

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.,
Petitioner,

v.

AUTONOMOUS DEVICES, LLC
Patent Owner

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 11,055,583**

TABLE OF CONTENTS

	Page
I. MANDATORY NOTICES	1
A. Real Party-in-Interest	1
B. Related Matters.....	1
C. Lead and Back-up Counsel and Service Information	1
II. GROUNDS FOR STANDING.....	2
III. INTRODUCTION	2
IV. STATUTORY GROUNDS FOR THE CHALLENGES	3
V. LEVEL OF ORDINARY SKILL IN THE ART	3
VI. CLAIM CONSTRUCTION	4
A. “means for receiving or generating”; “means for learning” (claim 34).....	4
VII. OVERVIEW OF THE ’583 PATENT	5
VIII. OVERVIEW OF THE PRIOR ART	6
A. U.S. 9,604,359 (“Grotmol”)	7
IX. IDENTIFICATION OF HOW THE CHALLENGED CLAIMS ARE UNPATENTABLE.....	8
A. Ground 1: Claims 1-9, 11-24, 26, 28-29, and 31-34 Are Anticipated by Grotmol.....	8
1. Claim 1	8
a. [1.1] A system comprising:	8
b. [1.2] one or more processors;	9

- c. [1.3] one or more non-transitory machine readable media storing machine readable code that, when executed by the one or more processors, causes the one or more processors to perform at least:.....10
 - d. [1.4] receiving or generating a first one or more digital pictures, wherein the first one or more digital pictures depict at least a portion of a first device's surrounding; 10
 - e. [1.5] receiving or generating a first one or more instruction sets for operating the first device; and 11
 - f. [1.6] learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.....12
2. Claim 2 14
- a. [2.1] The system of claim 1, wherein the receiving the first one or more digital pictures includes receiving the first one or more digital pictures from a picture capturing apparatus, ...15
 - b. [2.2] and wherein the receiving the first one or more instruction sets for operating the first device includes receiving the first one or more instruction sets for operating the first device from: an application for operating the first device, a system for operating the first device, one or more microcontrollers, another one or more processors, or one or more actuators, ...15
 - c. [2.3] and wherein the learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device includes storing the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device into or onto at least one of: at least one non-transitory machine readable medium of the one or more non-transitory machine

	readable media, another one or more non-transitory machine readable media, one or more volatile memories, one or more non-volatile memories, one or more storage devices, or one or more storage systems.....	17
3.	Claim 3	18
4.	Claim 4	19
	a. [4.1] The system of claim 1, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:	19
	b. [4.2] receiving or generating a new one or more digital pictures;	19
	c. [4.3] determining the first one or more instruction sets for operating the first device based on at least partial match between the new one or more digital pictures and the first one or more digital pictures; and.....	20
	d. [4.4] at least in response to the determining, causing the first device or a second device to perform one or more operations defined by the first one or more instruction sets for operating the first device.	22
5.	Claim 5	24
6.	Claim 6	25
7.	Claim 7	28
8.	Claim 8	29
9.	Claim 9	30
10.	Claim 11	32

11.	Claim 12.....	33
12.	Claim 13.....	33
13.	Claim 14.....	35
14.	Claim 15.....	37
	a. [15.1] The system of claim 4, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:	37
	b. [15.2] modifying: the first one or more instruction sets for operating the first device, or a copy of the first one or more instruction sets for operating the first device, ...	37
	c. [15.3] and wherein the determining ... includes: determining the modified the first one or more instruction sets for operating the first device based on the at least partial match between the new one or more digital pictures and the first one or more digital pictures, or ..., ...	38
	d. [15.4] and wherein the causing ... includes: causing the first device or the second device to perform one or more operations defined by the modified the first one or more instruction sets for operating the first device, or	38
15.	Claim 16.....	39
16.	Claim 17.....	40
	a. [17.1] The system of claim 4, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:	40

b.	[17.2] modifying: the first one or more digital pictures, or ..., ...	41
c.	[17.3] and wherein the learning ... includes: learning the modified the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device, or ...,.....	42
d.	[17.4] and wherein the determining ... includes: determining the first one or more instruction sets for operating the first device based on at least partial match between the new one or more digital pictures and the modified the first one or more digital pictures, or	42
17.	Claim 18.....	43
a.	[18.1] The system of claim 4, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:	43
b.	[18.2] modifying: the first one or more instruction sets for operating the first device, or ..., ...	43
c.	[18.3] and wherein the learning ... includes: learning the first one or more digital pictures correlated with the modified the first one or more instruction sets for operating the first device, or ...,.....	45
d.	[18.4] and wherein the determining ... includes: determining the modified the first one or more instruction sets for operating the first device based on the at least partial match between the new one or more digital pictures and the first one or more digital pictures, or ..., ...	45
e.	[18.5] and wherein the causing ... includes: causing the first device or the second device to perform one or more operations defined by the modified the first	

	one or more instruction sets for operating the first device, or	46
18.	Claim 19	47
19.	Claim 20	49
20.	Claim 21	52
21.	Claim 22	52
22.	Claim 23	52
23.	Claim 24	53
24.	Claim 26	54
a.	[26.1] The system of claim 4, wherein at least some elements of the system are included in: a single device, or multiple devices, ...	54
b.	[26.2] and wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices, ...	54
c.	[26.3] and wherein the first device includes a robot, a vehicle, ... or a mechanical machine, ...	55
d.	[26.4] and wherein the second device includes a robot, a vehicle ... or a mechanical machine, ...	56
e.	[26.5] and wherein the first one or more digital pictures include: one or more still digital pictures, or one or more motion digital pictures, and wherein the new one or more digital pictures include: one or more still digital pictures, or one or more motion digital pictures, ...	57
f.	[26.6] and wherein an instruction set of the first one or more instruction sets for operating the first device	

	includes at least one of: only one instruction, a plurality of instructions, ... one or more commands, ... or one or more information.....	57
25.	Claim 28	57
26.	Claim 29	59
27.	Claim 31: The system of claim 4, wherein the generating the first one or more instruction sets for operating the first device includes generating the first one or more instruction sets for operating the first device as one or more representations of a second one or more instruction sets for operating the first device.....	61
28.	Claim 32: The system of claim 4, wherein the first one or more digital pictures that depict the at least the portion of the first device's surrounding include one or more digital pictures that depict a representation of the at least the portion of the first device's surrounding.....	62
29.	Claim 33	63
30.	Claim 34	64
X.	CONCLUSION.....	65

PETITIONER'S EXHIBIT LIST

Exhibit No.	Description
Ex. 1001	U.S. Patent No. 11,055,583 (the “’583 patent”)
Ex. 1002	Prosecution File History of U.S. Patent No. 11,055,583
Ex. 1003	Declaration of Jason Janet, Ph.D.
Ex. 1004	Curriculum Vitae of Jason Janet, Ph.D.
Ex. 1005	U.S. Patent No. 9,604,359 (“Grotmol”)
Ex. 1006	<i>Reserved</i>
Ex. 1007	U.S. Patent No. 8,639,644 (“Hickman”)
Ex. 1008	Service, Robert F., <i>Minds of Their Own</i> , Sci. Magazine. Oct. 10, 2014, 346(6206):182-183, available at, https://www.usna.edu/AcResearch/_files/documents/Science_Minds_of_Their_own_182.full1.pdf
Ex. 1009	U.S. Patent No. 9,283,674 (“Hoffman”)

I. MANDATORY NOTICES

A. Real Party-in-Interest

Tesla, Inc. is the sole real party-in-interest.

B. Related Matters

The '583 patent is the subject of a civil action, *Autonomous Devices LLC v. Tesla, Inc.*, No. 22-cv-1466 (MN) (D. Del.), filed November 7, 2022. Tesla has also filed petitions for *inter partes* review involving U.S. Patent No. 10,452,974, U.S. Patent No. 11,238,344, U.S. Patent No. 10,102,449, U.S. Patent No. 10,607,134, and U.S. Patent No. 11,113,585 that were asserted in the Delaware litigation.

C. Lead and Back-up Counsel and Service Information

Anish Desai (lead counsel)

Reg. No. 73,760

Ian Moore (backup counsel)

Reg. No. 77,395

Weil, Gotshal & Manges LLP

767 Fifth Ave.

New York, NY 10153

T: 212-310-8000

F: 212-310-8007

anish.desai@weil.com

ian.moore@weil.com

Adrian Percer (backup counsel)

Reg. No. 46,986

Weil, Gotshal & Manges LLP

201 Redwood Shores Parkway

Redwood Shores, CA 94065

T: 650-802-3000

F: 650-802-3100

adrian.percer@weil.com

Christopher Pepe (backup counsel)

Reg. No. 73,851

Matthew Sieger (backup counsel)

Reg. No. 76,051

Taylor J. Lawrence (backup counsel)

Reg. No. 79,817

Allison D. Herzig (backup counsel)

Reg. No. 80,747

Weil, Gotshal & Manges LLP

2001 M Street NW, Suite 600

Washington, DC 20036

T: 202-682-7000

F: 202-857-0940

christopher.pepe@weil.com

matthew.sieger@weil.com

taylor.lawrence@weil.com

allison.herzig@weil.com

II. GROUNDS FOR STANDING

Petitioner certifies that the '583 patent is available for *inter partes* review and that Petitioner is not barred or estopped from requesting *inter partes* review challenging the patent claims on the grounds identified in this Petition.

III. INTRODUCTION

The '583 patent was filed on September 26, 2019 and claims a system in which: (i) stored pictures correlated with instructions for operating a first device are learned; (ii) new pictures are obtained; (iii) determining instructions for operating the first device based on a least a partial match between the new pictures and the stored pictures; and (iv) causing a first device or a second device to perform operations defined by the instructions for operating the first device. As will be demonstrated in this Petition, the prior art anticipates claims 1-9, 11-24, 26, 28-29, and 31-34 of the '583 patent.

Petitioner requests *inter partes* review based upon U.S. Patent No. 9,604,359 ("Grotmol"). Grotmol was not before the Examiner during prosecution. Grotmol, filed more than four years before the '583 patent, discloses the same approach to autonomous operation claimed in the '583 patent:

A robotic vehicle may utilize a camera and stores the sequence of images of a visual scene seen when following a trajectory during training in an ordered buffer. Motor commands associated with a given image may be stored. During

autonomous operation, an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.

Ex. 1005, Abstract. Grotmol further teaches that a robot’s “trained configuration may be stored” and “loaded [in]to one or more other robots in order to provide learned behaviors.” *Id.*, 33:12-35.

Grotmol anticipates claims 1-9, 11-24, 26, 28-29, and 31-34 of the ’583 patent. Petitioner respectfully requests that the Board institute review.

IV. STATUTORY GROUNDS FOR THE CHALLENGES

Ground 1	Anticipation of Claims 1-9, 11-24, 26, 28-29, and 31-34 based on U.S. 9,604,359 (“Grotmol”) (Ex. 1005)
----------	--

V. LEVEL OF ORDINARY SKILL IN THE ART

A person having ordinary skill in the art (“POSITA”) relevant to the ’583 patent as of November 26, 2017—the earliest possible priority date of the ’583 patent—would have had at least: (1) a bachelor’s degree in electrical engineering, computer engineering, computer science, systems engineering, or equivalent course work with two to three years of work experience in computer vision, image/video processing, and/or machine learning; or (2) a master’s degree in electrical engineering, computer engineering, computer science, systems engineering, or

equivalent course work with a focus in computer vision, image/video processing, and/or machine learning. Decl. ¶ 41.

VI. CLAIM CONSTRUCTION

The Petition applies the ordinary and customary meaning of the claim terms. Where appropriate, Petitioner explains the meaning of claim terms in its analysis of how the prior art anticipates and/or renders obvious the challenged claims.

