdf)
1013	Declaration of Jessica Benner regarding "Zhao", Carnegie Mellon University (April 4, 2023)

Petitioners' Reply to Patent Owner's Response

Exhibit No.	Description
1014	Declaration of Carl Megeski regarding “Zhao,” ProQuest (March 2, 2023)
1015	Leonid Sigal, Michael Isard, Benjamin H. Sigelman & Michael J. Black, <i>Attractive people: Assembling loose-limbed models using non-parametric belief propagation</i> , Advances in Neural Information Processing Systems 16, NIPS 1539–46 (MIT Press 2003)
1016	Hedvig Sidenbladh, Michael J. Black & David Fleet, <i>Stochastic tracking of 3D human figures using 2D image motion</i> , European Conference on Computer Vision, 702–18, LNCS 1843 (Springer-Verlag 2000)
1017	Noel A. C. Cressie, <i>Statistics for Spatial Data</i> , Chapters 1, 9, John Wiley and Sons (1993)
1018	Yaakov Bar-Shalom, X. Rong Li & Thiagalingam Kirubarajan, <i>Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software</i> , Chapters 1–2, 11, John Wiley and Sons (2001)
1019	M. Isard, <i>PAMPAS: Real-valued graphical models for computer vision</i> , CVPR, vol. 1, at 613–20 (2003)
1020	D. L. Alspach & H. W. Sorenson. <i>Nonlinear Bayesian estimation using Gaussian sum approximations</i> , IEEE Trans. AC, 17(4): at 439–48 (August 1972)
1021	James M. Coughlan & Sabino J. Ferreira, <i>Finding deformable shapes using loopy belief propagation</i> , ECCV 453–68 (2002)
1022	F. De la Torre & M. J. Black., <i>Robust parameterized component analysis: Theory and applications to 2D facial modeling</i> , ECCV 653–69 (2002)
1023	S. L. Lauritzen, <i>Graphical Models</i> , Chapters 1, 3, Oxford University Press, Oxford (1996)

Petitioners' Reply to Patent Owner's Response

Exhibit No.	Description
1024	A. Doucet, N. de Freitas, & N. Gordon, editors. <i>Sequential Monte Carlo Methods in Practice</i> , Chapters 5, 16, Springer-Verlag, New York (2001)
1025	Vladimir Pavlovic, James M. Rehg, Tat-Jen Cham & Kevin Murphy, <i>A dynamic Bayesian network approach to figure tracking using learned dynamic models</i> , Int'l. Conf. on Computer Vision at 94–101 (1999)
1026	Y. Wu, G. Hua & T. Yu, <i>Tracking articulated body by dynamic Markov network</i> , Int'l. Conf. on Computer Vision at 1094–101 (2003)
1027	Christopher Bregler & Jitendra Malik, <i>Tracking people with twists and exponential maps</i> , CVPR, at 8–15 (1998)
1028	J. Deutscher, A. Davison & I. Reid, <i>Automatic partitioning of high dimensional search spaces associated with articulated body motion capture</i> , IEEE Computer Society Conference on Computer Vision and Pattern Recognition, at 669–76 (2001)
1029	Cristian Sminchisescu & Bill Triggs, <i>Covariance scaled sampling for monocular 3D body tracking</i> , IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, at 447–54 (2001)
1030	John MacCormick & Michael Isard, <i>Partitioned sampling, articulated objects, and interface-quality hand tracking</i> , European Conference on Computer Vision (2), at 3–19 (2000)
1031	Yacov Hel-Or & Michael Werman, <i>Constraint Fusion for Recognition and Localization of Articulated Objects</i> , International Journal of Computer Vision 19(1), at 5–28 (1996)
1032	Y. Weiss & W. T. Freeman, <i>Correctness of belief propagation in Gaussian graphical models of arbitrary topology</i> , Neural Comp., 13:2173–2200 (2001)

Petitioners' Reply to Patent Owner's Response

Exhibit No.	Description
1033	M. Fischler & R. Elschlager, <i>The representation and matching of pictorial structures</i> , IEEE. Trans. Computers, 22(1):67–92 (1973)
1034	M. Burl, M. Weber & P. Perona, <i>A probabilistic approach to object recognition using local photometry and global geometry</i> , European Conference on Computer Vision at 628–41 (1998)
1035	P. Felzenszwalb & D. Huttenlocher, <i>Efficient matching of pictorial structures</i> , IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, at 66–73 (2000)
1036	D. Ramanan & D. Forsyth, <i>Finding and tracking people from the bottom up</i> , IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. II, at 467–716 (2003)
1037	Andrew Blake & Michael Isard, <i>Active Contours</i> , Chapters 7, 8, Springer-Verlag (1998)
1038	D. L. Alspach, <i>Gaussian Sum Approximations in Nonlinear Filtering and Control</i> , Information Sciences 7, 271–290 (1974)
1039	A. R. Tilley, <i>The Measure of Man and Woman: Human Factors in Design</i> H.D. Associates (1993)
1040	Paul Viola & Michael Jones, <i>Rapid Object Detection using a Boosted Cascade of Simple Features</i> , Conference on Computer Vision and Pattern Recognition (2001)
1041	Hedvig Sidenbladh & Michael J. Black, <i>Learning Statistics Learning the Statistics of People in Images and Video</i> , International Journal of Computer Vision 54 (1/2/3), at 183–209 (2003)
1042	Jonathan S. Yedidia, William T. Freeman & Yair Weiss, <i>Generalized Belief Propagation</i> , Advances in Neural Info. Proc. Sys. 13, at 689–95 (2000)

Exhibit No.	Description
1043	<i>VideoLabs, Inc. et al v. Meta Platforms, Inc. et al</i> , Case No. 1-22-cv-00680, Dkt. 15 (D. Del. Aug. 29, 2022)
1044	Patent Assignment for U.S. Patent No. 7,436,980
1045	Patent Assignment for U.S. Patent No. 7,436,980
1046	Application Data Sheet 37 C.F.R. 1.76 for U.S. Patent No. 9,189,886
1047	Patent Assignment for U.S. Patent No. 7,436,980
1048	<i>CVPR Technical Program: Call for Papers</i> , IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 27 – July 2, 2004
1049	ResearchGate Web Page for TLP
1050	Declaration of Dr. Sylvia Hall-Ellis relating to TLP
1051	Deposition Transcript of Eli Saber, Ph.D. (May 30, 2024)
1052	Supplemental Declaration of Dr. Chandrajit Bajaj
1053	U.S. Patent No. 6,597,801 to Cham <i>et al.</i>

