

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.,
Petitioner,

v.

AUTONOMOUS DEVICES, LLC
Patent Owner

U.S. PATENT NOS. 10,452,974 and 11,238,344

DECLARATION OF JASON JANET

I.	Introduction.....	1
II.	Summary of Opinions.....	1
III.	Qualifications and Compensation.....	2
IV.	Relevant Legal Standards	10
A.	Claim Construction	11
B.	Invalidity	12
1.	Anticipation.....	13
2.	Obviousness	14
C.	Person of Ordinary Skill in the Art	17
V.	Background of the Technology	18
VI.	The Prior Art.....	20
A.	U.S. 9,604,359 (“Grotmol”)	20
B.	U.S. 8,639,644 (“Hickman”).....	22
VII.	Claim Construction.....	26
A.	“means for generating or receiving”; “means for anticipating”; and “means for executing” (’974 patent claim 19)	27
B.	“means for processing” (’344 patent claim 20).....	28
VIII.	The ’974 Patent.....	30
A.	The Claims	30
B.	Ground 1: Claims 1-9 and 12-20 Are Anticipated by Grotmol	38
1.	Claim 1	38
a.	[1.1] A method implemented using a computing system that includes one or more processor circuits, the method comprising:	38

- | | | |
|----|--|----|
| b. | [1.2] accessing a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device and a second correlation including a second circumstance representation correlated with a second one or more instruction sets for operating the first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device and the second circumstance representation represents a second circumstance detected at least in part by the one or more sensors of the first device,..... | 40 |
| c. | [1.3] and wherein at least a portion of the first correlation and at least a portion of the second correlation are learned in a learning process that includes operating the first device at least partially by a user;..... | 45 |
| d. | [1.4] generating or receiving a third circumstance representation, wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device or at least in part by one or more sensors of a second device;..... | 46 |
| e. | [1.5] anticipating the first one or more instruction sets for operating the first device learned in the learning process based on at least partial match between the third circumstance representation and the first circumstance representation; and: | 50 |
| f. | [1.6] at least in response to the anticipating, executing the first one or more instruction sets for operating the first device learned in the learning process, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process. | 52 |
| 2. | Claim 2: The method of claim 1, wherein the knowledgebase further includes: a first knowledge cell | |

	and a second knowledge cell, and wherein the first knowledge cell includes the first correlation and the second knowledge cell includes the second correlation.	54
3.	Claim 3: The method of claim 1, wherein the learning process includes: creating, inserting, deleting, modifying, or manipulating an element of the first correlation, or creating, inserting, deleting, modifying, or manipulating an element of the second correlation.	57
4.	Claim 4.....	58
a.	[4.1] The method of claim 1, wherein the learning process includes: generating or receiving the first circumstance representation, and generating or receiving the second circumstance representation; and,.....	59
b.	[4.2] obtaining or receiving the first one or more instruction sets for operating the first device, and obtaining or receiving the second one or more instruction sets for operating the first device, ...	59
c.	[4.3] and wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes: generating or receiving the fourth circumstance representation; and obtaining or receiving the fourth one or more instruction sets for operating the second device.....	60
5.	Claim 5: The method of claim 1, wherein the anticipating the first one or more instruction sets for operating the first device learned in the learning process based on the at least partial match between the third circumstance representation and the first circumstance representation includes:	62
6.	Claim 6: The method of claim 1, wherein the at least the portion of the first correlation and the at least the portion	

	of the second correlation are learned in the learning process while the user operates the first device, and wherein the one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process correspond to the user's methodology of operating the first device learned in the learning process.....	64
7.	Claim 7	66
	a. [7.1] The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating: the first device, the second device, or a third device,.....	66
	b. [7.2] and wherein a first connection is generated to connect the first correlation with the second correlation, and wherein a second connection is generated to connect the second correlation with the fourth correlation.	67
8.	Claim 8	70
	a. [8.1] The method of claim 1, further comprising: modifying the first one or more instruction sets for operating the first device learned in the learning process or a copy of the first one or more instruction sets for operating the first device learned in the learning process, ...	70
	b. [8.2] and wherein the executing the first one or more instruction sets for operating the first device learned in the learning process includes executing the modified the first one or more instruction sets for operating the first device learned in the learning process or the modified the copy of the first one or more instruction sets for operating the first device learned in the learning process, and wherein the performing, by the first device or by the second device, the one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process includes performing, by the first device or by the second device, one or more operations	

	defined by the modified the first one or more instruction sets for operating the first device learned in the learning process or by the modified the copy of the first one or more instruction sets for operating the first device learned in the learning process.	72
9.	Claim 9: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the first device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes operating the first device at least partially by the user.	73
10.	Claim 12	75
	a. [12.1] The method of claim 1, wherein the first circumstance includes one or more objects detected at least in part by the one or more sensors of the first device at a first time or during a first time period, and wherein the second circumstance includes one or more objects detected at least in part by the one or more sensors of the first device at a second time or during a second time period, ...	75
	b. [12.2] and wherein the third circumstance includes: one or more objects detected at least in part by the one or more sensors of the first device at a third time or during a third time period, or one or more objects detected at least in part by the one or more sensors of the second device at a third time or during a third time period.....	78
11.	Claim 13: The method of claim 1, wherein the first circumstance representation is a data structure including one or more data about the first circumstance of the first device, and wherein the second circumstance representation is a data structure including one or more data about the second circumstance of the first device, and wherein the third circumstance representation is a	

	data structure including one or more data about: the third circumstance of the first device, or the third circumstance of the second device.....	79
12.	Claim 14: The method of claim 1, wherein the first circumstance representation includes: one or more object representations, or one or more collections of object representations, and wherein the second circumstance representation includes: one or more object representations, or one or more collections of object representations, and wherein the third circumstance representation includes: one or more object representations, or one or more collections of object representations.	82
13.	Claim 15	85
	a. [15.1] The method of claim 1, wherein, to correlate the first circumstance representation with the first one or more instruction sets for operating the first device, a determination is made that the first one or more instruction sets for operating the first device temporally correspond to the first circumstance representation,.....	85
	b. [15.2] and wherein the first circumstance representation includes: one or more object representations, or one or more collections of object representations.	86
14.	Claim 16: The method of claim 1, wherein the anticipating the first one or more instruction sets for operating the first device learned in the learning process based on the at least partial match between the third circumstance representation and the first circumstance representation includes determining the first one or more instruction sets for operating the first device learned in the learning process based on the at least partial match between the third circumstance representation and the first circumstance representation.	86
15.	Claim 17	87
	a. [17.1] The method of claim 1, wherein elements of the computing system are included in: a single device, or multiple devices, ...	87

b.	[17.2] and wherein the one or more processor circuits include: one or more microcontrollers, one or more computing circuits, or one or more electronic circuits,.....	88
c.	[17.3] and wherein the memory includes: a volatile memory, or a non-volatile memory, ...	88
d.	[17.4] and wherein the first device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine, ...	88
e.	[17.5] and wherein the second device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine, ...	89
f.	[17.6] and wherein an instruction set of the first one or more instruction sets for operating the first device includes at least one selected from the group comprising: only one instruction, a plurality of instructions, one or more inputs, one or more commands, one or more computer commands, one or more keywords, one or more symbols, one or more operators, one or more variables, one or more values, one or more objects or object references, one or more data structures or data structure references, one or more functions or function references, one or more parameters, one or more signals, one or more characters, one or more digits, one or more user operating directions, one or more user directions, one or more user inputs, one or more representations of one or more user actions, one or more representations of one or more user clicks, one or more binary bits, one or more assembly language commands, one or more states, one or more codes, one or more data, and one or more information, ...	89
g.	[17.7] and wherein the one or more sensors of the first device include at least one selected from the group comprising: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, and one or more apparatuses for detecting objects or object properties, and wherein the one or more sensors of	

	the second device include at least one selected from the group comprising: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, and one or more apparatuses for detecting objects or object properties, ...	91
h.	[17.8] and wherein the at least the portion of the first correlation includes: one portion of the first correlation, multiple portions of the first correlation, all portions of the first correlation, or the entire first correlation, and wherein the at least the portion of the second correlation includes: one portion of the second correlation, multiple portions of the second correlation, all portions of the second correlation, or the entire second correlation, ...	94
i.	[17.9] and wherein an object of the first circumstance is the same as an object of the third circumstance, or multiple objects of the first circumstance are the same as multiple objects of the third circumstance, or all objects of the first circumstance are the same as all objects of the third circumstance, or all objects of the first circumstance are different than all objects of the third circumstance.....	95
16.	Claim 18	97
17.	Claim 19	99
18.	Claim 20	101
C.	Ground 2: Claims 10 and 11 Are Obvious Based on Grotmol in View of the Knowledge of a POSITA	102
1.	Claim 10: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the first device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the fourth correlation is learned in another learning process that	