A. “means for receiving or generating”; “means for learning” (claim 34)

Claim 34 requires “means for receiving or generating a first one or more digital pictures,” “means for receiving or generating a first one or more instruction sets for operating the first device,” and “means for learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.” Ex. 1001, Claim 34. The ’583 patent specification describes that a corresponding structure for performing these claimed functions is one or more processor circuits. Ex. 1001, 16:8-35 (stating that the system includes “one or more processor circuits” that perform the functions of receiving digital pictures from a picture capturing apparatus, receiving one or more instruction sets for operating the device, and learning the digital pictures are correlated with the one or more instructions sets); *see also id.*, 16:36-60, claim 1 (requiring that one or more processor circuits perform the receiving/generating pictures, receiving/generating instructions, and learning steps).

VII. OVERVIEW OF THE '583 PATENT

The '583 patent was filed on September 26, 2019 and is a continuation of application no. 15/822,150, filed on November 26, 2017, now U.S. Patent No. 10,474,934. Jasmin Cosic is the sole named inventor.

The disclosure of the '583 patent generally relates to “computing enabled systems and/or devices” including embodiments where “the system captures one or more digital pictures, receives one or more instruction sets, and learns correlations between the captured pictures and the received instruction sets.” Ex. 1001, Abstract. More specifically, the '583 patent describes a method of automation comprising “learning one or more digital pictures of a device’s surrounding along with correlated instruction sets for operating the device, storing this knowledge in a knowledgebase (i.e. neural network, graph, sequences, etc.), and autonomously operating a device.” *Id.*, 59:14-18. The '583 patent refers to this approach as “visual surrounding for autonomous device operation,” or “VSADO.” *Id.*, 58:27-29. Figure 31, copied below, illustrates this functionality.

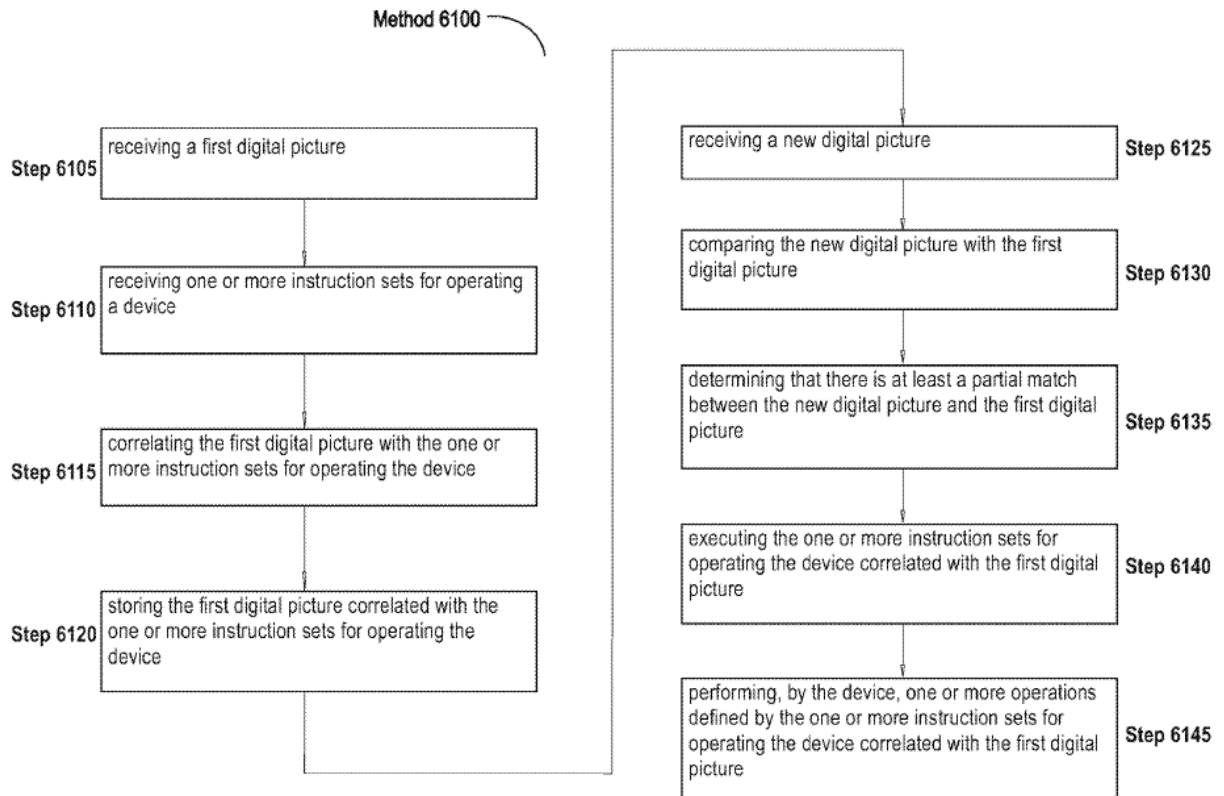


FIG. 31
Ex. 1001, Fig. 31

The '583 patent further describes that the VSADO functionality can be implemented on a “Server 96 accessible over network 95,” such that “[a]ny number of Devices 98 may connect to such remote VSADO Unit 100 and the remote VSADO Unit 100 may learn their operations in various visual surroundings. In turn, any number of Devices 98 can utilize the remote VSADO Unit 100 for autonomous operation.” *Id.*, 85:54-60.

VIII. OVERVIEW OF THE PRIOR ART

This Petition relies on Grotmol. Grotmol was not considered during prosecution or cited on the face of the '583 patent.

A. U.S. 9,604,359 (“Grotmol”)

Grotmol is a U.S. patent filed on January 27, 2015, which issued on March 28, 2017. Grotmol is prior art to the ’583 patent under at least 35 U.S.C. §§ 102(a) and 102(d).

Grotmol relates to “[a]n apparatus and methods for training and/or operating a robotic device to follow a trajectory.” Ex. 1005, Abstract. Grotmol describes that “conventional approaches” to training a robot entailed a user “first control[ing] the robot one time or multiple times to move along the desired path.” *Id.*, 1:62-67. That would constitute the “training of the robot,” and thereafter, the robot would perform the same navigation autonomously. *Id.*, 1:67-2:2. A “typical approach” was to store the motor commands executed by a user during training then “simply replay them.” *Id.*, 2:3-5. Grotmol then describes its own approach, whereby user commands are associated with camera images captured during training, such that during “autonomous operation, an acquired image may be compared with one or more images from the training buffer” to determine how to control the robot. *Id.*, Abstract.

More specifically, Grotmol describes a “listen mode,” in which a robot observes a teacher’s commands in order to “learn an association.” *Id.*, 23:4-12. Grotmol also describes an “override mode,” in which a robot will operate autonomously until a teacher sends a command, in which case the robot will apply the teacher’s command and “try to remember it for the next time the same situation

occurs.” *Id.*, 23:12-22. “During training, images (e.g., raw and/or pre-processed) may be stored in a memory buffer (training buffer),” along with “[m]otor commands issued by a trainer corresponding to time instances when the images are acquired.” *Id.*, 39:13-14, 39:17-18. The robot may then switch—either manually or automatically—into an “autonomous mode,” in which the robot performs tasks autonomously based on what it has learned. *Id.*, 23:46-48, 24:1-16. “During autonomous operation, control process of the robot may be configured to compare a given (e.g., the most recent, current) image with one or more [] images from the training buffer.” *Id.*, 39:23-28. “In some implementations, motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the best matching stored image.” *Id.*, 43:4-9.

Grotmol further teaches that a robot’s “trained configuration may be stored” and “loaded into one or more other robots in order to provide learned behaviors.” *Id.*, 33:12-35.

IX. IDENTIFICATION OF HOW THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. Ground 1: Claims 1-9, 11-24, 26, 28-29, and 31-34 Are Anticipated by Grotmol

Grotmol anticipates each element of claims 1-9, 11-24, 26, 28-29, and 31-34.

1. Claim 1

a. [1.1] A system comprising:

Grotmol discloses “a system” implemented as a “self-contained robotic apparatus.” Ex. 1005, 3:27-30; *see also id.*, Abstract (“An apparatus and methods for training and/or operating a robotic device to follow a trajectory.”); 10:45-48. Accordingly, Grotmol discloses “a system.” Decl. ¶¶ 178-179.

b. **[1.2] one or more processors;**

Grotmol discloses limitation [1.2]. Decl. ¶¶ 180-183.

Grotmol includes a “processing” unit that is connected to at least a sensory unit and a memory unit. Ex.1005, Figure 7. Grotmol defines “processor”:

As used herein, the terms “processor”, “microprocessor” and “digital processor” are meant generally to include all types of digital processing devices including, without limitation, digital signal processors (DSPs), reduced instruction set computers (RISC), general-purpose (CISC) processors, microprocessors, gate arrays (e.g., field programmable gate arrays (FPGAs)), PLDs, reconfigurable computer fabrics (RCFs), array processors, secure microprocessors, and application-specific integrated circuits (AS I Cs). Such digital processors may be contained on a single unitary IC die, or distributed across multiple components.

Id., 8:5-15. Grotmol discloses “a self-contained robotic apparatus, the apparatus comprising...one or more physical processors configured to operate [the system].”

Id., claim 12; *see also id.*, 3:27-30, claims 1, 17. Accordingly, Grotmol discloses a system comprising “one or more processors.” Decl. ¶¶ 180-183.

- c. **[1.3] one or more non-transitory machine readable media storing machine readable code that, when executed by the one or more processors, causes the one or more processors to perform at least:**

Grotmol discloses that its system includes “one or more non-transitory machine readable media storing machine readable code that, when executed by the one or more processors, causes the one or more processors to perform at least” limitations [1.4]-[1.6]. Decl. ¶¶ 184-186; Ex. 1005, claim 17 (“A non-transitory computer-readable storage medium having instructions embodied thereon, the instructions being executable by a processor ...”); *id.*, claim 12 (“a self-contained robotic apparatus, the apparatus comprising...one or more physical processors configured to operate [the system]”); *id.*, 33:14-30, 4:13-17, claim 1; *see also infra* limitations [1.4]-[1.6].

- d. **[1.4] receiving or generating a first one or more digital pictures, wherein the first one or more digital pictures depict at least a portion of a first device’s surrounding;**

Grotmol discloses limitation [1.4]. Decl. ¶¶ 187-189.

Grotmol discloses processor(s) receiving “training sets,” where each training set “may comprise an instance of the video frame,” *i.e.*, *a first one or more digital pictures*. Ex. 1005, 3:27-39.

Grotmol further describes that the images can be “of a visual scene seen when following a trajectory,” *i.e.*, *first one or more digital pictures ... depict[ing] at least a portion of a first device’s surrounding*, and are received and stored by the robot during training. *Id.*, Abstract (“[a] robotic vehicle may utilize a camera and stores the sequence of images of a visual scene seen when following a trajectory during training in an ordered buffer.”); *id.*, 10:16-22 (“In some implementations of robotic navigation in an arbitrary environment, the sensor component 166 may comprise a camera configured to provide an output comprising a plurality of digital image frames ...”); *id.*, 8:57-9:3, 38:28-60, 39:13-22, Figure 20A.

Accordingly, Grotmol discloses processor(s) receiving or generating a first one or more digital pictures, wherein the first one or more digital pictures depict at least a portion of a first physical device’s surrounding. Decl. ¶¶ 187-189.

e. **[1.5] receiving or generating a first one or more instruction sets for operating the first device; and**

Grotmol discloses limitation [1.5]. Decl. ¶¶ 190-192.

As described above in limitation [1.4], Grotmol discloses processor(s) receiving “training sets,” where each training set “may comprise ... an instance of a motor control indication configured to cause the robot to execute an action,” *i.e.*, *a first one or more instruction sets for operating the first device*. Ex. 1005, 3:27-39; *id.*, 39:17-19 (“Motor commands issued by a trainer corresponding to time instances when the images are acquired may be stored.”); *id.*, 33:14-30, Abstract.

Accordingly, Grotmol discloses receiving or generating a first one or more instruction sets for operating the first physical device. Decl. ¶¶ 190-192.

f. **[1.6] learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.**

Grotmol discloses limitation [1.6]. Decl. ¶¶ 193-200.