** Unless otherwise indicated, all emphasis added.

I. INTRODUCTION

Patent Owner's ("PO") Reply ("POR") only argues that claims 1/9[e], 1/9[d], 3/11, 4/12, and 8/16 are not rendered obvious by Grounds I-IV, conceding the remaining Challenged Claims are obvious as set forth in the Petition. PO's last-ditch effort to rescue the '980 Patent by proffering a "new" claim construction is an unavailing attempt to repackage the same argument this Board previously rejected as to element 1/9[e]—that 1/9[e] requires comparing multiple probabilities from multiple images to a threshold. Paper-11, 26-28. PO further impermissibly reads limitations from the specification into independent claims 1/9, which do ***not*** require (i) a two-layer graphical model, (ii) use of the spatio-temporal model to detect the object (*see* Paper-11, 37), or (iii) motion information to be based on probability of the presence of the object. PO's arguments lack merit and should be rejected. At bottom, the '980 Patent is an obvious extension of TLP (Ex1009), as acknowledged in the provisional application (Ex1003, 6), and rendered obvious by other known detecting/tracking technology to identify humans/cars in videos.

II. CLAIM CONSTRUCTION

A. 1/9[E] DOES NOT REQUIRE MULTI-IMAGE ANALYSIS

PO's confusing rewrite of 1/9[e] to require a multi-image analysis is unsupported by intrinsic evidence. The '980 Patent notes that "a need exists for a system and method of detecting and tracking an object" and "performing ***inference*** over space and time." Ex1001, 1:22-24. The specification further notes "[w]hile it

is possible to perform *inference* over the spatio-temporal model defined for the entire image sequence, there are many applications for which this is not an option due to the lengthy off-line processing needed.” *Id.*, 7:6-9. In such instances, a window-smoothing algorithm can be used where w is an odd integer ≥ 1 ; where $w=1$, the window-smooth algorithm resembles single frame component-based fusion. *Id.*, 7:10-22; Paper-11, 6-7. Ex1052, ¶¶6-22, 23-35.

Contrary to PO and its expert, Dr. Saber, the '980 Patent is *not* “distinguish[ing] the $w=1$ scenario from the invention” (POR, 22), but generally describing ways of “performing inference over space and time” with and without “lengthy off-line processing.” Ex1001, 1:22-24, 7:6-9; Ex1052, ¶¶23-35; Paper-11, 26-27; *Phillips v. AWH Corp.*, 415 F.3d 1303, 1324-27 (Fed. Cir. 2005); *see also Apple Inc. v. Corephotonics, Ltd.*, 81 F.4th 1353, 1359 (Fed. Cir. 2023) (no indication in claims, specification, or otherwise that patentee meant to claim invention more narrowly). Indeed, the '980 Patent itself discloses that the window-smoothing algorithm is used where w is an odd integer ≥ 1 —*i.e.*, greater than *or equal to* 1. PO's reliance on “Experiments” in the specification fares no better, particularly when read cohesively with the intrinsic record. Ex1001, 7:56-8:3; POR, 21-23; Paper-11, 27; *VLSI Tech. LLC v. Intel Corp.*, 53 F.4th 646, 652 (Fed. Cir. 2022) (claims should not be limited to specific examples from specification); *Raytheon Co. v. Roper Corp.*, 724 F.2d 951, 957 (Fed. Cir. 1983)

("[t]hat claims are interpreted in light of the specification does not mean that *everything* expressed in the specification must be read into all the claims").

As the Board aptly explained, a construction requiring comparing multiple probabilities from multiple images to a threshold—which is what PO's construction attempts to do by adding "of the plurality of images" after "each image" (POR, 18)—is inconsistent with the specification. Paper-11, 27.

Ex1001, Figure 5, 505	Claim 1/9[e]
detecting the object in an image upon comparing the probabilities for the image to a threshold for detecting the object	detecting the object in any image upon comparing the probabilities for each image to a threshold for detecting the object

Figure 5 at 505 discloses "**an** image" and "**the** image" in describing a computer implemented method for object detection. Claim 1/9[e] uses identical language—except for the phrases "**any** image" and "**each** image." PO's suggestion that "**each** image" must mean two or more images when the specification's parallel disclosure describes "**the** image" in Figure 5 makes no sense. POR, 22; Ex1051, 114:16-129:12.

Moreover, in the Notice of Allowance, the Examiner noted that close prior art Cham (Ex1053) teaches providing a kinematic 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, performing feature search operations for object features by comparing "a total probability with

a predetermined threshold.” Ex1002, 163; Ex1008, 7:37-42. The Examiner (incorrectly) stated that Cham “does not teach the claimed ‘spatio-temporal model.’” Ex1002, 163. Therefore, the Examiner understood Cham to disclose element 1/9[e] as claimed, and that the Examiner therefore understood 1/9[e] to require a frame by frame comparison of a threshold to the probability that the object is in the respective frame. Ex1052, ¶¶6-22; *cf. Neville v. Found. Constructors, Inc.*, 972 F.3d 1350, 1356 (Fed. Cir. 2020) (affirming claim construction because mutual understanding reached with examiner reinforced meaning, which was intended to distinguish prior art).¹

B. CLAIM 1/9’S “SPATIO-TEMPORAL MODEL” IS NOT A GRAPHICAL MODEL WITH NODES CORRESPONDING TO EACH OF COLLECTION OF COMPONENTS AND TO OBJECT

Although not explicitly offered as a claim construction position, PO nevertheless characterizes “spatio-temporal model” in independent claims 1/9 as a “two-layer graphical model,” where a coarse layer node represents the whole object and fine layer nodes represent multiple components of the object. POR, 1-5; Ex1001, 2:55-59. Again, this impermissibly reads limitations from the

¹ PO additionally and unnecessarily replaces “upon” with “based on” in 1/9[e]. POR, 24-25. PO cites *no* support (in the specification or in caselaw) for “based on,” or why or how it should override the unambiguous word “upon.” *Phillips*, 415 F.3d at 1327-28 (cannot override an unambiguous construction).

specification into the claims. *Phillips*, 415 F.3d at 1324-27; *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 904 (Fed. Cir. 2004). As Dr. Saber acknowledged, other spatio-temporal models existed before the priority date of the '980 Patent. Ex1051, 59:6-62:13; *see also* Ex1052, ¶¶6-22. Moreover, PO's equating "spatio-temporal model" as a two-layer graphical model with a coarse layer node and a fine layer node effectively reads dependent claims 3/11 into independent claims 1/9. Ex1001, 8:30-43, 8:48-51 ("wherein the spatio-temporal model is a graphical model comprising nodes corresponding to each of the collection of components and to the object"). But claim differentiation disfavors reading a limitation from a dependent claim into an independent claim. *Virnetx, Inc. v. Cisco Sys., Inc.*, 767 F.3d 1308, 1316 (Fed. Cir. 2014).