	includes operating the first device at least partially by another user.	102
2.	Claim 11: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes operating the second device at least partially by another user.	104
D.	Ground 3: Claims 4 and 11 Are Obvious Based on Grotmol in View of Hickman	105
1.	Motivation to Combine Grotmol and Hickman.....	105
2.	Claim 4: The method of claim 1, wherein the learning process includes: generating or receiving the first circumstance representation, and generating or receiving the second circumstance representation; and obtaining or receiving the first one or more instruction sets for operating the first device, and obtaining or receiving the second one or more instruction sets for operating the first device, and wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes: generating or receiving the fourth circumstance representation; and obtaining or receiving the fourth one or more instruction sets for operating the second device.	111
3.	Claim 11: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating	

	the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes operating the second device at least partially by another user.	116
IX.	The '344 Patent.....	118
A.	The Claims	118
B.	Ground 1: Claims 1-5, 7-9, and 13-20 Are Anticipated by Grotmol	126
1.	Claim 1	126
a.	[1.1] A system comprising:	126
b.	[1.2] one or more processors configured to perform at least:	127
c.	[1.3] accessing a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device;	128
d.	[1.4] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;.....	134
e.	[1.5] generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;.....	135
f.	[1.6] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance	

	representation and the first circumstance representation;.....	138
g.	[1.7] and at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.	140
2.	Claim 2: The system of claim 1, wherein the first one or more instruction sets for operating the first device include one or more information about one or more states of: the first device, or a portion of the first device.	143
3.	Claim 3: The system of claim 1, wherein the first circumstance representation includes a first one or more object representations, and wherein the second circumstance representation includes a second one or more object representations.	144
4.	Claim 4: The system of claim 3, wherein the second circumstance representation represents the second circumstance detected at least in part by the one or more sensors of the first device, and wherein the first device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.....	146
5.	Claim 5: The system of claim 3, wherein the second circumstance representation represents the second circumstance detected at least in part by the one or more sensors of the second device, and wherein the second device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.....	147
6.	Claim 7	148
a.	[7.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: 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, ...	148
b.	[7.2] and wherein the anticipating the first one or more instruction sets for operating the first device	

	based on the at least partial match between the second circumstance representation and the first circumstance representation includes: anticipating the modified the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, or anticipating the modified the copy of the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, and.....	150
c.	[7.3] wherein the executing the first one or more instruction sets for operating the first device includes: executing the modified the first one or more instruction sets for operating the first device, or executing the modified the copy of the first one or more instruction sets for operating the first device, and wherein the autonomously performing, by the first device or by the second device, the one or more operations defined by the first one or more instruction sets for operating the first device includes: autonomously performing, by the first device or by the second device, one or more operations defined by the modified the first one or more instruction sets for operating the first device, or autonomously performing, by the first device or by the second device, one or more operations defined by the modified the copy of the first one or more instruction sets for operating the first device.	151
7.	Claim 8.....	152
a.	[8.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: modifying at least one of: the first circumstance representation, a copy of the first circumstance representation, the second circumstance representation, or a copy of the second circumstance representation, ...	152

- b. [8.2] and wherein the at least partial match between the second circumstance representation and the first circumstance representation includes: (i) at least partial match between the modified the second circumstance representation and the first circumstance representation, (ii) at least partial match between the modified the copy of the second circumstance representation and the first circumstance representation, (iii) at least partial match between the second circumstance representation and the modified the first circumstance representation, (iv) at least partial match between the second circumstance representation and the modified the copy of the first circumstance representation, (v) at least partial match between the modified the second circumstance representation and the modified the first circumstance representation, (vi) at least partial match between the modified the copy of the second circumstance representation and the modified the copy of the first circumstance representation, (vii) at least partial match between the modified the second circumstance representation and the modified the copy of the first circumstance representation, or (viii) at least partial match between the modified the copy of the second circumstance representation and the modified the first circumstance representation. 154
8. Claim 9: The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating the first device, and wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the third circumstance representation or at least a portion of the second one or more instruction sets for operating the first device is learned in another learning process that includes operating the first device at least partially by the user.....156

9.	Claim 13: The system of claim 3, wherein the system further comprising: an artificial intelligence system, and wherein the learning the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device includes learning, at least in part by the artificial intelligence system, the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device, and wherein the anticipating the first one or more instruction sets for operating the first device includes anticipating, at least in part by the artificial intelligence system, the first one or more instruction sets for operating the first device.....	158
10.	Claim 14	162
	a. [14.1] The system of claim 13, wherein the artificial intelligence system includes: one or more inputs for receiving one or more object representations, and one or more outputs for providing one or more instruction sets, ...	162
	b. [14.2] and wherein the learning, at least in part by the artificial intelligence system, the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device includes applying the first one or more object representations to the one or more inputs of the artificial intelligence system,.....	165
	c. [14.3] and wherein the anticipating, at least in part by the artificial intelligence system, the first one or more instruction sets for operating the first device includes: applying the second one or more object representations to the one or more inputs of the artificial intelligence system; and receiving the first one or more instruction sets for operating the first device from the one or more outputs of the artificial intelligence system.	166
11.	Claim 15: The system of claim 13, wherein the artificial intelligence system includes at least one of: the knowledgebase, or a neural network.....	167

12. Claim 16: The system of claim 3, wherein the first circumstance representation correlated with the first one or more instruction sets for operating the first device include at least a portion of the first one or more object representations connected, using at least one or more connections, with at least a portion of the first one or more instruction sets for operating the first device. 168
13. Claim 17: The system of claim 3, wherein the anticipating the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation includes determining the first one or more instruction sets for operating the first device based on at least partial match between the second one or more object representations and the first one or more object representations. 169
14. Claim 18 171
 - a. [18.1] The system of claim 3, wherein at least some elements of the system are included in: a single device, or multiple devices, and.....171
 - b. [18.2] wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices, and.....171
 - c. [18.3] wherein the memory includes: a non-transitory machine readable medium, a volatile memory, a non-volatile memory, a storage device, or a storage system, and.....171
 - d. [18.4] wherein the one or more sensors of the first device include at least one of: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, or one or more apparatuses for detecting objects or object properties, and wherein the one or more sensors of the second device include at least one of: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, or one or more apparatuses for detecting objects or object properties, and.....172

- e. [18.5] wherein the first device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine, and.....175
- f. [18.6] wherein the second device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine, and.....175
- g. [18.7] wherein the first circumstance representation is a data structure that includes one or more data about the first circumstance, and wherein the second circumstance representation is a data structure that includes one or more data about the second circumstance, ...176
- h. [18.8] and wherein the first circumstance includes one or more objects detected at least in part by the one or more sensors of the first device at a first time or during a first time period, and wherein the second circumstance includes: one or more objects detected at least in part by the one or more sensors of the first device at a second time or during a second time period, or one or more objects detected at least in part by the one or more sensors of the second device at a second time or during a second time period, and 178
- i. [18.9] 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, multiple instructions, one or more inputs, one or more commands, one or more keywords, one or more symbols, one or more operators, one or more variables, one or more parameters, one or more characters, one or more numbers, one or more values, one or more signals, one or more binary bits, one or more functions, one or more function references, one or more objects, one or more object references, one or more data structures, one or more data structure references, one or more states, one or more representations of one or more states, one or more representations of one or more user inputs, one or more codes, one or more data, or one or more information, and 181

j.	[18.10] wherein the first one or more object representations include: one object representation, ... and wherein the second one or more object representations include: one object representation, ... and wherein the first one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the second one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the at least the portion of the first circumstance representation includes: at least a portion of one object representation ... and	181
k.	[18.11] wherein the at least the portion of the first one or more instruction sets for operating the first device includes: at least a portion of one instruction set of the first one or more instruction sets for operating the first device	182
l.	[18.12] and wherein: an object of the first circumstance is the same as an object of the second circumstance, multiple objects of the first circumstance are the same as multiple objects of the second circumstance, all objects of the first circumstance are the same as all objects of the second circumstance, or all objects of the first circumstance are different than all objects of the second circumstance.	182
15.	Claim 19	183
16.	Claim 20	185