Grotmol discloses processor(s) generating and storing “training sets,” including “a memory component,” which “may be configured to store training sets,” where each training set “may comprise an instance of the video frame,” *i.e., a first one or more digital pictures*, “... and an instance of a motor control indication configured to cause the robot to execute an action,” *i.e., a first one or more instruction sets for operating the first device*. Ex. 1005, 3:27-39; *see supra* limitations [1.4] and [1.5]; Decl. ¶¶ 187, 190-191. The video frames in Grotmol are “associated with” or “correspond to”—*i.e., are correlated with*¹—motor control indications issued around the same time. Decl. ¶ 195. Grotmol describes training behaviors in a robotic device by “learning of *associations* between sensory context ... and a respective action ... during training.” Ex. 1005, 33:59-62; *see also id.*, Abstract (“Motor commands *associated with* a given image may be stored.”). This

¹ The ’583 patent describes that a device may “correlate Digital Pictures 525 from the stream of Digital Pictures 525 with *temporally corresponding* Instruction Sets”). Ex. 1001, 91:52-56

association may be based on the timing of the images and the motor commands. Ex. 1005, 39:17-19 (“Motor commands issued by a trainer *corresponding*² to time instances when the images are acquired may be stored.”); *see also id.*, 43:6-9 (“[M]otor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may *correspond to* the best matching stored image.”). Decl. ¶ 195.

Grotmol further discloses that a human (“user”³) can, by manually operating the robot, train a robot to learn pictures that are correlated with instructions (a “learning process”⁴):

Trained behaviors of the robotic device may be configured based on *learning* of associations between sensory context (e.g., presence of an obstacle in front of the robotic vehicle) and a respective action (e.g., right turn) during training.

² Emphasis added unless otherwise noted.

³ “User 50 (also referred to simply as user or other similar name or reference) comprises a human user or non-human user.” Ex. 1001, 69:11-13.

⁴ The ’583 patent does not limit “learning process” to a particular technique. *See* Ex. 1001, 95:53-64 (“The disclosed systems, devices, and methods are independent of the artificial intelligence model and/or technique used and any model and/or technique can be utilized to facilitate the functionalities described herein. Examples of these models and/or techniques include deep learning, supervising learning, unsupervised learning...”).

Grotmol, 33:59-62. In both the “listen mode” and “override mode,” Grotmol teaches that the robot will listen to commands sent by a teacher (user) “and try to remember it for next time the same situation occurs.” *Id.*, 23:4-29; *see also id.*, 24:57-61 (“While training a particular behavior, the *user* may concentrate on a particular behavior and may not show any examples of other behaviors thus making a prediction task for the classifier substantially easier (e.g., less computationally intensive).”), 34:50-55 (“In some implementations of supervised training of robots, control instructions (also referred to as corrections) produced by the *trainer* (e.g., human) may be combined with control instructions produced by the robot controller instructions (predictions).”); *see also id.*, 33:14-30; Decl. ¶¶ 196-197.

Grotmol describes this training as a learning process. *E.g., id.*, 24:17-20 (“Apparatus and methods for hierarchical *learning* are disclosed herein, in accordance with one or more implementations. In *supervised learning*, a user may train simple tasks by demonstrating example tasks to the system.”), claim 12 (reciting “one or more physical processors configured to operate a random k-nearest neighbors *learning process*”). Decl. ¶ 198.

Accordingly, Grotmol discloses “learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.” Decl. ¶¶ 193-200.

2. Claim 2

- a. **[2.1] The system of claim 1, wherein the receiving the first one or more digital pictures includes receiving the first one or more digital pictures from a picture capturing apparatus,**

Grotmol discloses limitation [2.1]. Decl. ¶¶ 201-203.

For example, Grotmol discloses that “[a] robotic vehicle may utilize a camera,” *i.e.*, picture capturing apparatus, “and stores the sequence of images of a visual scene seen when following a trajectory during training in an ordered buffer.” Ex. 1005, Abstract; *see also id.*, 10:16-22 (“In some implementations of robotic navigation in an arbitrary environment, the sensor component 166 may comprise a camera configured to provide an output comprising a plurality of digital image frames ...”); *id.*, 38:28-60; *see supra* limitation [1.4]; Decl. ¶¶ 187-188.

- b. **[2.2] and wherein the receiving the first one or more instruction sets for operating the first device includes receiving the first one or more instruction sets for operating the first device from: an application for operating the first device, a system for operating the first device, one or more microcontrollers, another one or more processors, or one or more actuators,**

Grotmol discloses limitation [2.2]. Decl. ¶¶ 204-206.

As I described above in limitations [1.1]-[1.3], Grotmol discloses a system implemented as a “self-contained robotic apparatus” that includes a memory component that stores the training sets that include digital pictures correlated with instructions. Ex. 1005, Abstract, 3:28-39, Figure 7. In Grotmol, the processor receives the instructions from the memory, which is part of the system for operating

the self-contained robotic device. Thus, the instructions are received from a system for operating the first device. Decl. ¶ 204.

Grotmol further discloses that the robotic device may “comprise a controller operating a software component (e.g., the BrainOS® software platform),” and therefore, discloses that the instructions are received from an application for operating the first device: “configured to enable training”:

In some implementations, a robotic device may comprise a controller operating a software component (e.g., the BrainOS® software platform) configured to enable training. A user may control/training the robot with a remote device (e.g., comprising a Gamepad® controller and *an iOS® application*, and/or a handset device (e.g., a smartphone)). Training of the robot’s controller may be based on the user observing robot’s actions and *sending one or more target control commands to the robot* via the training handset. The trained controller of the robot may comprise a trained configuration configured to enable autonomous operation (e.g., without user input) by the robotic device. The trained configuration may be stored. A saved configuration may be loaded into the robot being trained thereby providing one or more trained behaviors to the robot

Ex. 1005, 33:14-30, *see also id.*, 2:13-18 (“One aspect of the disclosure relates to a method of determining a control signal for a robot. The method may be performed by a special purpose computing platform having one or more processors executing

instructions stored by a non-transitory computer-readable storage medium.”).

Decl. ¶ 205.

- c. **[2.3] and wherein the learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device includes storing the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device into or onto at least one of: at least one non-transitory machine readable medium of the one or more non-transitory machine readable media, another one or more non-transitory machine readable media, one or more volatile memories, one or more non-volatile memories, one or more storage devices, or one or more storage systems.**

Grotmol discloses limitation [2.3]. Decl. ¶¶ 207-210.

As described above, Grotmol discloses “a memory component,” which “may be configured to store training sets,” where each training set “may comprise an instance of the video frame,” *i.e., a first one or more digital pictures*, “... and an instance of a motor control indication configured to cause the robot to execute an action,” *i.e., a first one or more instruction sets for operating the first device*. Ex. 1005, 3:27-39; *see supra* limitations [1.4]-[1.6]; Decl. ¶¶ 187, 190. Grotmol discloses that its “memory component” may comprise a “non-volatile memory.” Ex. 1005, 38:39-42 (“In some implementations, the images may be stored in an ordered buffer in a non-volatile memory.”); *id.*, 38:33-42 (disclosing training data stored as a “sequence” in “an ordered buffer in a non-volatile memory”). Decl. ¶¶ 208-209.

3. Claim 3

Grotmol discloses claim 3 which requires that learning of the digital picture correlated with an instruction includes determining that the instruction temporally corresponds to the digital picture. Decl. ¶¶ 211-213.

As described above with respect to limitation [1.6], the pictures in Grotmol are correlated with motor control indications issued around the same time. Grotmol describes training behaviors in a robotic device by “learning of *associations* between sensory context ... and a respective action ... during training.” Ex. 1005, 33:59-62; *see also id.*, Abstract (“Motor commands *associated with* a given image may be stored.”). This association may be based on the timing of the images and the motor commands. *Id.*, 38:39-72 (“The robot may comprise a controller configured to store a sequence of images obtained during training phase. In some implementations, the images may be stored in an ordered buffer in a non-volatile memory (e.g., memory 814 in FIG. 8).”); 39:17-19 (“Motor commands issued by a trainer *corresponding to* time instances when the images are acquired may be stored.”); *see also id.*, 43:5-9 (“[M]otor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may *correspond to* the best matching stored image.”); Decl. ¶¶ 193-200. Accordingly, Grotmol discloses determining that the first one or more instruction sets for operating the first physical device temporally correspond to the first one or more digital pictures. Decl. ¶¶ 211-213.

4. Claim 4

- a. **[4.1] The system of claim 1, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:**

Grotmol discloses “wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least” limitations [4.2]-[4.4]. *See infra* limitations [4.2]-[4.4]; Ex. 1005, 33:14-30, 4:13-17; Decl. ¶ 214.

- b. **[4.2] receiving or generating a new one or more digital pictures;**

Grotmol discloses limitation [4.2]. Decl. ¶¶ 215-218.

Grotmol discloses that its processor receives new digital pictures. Ex. 1005, Abstract (“During autonomous operation, an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.”). As Grotmol’s robotic vehicle operates, it requires receiving a new digital picture. *Id.*, 39:23-26 (“During autonomous operation, control process of the robot may be *configured to compare a given (e.g., the most recent, current) image* with one or more the images from the training buffer.”); 10:15-20, 38:51-58; Decl. ¶ 216. In Grotmol, “the most recent, current” image corresponds to the claimed “new one or more digital pictures” and it is compared to the stored images in the training buffer (the first one or more digital pictures). Decl. ¶ 217.

- c. **[4.3] determining the first one or more instruction sets for operating the first device based on at least partial match between the new one or more digital pictures and the first one or more digital pictures; and**

Grotmol discloses limitation [4.3]. Decl. ¶¶ 219-223.

Grotmol discloses that “motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the *best matching* stored image.” Ex. 1005, 43:5-9; *see also id.*, Abstract (“During autonomous operation an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.”). Grotmol discloses that the matching step may comprise calculating a similarity measure between a new image and an old image then determining whether that similarity measure breaches a threshold (*i.e.*, “determining ... based on at least partial match between the new one or more digital pictures and the first one or more digital pictures”).

The method may comprise, based on the comparison, determining a *similarity measure* for a given training set of the plurality of training feature sets. The similarity measure may characterize similarity between features of the subset and features of the given training set. The method may comprise, responsive to the similarity measure *breaching a threshold*, selecting one or more training sets from the plurality of training sets.

Ex. 1005, 4:30-38; *see also id.*, claim 1 (reciting “*responsive to the similarity measure breaching a threshold*, selecting one or more training sets from the plurality of training feature sets”). As used in Grotmol, “the term feature may be used to describe one or more integer or floating point values characterizing the input, e.g., the presence or absence of an edge, corner, shape, texture, color, object, at particular locations in the image, *values of pixels in an image*, patches of color texture, brightness in the image, and/or in the image as a whole ...” *id.*, 28:19-24. As used in Grotmol, a “training set” may comprise “an *instance of the video frame*, an instance of the binary sensor signal, and an instance of a motor control indication configured to cause the apparatus to execute an action.” *id.*, claim 12. Decl. ¶¶ 219-221.

Both Grotmol and the '583 patent disclose that determining “matching portions” may comprise pixel-wise difference analysis between the two digital pictures. *Compare* Ex. 1001, 21:29-46 (discussing “portion”), 67:41-45 (discussing “matching”), 112:38-113:2 (discussing RGB pixel matching) *with* Ex. 1005, 3:14-16 (“The first input type may comprise a digital image comprising a plurality of pixel values.”), 41:8-13 (“A similarity metric ... may be determined by performing pixel-wise difference between the overlapping regions (e.g., the frames 2020, 2040 in FIG. 20B) of the given image and the matching image.”). Decl. ¶ 222.

Accordingly, Grotmol discloses determining the first one or more instruction sets for operating the first physical device based on at least partial match between the new one or more digital pictures and the first one or more digital pictures. Decl. ¶¶ 219-223.

- d. **[4.4] at least in response to the determining, causing the first device or a second device to perform one or more operations defined by the first one or more instruction sets for operating the first device.**

Grotmol discloses limitation [4.4]. Decl. ¶¶ 224-227.