III. TLP WAS A PUBLICLY AVAILABLE PRINTED PUBLICATION BEFORE THE '980 PATENT'S EARLIEST PRIORITY DATE

PO argues as of the filing date of the provisional application (Ex1003) on May 27, 2004, the earliest priority date of the '980 Patent, "the inventors had possession of the claimed subject matter." POR, 7. Petitioners reserved the right to challenge the May 27, 2004 priority date or any evidence of prior invention. Pet., 5. Under PO's interpretation and claim construction of 1/9[d]-[e], the provisional does not provide written description under pre-AIA §112. *Parus Holdings, Inc. v. Google LLC*, 70 F.4th 1365, 1374 (Fed. Cir. 2023) (patentee had failed to prove priority claim because its published patent application did not have

adequate written description support for challenged claims). Likewise, PO cannot demonstrate that the provisional discloses 1/9[e] or 1/9[d]. POR, 8-9; Ex2004, ¶¶35-36; Ex1051, 39:17-66:14.

For example, PO points to the provisional Abstract, which discusses “an inference over directed loopy graph,” solved with “non-parametric belief propagation that uses a variation of particle filtering.” Ex1003, 3, POR, 9. But beyond Dr. Saber’s conclusory statement that “a POSITA would have understood that the inventors were in possession of this element” 1/9[e], PO offers no other evidence. Ex2004, ¶¶35-36; Ex1051, 39:12-60:10. Indeed, the provisional provides no written support for the flowchart at ’980 Patent, Figure 5. Ex1051, 39:17-40:2; Ex1003, 9-10. The only discussion about “threshold” in the provisional is with regards to AdaBoost, but without further disclosure of “detecting the object in an image upon comparing the probabilities for the image to a threshold for detecting the object.” Ex1003, 9-10; Ex1001, Figure 5. Similarly, the additional disclosures in the provisional that PO and Dr. Saber identify also do not have sufficient written description support for 1/9[e] or 1/9[d]. POR, 8-9; Ex1003, 8-10, Ex2004, ¶¶35-36. Given the provisional lacks sufficient written description support as to at least 1/9[e] and 1/9[d], PO is *not* entitled to the benefit of the May 27, 2004 provisional filing date as the earliest priority date of the ’980 Patent.

Petitioners' Reply to Patent Owner's Response

Even assuming the '980 Patent can claim priority to May 27, 2004, TLP was still a printed publication prior art. TLP was indisputably published in 2004, and cited to in the provisional. Ex1009; Ex1050, ¶¶42-44; Ex1003, 12; Ex1051, 106:22-107:14. Petitioners' evidence of record shows that not only was TLP publicly available at least as of April 9, 2004 (Ex1004, ¶¶65-66; Ex1010), TLP was available even earlier as of January 31, 2004 (Ex1050, ¶¶40-48; Paper-28). TLP is accordingly prior art to the '980 Patent.

IV. '980 PATENT IS OBVIOUS

PO argues that the '980 Patent's approach to object detection and tracking is different than Ground I (Zhao), Ground II (Zhao and Steffens), Ground III (Zhao and Ozer), and Ground IV (TLP) because each Ground fails to teach element 1/9[e] and claims 8/16 under PO's construction and interpretation. However, as explained above, PO's construction is wrong and entirely unsupported by the intrinsic record. *See supra* Section II. For Ground IV in particular, TLP teaches a graphical model of the body comprising multiple components, despite PO's incomplete and incorrect analysis. The claims and claim elements raised by PO under Ground IV do not require that an object is distinct from its corresponding components.

**A. GROUND I: ZHAO RENDERS THE CHALLENGED CLAIMS
UNPATENTABLE**

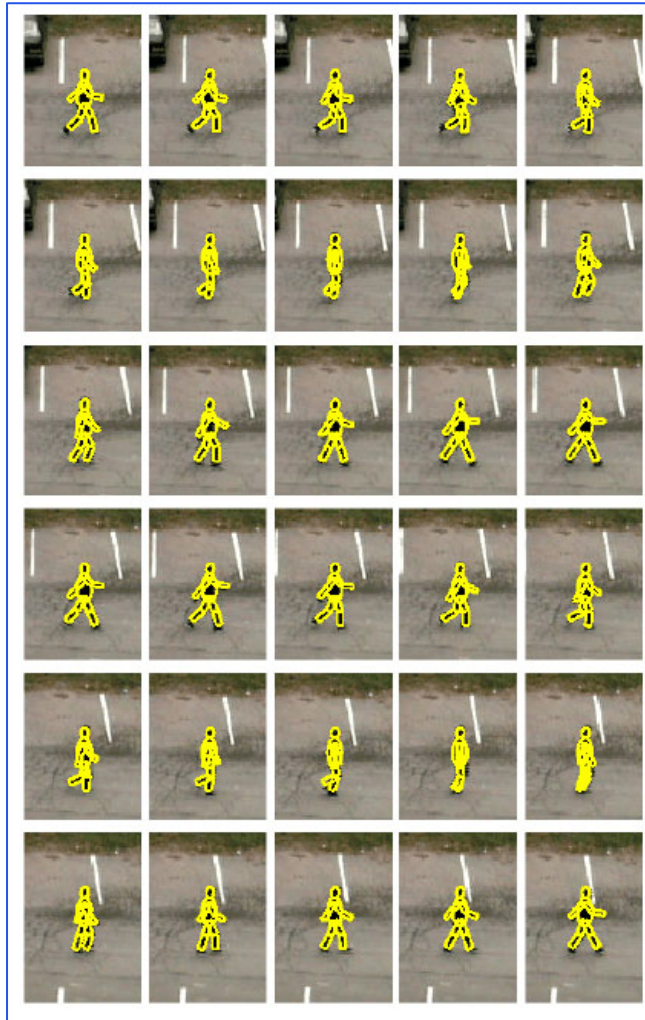
1. Claim 1/9[e]

PO argues at the onset, but merely in passing, that Petitioners fail to address element 1/9[e] because Petitioners' analysis of 1/9[e] was "based on their analysis" of element 1/9[d]. POR, 26. PO has not actually articulated any deficiency in Petitioners' analysis of 1/9[d], and PO is incorrect in arguing that Petitioners never explained "how comparing the probability that a contour corresponds to a person to a threshold." POR, 27; *but see* Pet., 25-27; Ex1004, ¶¶91-100; Paper-11, 26-29. In short, PO's entire argument as to 1/9[e] is based on its new claim construction that 1/9[e] requires multi-image analysis because this limitation purportedly requires "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," *i.e.*, comparing multiple probabilities from multiple images to a threshold. POR, 26-30.