C. Ground 2: Claims 10 and 11 Are Obvious Based on Grotmol in View of the Knowledge of a POSITA 187

1.	Claim 10: The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating the first device, and wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the third circumstance representation or at least a
----	---

	portion of the second one or more instruction sets for operating the first device is learned in another learning process that includes operating the first device at least partially by another user.....	187
2.	Claim 11: The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating a third device, and wherein the third circumstance representation represents a third circumstance detected at least in part by one or more sensors of the third device, and wherein at least a portion of the third circumstance representation or at least a portion of the second one or more instruction sets for operating the third device is learned in another learning process that includes operating the third device at least partially by: the user, or another user.	189
D.	Ground 3: Claims 6, 11, and 12 Are Obvious Based on Grotmol in View of Hickman	191
1.	Motivation to Combine Grotmol and Hickman.....	191
2.	Claim 6: The system of claim 3, wherein the system further comprising: a server that receives from the first device at least one of: the first circumstance representation, or the first one or more instruction sets for operating the first device, and wherein the second device receives from the server at least one of: the first circumstance representation, or the first one or more instruction sets for operating the first device; and wherein the second device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.....	192
3.	Claim 11	198
4.	Claim 12	201
a.	[12.1] The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating a third device, and wherein the third circumstance representation represents a third circumstance detected at least in part by one or more sensors of	

	the third device, and wherein the one or more processors are further configured to perform at least:	201
b.	[12.2] generating or receiving a fourth circumstance representation, wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by one or more sensors of a fourth device; anticipating the second one or more instruction sets for operating the third device based on at least partial match between the fourth circumstance representation and the third circumstance representation; and at least in response to the anticipating the second one or more instruction sets for operating the third device, executing the second one or more instruction sets for operating the third device, wherein the fourth device autonomously performs one or more operations defined by the second one or more instruction sets for operating the third device;	202
X.	Declaration.....	205

I. Introduction

1. I have been retained by Petitioner Tesla, Inc. (“Petitioner” or “Tesla”), as an independent expert witness in the above-captioned *inter partes* reviews (“IPRs”), in which Tesla has requested that the U.S. Patent and Trademark Office cancel as unpatentable all claims of U.S. Patent Nos. 10,452,974 (“the ’974 patent”) and 11,238,344 (“the ’344 patent”) (together, “the Object Representation Patents”).

2. This declaration sets forth my analyses and opinions based on my knowledge, experience, and the materials I have considered. As I explain below, it is my opinion that all claims of the ’974 patent and ’344 patent are directed to subject matter that was routine, conventional, and well known in the art before the patents’ priority dates. As would be readily appreciated by one of skill in the art, the claims of the ’974 patent and the ’344 patent are anticipated and/or rendered obvious by the prior art references discussed herein.

3. I have reviewed the documents referenced in this declaration. I understand they have been submitted as exhibits in conjunction with Tesla’s Petitions for IPR.

II. Summary of Opinions

4. Based on my knowledge, experience, and the materials that I have reviewed, it is my opinion that claims 1-20 of the ’974 patent and claims 1-20 of the ’344 patent are anticipated and/or obvious. Specifically:

- (i) Claims 1-9 and 12-20 of the '974 patent are anticipated by U.S. Patent No. 9,604,359 ("Grotmol") (Ex. 1004);
- (ii) Claims 10 and 11 of the '974 patent are obvious based on Grotmol in view of the Knowledge of a POSITA;
- (iii) Claims 4 and 11 of the '974 patent are obvious based on Grotmol in view of U.S. Patent No. 8,639,644 ("Hickman") (Ex. 1005);
- (iv) Claims 1-5, 7-9, and 13-20 of the '344 patent are anticipated by Grotmol;
- (v) Claims 10 and 11 of the '344 patent are obvious based on Grotmol in view of the Knowledge of a POSITA; and
- (vi) Claims 6, 11, and 12 of the '344 patent are obvious based on Grotmol in view of Hickman.

III. Qualifications and Compensation

5. I have a Ph.D. in Electrical & Computer Engineering from North Carolina State University (or "NCSU"), which I earned in 1998. Prior to that I received a M.S. in Integrated Manufacturing Systems Engineering from North Carolina State University in 1994 and a B.S. in Mechanical Engineering from the University of Virginia in 1990.

6. Since 1991, I have been active in the robotics, sensors, artificial intelligence, pattern analysis, computer vision, image/video processing, and

automation fields. I have authored numerous publications and have co-authored a textbook entitled *Computational Intelligence*. My PhD dissertation is entitled *Pattern Analysis, Tracking and Control for Autonomous Vehicles Using Neural Networks*. It described the use of unsupervised training (i.e., the Hyper-Ellipsoid Clustering Neural Network, or "HECNN"), and supervised training (i.e., the Region-Feature Neural Network, or "RFNN") for dissimilar mobile robots to learn aspects of their common environments through multiple types of sensors (e.g., acoustic rangefinding, infrared proximity, tactile and imaging), build knowledge libraries, and then transfer knowledge between for operational use and/or refinement. Among other outputs, the robots learned to recognize spaces, landmarks within those spaces, and then plan and track motion from the accumulated and shared knowledge. Efficiency and accuracy were shown to increase significantly with the combined unsupervised and supervised learning, and knowledge transfers.

7. I have designed, built, and marketed robots, automated systems, and components thereof, including ground mobile robots, unmanned aerial vehicles (UAV), unattended sensors automated storage and retrieval systems (ASRS), submersible mobile robots, proof-of-concept extra-terrestrial robots, manually manipulated ("twiddled") submersible foils, wall climbing mobile robots, and augmented reality systems.

8. I have taught the following courses: Introduction to Robotics and Automation (Duke and NCSU); Introduction to Control Theory (Duke and NCSU); Distributed Real-Time Controls (NCSU); and myriad independent studies courses in the areas of robotics, automation, artificial intelligence, autonomy, pattern analysis, and control systems. I have also served on several MS- and PhD-level graduate student committees, designed qualifying exam problems, served on the NCSU IMSEI Board, and participated in curriculum development for undergraduate and graduate level programs.

9. I have also started, sponsored, and advised multiple champion student teams for international robot competitions including, but not limited to, the NASA/ASCE Extra-Terrestrial Robotics Competition, the DARPA Grand Challenge, the AUVSI/ONR Autonomous Underwater Vehicle Competition, and the European CLAWAR Wall-Climbing Robot competition.

10. I have initiated multiple unmanned aerial vehicle (UAV) projects at both academic and industry levels. Academic UAVs include, but are not limited to, Quadcopters with Hybrid Remote and Autonomous Control, Marsupial UAVs that Deploy and Recover Unmanned Ground Vehicles, and Wall-Climbing UAVs. Additionally, the AngelFish Cross-Domain Submersible UAV, a DARPA ASW program that I incubated and procured funding for at Teledyne, included a partnership with North Carolina State University.

11. In 1999, while I was employed by Nekton Research (“Nekton”), I captured and managed multiple programs sponsored by the Department of Defense (DoD) and private-sector companies that focused primarily on autonomous underwater vehicles (AUV), remotely operated vehicles (ROV), TwiddleFish toys, and indirect-fire projectiles. During my tenure, Nekton entered into a joint venture agreement with the founders of what eventually became known as Parata Systems – a pharmacy automation solutions company, which I supported launching.

12. After leaving Nekton in March 2002, I joined a then New Jersey-based company called Avionic Instruments, Inc. (“Avionic”). While at Avionic, I continued developing and marketing robots and supported engineering related to various design, manufacturing, quality and assembly issues on the core aerospace product lines. Avionic product lines include, but are not limited to, ducted fans, transformer rectifier units (TRU), regulated TRU (RTRU), auxiliary power unit (APU) control systems, power distribution systems (PDU), frequency converters, corner clamps, and VRAM attractors/thrusters. Customers included DoD, NASA, Boeing, Sikorsky, Augusta-Westland, Dassault, and Lockheed-Martin.

13. In late 2004, soon after Avionic was sold, I, along with others from Avionic, secured funding to develop an alpha-level multi-dispenser robotic system and afforded me time to write the RxMedic business plan and raise multiple rounds of venture capital. In late 2006, RxMedic (called “APDS” until November 2006)

was launched as a stand-alone, sole-focus venture. After the operational launch of RxMedic, I served as General Manager and eventually Chief Technical Officer. Through my roles at RxMedic, I oversaw the development of the RxMedic ADS robot, managed the intellectual property portfolio, coordinated sales and marketing, and provided strategic, fiscal, and operational leadership. In May 2010, J.M Smith Corporation acquired RxMedic, and in an effort to assist in the change of ownership, I served as a director of RxMedic until May 2011.

14. Also in late 2004, after Avionic was sold, I, along with others from Avionic, secured funding for Vortex HC through DoD contracts and robot sales, to continue developing the VRAM Mobile Robot Platform (VMRP – a wall-climbing robot), the ARTEMIS AUV (a holonomic submersible robot for counter-mine and counter-obstacle operations), the submersible crawler, the nuclear-grade boiler water reactor (BWR) inspection robot, and other robot products centered around the VRAM. Some DoD programs were/are classified, for which I maintained a SECRET clearance at both the personal and facilities level, and served as the facilities security officer (FSO). In late 2006, corresponding to the full launch of RxMedic, Vortex HC technologies were largely licensed to Teledyne SeaBotix, SeaRobotics, and HDT.