As described above for limitation [4.3], Grotmol discloses determining the learned instruction sets for operating the device by comparing a new image with a stored image at least by matching portions of the images and determining whether those matching portions exceed a threshold number. Ex. 1005, 43:5-9, Abstract, 28:19-40; Decl. ¶¶ 219-222.

Grotmol discloses that after it is determined that the similarity measure between the new image and image from the training set breaches a threshold, one or more potential control signals corresponding to the image from the training set that were obtained during a training operation of the robot (the instruction sets correlated with the first digital picture) are determined and executed:

The method may comprise, responsive to the similarity measure breaching a threshold, selecting one or more training sets from the plurality of training sets. The method may comprise

determining one or more potential control signals for the robot. Individual ones of the one or more potential control signals may be associated with a corresponding training set of the plurality of training sets ... In some implementations, the control signal may be configured to *cause the robot to execute the action*.

Ex. 1005, 2:34-41, 3:12-13; *see also id.*, Abstract (“During autonomous operation, an acquired image may be compared with one or more images from the training buffer in order to determine the most likely match.”); 43:4-11 (“In some implementations, motor commands at a given time step may be obtained by *taking the stored motor commands from the training buffer* that may correspond to the best matching stored image.”).⁵ Decl. ¶ 225.

Grotmol also discloses that a second robot may execute actions learned on a first robot. Specifically, Grotmol discloses that training sets learned on a first robot may be loaded into a second robot. *See* Ex. 1005, 33:27-30 (“A saved configuration may be loaded into the robot being trained thereby providing one or more trained

⁵ The ’583 patent describes that this step may comprise simply causing an application to execute an instruction. Ex. 1001, 11:1-7 (“In certain embodiments, the causing the processor circuit to execute the one or more instruction sets for operating the device correlated with the first digital picture includes causing an application for operating the device to execute the one or more instruction sets for operating the device correlated with the first digital picture, the application running on the processor circuit.”).

behaviors to the robot.”). In such a case, autonomous operation of the second robot executes actions defined by the training set learned on the first robot. Decl. ¶ 226.

Accordingly, Grotmol discloses that at least in response to the determining, causing the first device or a second device (Grotmol discloses both) to perform one or more operations defined by the first one or more instruction sets for operating the first device. Decl. ¶¶ 224-227.

5. Claim 5

Grotmol discloses claim 5 which requires causing the first or second device to perform an operation defined by an instruction for operating the first device includes executing the instruction. Decl. ¶¶ 228-231.

As described above for limitation [4.4], Grotmol discloses executing instruction sets learned on a first robot on both that first robot and a second robot. Ex. 1005, 33:21-32; *see also* 2:33-41, 3:12-13, 43:4-11; Decl. ¶¶ 224-226. The ’583 patent similarly describes that this step may comprise simply causing an application to execute an instruction. Ex. 1001, 11:1-7 (“In certain embodiments, the causing the processor circuit to execute the one or more instruction sets for operating the device correlated with the first digital picture includes causing an application for operating the device to execute the one or more instruction sets for operating the device correlated with the first digital picture, the application running on the processor circuit.”); Decl. ¶ 230.

Accordingly, Grotmol discloses wherein the causing the first physical device or the second physical device to perform the one or more operations defined by the first one or more instruction sets for operating the first physical device includes executing the first one or more instruction sets for operating the first physical device. Decl. ¶¶ 228-231.

6. Claim 6

Grotmol discloses claim 6 which requires determining at least a partial match between a portion of the new pictures that represent one or more objects and a portion of the stored pictures that represent one or more objects. Decl. ¶¶ 232-237.

Grotmol discloses determining the learned instruction sets for operating the device by comparing a new image with a stored image, which includes by matching portions of the images. *See supra* limitation [4.3], Ex. 1005, 43:5-9, Abstract, 2:29-37; Decl. ¶¶ 219-221.

The '583 patent provides examples of objects in pictures such as “car, pedestrian, door, building, animal, person, etc.” Ex. 1001, 88:32-35. The '583 patent further describes various approaches to object recognition, such as “identifying and/or analyzing object features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc.” *Id.*, 87:14-19. Decl. ¶ 233.

Similarly, Grotmol discloses images that may “comprise *representations* of one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.”

Ex. 1005, 38:56-65; Decl. ¶ 234.

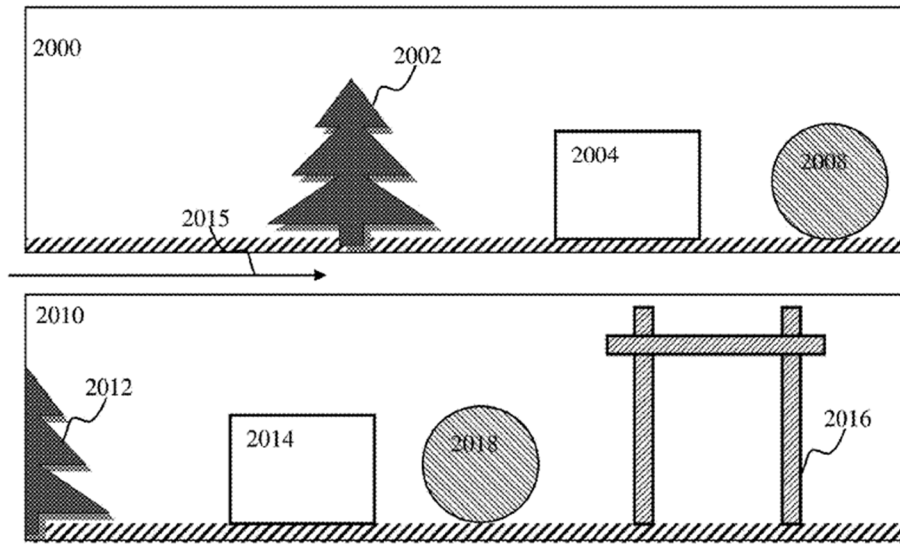


FIG. 20A

Ex. 1005, Fig. 20A

Grotmol discloses that the sensor used to capture these images “may provide information associated with objects within the field-of-view ... In one or more implementations, such as *object recognition*, and/or obstacle avoidance, the output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.” Ex. 1005, 8:65-9:3. *See, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented ...”); 10:23-26. Grotmol discloses calculating a similarity measure between one or more portions of new and old images represent one or more objects then determining whether that similarity measure breaches a threshold:

The method may comprise, based on the comparison, determining a *similarity measure* for a given training set of the plurality of training feature sets. *The similarity measure may characterize similarity between features of the subset and features of the given training set.* The method may comprise, responsive to the similarity measure breaching a threshold, selecting one or more training sets from the plurality of training sets.

Ex. 1005, 2:29-37; *see also id.*, claim 1 (reciting “*responsive to the similarity measure breaching a threshold*, selecting one or more training sets from the plurality of *training feature sets*”). As used in Grotmol, “the term feature may be used to describe one or more integer or floating point values characterizing the input, e.g., the presence or absence of an *edge, corner, shape*, texture, color, *object*, at particular locations in the image, values of pixels in an image, patches of color texture, brightness in the image, and/or in the image as a whole” *id.*, 28:19-24.

Accordingly, Grotmol discloses “determining at least partial match between one or more portions of the new one or more digital pictures that represent one or more objects and one or more portions of the first one or more digital pictures that represent one or more objects.” Decl. ¶¶ 232-237.

7. Claim 7

Grotmol discloses claim 7 which requires determining at least a partial match between one or more detected objects in a new picture and one or more detected objects in the stored picture. Decl. ¶¶ 238-239.

As described above for limitation [4.3] and claim 6, Grotmol discloses “determining the first one or more instruction sets for operating the first device based on at least partial match between the new one or more digital pictures and the first one or more digital pictures” and “determining at least partial match between one or more portions of the new one or more digital pictures that represent one or more objects and one or more portions of the first one or more digital pictures that represent one or more objects.” Ex. 1005, 43:5-9, Abstract, 2:29-37, 38:56-65, 8:57-9:13, 9:31-52; Decl. ¶¶ 224-227, 232-237.

Grotmol discloses calculating a similarity measure between one or more portions of new and old images where the similarity measure may characterize similarity between features of the subset and features of the given training set and features may be used to “describe one or more integer or floating point values characterizing the input, e.g., the *presence or absence* of an *edge*, *corner*, *shape*, texture, color, *object*, at particular locations in the image, values of pixels in an image, patches of color texture, brightness in the image, and/or in the image as a

whole” Ex. 1005, 28:19-24; *id.*, 2:29-37; Decl. ¶ 238; *see supra* limitation [4.3] and claim 6.

Accordingly, Grotmol discloses that the “determining ... includes determining at least partial match between one or more objects detected in the new one or more digital pictures and one or more objects detected in the first one or more digital pictures.” Decl. ¶¶ 238-239.

8. Claim 8

Grotmol discloses claim 8 which requires that the instruction set for operating the first device includes information about the state of the first device or a portion of the first device. Decl. ¶¶ 240-242.

Grotmol discloses training sets (that include instruction sets) that contain state information of the first device. As described above for limitation [1.6], Grotmol discloses a memory component configured to store training sets comprising “an instance of the video frame, an instance of the binary signal, and an instance of a motor control indication configured to cause the robot to execute an action.” Ex. 1005, 3:36-39; Decl. ¶ 194. Grotmol further discloses that “[a]dditional sensory information (e.g., vehicle motion information, ambient environment information, vehicle operational parameters) corresponding to time instances when the images are acquired may be stored,” which constitutes information about the state of the robot (i.e., first device). Ex. 1005, 39:16-22; Decl. ¶¶ 240-242.

9. Claim 9

Grotmol discloses claim 9 which requires “wherein the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device are included in a knowledgebase, and wherein the knowledgebase further includes a second one or more digital pictures correlated with a second one or more instruction sets for operating the first device, and wherein at least a portion of the first one or more digital pictures or at least a portion of the first one or more instruction sets for operating the first device is learned in a first learning process that includes operating the first device at least partially by a first user, and wherein at least a portion of the second one or more digital pictures or at least a portion of the second one or more instruction sets for operating the first device is learned in a second learning process that includes operating the first device at least partially by the first user.” Decl. ¶¶ 243-246.

As described above with respect to claim 1, Grotmol discloses a “memory component may be configured to store training sets,” where “[a] given training set may comprise an instance of the video frame, an instance of the binary signal, and an instance of a motor control indication configured to cause the robot to execute an action.” Ex. 1005, 3:27-39; Abstract (“R]obotic vehicle may utilize a camera and stores the sequence of images of a visual scene seen when following a trajectory during training in an ordered buffer. Motor commands associated with a given image

may be stored.”); *see also id.*, 3:31-39, 8:66-9:3, 10:23-26, 38:36-58, 39:13-22; Decl. ¶¶ 187, 190. Thus, Grotmol discloses a memory that stores a knowledgebase⁶ including correlations (training sets), where each correlation includes a digital picture correlated with an instruction sets for operating the first device. Decl. ¶¶ 243-244.

Grotmol describes that multiple images may be acquired, with images respectively stored in association with corresponding motor commands issued at the time the image was acquired. Grotmol describes that multiple training processes can be conducted to learn multiple tasks. Ex. 1005, 24:17-67 (describing training of “simple tasks” and “complex tasks,” including behaviors such as “fetch” and “grasping.”); 26:1-20 (describing that “[t]he system may be configured to learn the *second* behavior from the user.”). Decl. ¶ 245.

Thus, Grotmol discloses the knowledgebase includes a first correlation learned in a first learning process and a second correlation learned in a second learning process that includes operating the first device at least partially by a first user. Decl. ¶¶ 243-246.

⁶ The ’583 patent states that a “[k]nowledgebase 530 may be or include *any data structure or arrangement* capable of storing the knowledge of a device’s operation in various visual surroundings.” Ex. 1001, 95:20-23.

10. Claim 11

Grotmol discloses claim 11 which requires: (i) first digital pictures correlated to first instruction set for operating a first device (i.e., first correlation) are included in a knowledgebase; (ii) second digital pictures correlated to second instruction set for operating another (third) device (i.e., second correlation) are included in a knowledgebase; (iii) first correlation is learned in a first learning process that includes operating the first device at least partially by a first user; and (iv) second correlation is learned in a second learning process that includes operating the other (third) device at least partially by a first user (or a second user). Decl. ¶¶ 247-249.