Zhao is directed to much more than a "simple fusion scheme" as PO suggests. POR, 30-34. As noted in Section II.B *supra*, independent claims 1/9 do not require a graphical model or a two-layer model. And despite what PO argues, Zhao *does* apply a temporal analysis from frame to frame. Ex1052, ¶¶6-22, 23-35. PO concedes that Zhao determines the probability that a human is in a single image, and compares that probability to a threshold for detecting the human in that

image. POR, 28-29. However, PO argues that Zhao's RCR algorithm "does not detect a human in the image under consideration based on the probability of the human being detected in *other* images" because Zhao's algorithm uses a contour updating procedure to predict the contours of missed parts and align them to generate redefined contours. *Id.*, 31-32. Not so, and PO ignores Zhao's teachings as well as the knowledge of a POSITA. Ex1052, ¶¶6-22, 23-35.

Zhao notes that human motion capture plays a role in "a wide spectrum of applications," and that the earliest computer vision attempted to recognize human movements. Ex1006, 89. Zhao also knows that the RCR algorithm is useful for vision tasks such as object tracking as well by incorporating motion information into the model. *Id.*, 99. Such an object being tracked in a video could be a person walking or a vehicle moving. Zhao further describes experimental results concerning locating the body parts of a person walking in a parking lot. Zhao explicitly notes that "[m]otion information obtained from previous frames can be used to predict the orientations of the limbs and to solve the ambiguity." Ex1006, 91 ("Fig. 5.6 presents a full cycle of a walking person."):



In the cycle of a person walking, the first half of the cycle involves no motion information, while the second half involves “the *prediction from previous frames* (constant angular velocity is assumed)” in order to get better results of body part localization. *Id.*, 92. PO ignores this disclosure in Zhao and states that “motion information” does not include the probability of the object being present in another image, or comparing the probability to a detection threshold. POR, 33. A POSITA would understand that if the cycle of a person walking involves prediction from previous frames, not only is there a temporal component, the RCR

algorithm contour updating procedure and prediction procedure would contain comparison of that probability in the previous frame to a detection threshold. The RCR algorithm is then repeated for every frame. Ex1006, 56, 91-95. Each body part is not identified independently in Zhao (nor do the elements 1/9[e] require it). Further, Dr. Saber acknowledged that Zhao has a temporal component. *See generally* Ex1051, 167-251.

As discussed earlier, the '980 Patent's $w=1$ temporal window embodiment is nevertheless part of the window smoothing algorithm discussed in the specification, where w is an odd integer ≥ 1 (greater than *or equal to* 1). It is therefore incorrect that PO argues a scenario where $w=1$ is excluded from the purported invention. Dr. Bajaj's opinion, unlike PO's, is not based on an incorrect interpretation of the $w=1$ embodiment. Ex1052, ¶¶23-35.

2. Claim 4/12

PO argues claims 4/12 require a three-frame analysis, and Zhao's MAP estimation disclosure is merely transforming coordinate systems. PO, however, ignores Zhao's disclosure that in analyzing data in image frames from a video using measurements from adjacent frames, that necessarily includes previous and later frames. As Dr. Bajaj further clarifies, it would have been obvious to extrapolate from later frames. Ex1052, ¶¶36-39. Even Zhao notes that a "common feature is that the uncertainty information has been used for feature matching and

for determining transformation uncertainty with which to predict the positions of adjacent features.” Ex1006, 90. A POSITA would have understood that adjacent features would involve looking at prior and later images in order to predict positions associated with said features. It was well known to extrapolate from adjacent frames, and accordingly, determine probability that the object is in each image based on a prior image and a later image relative to the current image. Ex1052, ¶¶36-39; *see also generally* Ex1051, 192-226.

3. Claim 8/16

For claims 8/16, PO argues that Zhao's framework lacks certain elements of a spatio-temporal model as required by the claims, *e.g.*, object node and distinct component node at distinct points in time; a temporal dimension representing distinct points in time. Not so. Ex1004, ¶121; Ex2006, 115:5-117:20; Ex1052, ¶¶65-85. As Dr. Bajaj previously explained, which Dr. Saber has ignored, the equation of claim 8/16 represents a typical factorized form of the joint probability distribution $P(\dots)$ for a graphical model G of an object O with state X_i^O at time frame I of the video and $\{k = 1 \dots N\}$ components state X_i^{G-k} . Ex1004, ¶¶120-122. Petitioners are not only arguing obviousness of claims 8/16 in view of Zhao and the knowledge of POSITA. Zhao's teachings disclose a joint probability represented by the claimed equation, and any differences in the representation of the information and relationship would have been obvious. Ex1052, ¶¶65-85.

A joint probability distribution for the spatio-temporal model with N components is:

$$P(X_0^O, X_0^{C_0}, X_0^{C_1}, \dots, X_0^{C_N}, \dots, X_T^O, X_T^{C_0}, X_T^{C_1}, \dots, X_T^{C_N}) = \frac{1}{Z} \prod_{ij} \psi_{ij}(X_i^O, X_j^O) \prod_{ik} \psi_{ik}(X_i^O, X_i^{C_k}) \prod_{ikl} \psi_{kl}(X_i^{C_k}, X_i^{C_l}) \prod_i \phi_i(X_i^O) \prod_{ik} \phi_i(X_i^{C_k}).$$

Contrary to Dr. Saber's statement, the lack of a certain specific node or having extra nodes does not affect the formula of the joint probability distribution. This joint probability distribution is disclosed in the prior art. Ex1009, 9 ("we use a form of non-parametric belief propagation"); Ex1019, Ex1012; Ex1052, ¶¶65-85.

$\mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$. The joint probability over $\mathcal{X}$ is given by

$$P(\mathcal{X}) = \frac{1}{Z} \left(\prod_{(i,j) \in \mathcal{P}} \psi_{ij}(x_j, x_i) \right) \left(\prod_i \phi_i(x_i) \right). \quad (1)$$

Specifically, Sudderth teaches the same joint probabilistic distribution formulas of claims 8/16, using the same belief propagation approach as the '980 Patent. Although the joint probability distribution equation of claim 8/16 appears to have more variables, a POSITA would understand that the information conveyed is identical. Therefore, a POSITA would also understand that the joint probability

distribution P (...) of Zhao would similarly convey this information, which could be rewritten as shown in claim 8/16. Ex1052, ¶¶65-85.