15. However, I have continued to support Vortex HC licensees and customers to date. After the sale of RxMedic, and after fulfilling my 12-month employment obligation, I joined Teledyne Technologies in summer 2011. I served

as the Senior Manager for the RTP division of Teledyne Scientific, and supported multiple DoD sponsored robotic-focused programs. Some of these programs were/are classified, for which I maintained a personal SECRET clearance. Among these programs, were cargo unmanned ground vehicles (CUGV), squad-level autonomous unmanned ground vehicles (UGV), autonomous underwater vehicles (AUV), unmanned underwater vehicles (UUV), and a cross-domain autonomous vehicle capable of transitioning between air-, surface- and underwater-domains. Cross-domain vehicles that were evaluated, designed and prototyped included, but were not limited to, the AngelFish (later called "EagleRay") submersible AUV for anti-submarine warfare, and a ball-shaped robot for countermine operations on the ground, in the beach zone and in the surf zone. Sensor design, refinement and signal processing was a major component of each program. Sensors employed include, but are not limited to: proximity sensors; ranging sensors; electro-optical imaging; long-, short- and mid-wave infrared (IR); inertial measurement units (IMU); optical flow; and radiofrequency (RF). Additionally, control systems were designed, refined and integrated into the aforementioned systems. Most control systems were closed-loop, in that they utilized sensor-based feedback; others were open-loop, where states were estimated with little or no feedback.

16. In late 2013, Avionic, a Transdigm business unit at that point, requested that I return to turn around a supply-chain issue, and assume management of its

Sikorsky S97 Raider helicopter program. The S97 Raider is purported to be the fastest, most maneuverable helicopter, due to its coaxial, counter-rotating variable pitch wings, and an aft-based push-propeller. Avionic also controlled two business units named Acme Aerospace (Acme) and Aerospace Cooling Solutions (ACS). Avionic, Acme and ACS supported the S97 program, which has met milestones and continues to produce multiple successful demonstrations. In 2013, Transdigm expanded my role to include directorship of the Avionic, Acme and ACS sales and marketing team, and to report operational and financial status at six-week intervals. Transdigm required that I move my family to New Jersey in late 2014, which influenced my decision to resign and assume the CEO role of Delta Five Systems in Raleigh, NC.

17. As CEO of Delta Five, I led a private-equity backed, hospitality-focused venture, coordinating the company's launch and strategic path including an early stage pivot. Initially focused on back-end robotics and automation, Delta Five shifted to address the rampant bed bug problem—a top priority for hoteliers—with a novel unattended sensor and trap, called the Telemetered Pest Monitoring System (TPMS), which is based on computer vision and internet of things (IoT), and has proven capable of scaling to other pests. In addition to the TPMS, Delta Five developed a novel means to mass produce a natural, unscented aggregation pheromone that, in concert with placement and heat, lured invertebrates to the

TPMS. After leading Delta Five and serving on its Board for four years, I resigned to join ARA, but continue serving Delta Five as an investor and supportive transition agent.

18. In 2018, I joined ARA as sector CTO. In this role, my primary responsibilities included leading the commercialization and technology transition efforts, which require technical, business and corporate development, as well as the raising of capital and managing matrixed rosters. I assumed my other current role at ARA, Corporate Director of IP and Product Commercialization, in 2020. In this role, I evaluate, quantify, present, and recommend portfolios of IP, product lines, and ventures on the basis of competitive landscapes, addressable markets, sales strategy, resource availability and technology readiness. I also spearhead investments and provide general venture oversight.

19. My curriculum vitae is attached as Ex. 1003 to the IPR petition, and provides further information about my experience, expertise, and presentations.

20. Through my professional experience, I have gained extensive expertise in robotics, sensors, artificial intelligence, automation, computer vision, and image/video processing.

21. I am being compensated at my standard rate of \$500/hour. My compensation is in no way contingent upon my opinions or the outcome of the proceeding.

IV. Relevant Legal Standards

22. I am not an attorney, and therefore my understanding of patent law and the legal standards set forth in this report is based on explanations provided to me by counsel.

23. I understand that for any claim of a patent to claim priority to an earlier application (i.e., to benefit from the earlier application's filing date), the claims of the later patent must be fully supported by the disclosure of the earlier patent application to which priority is claimed. I understand that in order for the claims to be supported, the earlier application's disclosure must be sufficient to allow a person of ordinary skill in the art to reasonably conclude that the inventors were in possession of the claimed invention.

24. I understand that both the '974 patent and '344 patent claim priority to U.S. Patent Application No. US 15/340,991, filed on November 2, 2016. For the purposes of my opinions, I have considered the state of the art as of and shortly before this claimed priority date, and the level of knowledge that a POSITA would have possessed at that time. Unless I state otherwise, whenever I refer to any principle or technical subject matter as having been known or understood, this is meant to denote the knowledge and understanding of a POSITA at or prior to November 2, 2016.

A. Claim Construction

25. It is my further understanding that the numbered paragraphs at the end of the disclosure of a U.S. Patent are the patent “claims” that define the metes and bounds of the alleged invention. I understand that the claims of the Object Representation Patents are what is being challenged in the present IPR proceedings.

26. I have been informed that, in this proceeding, the Board must determine the scope of the claims by giving the claims their ordinary and customary meaning in light of the specification, as the claims would be interpreted by one of ordinary skill in the art.

27. I understand that patent claims generally include a “transitional” term or phrase, such as “consisting” or “comprising,” which may connect the preamble of the claim to the body of the claim. I have been informed that if a claim uses the term “consisting” as a transition term, that means that the claim is a “closed” claim, which means that the claim is limited to the claim features that follow the transition term and nothing else. On the other hand, I understand that the transition term “comprising” denotes an “open” claim, which means that the claim is not limited to only the features recited in the claim, and could encompass the listed elements as well as other unrecited elements.

28. I understand that 35 U.S.C. § 112 (f) created an exception to the general rule of claim construction called a “means plus function” limitation. These types of

terms and limitations should be interpreted to cover only the corresponding structure described in the specification, and equivalents thereof. I also understand that a limitation is presumed to be a means plus function limitation if: (a) the claim limitation uses the phrase “means for”; (b) the “means for” is modified by functional language; and (c) the phrase “means for” is not modified by sufficient structure for achieving the specified function.

29. I understand that a structure is considered structurally equivalent to the corresponding structure identified in the specification only if the differences between them are insubstantial. For instance, if the structure performs the same function in substantially the same way to achieve substantially the same result. I further understand that a structural equivalent must have been available at the time of the issuance of the claim.

B. Invalidity

30. I understand that Tesla bears the burden of proving that the challenged claims of the Object Representation Patents are invalid, and must prove this by a preponderance of the evidence, which means that invalidity must be shown to be more likely than not.

1. Anticipation

31. I have been asked to consider the question of whether the claims of the Object Representation Patents would have been anticipated by certain prior art references.

32. I understand that determining anticipation of a patent claim requires a comparison of the claim language to the prior art on a limitation-by-limitation basis.

33. I understand that a prior art reference “anticipates” an asserted claim, and thus renders the claim invalid, if all elements of the claim are disclosed in that prior art reference, either explicitly or inherently (i.e., necessarily present or implied), as those elements are arranged in the claim so that the disclosure effectively puts the public in possession of the invention. I understand that a prior art reference anticipates an asserted claim when each and every element of the claim is disclosed in a single embodiment as arranged in the claim. The mere fact that a certain outcome may result from a given set of circumstances is insufficient for inherency; rather, inherency requires that the claimed product or process at issue must be a necessary result from the operation as taught by the prior art reference. I understand that, although references cannot be combined for anticipation, additional references may be used to interpret an anticipating reference and shed light on what it would have meant to those skilled in the art at the time of the invention.

34. In order to anticipate a claimed invention, a prior art reference must enable one of ordinary skill in the art to make the invention without undue experimentation, i.e., provide an enabling disclosure. A prior art reference need not enable its full disclosure; it only needs to enable the portions of its disclosure alleged to anticipate the claimed invention.

2. Obviousness

35. I have been asked to consider the question of whether the claims of the Object Representation Patents would have been obvious. I understand that this analysis must be conducted from the perspective of the person of ordinary skill in the art, and whether the skilled artisan would consider any differences between the prior art and what is claimed to have been obvious. To make this assessment, I have been informed that the concept of patent obviousness involves four factual inquiries: (1) the scope and content of the prior art; (2) the differences between the claimed invention and the prior art; (3) the level of ordinary skill in the art; and (4) secondary considerations of non-obviousness. I have been instructed that one must not engage in hindsight. Rather, I understand that one should instead consider what the person of ordinary skill in the art would have reason to pursue further, and steps that were routinely done, such as in response to known problems, or obstacles.

36. It is my understanding that the following is a non-exhaustive list of rationales that may support the obviousness of an invention: combining prior art

elements according to known methods to yield predictable results; simple substitution of one known element for another to obtain predictable results; use of a known technique to improve a similar device (method, or product) in the same way; applying a known technique to a known device (method, or product) ready for improvement to yield predictable results; choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success; and some teaching, suggestion, or motivation in the prior art that would have led a POSITA to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention.