As described above with respect to limitation [1.6] and claim 9, Grotmol discloses that multiple learning processes can be done to train multiple behaviors that are stored in a knowledgebase. Ex. 1005, 38:36-58, 39:13-22, 3:13-39, 26:1-20, 24:17-67; Decl. ¶ 245.

Grotmol further discloses that the training sets and training configurations from a robot (which include the instruction sets for operating first device) can be “stored” or “saved.” Ex. 1005, 33:12-58 (discussing saved training configurations). The saved training configuration may then be “loaded into the robot being trained thereby providing one or more trained behaviors to the robot.” *Id.*, 33:28-30. Further, the “trained configuration may be loaded to one or more *other robots* in order to provide learned behaviors.” *Id.*, 33:30-32. Thus, the knowledgebase of a

first robot can store additional correlations, including a second correlation (second digital pictures correlated to second instruction set for operating another device) learned in a second learning process that includes operating the other device by the user (i.e., first user). Decl. ¶¶ 247-249.

11. Claim 12

Grotmol discloses claim 12 which requires that the new picture depicts the first device's surroundings, and the first instruction is applied to the first device and the first device is caused to perform an operation defined by the instruction. Decl. ¶¶ 250-252.

As described above for limitation [1.4] and claim 4, Grotmol discloses receiving new digital pictures (*e.g.*, video frame instances) capturing the first device's surroundings. Ex. 1005, 8:57-9:3, 10:15-26, 38:52-58; Decl. ¶¶ 187-188, 216. As described above for claims 4, 5, and 6, Grotmol discloses the first one or more instruction sets for operating the first device are applied to the first device, and the first device is caused to perform the one or more operations defined by the first one or more instruction sets for operating the first device. Ex. 1005, 2:34-41, 3:12-13, 43:4-1; Decl. ¶¶ 220, 225-225, 229-230.

12. Claim 13

Grotmol discloses claim 13 which requires that the new picture depicts the second device's surroundings, and the first instruction is applied to the second device

and the second device is caused to perform an operation defined by the instruction. Decl. ¶¶ 253-255.

Grotmol discloses that the training sets and training configurations from a first robot (which include the instruction sets for operating the first device) can be “stored” or “saved.” Ex. 1005, 33:12-58 (discussing saved training configurations). The saved training configuration may then be “loaded into the robot being trained thereby providing one or more trained behaviors to the robot.” *Id.*, 33:28-30. Further, the “trained configuration may be loaded to one or more *other robots* in order to provide learned behaviors.” *Id.*, 33:30-32.

Thus, Grotmol discloses a second physical device (“one or more other robots”) that can autonomously perform one or more operations defined by the first one or more instruction sets for operating the first physical device (training configuration from first robot (e.g., training sets with images correlated to instruction sets) that is loaded into a second one or more other robots to “provide learned behaviors”). During operation, the second device would determine the similarity measure between the new image (new digital picture depicting at least a portion of the second device’s surrounding) and an image from the training set loaded to the second device (first digital picture). After it is determined that the similarity measure between the new image and image from the training set breaches a threshold (i.e., partial match), one or more potential control signals corresponding to the image from

the training set that were obtained during a training operation of the robot (first one or more instruction sets for operating the first device correlated with the first digital picture) are determined and applied to the second device, causing the second device to perform the one or more operations defined by the first one or more instruction sets for operating the first device. Decl. ¶ 254.

13. Claim 14

Grotmol discloses claim 14 which requires “wherein the new one or more digital pictures depict at least a portion of the second device’s surrounding, and wherein the first one or more instruction sets for operating the first device ... are modified and applied to the second device, and wherein the second device is caused to perform one or more operations defined by: the modified the first one or more instruction sets for operating the first device ...” Decl. ¶¶ 256-260.

The ’583 specification describes that the instruction set may be modified by modifying the underlying source code or hardware, or modifying the application itself:

In further embodiments, the causing the processor circuit to execute the one or more instruction sets for operating the device correlated with the first digital picture includes modifying one or more instruction sets of an application, the application running on the processor circuit. In further embodiments, the causing the processor circuit to execute the one or more instruction sets for operating the device correlated with the first digital picture includes *modifying a source code, a bytecode, an intermediate code, a compiled code, an interpreted code, a translated code, a runtime code, an assembly code, or a machine code*. In further

embodiments, the causing the processor circuit to execute the one or more instruction sets for operating the device correlated with the first digital picture includes *modifying at least one of: the memory unit, a register of the processor circuit, a storage, or a repository where instruction sets are stored or used*. In further embodiments, the causing the processor circuit to execute the one or more instruction sets for operating the device correlated with the first digital picture includes *modifying one or more instruction sets for operating an application or an object of the application, the application running on the processor circuit*.

See, e.g., Ex. 1001, 11:28-12:5; Decl. ¶ 257.

Grotmol discloses that autonomous operation of a robot may include modifying the learned instruction set to compensate for, e.g., “position mismatch” and “orientation mismatch.” Ex. 1005, 42:46-58; *see also id.*, Abstract (“An evaluation may be performed in order to determine if the image may correspond to a shifted (e.g., left/right) version of a stored image as previously observed. If the new image is shifted left, right turn command may be issued.”). Therefore, when a training configuration is loaded into a second physical device – *see supra* claim 13, and the second robot encounters “position mismatch” or “orientation mismatch” from the first robot, the instruction set is shifted. Thus, Grotmol discloses that the “first one or more instruction sets” are “are modified and applied to the second device, wherein the second device is caused to perform one or more operations defined by: the modified the first one or more instruction sets for operating the first device” – *see supra* claim 4. Decl. ¶¶ 258-259.

14. Claim 15

- a. **[15.1] The system of claim 4, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:**

Grotmol discloses “wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least” limitations [15.2]-[15.4]. *See infra* limitations [15.2]-[15.4]; *See supra* limitations [1.1]-[1.3]; Ex. 1005, Abstract, 3:27-30, 8:5-15, 33:14-30, 4:13-17; Decl. ¶¶ 178-186, 261.

- b. **[15.2] modifying: the first one or more instruction sets for operating the first device, or a copy of the first one or more instruction sets for operating the first device,**

Grotmol discloses processor(s) modifying the first one or more instruction sets for operating the first device. Decl. ¶¶ 262-263.

Grotmol discloses that autonomous operation of a robot may include modifying the instruction set to compensate for, e.g., “position mismatch” and “orientation mismatch.” Ex. 1005, 42:46-43:3; *see also id.*, Abstract (“An evaluation may be performed in order to determine if the image may correspond to a shifted (e.g., left/right) version of a stored image as previously observed. If the new image is shifted left, right turn command may be issued. If the new image is shifted right then left turn command may be issued.”). For example, Grotmol discloses that the “motor commands at a given time step may be obtained by taking the stored motor

commands from the training buffer that may correspond to the best matching image,” and “[t]hose motor commands may be combined with the output from the PID controller in order to stabilize operation of the robot.” *Id.*, 43:9-11. Thus, Grotmol discloses that the “first one or more instruction sets for operating the first device” are “modified.” Decl. ¶¶ 262-263.

- c. **[15.3] and wherein the determining ... includes: determining the modified the first one or more instruction sets for operating the first device based on the at least partial match between the new one or more digital pictures and the first one or more digital pictures, or**

As described above for limitations [4.3] and [15.2] and claim 14, Grotmol discloses that the motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the *best matching* stored image and the motor commands may be modified to compensate for, e.g., position mismatch and orientation mismatch. Ex.1005, 43:5-9, Abstract, 2:29-37, 42:46-43:3; Decl. ¶¶ 219-222, 262. Accordingly, Grotmol discloses determining the modified the first one or more instruction sets for operating the first device based on the at least partial match between the new one or more digital pictures and the first one or more digital pictures. Decl. ¶¶ 264-265.

- d. **[15.4] and wherein the causing ... includes: causing the first device or the second device to perform one or more operations defined by the modified the first one or more instruction sets for operating the first device, or**

As described above for limitations [4.3], [4.4], [15.2], and [15.3], Grotmol discloses that the motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the *best matching* stored image and the motor commands may be modified to compensate for, e.g., position mismatch and orientation mismatch, and the robot (i.e., first device) can autonomously perform the motor commands, including the modified motor commands. Ex. 1005, 43:5-9, Abstract, 2:29-37, 42:46-43:3; Decl. ¶¶ 219-222, 225-226, 262-265. Accordingly, Grotmol discloses causing the first device to perform one or more operations defined by the modified the first one or more instruction sets for operating the first device. Decl. ¶¶ 266-267.

15. Claim 16

Grotmol discloses “wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least: modifying at least one of: the first one or more digital pictures ... and wherein the at least partial match between the new one or more digital pictures and the first one or more digital pictures includes: (iii) at least partial match between the new one or more digital pictures and the modified the first one or more digital pictures.” Decl. ¶¶ 268-270.

Grotmol discloses captured and stored digital pictures can be undergo preprocessing operations including “resampling, cropping, light balancing, and/or

feature extraction” (e.g., modifying the digital pictures). *See* Ex. 1005, 39:13-22 (“[d]uring training, images (e.g., raw and/or pre-processed) may be stored in a memory buffer (training buffer). In one or more implementations, preprocessing operations may comprise resampling, cropping, light balancing, and/or feature extraction. Motor commands issued by a trainer corresponding to time instances when the images are acquired may be stored.”). Grotmol also discloses that pixels from individual image instances are selected, scaled, normalized, and/or equalized. *See, e.g., id.*, 4:18-22 (“The method may comprise applying a transformation to an instance of the sensory input. The transformation may be configured to produce scaled input based on analysis of the history.”); 30:47-31:2 (discussing “normalizing” sensory input by “shifting and/or scaling input features”); 31:61-32:5 (discussing equalizing sensory data); 3:42-46; 4:58-63; 31:3-8. Accordingly, Grotmol discloses determining whether there is a partial match between the new image and modified first image. *See supra* limitation [4.3]. Decl. ¶¶ 219-222, 268-270.

16. Claim 17

- a. **[17.1] The system of claim 4, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:**

Grotmol discloses the system of claim 4, “wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to

further perform at least” limitations [17.2]-[17.4]. *See infra* limitations [17.2]-[17.4]; *See supra* limitations [1.1]-[1.3]; Ex. 1005, Abstract, 3:27-30, 8:5-15, 33:14-30, 4:13-17; Decl. ¶¶ 178-186, 271.

b. **[17.2] modifying: the first one or more digital pictures, or**

Grotmol discloses modifying captured and stored digital pictures. Decl. ¶¶ 272-274.

Grotmol discloses captured and stored digital pictures can undergo preprocessing operations including “resampling, cropping, light balancing, and/or feature extraction” (e.g., modifying the digital pictures). *See* Ex. 1005, 39:13-22 (“During training, images (e.g., raw and/or pre-processed) may be stored in a memory buffer (training buffer). In one or more implementations, preprocessing operations may comprise resampling, cropping, light balancing, and/or feature extraction. Motor commands issued by a trainer corresponding to time instances when the images are acquired may be stored.”). Grotmol also discloses that pixels from individual image instances are selected, scaled, normalized, and/or equalized. *See, e.g.*, Ex. 1005, 4:18-22 (“The method may comprise applying a transformation to an instance of the sensory input. The transformation may be configured to produce scaled input based on analysis of the history.”), 30:47-31:2 (discussing “normaliz[ing]” sensory input by “shifting and/or scaling input features”), 31:61-32:5 (discussing equalizing sensory data); *see also id.*, 3:41-46, 4:58-62, 31:3-8.

Thus, Grotmol discloses modifying the first one or more digital pictures. Decl. ¶¶ 272-275.