B. GROUND II: ZHAO + STEFFENS RENDER THE CHALLENGED CLAIMS UNPATENTABLE

1. Claim 1/9[d]

PO argues Ground II does not disclose element 1/9[d], PO argues no motivation to combine Zhao with Steffens because Steffens employed a wavelet-based approach, while Zhao was contour-based. POR, 41-43. PO further alleges that Zhao taught away from Steffens because Zhao noted that a wavelet-based approach would be computationally expensive. *Id.* PO is incorrect. Steffens' "similarity" measure is related to probability. Ex1052, ¶¶40-57.

Contrary to PO and Dr. Saber, the Bayesian Similarity model of Zhao is not incompatible with object detection using wavelet analysis. As Dr. Bajaj previously explained, probabilistic frameworks for object detection were widely known in the prior art. Ex1004, ¶144. Such probabilistic frameworks were frequently used as part of tracking objects in image video sequences. *Id.* Steffens uses a model graph 168 that is designed for finding a human face in an image. Ex1007, 9:1-4. Steffens further discloses the use of similarity functions for the identification of the human face in video data, and Figure 14 teaches a "coarse to fine" approach for detecting the face. *Id.*, 1:57-60, 10:3-7; Ex1004, ¶¶136-137; Ex1052, ¶¶40-57.

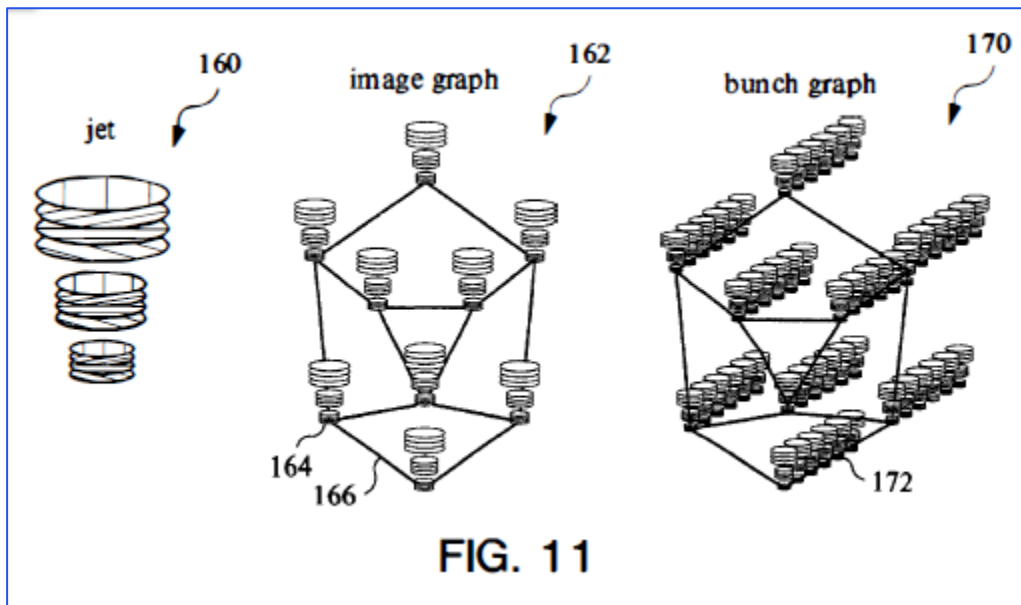
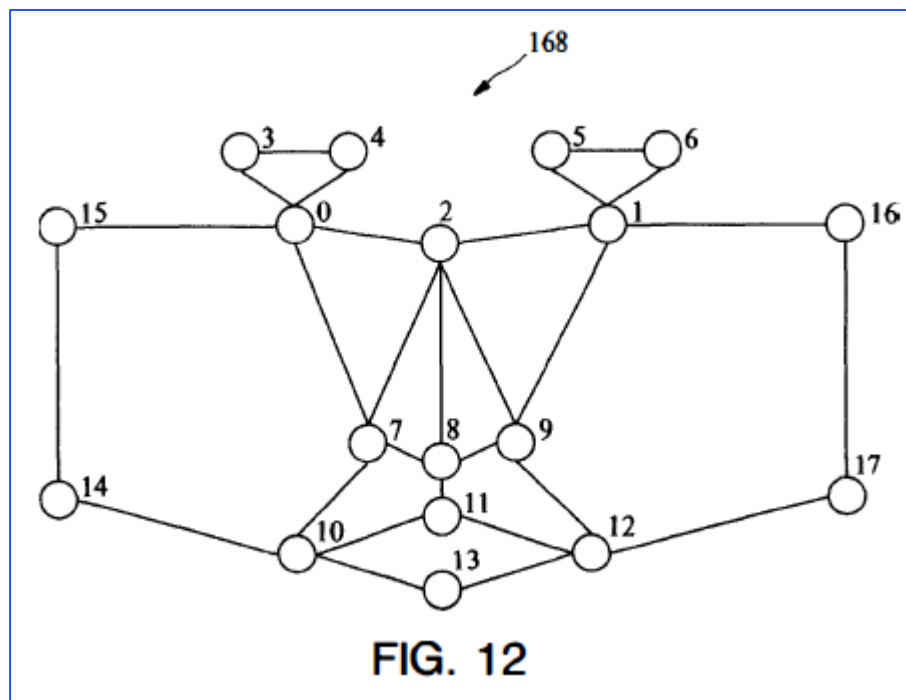


Figure 12 of Steffens additionally illustrates matching and tracking. *See also* Ex1004, ¶55.



Although Dr. Saber argues conclusorily that Zhao's contour-based approach would be "incompatible" with Steffens' wavelet-based approach (Ex2004, ¶129),

this is an overly rigid stance that would not be taken by a POSITA. Ex1052, ¶¶40-57.

Zhao broadly notes that the object for object detection is applicable to humans, and would shed light on general object detection and shape analysis. Ex1006, 19. Zhao performs “Part Identification” and “Human Detection” where decomposed parts are then matched to an object model to calculate the probability of the object’s presence in images. *Id.* Although Steffens focuses more on the human face for facial recognition, both Steffens and Zhao concern object detection and recognition in video images based on spatio-temporal models and probabilistic matching. Zhao explains that, for example, a human body can consist of six main body parts and twelve subparts. Ex1006, Figure 1.7. Steffens employs a bunch graph that breaks down a human face into a collection of components. Ex1007, Figure 12. Neither reference teaches away from the other. Just because wavelet analysis may be slow in certain instances, a POSITA would have known the benefit of Steffens’ wavelet approach, and why such approach would work specifically in recognizing facial features. Ex1052, ¶¶40-57.