37. It is my understanding that the motivation to combine prior art references may be implicit and may be found in the knowledge of one of ordinary skill in the art, or in the nature of the problem to be solved. Specifically, it is my understanding that an implicit motivation to combine exists not only when a suggestion may be gleaned from the prior art as a whole, but when the “improvement” is technology-independent and the combination of references results in a product or process that is more desirable, for example because it is stronger, cheaper, cleaner, faster, lighter, smaller, more durable or more efficient. It is my further understanding that the motivation to combine references may be found in the nature of the problem to be solved where prior art references are directed to precisely the same problem.

38. I also understand that prior art may be relied on for its express disclosure and teachings. I also understand that the prior art may be relied upon for a teaching of features that are necessarily present in the prior art reference even if that specific feature is not expressly or explicitly disclosed.

39. I understand that before reaching any final conclusion on obviousness, the obviousness analysis requires consideration of objective indicia of non-obviousness, if any such indicia are offered. These must be considered to ensure that, for example, there were not some unanticipated problems, obstacles or hurdles that may seem easy to overcome in hindsight, but which were not readily overcome prior to the relevant invention date of the patents/claims at issue here. I understand that these objective indicia are also known as “secondary considerations of non-obviousness,” and may include long-felt but unmet need and unexpected results, among others. I also understand, however, that any offered evidence of secondary considerations of non-obviousness must be commensurate with the scope of the challenged claims. This means that for any offered evidence of secondary considerations of non-obviousness to be given substantial weight, I understand the proponent of that evidence must establish a “nexus,” or a sufficient connection or tie between that evidence and the merits of the claimed invention, which I understand specifically incorporates any novel element(s) of the claimed invention. If the secondary consideration evidence offered actually results from something other than

the merits of the claim, then I understand that there is no nexus or tie to the claimed invention. I also understand it is the Patent Owner who has the burden of proving that a nexus exists, and I understand that secondary considerations will not overcome a strong showing of obviousness.

C. Person of Ordinary Skill in the Art

40. I have been asked to review the Object Representation Patents from the perspective of a person of ordinary skill in the art (“POSITA”) as of the earliest claimed priority date for the patents— November 2, 2016.

41. It is my opinion that a POSITA relevant to the Object Representation Patents would have had at least: (1) a bachelor’s degree in electrical engineering, computer engineering, computer science, 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, or equivalent course work with a focus in computer vision, image/video processing, and/or machine learning.

42. Based on my education, training and experience, it is my opinion that I can accurately represent the views of a POSITA as of the earliest claimed priority date of November 2, 2016 for the ’974 patent and the ’344 patent. The opinions I provide in this declaration are provided using the viewpoint of the POSITA as of that date.

V. Background of the Technology

43. The asserted claims of the Object Representation Patents are directed to the well-known practice of employing “computing enabled devices and/or systems, ... generally directed to devices, systems, methods, and/or applications for learning a device's operation in various circumstances, storing this knowledge in a knowledgebase (i.e., neural network, graph, sequences, etc.), and enabling autonomous operation of the device.” ’974 patent, Abstract. Devices’ operations have been learned by intelligent systems for decades, at least as far back as several decades. Teach pendants, for example, have been integral to robots dating back to at least the 1990’s, where human operators would guide robots through various tasks based on sensory feedback, and the robots would learn to operate autonomously (i.e., without constant operator control) based on perceived circumstances.

44. Other types of supervised, unsupervised, and/or human-in-the-loop (HITL) training have been applied to mobile robots for decades, as well. For example, AAAI and IJCAI competitions from the mid 1990’s demonstrated mobile robots that would learn to recognize objects and features within perceived environments and, based on supervised training, initiate one or more actions including avoidance, pursuit, and manipulation. These mobile robots utilized computer vision, whiskers, ultrasonic, and infrared sensors for perception. Many would also evaluate the scene(s) utilizing a variety of well-known sensor processing

techniques such as color histograms, transforms, moments, codes, masks, and neural networks. Feedback from operator(s) in the loop might adjust, refine and optimize parameters (e.g., neural network synaptic weights and biases), and then be stored as “knowledge.” Multi-mobile robot systems, sometimes referred to as “multi-agent systems,” from as far back as the 1990’s were also shown to exchange knowledge and/or compile their respective observations to create a universal knowledge base that can be broadly used by any of a multitude of mobile robots. During subsequent operation, “knowledge” can be locked, deleted, added, and/or modified, depending on the demand for adaptability.

45. At its core, both supervised and unsupervised learning utilizes pattern analysis from well-known machine learning models, including statistics, Euclidean distances, moments, discriminators, masks, and the like. Unsupervised learning employs algorithms that learn patterns that were not previously “labeled” or “identified”; and learning features or other aspects about input data (e.g., images) can be used for clustering and/or certain aspects of supervised learning. In contrast, supervised learning maps input data, including meta-level features and topographies gleaned from unsupervised learning, to target outputs. For example, the target output of a mobile robot vehicle might be to avoid obstacles and/or react to traffic guidance mechanisms, and a human in the loop would provide the mapping from input data (e.g., images) to safe mobile robot vehicle operation. Supervised and unsupervised

learning models that integrate HITL allow humans to provide direct feedback to the model(s) for predictions, identification and decision making. As with most control models, training and operation can be based on individual images, image sequences, accumulated image aspects and even the rate at which aspects of images change. Through single and multiple events, the model refines its understanding of the world around it, and significantly improves its abilities to, for example, safely control a driver-optional mobile robot. Examples of HITL AI models abound, including supervised and unsupervised models, and were well understood at least as far back as the 1990s.

VI. The Prior Art

A. U.S. 9,604,359 (“Grotmol”)

46. Grotmol is a U.S. patent filed on January 27, 2015, which issued on March 28, 2017. Grotmol discloses robotic devices that can be trained to operate autonomously to follow a trajectory. I understand that Grotmol was not of record during prosecution of the applications that became the '974 patent or the '344 patent.

47. Grotmol discloses an apparatus and methods for training and operating a robotic device to follow a trajectory. Ex. 1004, Abstract. As Grotmol notes, a “typical approach” to autonomous navigation training entailed the recording of user controlled movements along a desired path and then a “simpl[e] replay” of the same navigation autonomously. *Id.*, 1:62-2:5. The training approach described in Grotmol

is more dynamic. Grotmol utilizes associations between user commands and camera images captured during training, such that during subsequent autonomous operation, a newly acquired image “may be compared with one or more images from the training buffer” to determine how to control the robot. Ex. 1004, Abstract.

48. Grotmol’s training process first involves a “listen mode,” in which a robot observes a user’s commands in order to “learn an association.” Ex. 1004, 23:3-11. Grotmol discloses that the robot may then switch—either manually or automatically—into an “autonomous mode,” in which the robot performs tasks autonomously based on learned associations. Ex. 1004, 23:46-48, 24:1-16. Additionally, Grotmol describes “override” and “control” modes that use varying amounts of user operation to facilitate the learning process. In “override mode,” a robot will operate autonomously until it receives a user’s command, then apply the command, and “try to remember it for the next time the same situation occurs.” *Id.*, 23:12-22. In “correct mode,” a user may provide “small corrections” to otherwise autonomous operation, where the robot “may sum the teacher’s command with what the robot already knows to get a final motor command.” *Id.*, 23:23-29, 64-67.

49. Grotmol discloses that the associations are learned by storing images captured during training 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.*, 39:13-22. Then, “[d]uring 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.” *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.” Ex. 1004, 43:5-9.

50. Grotmol further teaches that a robot’s “trained configuration may be stored” and “loaded to one or more other robots in order to provide learned behaviors.” Ex. 1004, 33:12-35. Thus, a POSITA would have understood Grotmol to disclose that the memory component of one robot can store behaviors learned by another robot.

51. Grotmol discloses that multiple learning processes can be performed to train multiple behaviors, including “simple tasks” and “complex tasks.” *Id.*, 24:17-67, 26:1-20. Grotmol does not explicitly limit the number of users that can train the multiple behaviors. Nevertheless, a POSITA would have understood the advantages of utilizing different users, of varying skills and knowledge, to train different behaviors appropriately.

B. U.S. 8,639,644 (“Hickman”)

52. Hickman is a U.S. patent filed on May 4, 2012, which issued on January 28, 2014. Hickman discloses a shared robot knowledge base that can instruct robots to interact with objects and learn based on feedback from the robots. I understand

that Hickman was not of record during the prosecution of the applications that became the '974 patent or the '344 patent. Hickman is cited on page 2 of Ex. 1004 in the list of "References Cited."