- c. **[17.3] and wherein the learning ... includes: learning the modified the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device, or**

As described with respect to limitation [1.6], Grotmol discloses a learning process that includes learning correlations between digital pictures and instruction sets for operating the robot (first device). *See supra* limitation [1.6]; Ex. 1005, 33:59-62, Abstract, 43:5-9, 39:17-19; Decl. ¶¶ 195-199. As described above with respect to limitation [17.2], Grotmol discloses modifying digital pictures, including the first one or more digital pictures that are received during the training process. *See supra* limitation [17.2]; Ex. 1005, 39:13-22, 4:8-22, 30:37-31:2, 31:65-32:5; Decl. ¶¶ 272-273. Accordingly, Grotmol discloses “learning the modified the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.” Decl. ¶¶ 275-277.

- d. **[17.4] and wherein the determining ... includes: determining the first one or more instruction sets for operating the first device based on at least partial match between the new one or more digital pictures and the modified the first one or more digital pictures, or**

As described above for limitation [4.3], Grotmol discloses motor commands at given step may be obtained from the training buffer based on the best matching

image. Ex. 1005, 43:5-9, Abstract, 2:30-37, 28:19-40; Decl. ¶¶ 219-222. A POSITA would understand that where the first one or more digital pictures stored during the learning process had been modified, the partial match would be determined using the current image (i.e., new one or more digital pictures) and the modified stored image (i.e., modified first one or more digital pictures). Decl. ¶ 279. Thus, Grotmol discloses “determining the first one or more instruction sets for operating the first device based on at least partial match between the new one or more digital pictures and the modified the first one or more digital pictures.” Decl. ¶¶ 278-280.

17. Claim 18

- a. **[18.1] The system of claim 4, wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least:**

Grotmol discloses “wherein the machine readable code, when executed by the one or more processors, causes the one or more processors to further perform at least” limitations [18.2]-[18.5]. *See infra* limitations [18.2]-[18.5]; *See supra* limitations [1.1]-[1.3]; Ex. 1005, Abstract, 3:27-30, 8:5-15, 33:14-30, 4:13-17; Decl. ¶¶ 178-186, 281.

- b. **[18.2] modifying: the first one or more instruction sets for operating the first device, or**

Grotmol discloses processor(s) “modifying: the first one or more instruction sets for operating the first device.” Decl. ¶¶ 282-285.

Grotmol discloses that autonomous operation of a robot may include modifying the instruction set to compensate for, e.g., “position mismatch” and “orientation mismatch.” Ex. 1005, 42:46-43:3; *see also id.*, Abstract (“An evaluation may be performed in order to determine if the image may correspond to a shifted (e.g., left/right) version of a stored image as previously observed. If the new image is shifted left, right turn command may be issued. If the new image is shifted right then left turn command may be issued.”). For example, Grotmol discloses that the “motor commands at a given time step may be obtained by taking the stored motor commands from the training buffer that may correspond to the best matching stored image,” and “[t]hose motor commands may be combined with the output from the PID controller in order to stabilize operation of the robot.” *Id.*, 43:5-11; *see also id.*, 42:37-40 (“The best match image obtained using image match methodology (e.g., such as described herein) may be used to determine changes (corrections) to motor commands during path navigation by a robotic vehicle.”).

Further, Grotmol describes a “listen mode,” in which a robot observes a teacher’s commands in order to “learn an association.” Ex. 1005, 23:4-12. Grotmol also describes an “override mode,” in which a robot will operate autonomously until a teacher sends a command, in which case the robot will apply the teacher’s command and “try to remember it for the next time the same situation occurs.” *Id.*, 23:12-22. When the robot remembers the new command, overriding the previous

one, the robot learns the first digital picture correlated with the modified instruction set. Decl. ¶¶ 283-284.

Thus, Grotmol discloses “modifying: the first one or more instruction sets for operating the first device.” Decl. ¶¶ 282-285.

- c. **[18.3] and wherein the learning ... includes: learning the first one or more digital pictures correlated with the modified the first one or more instruction sets for operating the first device, or ...,**

As described above for limitations [1.6], [4.3]-[4.4], and [18.2], Grotmol discloses that motor commands at given step may be obtained from the training buffer based on best matching image, the motor commands may be modified, the motor commands are executed to perform autonomous operation of the robotic vehicle, and the robotic vehicle learns the new correlation. Ex. 1005, Abstract, 3:27-39, 33:59-62, 42:37-40, 42:46-58, 43:5-11; *see supra* limitations [1.6], [4.3]-[4.4], [18.2]; Decl. ¶¶ 219-222, 283-284.

Accordingly, Grotmol discloses “wherein the learning ... includes: learning the first one or more digital pictures correlated with the modified the first one or more instruction sets for operating the first device.” Decl. ¶¶ 286-287.

- d. **[18.4] and wherein the determining ... includes: determining the modified the first one or more instruction sets for operating the first device based on the at least partial match between the new one or more digital pictures and the first one or more digital pictures, or ...,**

As described above for limitations [4.3], [18.2], and [18.3], Grotmol discloses that motor commands at a given step may be obtained from the training buffer based on the best matching image, and motor commands can be modified to compensate for, e.g., “position mismatch” and “orientation mismatch” between the new and stored images, and the new correlation is learned. Ex. 1005, Abstract, 3:27-39, 33:59-62, 42:37-40, 42:46-58, 43:5-11; *see supra* limitations [1.6], [4.3]-[4.4], [18.2]; Decl. ¶¶ 219-222, 283-287.

Accordingly, Grotmol discloses “determining the modified the first one or more instruction sets for operating the first device based on the at least partial match between the new one or more digital pictures and the first one or more digital pictures.” Decl. ¶¶ 288-289.

- e. **[18.5] and wherein the causing ... includes: causing the first device or the second device to perform one or more operations defined by the modified the first one or more instruction sets for operating the first device, or**

As described above for limitations [4.4] and [18.2]-[18.4], Grotmol discloses that the system causes the robot (i.e., first device) to perform operations defined by the motor commands and the motor commands can be modified to compensate for, e.g., “position mismatch” and “orientation mismatch” between the new and stored images, and the robotic vehicle executes the instruction set. Ex. 1005, Abstract, 3:27-39, 33:59-62, 42:37-40 42:46-58, 43:5-11; Decl. ¶¶ 219-222, 283-289.

Accordingly, Grotmol discloses “wherein the causing ... includes: causing the first device or the second device to perform one or more operations defined by the modified the first one or more instruction sets for operating the first device.” Decl. ¶¶ 290-291.

18. Claim 19

Grotmol discloses claim 19 that depends from claim 4, and requires the “learning” of the first correlation and that the determining of the first instruction set is done at least in part by an artificial intelligence system. Decl. ¶¶ 292-299.

The ’583 patent broadly describes an “artificial intelligence unit” as a generic piece of hardware or software. *See e.g.*, Ex. 1001, 2:47-3:22. Neither claim 19 nor the specification of the ’583 patent limits the artificial intelligence system to a specific artificial intelligence model or technique. *Id.*, 95:49-64. The ’583 patent identifies “supervised learning” and “unsupervised learning” as examples of artificial intelligence models or techniques that can be used, and for determining based on partial match “any level of similarity, however high or low, maybe be utilized” and “portions (i.e., regions, features, pixels, etc.)” can be compared to determine similarity. Ex. 1001, 95:49-64, 102:29-103:5, 103:6-18; Decl. ¶¶ 294-295.

As recited in claim 19 of the ’583 patent, the “artificial intelligence system” must perform the following functions:

(i) learning the first one or more digital pictures correlated with the first one or more instruction sets; and

(ii) determining the first one or more instruction sets based on at least a partial match between the new digital image and the first digital image.

Grotmol discloses an “artificial intelligence unit” as a part of, or operating on, its processor circuit. *See, e.g.*, Ex. 1005, claim 12 (reciting “one or more physical processors configured to operate a random k-nearest neighbors learning process to determine a motor control indication”); Decl. ¶ 297.

As described above for limitation [1.6], Grotmol discloses a system in which processors perform the operation of learning correlations between digital pictures and instruction sets for operating the first device. Ex. 1005, 23:3-29, 33:59-62, 34:49-55; Decl. ¶¶ 193-199. Grotmol further describes that either supervised or unsupervised learning may be utilized. Ex. 1005, 16:8-12 (“[T]he control entity may comprise a computerized agent such as a multifunction adaptive controller operable using reinforcement and/or unsupervised learning and capable of training other robotic devices for one and/or multiple tasks.”), 24:17-20 (“Apparatus and methods for hierarchical learning are disclosed herein, in accordance with one or more implementations. In supervised learning, a user may train simple tasks by demonstrating example tasks to the system.”). And, as described above for limitation [4.3], Grotmol discloses a system in which one or more processors

perform the operation of determining an instruction by calculating a similarity measure between a current image and a stored image, then determining whether that similarity measure breaches a threshold. Ex. 1005, 43:5-9, Abstract, 2:29-38, 28:19-40; Decl. ¶¶ 219-222.

Accordingly, Grotmol discloses an artificial intelligence system that performs the learning and determining steps. Decl. ¶¶ 292-299.

19. Claim 20

Grotmol discloses the first portion of claim 20 which requires an “artificial intelligence system” with “one or more inputs for receiving one or more digital pictures, and one or more outputs for providing one or more instruction sets.” Decl. ¶¶ 300-303.

As depicted in Figure 8, Grotmol discloses the “sensory” functionality, which may include camera input. *See* Ex. 1005, 9:1-3 (“[T]he output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.”). Both Grotmol and the ’583 patent disclose that the digital picture data may come from, for example, a CCD image sensor. *Compare* Ex. 1001, 70:2-11; *with* Ex. 1005, 9:8-22.

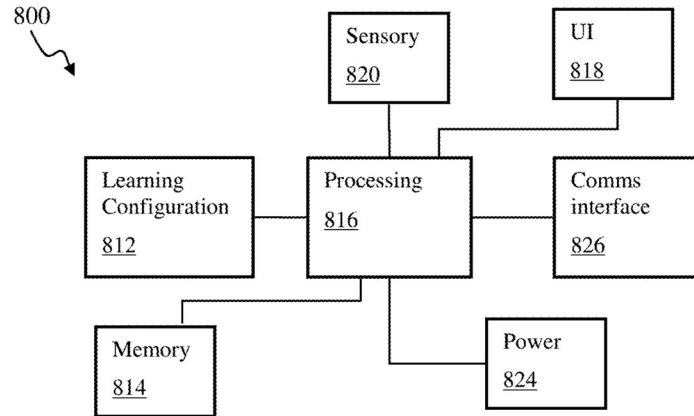


FIG. 8

Ex. 1005, Fig. 8

As further depicted in Figure 4A of Grotmol, the adaptive control system (400) receives sensory input (402). Ex. 1005, 15:64-66.

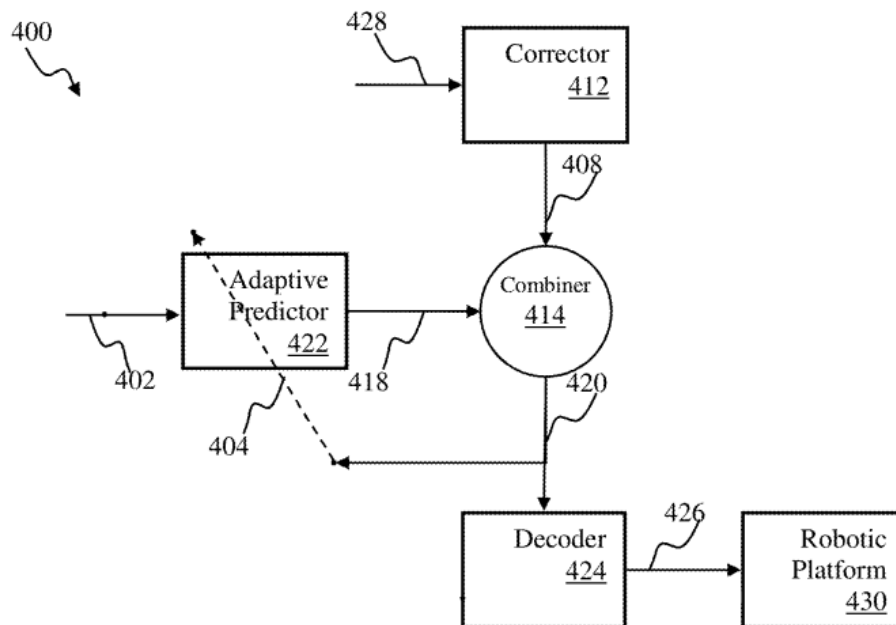


FIG. 4A

Ex. 1005, Fig. 4A

Grotmol discloses that the adaptive control system's corrector, predictor, and combiner "may cooperate to produce a control signal 420 for the robotic platform 410. In one or more implementations, the control signal 420 may comprise one or more motor commands (e.g., pan camera to the right, turn right wheel forward), *sensor acquisition commands (e.g., use high resolution camera mode)*, and/or other commands." *Id.*, 15:26-33. Thus, Grotmol discloses one or more inputs for receiving one or more digital pictures, and one or more outputs for providing one or more instruction sets. Decl. ¶¶ 300-301.