2. Claim 1/9[e]

For element 1/9[e] (like in Ground I), PO makes a procedural argument that the Petition failed to address “every limitation,” and that even if both Zhao and Steffens use a “similarity measure,” a POSITA would not have combined Steffens’

wavelet-based approach with Zhao's contour-based approach. POR, 43-46. But as explained Section IV.B.1 *supra*, Petitioners and Dr. Bajaj addressed why a POSITA would have been motivated to combine the teachings of Zhao and Steffens, contrary to PO's argument that a POSITA would not have found the specific requirements of element 1/9[e] to be obvious. POR, 44-45. Again, PO's position for element 1/9[e] in Ground II is premised on its flawed and unsupported claim construction requiring multi-image analysis. Section II.A *supra*. For the reasons already explained as to Ground I, the combination of Zhao and Steffens renders obvious element 1/9[e]. *See* IV.A.

3. Claim 4/12

PO also argues Ground II does not render dependent claims 4/12 obvious. PO argues Steffens' disclosure about "between frames" does not involve detecting an object in a current frame according to a prior frame and later frame. POR, 46-49. According to PO, Steffens' "elastic graph matching technique" is for a single image. Dr. Saber, however, recognizes that Steffens uses a tracking technique involving a linear motion prediction. *See generally* Ex1051, 252-266, 259:15-22. However, linear motion prediction necessarily infers object detection using multiple frames. Ex1007, 10:62-11:12 (discussing predicting position $X_{(n+1)}$ from position X_n). A POSITA would have known the benefits of extrapolating motion information from prior and later frames. Ex1052, ¶¶36-39. Accordingly, a

POSITA reading Steffens would know to extrapolate from a linear motion prediction, which necessarily involves looking at a prior frame and a later frame relative to a current frame.

4. Claim 8/16

For claims 8/16, PO argues because Ground II did not show elements 1/9[d]-[e] were rendered obvious, it similarly could not show that claims 8/16 were obvious. POR, 49. As described above in Ground I, and with respect to elements 1/9[d]-[e] in Ground II, Zhao and Steffens would render obvious claims 8/16. *See* Sections II.A, IV.A, IV.B.1-IV.B.2.

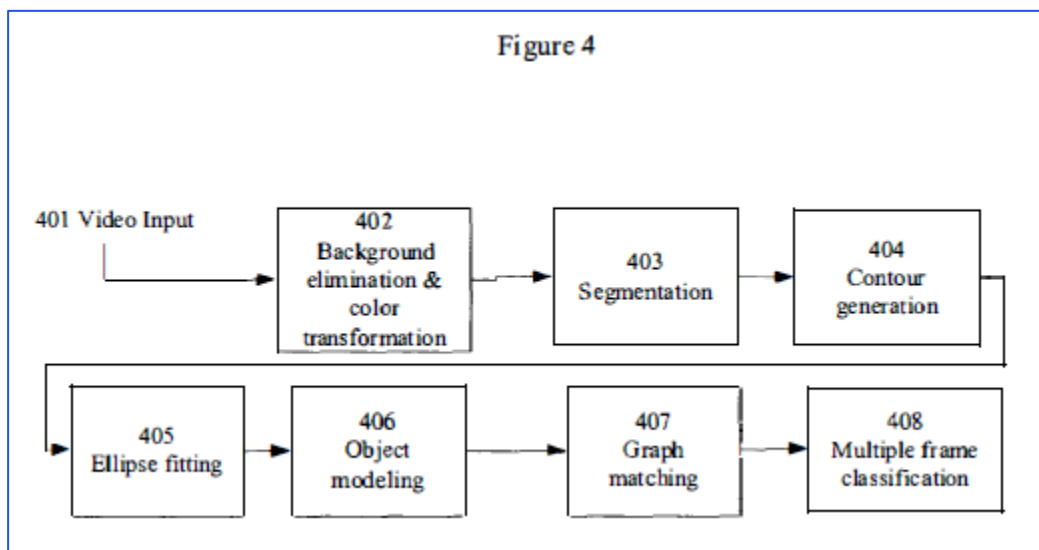
C. GROUND III: ZHAO + OZER RENDER THE CHALLENGED CLAIMS UNPATENTABLE

1. Claim 1/9[e]

For element 1/9[e], PO refers back to its argument as to deficiencies in Zhao (Ground I), but makes no additional argument. POR, 50. PO otherwise does not separately argue why Ozer or Zhao in view of Ozer does not render element 1/9[e] obvious. As Petitioners address above, PO's contention that element 1/9[e] are not rendered obvious is based on a flawed and unsupported claim construction. Accordingly, the analysis under Ground I and with respect to element 1/9[e] generally is incorporated by reference, and this element is disclosed and at least obvious for the same reasons. *See* Sections II.A, IV.A.

2. Claim 4/12

For claims 4/12, PO argues Ozer's method of classifying multiple frames is not detecting an object, but rather describing a process of checking movements and calculating probabilities of activities. POR, 50-53. Further, Ozer does not teach frames before or after, and instead talks about frames being processed in real time. *Id.* A POSITA would have understood that "activity" classification encompasses object tracking. Ex1052, ¶¶58-64. Ozer explicitly describes a "system of the present invention can **detect** non-rigid (e.g. human body) and rigid object parts and recognize their activities." Ex1008, 3:55. In Ozer, the non-rigid object detection infers that the object is moving. For instance, Figure 4 illustrates a video input involving multiple frames ("multiple frame classification"):



Ozer further explains that the object "is decomposed for its presentation as a combination of component shapes. The result will be unaffected by a partial

occlusion of the object.” Ex1008, 4:2-6. Ozer thus describes how the low level and high-level processes work together to detect object parts.

Dr. Saber acknowledges that Ozer describes a system that “builds a model of the object of interest on each video frame.” Ex2004, ¶¶149-150; Ex1008, 6:19-27. Ozer notes that the disclosed system is “more accurate because it easily rejects many elements in the scene that may be moving but are not objects of interest.” If an object is moving and is rejected as not of interest, a POSITA would have understood that Ozer is describing more than just a single current frame analysis. Ozer further “relies on the combination of the modular parts that form logical semantics at each stage ... The combination of object parts *defines the object*.” Ex1008, 7:32-59; Ex1052, ¶¶58-64. Accordingly, Zhao in view of Ozer renders claim 4/12 obvious.