53. Hickman discloses "shared robot knowledge bases for use with cloud computing systems." Ex. 1005, Abstract. Hickman teaches that "information in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other robots." *Id.* One embodiment of Hickman's shared robot knowledge base is illustrated in Figure 4, copied below, in which the results of one robot's learning process for grasping a cup with the appropriate force can be shared with other robots:

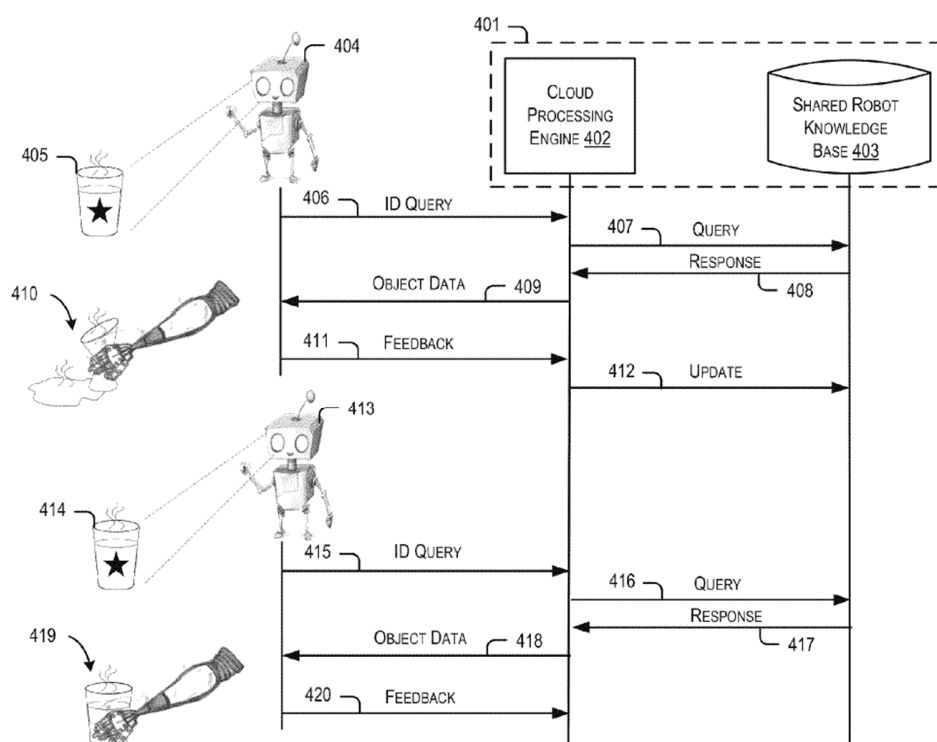


FIGURE 4

Ex. 1005, Fig. 4, 25:3-27:26.

54. As Hickman shows in Figure 4 above, a first robot obtains image data associated with a first object, shown as a cup, communicates with the shared knowledge base to figure out how to grasp the cup, and then attempts to grasp the cup but crushes it. *Id.*, 26:38-56.

In operation, the robot 404 obtains identification information associated with an object 405 via one or more sensors associated with the robot 404. For example, in FIG. 4, **the robot 404 obtains image data (e.g., an image, a 3D image, a point cloud, etc.) associated with the object 405 via optical sensors.**

After obtaining image data associated with the object 405, the robot 404 sends an identification query 406 to the cloud processing engine 402. **The identification query 406 may include at least some of the image data associated with the object 405.** After receiving image data associated with the object 405 from the robot 404, the cloud processing engine 402 may analyze the image data.

After determining that the object is a cup 405 and obtaining information about the cup 405, the cloud computing system 401 can send object data 409 associated with the cup 405 to the robot 404 in response to the identification query 406 received from the robot 404. The object data 409 may include both the identity of the cup 405 and instructions for interacting with the cup 405 based on the application or task that the robot 404 has been instructed to execute or perform.

After receiving the instructions in the object data 409 from the cloud computing system 401, the robot 404 can execute the

instructions to grasp the cup 405. However, in this example, when the robot 404 grasps the cup 405 according to the instructions it received in the object data 409 from the cloud computing system 401, **the robot 404 crushes the cup 405** and causes the coffee to spill 410. In this example interaction, the grasping force specified in the instructions was too high, and as a result, the robot crushed the cup and caused a mess.

Ex. 1005, 25:14-19, 25:24-30, 26:21-28, 26:38-47¹.

55. Hickman further explains that a human can then take control over the first robot to show it the appropriate force to use when grasping the cup, which will in turn update the shared knowledge base with the appropriate force for grasping the cup:

When an accident happens, a human can “coach” the robot on how much force to use by, for example, manually controlling the robot's hand to grasp a new cup. The robot can capture the grasping force that it used to successfully grasp the cup under the manual control of the human, and then send feedback 411 to the cloud processing system 402 with the modified grasping force. **The cloud processing system 402 can then update 412 task and object data in the shared robot knowledge base 403** to improve how the “grasp” task is applied this particular “cup” object.

Ex. 1005, 26:47-56.

56. If a second robot later encounters a second cup similar to the first, the second robot can grasp the cup with the appropriate amount of force learned based

¹ Emphasis added unless otherwise stated.

on the updated training data shared by the first robot, and confirm the accuracy of the learned force to the shared knowledge base:

Later, a second robot 413 may encounter an object 414 while performing the same (or a similar) application that requires grasping the same type of cup. The second robot 413 may send image data associated with the object 414 in an identification query 415 to the cloud computing system 401.

The robot 413 may be performing a task that requires the robot 413 to grasp the cup 414. The instructions for interacting with the cup 414 may include instructions for positioning the robot hand around the cup 414, the amount of force to use when grasping the cup 414, and how tightly to close the fingers of the robot hand around the cup 414. Importantly, **the cloud computing system 401 updated the amount of force to use when grasping the particular type of cup that the robot 413 must now grasp (i.e., cup 414) based on the feedback 411 previously received from robot 404 regarding the same type of cup.** As a result, when the robot 413 applies the “grasp” task to the type of “cup” object here (i.e., when robot 413 grasps cup 414), the robot 413 successfully grasps the cup without crushing the cup 419. Thus, **the second robot 413 has in effect learned from the experience of the earlier robot 404.**

After successfully grasping the cup 419, the robot 413 may send feedback 420 to the cloud processing engine 402 that confirms the accuracy of the instructions that the robot 413 received in the object data 418 from the cloud computing system 401.

Ex. 1005, 26:57-61, 27:7-26.

VII. Claim Construction

57. I understand that the words of the claims of a patent are generally given the ordinary and customary meaning as interpreted by a person of ordinary skill in

the art at the time of invention. I have been informed that patent claims are construed in light of the claim language, the specification (including the drawings), and the prosecution history (including the references cited during the prosecution of the patent).

A. “means for generating or receiving”; “means for anticipating”; and “means for executing” (’974 patent claim 19)

58. Claim 19 of the ’974 patent requires a “means for generating or receiving a third circumstance representation,” a “means for anticipating the first one or more instruction sets,” and a “means for executing ... the first one or more instruction sets.” In determining the corresponding structure for these claimed functions, a POSITA would have looked to the ’974 specification, which describes one or more processor circuits for performing all three:

In some aspects, the disclosure relates to a non-transitory computer storage medium having a computer program stored thereon, the program including instructions that when executed by **one or more processor circuits** cause the one or more processor circuits to perform operations comprising: **receiving a first collection of object representations**, the first collection of object representations including one or more representations of objects detected by a sensor. The operations may further include: receiving a first one or more instruction sets for operating a device. The operations may further include: learning the first collection of object representations correlated with the first one or more instruction sets for operating the device. The operations may further include: receiving a new collection of object representations, the new collection of object representations including one or more representations of objects detected by the sensor. The operations may further include: **anticipating the first one or more instruction sets** for operating the device correlated with the first collection of object representations based on at least a partial match between the new

collection of object representations and the first collection of object representations. The operations may further include: **causing an execution of the first one or more instruction sets for operating the device correlated with the first collection of object representations**, the causing performed in response to the anticipating the first one or more instruction sets for operating the device correlated with the first collection of object representations based on at least a partial match between the new collection of object representations and the first collection of object representations, wherein the device performs one or more operations defined by the first one or more instruction sets for operating the device correlated with the first collection of object representations, the one or more operations by the device performed in response to the execution.

'974 patent, 16:63-17:29; *see also id.*, 17:40-18:8; claim 1 (requiring that one or more processor circuits perform the generating/receiving, anticipating, and executing steps). A POSITA would have also looked to dependent claim 20 of the '974 patent for its confirmation that a corresponding structure for these functions is one or more processor circuits. *Id.*, 174:48-59 ("The system of claim 19, wherein the means for generating or receiving the third circumstance representation includes one or more processor circuits, and wherein the means for anticipating the first one or more instruction sets for operating the first device ... includes one or more processor circuits, and wherein the means for executing ... the first one or more instruction sets ... includes one or more processor circuits.").