Grotmol further discloses the second portion of claim 20 which requires "wherein the learning ... includes applying the first one or more digital pictures to the inputs of the artificial intelligence system, and wherein the determining ... includes: applying the new one or more digital pictures to the inputs of the artificial intelligence system; and receiving the first one or more instruction sets for operating the first device from the outputs of the artificial intelligence system." Decl. ¶ 302. Grotmol discloses that the learning process includes receiving new digital pictures that are compared to stored digital pictures, and further that the processor includes an input for receiving images, and that instruction sets are determined and cause the robot to perform. Thus, Grotmol discloses that the learning process includes applying the first one or more digital pictures to the inputs of the artificial intelligence system, and wherein the determining ... includes: applying the new one

or more digital pictures to the inputs of the artificial intelligence system; and receiving the first one or more instruction sets for operating the first device from the outputs of the artificial intelligence system. Decl. ¶¶ 300-303.

20. Claim 21

Grotmol discloses claim 21 which requires “wherein the one or more inputs for receiving the one or more digital pictures include: one input for one digital picture.” *See supra* claim 20; Decl. ¶¶ 300-304; Ex. 1005, Figure 8, 8:66-9:3, 10:16-22, 15:19-33, Figure 4A, 115:64-66.

21. Claim 22

Grotmol discloses claim 22 which requires “the artificial intelligence system includes: a knowledgebase.” As described above in claim 9, Grotmol discloses a knowledgebase stored in a memory unit that is coupled to the processor. Ex. 1005, Abstract, 3:27-39, 8:66-9:3, 10:23-26, 38:36-58, 39:13-22. Thus, Grotmol discloses an artificial intelligence system that includes a knowledgebase. Decl. ¶¶ 244, 305; *see supra* claims 9, 19; Ex. 1005, Abstract, 38:36-58, 39:13-22, 3:31-39, 8:66-9:3, 10:23-25.

22. Claim 23

Grotmol discloses claim 23 which requires that the first digital picture correlated with the first instruction includes a connection between the digital picture and the instruction set. As described above with respect to claim 1, Grotmol discloses storing correlated digital pictures and instruction sets as training sets. Ex.

1005, 33:24-27 (“The trained controller of the robot may comprise a trained configuration configured to enable autonomous operation (e.g., without user input) by the robotic device.”); Decl. ¶¶ 194-195. The training sets “may be configured based on learning of associations between sensory context (e.g., presence of an obstacle in front of the robotic vehicle) and a respective action (e.g., right turn) during training.” Ex. 1005, 33:59-62. An association between a digital picture and an action (i.e., instruction set) constitutes a “connection” between the two.⁷ Decl. ¶ 307. Thus, Grotmol discloses a system “wherein the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device include the first one or more digital pictures connected, using at least one or more connections, with the first one or more instruction sets for operating the first device.” Decl. ¶¶ 306-308.

23. Claim 24

Grotmol discloses claim 24 which requires that the first digital picture is a whole digital picture, a collection of pixels, or a collection of values. *See supra* limitations [1.4], [4.2] and claim 20; Decl. ¶¶ 188, 217, 300, 309; Ex. 1005, Abstract, 8:66-9:3, 9:8-22.

⁷ The ’583 patent describes that the connection between the digital picture and the instruction set “may include or is associated with at least one of: an occurrence count, a weight, a parameter, or a data.” Ex. 1001, 15:12-14.

24. Claim 26

- a. **[26.1] The system of claim 4, wherein at least some elements of the system are included in: a single device, or multiple devices,**

Grotmol discloses “wherein at least some elements of the system are included in: a single device, or multiple devices.” Decl. ¶ 310.

Grotmol discloses the system of claim 4 implemented using a “self-contained robotic apparatus,” and thus, discloses that the elements of the system are included in a single device. Ex. 1005, 3:27-30; *see supra* claims 1, 4.

- b. **[26.2] and wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices,**

Grotmol discloses “wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices.” Decl. ¶¶ 311-314. Grotmol states that the terms “‘processor’, ‘microprocessor’ and ‘digital processor’ are meant generally to include all types of digital processing devices.” Ex. 1005, 8:5-15. Grotmol gives examples of “digital processing devices” such as digital signal processors (DSPs), field programmable gate arrays (FPGAs), programmable logic devices (PLDs), and application-specific integrated circuits (ASICs). *Id.* Grotmol further states that “the terms ‘computer’, ‘computing device’, and ‘computerized device’, include . . . literally any other device capable of executing a set of instructions and processing an incoming data signal.” *Id.*, 7:30-42. A microcontroller is both a “digital processing device” and a “device

capable of executing a set of instructions and processing an incoming data signal.” Decl. ¶ 312. Thus, a POSITA would understand that Grotmol discloses a system in which the “one or more processor circuits” includes “one or more microcontrollers.”

In addition, Grotmol discloses several processor circuits that a POSITA would have understood include one or more microcontrollers. For example, Grotmol discloses adaptive controllers that are designed to process and execute specific tasks. Grotmol discloses an “‘Adaptive controller’ of type ‘1’” that may learn to approach the target autonomously generating correction torques on robot’s wheels and may not control the gripper (“output N”). Ex. 1005, 25:38-42. Grotmol also discloses an “‘Adaptive controller’ of type ‘2’” that is specifically designed to learn to turn the robot and appropriately grasp targets by autonomously effectuating correct torques on the robot’s wheels and gripper while not controlling another wheel. *Id.*, 26:14-19. Grotmol further discloses a “learning controller” that is used to process images that are outputted from the camera. *Id.*, 10:15-21. Grotmol further discloses a “predictor” that is specifically programmed and used to assign priorities to available tasks. *Id.*, 10:45-53. A POSITA would have understood that Grotmol’s “adaptive controllers,” “learning controller,” and “predictor” include one or more microcontrollers. Decl. ¶¶ 312-313.

- c. **[26.3] and wherein the first device includes a robot, a vehicle, ... or a mechanical machine,**

Grotmol discloses the first device includes a robotic vehicle. Ex. 1005, 3:27-30, 38:36-39 (“In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1), Abstract; *see supra* claims 1, 4; Decl. ¶ 315.

d. **[26.4] and wherein the second device includes a robot, a vehicle ... or a mechanical machine,**

Grotmol discloses that second device includes a robotic vehicle. Decl. ¶¶ 315-318.

Grotmol discloses that the training sets and training configurations from a first robot (which include the instruction sets for operating first device) can be “stored” or “saved.” Ex. 1005, 33:12-58 (discussing saved training configurations). The saved training configuration may then be “loaded into the robot being trained thereby providing one or more trained behaviors to the robot.” *Id.*, 33:27-30. Further, the “trained configuration may be loaded to one or more *other robots* in order to provide learned behaviors.” *Id.*, 33:30-32. Thus, Grotmol discloses a second device (“one or more other robots”) that can autonomously perform one or more operations defined by one or more instruction sets for operating the first device training configuration from first robot (e.g., training sets with circumstance representations correlated to instruction sets) that is loaded into a second one or more other robots to “provide learned behaviors.” Decl. ¶¶ 316-318.

- e. **[26.5] and wherein the first one or more digital pictures include: one or more still digital pictures, or one or more motion digital pictures, and wherein the new one or more digital pictures include: one or more still digital pictures, or one or more motion digital pictures,**

Grotmol discloses the first digital picture and the new digital picture includes one or more digital pictures. *See supra* limitations [1.4], [1.6], [2.1], [4.2]; Ex. 1005, Abstract, 3:27-39, 10:16-22, 8:57-9:3; Decl. ¶¶ 187-188, 202, 216, 319.

- f. **[26.6] and wherein an instruction set of the first one or more instruction sets for operating the first device includes at least one of: only one instruction, a plurality of instructions, ... one or more commands, ... or one or more information.**

Grotmol discloses an instruction set for operating the first device, where the instruction set includes, for example, one or more instructions, one or more commands, and one or more information. *See supra* limitations [1.5], [1.6], [2.2], [4.3]; Decl. ¶¶ 190-191, 196-197, 320; Ex. 1005, 3:27-39, 39:13-22, 33:14-30.

25. Claim 28

Grotmol discloses “wherein the at least partial match between the new one or more digital pictures and the first one or more digital pictures is at least in part based on a product of one or more values that represent one or more portions of the new one or more digital pictures and one or more values that represent one or more portions of the first one or more digital pictures.” Decl. ¶¶ 321-325.

Grotmol discloses, for example, that the similarity between the new and stored images can be calculated using spatial frequency representation. Ex. 1005, 40:34-63. Grotmol describes that each image is processed using a windowing function and a fast-Fourier Transform (FFT), and from the resulting frequency representations, a normalized cross-power spectrum is obtained. *Id.*, 40:34-41; Decl. ¶¶ 322. A POSITA would understand that the cross-power spectrum is calculated by multiplying the FFT of one image with the complex conjugate of the FFT of the other image. Decl. ¶ 322. Grotmol discloses that the inverse FFT is then applied to the cross-power spectrum to change it back from the frequency domain to the spatial domain, resulting in a cross-correlation function that shows the degree of similarity between the two images for various amounts of shift. Decl. ¶ 323. The maximum value (argmax) of the cross-correlation function is then found, which corresponds to the shift (x,y values) that gives the best match between the two images. Ex. 1005, 40:42-63; Decl. ¶ 323. A POSITA would understand that this process “is at least in part based on a product of one or more values that represent one or more portions of the new one or more digital pictures and one or more values that represent one or more portions of the first one or more digital pictures” — i.e., when calculating the cross-power spectrum by multiplying the Fourier transform of one image with the complex conjugate of the Fourier transform of the other image. Decl. ¶ 324.

26. Claim 29

Grotmol discloses “wherein the generating the new one or more digital pictures includes: detecting one or more objects in the first device’s surrounding or the second device’s surrounding; generating one or more representations of the one or more objects; and generating the new one or more digital pictures that include the one or more representations of the one or more objects.” Decl. ¶¶ 326-330.

As noted above, the ’583 patent provides examples of objects in pictures such as “car, pedestrian, door, building, animal, person, etc.” Ex. 1001, 88:22-37. The ’583 patent further describes various approaches to object recognition, such as “identifying and/or analyzing object features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc.” *Id.*

Similarly, Grotmol discloses that images 2000 and 2010 may “comprise representations [of] one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.” Ex. 1005, 38:56-65.