3. Claim 8/16

For claims 8/16, PO merely argues because element 1/9[e] is not met, neither are dependent claims 8/16. PO also references element 1/9[d], but makes no discussion as 1/9[d] for Ground III. POR, 53-54. As described above in Ground I, and with respect to elements 1/9[d]-[e] in Ground II, Zhao and Ozer here too would render obvious claims 8/16. *See* Sections II.A, IV.A, IV.B.1-IV.B.2.

**D. GROUND IV: TLP RENDERS THE CHALLENGED CLAIMS
UNPATENTABLE**

**1. Patent Owner Failed to Question Dr. Bajaj Regarding
Public Availability of TLP Based on CVPR 2004**

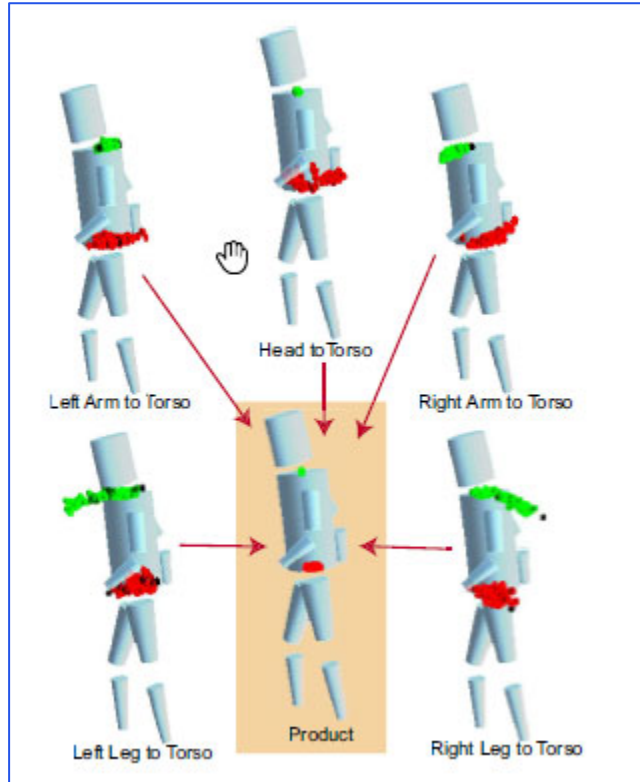
As noted above in Section III, PO has not shown the provisional provided written description support with respect to at least element 1/9[e] and 1/9[d]. Accordingly, claim 1/9, as well as all dependent claims of the '980 Patent, are not entitled to the priority date of the May 27, 2004 filing date of the provisional. Further, Petitioners put forth supplemental evidence (*see* Paper-28) showing that TLP was a publicly available printed publication as of at least January 31, 2004. Ex1050, ¶¶44-48.

PO now disputes Dr. Bajaj's Declaration and argues that Dr. Bajaj's "testimony" "fails to show that TLP was disseminated or publicly available before May 27, 2004 as part of the 2004 CVPR peer review process or otherwise." POR, 54-58. Yet when PO had the opportunity to further question Dr. Bajaj concerning his Declaration, PO failed to do so. *See* Ex2006, 118-133 (never discussing Ex1004, ¶¶65-66). Further, Petitioners' supplemental evidence (Ex1049, Ex1050) shows that the first-named author of TLP, Michael J. Black, uploaded TLP to ResearchGate and indicated a January 2004 date of publication. Based on Mr. Black's course of conduct, each of his uploads to ResearchGate and his identification of publication date closely aligned with an official publication date

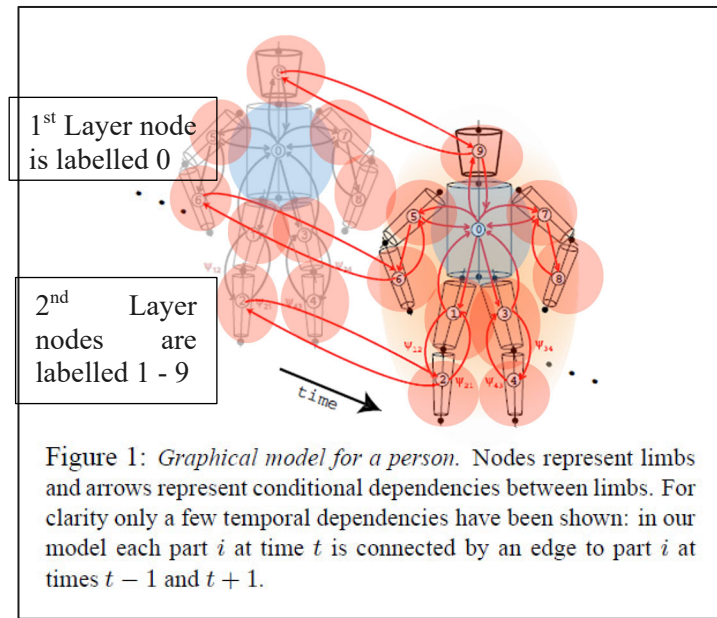
for the document. *See* Ex1050, ¶47. PO points to *SRI Int'l* and argues the facts and evidence presented here were “insufficient to meet the legal standard.” POR, 55. However, Dr. Bajaj already stated that as a presenter and fellow conference attendee at CVPR2004, papers that were selected for posters were widely disseminated and made publicly available by at least April 9, 2004. Together with Petitioners' additional evidence, the facts here are different and distinguishable from *SRI Int'l*.

2. Claim 1/9[d]

The provisional application to the '980 Patent does not provide written support description for element 1/9[d]. PO cites to the same evidence and excerpts of the provisional for element 1/9[d] and element 1/9[e]. Yet these portions of the provision do not provide written description support for “determining a probability that the object is in each image.” Ex1004, 8-10; POR, 8; Ex2004, ¶35; Section III *supra*. In any event, TLP determines probability that an object is in an image based on the probability that components are in the image. Ex1009, Figure 4 (predicting torso location). Element 1/9[d] does not require a body node, and PO's argument again attempts to read limitations into claim 1 that do not exist.



TLP presents a probabilistic method for human detection and tracking in 3D. TLP uses a loose-limbed body model in which the limbs are connected via learned probabilistic constraints. Specifically, TLP describes a graphical model for detecting objects, such as a human, that comprise of nodes corresponding to the object's constituent parts. Edges in the graphical model correspond to position and angle relationships between adjacent body parts in space and time, as illustrated in Figure 1:



Tracking in TLP is formulated as “one of inference in a graphical model and belief propagation is used to estimate the pose of the body at each time step.”