B. "means for processing" ('344 patent claim 20)

59. Claim 20 of the '344 patent requires a "means for processing" that performs the following functions: "generating or receiving a second circumstance

representation ...,” “anticipating the first one or more instruction sets for operating the first device ...,” and “executing the first one or more instruction sets” The ’344 patent specification describes that the corresponding structure for the claimed functions comprises a processor or processor circuit that performs operations including receiving circumstance representations, receiving a new digital picture, anticipating the instruction set, and executing the instruction set:

In some aspects, the disclosure relates to a non-transitory computer storage medium having a computer program stored thereon, the program including instructions that **when executed by one or more processor circuits cause the one or more processor circuits to perform operations comprising: receiving a first collection of object representations**, the first collection of object representations including one or more representations of objects detected by a sensor. The operations may further include: receiving a first one or more instruction sets for operating a device. The operations may further include: learning the first collection of object representations correlated with the first one or more instruction sets for operating the device. The operations may further include: **receiving a new collection of object representations**, the new collection of object representations including one or more representations of objects detected by the sensor. The operations may further include: **anticipating the first one or more instruction sets for operating the device correlated with the first collection of object representations based on at least a partial match between the new collection of object representations and the first collection of object representations**. The operations may further include: **causing an execution of the first one or more instruction sets for operating the device correlated with the first collection of object representations**, the causing performed in response to the anticipating the first one or more instruction sets for operating the device correlated with the first collection of object representations based on at least a partial match between the new collection of object representations and the first collection of object representations, wherein the device performs one or more operations defined by the first one or more instruction sets for operating the device correlated with the

first collection of object representations, the one or more operations by the device performed in response to the execution.

'344 patent, 17:9-42. This disclosure coincides with the '344 patent's recitation of a processor or processor circuit that performs the same receiving, anticipating, and executing steps in claim 1, which is otherwise nearly identical to claim 20. *See id.*, claim 1. Given the similarity between claims 1 and 20 in all other limitations, as well as the '344 patent's disclosures of a processor or processor circuit that performs the functions of claim 1, a POSITA would have understood that the corresponding structure for the means for processing comprises a processor circuit or processor. *See also id.*, 58:64-59:41 ("Processor 11 includes one or more circuits or devices that can execute instructions fetched from memory 12 and/or other element. Processor 11 may include any combination of hardware and/or processing techniques or capabilities for executing or implementing logic functions or programs.").

VIII. The '974 Patent

A. The Claims

60. I understand that the petition challenges claims 1-20 of the '974 patent. Independent claim 1 of the '974 patent is directed to a method in which a device anticipates and autonomously executes instruction(s) corresponding to a newly detected circumstance based on comparison to a previously detected and stored circumstance. As I explain in this Declaration, a POSITA, in and prior to

2016, would have understood that there was nothing inventive about the steps of this method, which were all well known in the art at the time. Claim 1 is reproduced below:

1. A method implemented using a computing system that includes one or more processor circuits, the method comprising:
 - (a) accessing a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device and a second correlation including a second circumstance representation correlated with a second one or more instruction sets for operating the first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device and the second circumstance representation represents a second circumstance detected at least in part by the one or more sensors of the first device,
 - (b) and wherein at least a portion of the first correlation and at least a portion of the second correlation are learned in a learning process that includes operating the first device at least partially by a user;
 - (c) generating or receiving a third circumstance representation, wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device or at least in part by one or more sensors of a second device;

- (d) anticipating the first one or more instruction sets for operating the first device learned in the learning process based on at least partial match between the third circumstance representation and the first circumstance representation; and
- (e) at least in response to the anticipating, executing the first one or more instruction sets for operating the first device learned in the learning process, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process.

61. Independent claim 18 is nearly identical to claim 1, but recites a different preamble – “One or more non-transitory machine readable media storing machine readable that when executed by one or more processor circuits causes the one or more processor circuits to perform at least”

62. Independent claim 19 is nearly identical to claim 1, but requires “means for generating or receiving,” “means for anticipating,” and “means for executing,” as compared to claim 1 which specifies that the generating/receiving, anticipating, and executing steps are performed by “one or more processor circuits.” Dependent claim 20 recites that the “means for generating,” “means for anticipating,” and “means for executing” includes “one or more processor circuits.”

63. Dependent claims 2-17, which all depend from claim 1, can be summarized as follows in the table below based on the additional steps or features that they require. As I further explain in this Declaration, the steps and features added by these dependent claims were also well-known at or prior to 2016, and a POSITA would have understood that there was nothing inventive about the combination of steps and features recited in these claims.

Dependent Claim	Additional Features Claimed
Claim 2	In addition to the limitations of claim 1, claim 2 further requires the knowledgebase to include a first and second knowledge cell, which contain the first and second correlations, respectively. + 1st and 2nd knowledge cells • 1st knowledge cell includes 1st correlation • 2nd knowledge cell includes 2nd correlation
Claim 3	In addition to the limitations of claim 1, claim 3 further requires the learning process to include the creation, insertion, deletion, or modification of an element of one of the first or second correlations. + creation/deletion/modification of instruction or image while learning

Dependent Claim	Additional Features Claimed
Claim 4	In addition to the limitations of claim 1, claim 4 further requires the learning process to include a fourth correlation between instructions for, and a circumstance detected by a second device, learned in another learning process. + 4th correlation (4th CR and 4th instructions) (user undefined) • 4th CR detected by 2nd device • 4th instructions for 2nd device
Claim 5	In addition to the limitations of claim 1, claim 5 further requires the anticipation of instruction set(s) to include the determination that the number of partially matching portions of the third and first circumstance representation exceeds either a threshold number or percentage. + threshold number or % for partial match
Claim 6	In addition to the limitations of claim 1, claim 6 further requires the operations learned in the learning process to correspond to the user's methodology of operation during the learning process. + user's methodology of operation

Dependent Claim	Additional Features Claimed
Claim 7	In addition to the limitations of claim 1, claim 7 further requires the knowledgebase to include a fourth correlation; the generation of a first connection between the first and second correlations, and the generation of a second connection between the second and fourth correlations. + 4th correlation + 1st connection between 1st correlation and 2nd correlation + 2nd connection between 2nd correlation and 4th correlation
Claim 8	In addition to the limitations of claim 1, claim 8 further requires modification of the first instruction set(s) for operating the first device, and performance of those instruction(s) on either the first or second device. + modification of 1st instructions
Claim 9	In addition to the limitations of claim 1, claim 9 further requires the knowledgebase to include a fourth correlation learned in another learning process by the first device operated by the same user. + 4th Correlation (including 4th CR detected by 1st device) learned by 1st user.

Dependent Claim	Additional Features Claimed
Claim 10	In addition to the limitations of claim 1, claim 10 further requires the knowledgebase to include a fourth correlation learned in another learning process by the first device operated by a different user. + 4th correlation (4th CR and 4th instructions) learned with 2nd user • 4th CR detected by 1st device • 4th instructions for 1st device
Claim 11	In addition to the limitations of claim 1, claim 11 further requires the knowledgebase to include a fourth correlation learned in another process by the second device operated by a different user. + 4th correlation (4th CR and 4th instructions) learned with 2nd user • 4th CR detected by 2nd device • 4th instructions for 2nd device
Claim 12	In addition to the limitations of claim 1, claim 12 further requires the first, second, and third circumstances to include object(s) detected by the first device at first, second, and third times. + CRs include objects detected at specific times

Dependent Claim	Additional Features Claimed
Claim 13	In addition to the limitations of claim 1, claim 13 further requires the circumstance representations to be data structures including data about the circumstances of the device. + CRs = data structures
Claim 14	In addition to the limitations of claim 1, claim 14 further requires the circumstance representations to include one or more object representations. + CRs include object representations
Claim 15	In addition to the limitations of claim 1, claim 15 further requires determination that the first instruction set(s) temporally correspond to the first circumstance representation, which includes one or more object representations. + 1st instructions temporally correspond to 1st CR + 1st CR includes object representations
Claim 16	In addition to the limitations of claim 1, claim 16 further requires determination of the first instruction set(s) based on a partial match between the third and first circumstance representation. + determination of 1st instructions based on partially matching 1st CR and 3rd CR

Dependent Claim	Additional Features Claimed
Claim 17	In addition to the limitations of claim 1, claim 17 further requires all of the structures listed in claim 1 to include one of various listed embodiments. + various embodiments

B. Ground 1: Claims 1-9 and 12-20 Are Anticipated by Grotmol

64. As set forth in detail below, claims 1-9 and 12-20 of the '974 patent are anticipated by Grotmol. The discussion below specifies where each element of the claims is found in Grotmol, and includes a detailed explanation of the significance of the quotations and citations.

1. Claim 1

- a. **[1.1] A method implemented using a computing system that includes one or more processor circuits, the method comprising:**

Limitation [1.1]	Claimed Features
	method using computing system (including processor circuit(s))

65. Grotmol discloses a method for training and/or operating a robot comprising a computing system that includes “one or more processors”:

Another aspect of the disclosure relates to a self-contained robotic apparatus. The apparatus may comprise a platform, a first

sensor, a second sensor, a memory component, and **one or more physical processors**.