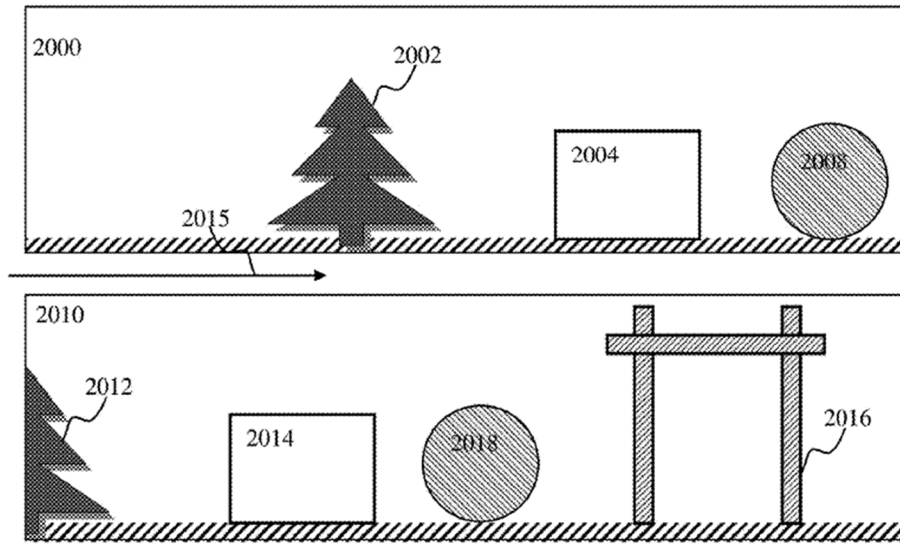


FIG. 20A

Ex. 1005, Fig. 20A

Grotmol likewise discloses that the sensor used to capture these images “may provide information associated with objects within the field-of-view []. In one or more implementations, such as *object recognition*, and/or obstacle avoidance, the output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.” Ex. 1005, 8:57-9:3. *See, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented”). Grotmol, like the ’583 patent, further includes edge detection. Ex. 1005, 7:60-63 (“[T]he term feature may refer to a representation of an object edge”).

Grotmol describes that the autonomous operation of a robot may be based in part on recognized objects and their characteristics and that the images may include representations of the objects detected in the device’s surroundings. *E.g., id.*, 10:23-

34 (noting that “tasks of the robot may be configured based on a context” such as “state of the environment (e.g., object size, location)”), 24:31-25:5 (describing embodiment in which a robot is trained to “find an object, grasp it, and bring it back to base”). Decl. ¶ 329.

27. Claim 31: The system of claim 4, wherein the generating the first one or more instruction sets for operating the first device includes generating the first one or more instruction sets for operating the first device as one or more representations of a second one or more instruction sets for operating the first device.

Grotmol discloses “wherein the generating the first one or more instruction sets for operating the first device includes generating the first one or more instruction sets for operating the first device as one or more representations⁸ of a second one or more instruction sets for operating the first device.” Decl. ¶¶ 331-333.

Grotmol discloses that the system may be configured to remove a learned behavior. Ex. 1005, 34:18-21 (“The BrainOS software platform may be configured to enable users to selectively remove a learned behavior (and/or a portion thereof) via an undo and/or time machine operation.”). In implementation, the system

⁸ The ’583 patent specification does not define or even contain the term “representations of a [first or second or any] instruction sets.” The most similar term used in the specification is “representations of that particular instruction set” which is described in the context of obtaining instruction sets implemented through assembly language. *See* Ex. 1001, 79:33-36 (“Syntax, addressing modes, operands, and/or other elements of an assembly language instruction set may translate directly into numeric (i.e. binary, etc.) representations of that particular instruction set.”).

reloads the previous configuration to unlearn a behavior, specifically, “[b]y loading at time t_1 the configuration saved at time $t_0 < t_1$, the robot controller effectively ‘forgets’ what it learned in time interval $t_0 < t < t_1$.” *Id.*, 34:25-27. *See also id.*, 34:43-45 (“A user may select a given bookmark in order to restore (undo) the learning configuration to the state corresponding to time of the bookmark.”). Decl. ¶ 332. When the system then uses the reloaded configuration, the system is generating the first instruction set as a representation of a second instruction set. Decl. ¶ 332.

28. Claim 32: The system of claim 4, wherein the first one or more digital pictures that depict the at least the portion of the first device’s surrounding include one or more digital pictures that depict a representation of the at least the portion of the first device’s surrounding.

Grotmol discloses claim 32 which requires the first digital picture depict a representation of the first device’s surrounding. *See supra* limitation [1.4], claim 29; Ex. 1005, Abstract, 10:16-22, 8:57-9:3, 38:28-60, Decl. ¶¶ 188, 327-329, 334-335.

Grotmol’s system discloses that, during the image matching process, in order to compare images, “the given image may be shifted by amount determined from the image matching operation,” where, for example, “image 2010 may be shifted to the right by amount depicted by arrow 2015 in FIG. 20A.” Ex. 1005, 40:55-63. Moreover, as described in claim 29, images 2000 and 2010 may “comprise representations [of] one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.” *Id.*, 38:56-65. *See supra* claim 29.

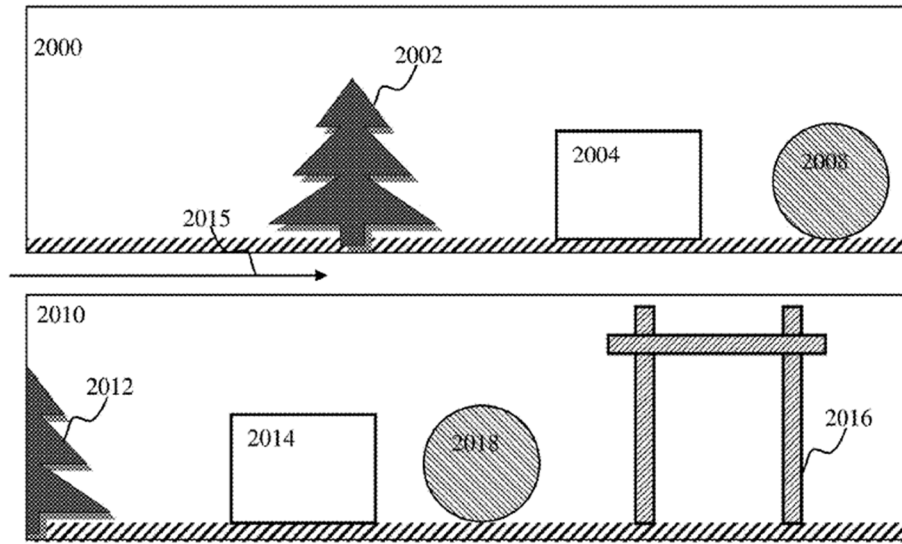


FIG. 20A
Ex. 1005, Fig. 20A

Accordingly, Grotmol discloses digital pictures that depict a representation of the at least the portion of the first device’s surrounding. Decl. ¶¶ 334-335.

29. Claim 33

Independent claim 33 is a method claim that corresponds to system claim 1. The analysis above explaining how Grotmol discloses the limitations of claim 1 applies to the limitations of claim 33 as set forth in the chart below. Grotmol anticipates claim 33. Decl. ¶ 336.

Claim 33	Disclosure of Grotmol
[33.1] A method implemented using a computing system that includes one or more processors, the method comprising:	To the extent the preamble is limiting, Grotmol discloses “[a] method implemented using a computing system that includes one or more processors.” Decl. ¶ 336; Ex. 1005, claim 1 (“A method of determining a control signal for a robot, the method being performed by a special purpose computing

	platform having one or more processors executing instructions stored by a non-transitory computer-readable storage medium”); <i>see supra</i> limitations [1.1]-[1.3]; Ex. 1005, 3:27-30, Abstract, 10:45-48, 8:5-15, claim 12, Figure 7; Decl. ¶¶ 178-186.
[33.2] receiving or generating a first one or more digital pictures, wherein the first one or more digital pictures depict at least a portion of a first device’s surrounding;	<i>See supra</i> limitation [1.4]; Decl. ¶¶ 187-189, 336; Ex. 1005, 3:27-39, Abstract, 10:16-22, 8:57-9:3, 3:28-60.
[33.3] receiving or generating a first one or more instruction sets for operating the first device; and	<i>See supra</i> limitation [1.5]; Decl. ¶¶ 190-192, 336; Ex. 1005, 3:27-39, 39:13-22, Abstract, 33:14-30.
[33.4] learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.	<i>See supra</i> limitation [1.6]; Decl. ¶¶ 193-200, 336; Ex. 1005, 3:27-39, 23:3-29, 24:17-30, 24:49-25:5, 33:13-30, 33:59-62, Abstract, 43:5-9, 39:17-19.

30. Claim 34

Independent claim 34 is a system claim that corresponds to system claim 1 and method claim 33. As explained in the claim construction section, a corresponding structure for the “means for receiving or generating” and “means for learning” is “one or more processor circuits.” Decl. ¶ 64. The analysis above explaining how Grotmol discloses the limitations of claims 1 and 33 applies to the limitations of claim 34 as set forth in the chart below. Grotmol anticipates claim 34. Decl. ¶ 337.

Claim 34	Disclosure of Grotmol
[34.1] A system comprising:	To the extent the preamble is limiting, Grotmol discloses a system implemented as a “self-contained robotic apparatus.” Ex. 1005, 3:27-30; <i>see also id.</i> , Abstract (“An apparatus and methods for training and/or operating a robotic device to follow a trajectory.”); <i>id.</i> , 10:45-48. Accordingly, Grotmol discloses a system. Decl. ¶¶ 178-179, 337; <i>see supra</i> limitations [1.1]-[1.3].
[34.2] means for receiving or generating a first one or more digital pictures, wherein the first one or more digital pictures depict at least a portion of a first device's surrounding;	<i>See supra</i> limitation [1.4]; Decl. ¶¶ 64, 188-189, 337; Ex. 1005, 3:27-39, Abstract, 10:16-22, 8:57-9:3, 3:28-60.
[34.3] means for receiving or generating a first one or more instruction sets for operating the first device; and	<i>See supra</i> limitation [1.5]; Decl. ¶¶ 64, 190-192, 337; Ex. 1005, 3:27-39, 39:13-22, Abstract, 33:14-30.
[34.4] means for learning the first one or more digital pictures correlated with the first one or more instruction sets for operating the first device.	<i>See supra</i> limitation [1.6]; Decl. ¶¶ 64, 193-200; 337; Ex. 1005, 3:27-39, 23:3-29, 24:17-30, 24:49-25:5, 33:13-30, 33:59-62, Abstract, 43:5-9, 39:17-19.

X. CONCLUSION

For the foregoing reasons, the *inter partes* review should be instituted and claims 1-9, 11-24, 26, 28-29, and 31-34 of the '583 patent should be cancelled as unpatentable.

Dated: June 30, 2023

Respectfully submitted,

/Anish R. Desai/

Anish R. Desai

Lead Counsel for Petitioner

Registration No. 73,760

Weil, Gotshal & Manges LLP

767 Fifth Ave.

New York, NY 10153

T: 212-310-8000

F: 212-310-8007

CERTIFICATE OF COMPLIANCE

1. The undersigned certifies that this brief complies with the type volume limitations of 37 CFR § 42.24(a)(1)(i). This brief contains 13,966 words (excluding the table of contents, the table of authorities, mandatory notices under 37 CFR § 42.8, the certificate of service, certificate of compliance, and appendix of exhibits), as calculated by the “Word Count” feature of Microsoft Word 2016, the word processing program used to create it.

2. The undersigned further certifies that this brief complies with the typeface requirements of 37 CFR § 42.6(a)(2)(ii) and typestyle requirements of 37 CFR § 42.6(a)(2)(iii). This brief has been prepared in a proportionally spaced typeface using Microsoft Word 2016 in Times New Roman 14 point font.

Dated: June 30, 2023

Respectfully submitted,

/Anish R. Desai/

Anish R. Desai

Lead Counsel for Petitioner

Registration No. 73,760

Weil, Gotshal & Manges LLP

767 Fifth Ave.

New York, NY 10153

T: 212-310-8000

F: 212-310-8007

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. § 42.105 and § 42.6(e), the undersigned certifies that on June 30, 2023, a complete and entire copy of this **PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 11,055,583 PETITIONER TESLA, INC.'S POWER OF ATTORNEY, and all supporting exhibits** were sent via FedEx next business day delivery to the following:

Jasmin Cosic
Autonomous Devices LLC
2271 Myla Lane
Melbourne, FL 32935

/Timothy J. Andersen/
Timothy J. Andersen
Case Manager
Weil, Gotshal & Manges LLP
2001 M Street, N.W., Suite 600
Washington, D.C. 20036
202-682-7000
timothy.andersen@weil.com