Ex1009, 1. TLP indicates that the body is treated as a graphical model, where each node in the graph corresponds to an independently parameterized body part. The spatial constraints between body parts are defined as directed edges in the graph. Each edge has an associated conditional distribution that models the probabilistic relationship between the parts. Each node in the graph also has a corresponding image likelihood function that models the probability of observing various image measurements conditioned on the position and orientation of the part. Person detection (or tracking) then exploits belief propagation to estimate the belief distribution over the parameters which accounts for the constraints and the observations, identical to the '980 Patent. Id.

A POSITA would have understood from TLP how to resolve the graphical inference problem using Non-parametric Belief Propagation (NBP) or the PAMPAS algorithm that allows the parameters of the individual parts to be modeled using continuous-valued random variables. *Id.* at 4. The body is represented by a graphical model in which each graph node corresponds to a body part (upper leg, torso, etc.). See Figure 1. Each part has an associated configuration vector defining the part's position and orientation in 3-space. Placing each part in a global coordinate frame enables the part detectors to operate independently while the full body is assembled by inference over the graphical model. Edges in the graphical model correspond to spatial and angular and temporal relationships between adjacent body parts, as illustrated in Figure 1. As is standard for graphical models, a POSITA would have understood that the variables in a node are conditionally independent of those in non-neighboring nodes, given the values of the node's neighbors. *Id.* at 2. Ex1052, ¶¶6-22, 23-25, 65-85.

A POSITA would have understood to select a component of the object (e.g., the torso) to be representative of the object node x_1 (i.e., a 1st layer), and to conditionally link all other components of the body (head, limbs, etc.) forming a 2nd layer of x_i nodes. See PAMPAS, EX1019, § 3.3 (finding a jointed object in clutter). Due to this two-layered model, even when the object node x_1 is occluded,

after one iteration the belief at that node has high weight at the correct location due to the agreement of the nearby arm nodes x_i . After a few iterations the belief has converged on the correct object position despite the lack of the central circle. *Id.* Thus, TLP teaches “determining a probability that the object is in each image.” Ex1052, ¶¶6-23, 24-34, 65-85.

3. Claim 1/9[e]

PO further argues element 1/9[e] are not disclosed by TLP because TLP talks about limb (i.e., component) probability, but not object probability. According to PO, TLP's non-parametric belief propagation with message passing is between body parts, not object nodes. POR, 60-64. PO's construction again is based on a flawed claim construction. Element 1/9[e] does not further require a distinction between component probability versus object probability. As Dr. Bajaj has clarified, TLP does teach object probability—because object probability can be determined based on component probability. Ex1052, ¶¶6-23, 24-34, 65-85.

4. Claim 3/11

PO argues that dependent claims 3/11 (“spatio-temporal model is a graphical model comprising nodes corresponding to each of the collection of components and to the object”) is not rendered obvious. PO argues the Petition did not identify where TLP identifies object nodes, or that TLP discloses both component and object. POR, 64. PO impermissibly reads limitations of Figures 2A, 2B into the

claims. Claims 3/11 claim a “graphical model comprising nodes corresponding to each of the collection of components and to the object.” However, claims 3/11 do not require a two-layer graphical model, only that there are nodes corresponding to collection(s) of components and nodes corresponding to the object.

5. Claim 8/16

PO also argues Ground IV does not render obvious claims 8/16 because TLP does not disclose object-level nodes. POR, 65. As the Petition originally and already set forth, the joint probability distribution would have been known to a POSITA and the same pairwise Markov field representation of a spatio-temporal graphical model. Pet., 75-76. As further elaborated by Dr. Bajaj, the joint probability distribution equation is nothing new, and the information was already disclosed in the prior art. Ex1052, ¶¶6-23, 24-34, 65-85. In essence, the joint probability distribution equation of claim 8/16 is identical to prior disclosures in the art.

V. SECONDARY CONSIDERATIONS

Unsurprisingly, PO and Dr. Saber (Ex1051, 30:21-31:4) put forth no evidence of any secondary considerations that would support non-obviousness of the '980 Patent.

Petitioners' Reply to Patent Owner's Response

June 17, 2024

Respectfully submitted,

/s/ W. Todd Baker

W. Todd Baker (No. 45,265)

todd.baker@kirkland.com

Postal and Hand-Delivery Address:

KIRKLAND & ELLIS LLP

1301 Pennsylvania Ave., N.W.

Washington, D.C. 20004

T: (202) 389-5000

F: (202) 389-5200

Ellisen Shelton Turner (No. 54,503)

ellisen.turner@kirkland.com

Postal and Hand-Delivery Address:

KIRKLAND & ELLIS LLP

2049 Century Park East, 37th Floor

Los Angeles, CA 90067

T: (310) 552-4200

F: (310) 552-5900

Jonathan D. Brit (No. 75,646)

jonathan.brit@kirkland.com

Postal and Hand-Delivery Address:

KIRKLAND & ELLIS LLP

601 Lexington Avenue

New York, NY 10022

T: (212) 390-4209

F: (212) 446-4900

Jennifer R. Bush (No. 50,784)

jbush@fenwick.com

Postal and Hand-Delivery Address:

FENWICK & WEST LLP

801 California Street

Mountain View, CA 94041

Telephone: (650) 988-8500

Attorneys for Petitioners

CERTIFICATE OF COMPLIANCE

This Reply complies with the type-volume limitations as mandated in 37 C.F.R. §42.24, totaling 5,219 words. Counsel has relied upon the word count feature provided by Microsoft Word.

/s/ W. Todd Baker
W. Todd Baker
(No. 45,265)

CERTIFICATE OF SERVICE

The undersigned hereby certifies that a copy of the foregoing document and Exhibits 1051 – 1053 were served on June 17, 2024 via electronic mail to counsel of record for the Patent Owner at the following:

Christine E. Lehman (Reg. No. 38,535)
clehman@reichmanjorgensen.com
Philip J. Eklem (Reg. No. 76,490)
peklem@reichmanjorgensen.com
REICHMAN JORGENSEN LEHMAN
& FELDBERG LLP
1909 K Street, NW, Suite 800
Washington, DC 20006

Jaime F. Cardenas-Navia
jcardenas-navia@reichmanjorgensen.com
Michael Matulewicz-Crowley (Reg. No. 72,220)
mmatulewicz-crowley@reichmanjorgensen.com
REICHMAN JORGENSEN LEHMAN & FELDBERG LLP
400 Madison Avenue, Suite 14D
New York, NY 10017

RJ_VideoLabs-Facebook@reichmanjorgensen.com

/s/ W. Todd Baker
W. Todd Baker
(No. 45,265)