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.

As used herein, the terms “computer”, “**computing device**”, and “computerized device”, include, but are not limited to, personal computers (PCs) and minicomputers, whether desktop, laptop, or otherwise, mainframe computers, workstations, servers, personal digital assistants (PDAs), handheld computers, embedded computers, programmable logic device, personal communicators, tablet or “phablet” computers, portable navigation aids, J2ME equipped devices, smart TVs, cellular telephones, smart phones, personal integrated communication or entertainment devices, or literally any other device capable of executing a set of instructions and processing an incoming data signal.

Ex. 1004, 3:27-30, 2:14-19, 7:30-42; *see id.*, claim 1.

66. Although Grotmol does not specifically use the term “processor circuits,” a POSITA would have understood the “one or more processors” of Grotmol to include “one or more processor circuits.” The ’974 patent only defines “processor circuits” as “includ[ing] a logic circuit,” which itself “may include a microcontroller.” ’974 patent, 2:21-24. Grotmol defines “processor” as follows:

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 (ASICs). Such digital processors may be contained on a single unitary IC die, or distributed across multiple components.

Ex. 1004, 8:5-15. Given this definition, a POSITA would have understood Grotmol's use of the term "processor" to include processor circuits that include logic circuits that may include a microcontroller.

- b. **[1.2] accessing a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device and a second correlation including a second circumstance representation correlated with a second one or more instruction sets for operating the first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device and the second circumstance representation represents a second circumstance detected at least in part by the one or more sensors of the first device,**

Limitation	Claimed Features
[1.2]	access memory including 1st correlation and 2nd correlation - Correlation = CR detected by sensors + associated instructions - CR = image/video frame

67. Grotmol discloses accessing a memory that stores multiple training sets (correlations), where each training set includes an image of a scene from a robotic vehicle's operation (circumstance representation) correlated with a motor control indication configured to cause the robot to execute an action (instruction set):

The **memory component** may be configured to store training sets. 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.

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.**

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.

In some implementations of robotic vehicle navigation, **output of the sensor 166 in FIG. 1 may comprise representations** of one or more objects (e.g., targets, and/or obstacles).

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.

Additional 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.

Ex. 1004, 3:35-39, Abstract, 8:66-9:3, 10:23-26, 39:13-22. Grotmol also teaches that one embodiment for this method is autonomous vehicle navigation. *Id.*, 38:36-58.

68. Grotmol discloses that the stored images are “associated with” or “correspond to” contemporaneously issued motor control indications. Ex. 1004, 33:59-62 (“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”); Abstract (“Motor commands *associated with* a given image may be stored.”). A POSITA would have understood the associations between the motor commands and images in Grotmol to correspond to the claimed “correlations.” Grotmol further explains that the associated motor commands may correspond to the time when the images are captured. Ex. 1004, 39:17-19 (“Motor commands issued by a trainer *corresponding to* time instances when the images are acquired may be stored.”); 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.”).

69. As an initial matter, a POSITA would have understood the “circumstance representations” of the ’974 patent to comprise digital images or

instances of a video frame. The '974 patent specification does not define the term “circumstance representation,” but instead defines a “collection of object representations,” as “one or more representations of objects detected by the sensor at a time.” '974 patent, 3:31-33. The '974 patent further explains that the “sensor” that detects the object representations may include a “[c]amera ... for capturing one or more pictures” or a “motion picture camera that can capture streams of pictures (i.e. motion pictures, videos, etc.).” *Id.*, 69:45-56. Consistent with the '974 patent’s description, Grotmol discloses that images from the robot’s sensors represent circumstances detected by a sensor of the first device. Specifically, Grotmol discloses the video frame instances are detected by a “sensor component” of the robotic vehicle. Ex. 1004, 3:31-33 (“The first sensor component may be configured to provide a signal conveying a video frame comprising a plurality of pixels.”); 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.”).

70. A POSITA would further understand the “circumstance representations” of the '974 patent to comprise the images from Grotmol due to additional similarities. Both Grotmol and the '974 patent discuss CCD image sensors that capture pictures in the form of color encoded pixels. '974 patent, 69:67-70:5, 112:24-32; Ex. 1004, 9:8-22. Grotmol further discloses that these pixels are used for

object (or circumstance) recognition. Ex. 1004, 8:66-9:3 (“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.”).

71. Grotmol shows in Figure 8, copied below, that the “memory” component is connected to the “sensory” functionality through the “processing” functionality, consistent with claim 1’s requirements that the correlations stored in the memory include images detected by the robot’s sensors:

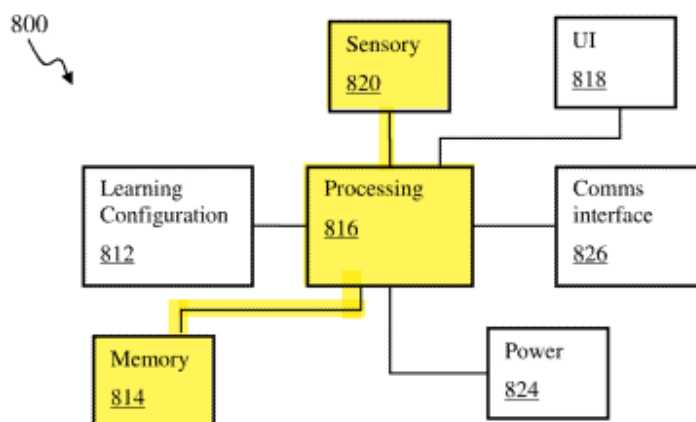


FIG. 8

Ex. 1004, Fig. 8 (annotated).

72. Grotmol further discloses accessing the training sets stored in the memory component to perform autonomous operation. Ex. 1004, 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.”); *see*

also 39:23-28 (“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.”).

73. Additionally, a POSITA would have understood the memory component of Grotmol to comprise a “knowledgebase” within the meaning of the ’974 patent. The ’974 patent states that a knowledgebase “may be or include *any data structure or arrangement* capable of storing the knowledge of a device’s operation in circumstances including objects with various properties.” ’974 patent, 102:8-12.

- c. [1.3] and wherein at least a portion of the first correlation and at least a portion of the second correlation are learned in a learning process that includes operating the first device at least partially by a user;

Limitation	Claimed Features
[1.2]	1st correlation and 2nd correlation learned with user operation

74. Grotmol discloses that a human² can operate the robot during a training process and “[t]rained behaviors of the robotic device may be configured based on

² The challenged claims of the ’974 patent recite “user(s),” which in my analysis I have considered to include humans. However, the ’974 patent does not limit the term “user” to only humans. ’974 patent, 68:22-24 (“User 50 (also referred to simply as

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. 1004, 33:59-62. In both the “listen mode” and “override mode,” Grotmol teaches that the robot will listen to commands sent by a user “and try to remember it for next time the same situation occurs.” Ex. 1004, 23:3-29; *see also id.*, 24:57-25:5 (“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).”). 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**”).

d. [1.4] generating or receiving a third circumstance representation, wherein the third circumstance

user or other suitable name or reference) comprises a human user or non-human user.”).

representation represents a third circumstance detected at least in part by the one or more sensors of the first device or at least in part by one or more sensors of a second device;

Limitation	Claimed Features
[1.4]	generation or receipt of 3rd CR detected by 1st or 2nd device

75. Grotmol discloses generating or receiving a new image (third circumstance representation) detected by the robot’s sensors in addition to the first two stored circumstance representations. Ex. 1004, 39:23-28 (“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.”). A POSITA would have understood this “most recent, current” image to correspond to the claimed “third circumstance representation,” and be detected by the sensors of a robot like all of the previous images taken during training. *See* Ex. 1004, 38:36-39 (“In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1) comprising a camera (e.g., 166). The robot may comprise a controller configured to store a *sequence of images* obtained during training phase.”).

76. Grotmol further discloses that training sets can be saved and loaded onto other robots (i.e., robots other than the one that was operated as part of the training). *Id.*, 33:27-32 (“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. In some implementations, *the trained configuration may be loaded to one or more other robots* in order to provide learned behaviors.”). Thus, a POSITA would have understood Grotmol to disclose that the most recent image can be detected by a camera of the first robot that was used during the training mode or by the camera of a second robot that has been loaded with a training set.

77. Grotmol further discloses that a sequence of video frames may be analyzed using a random k-nearest neighbors (RKNN) process in order to determine or (predict) motor control signals:

In some implementations, e.g., such as described above in connection with FIGS. 4 and/or 10A-10B a random k-nearest neighbors (RKNN) approach may be used for associating sensory context with one or motor actions. In some implementations, the RKNN methodology may comprise online learning of predicting output y (e.g., motor commands) based on the input (e.g., plurality of features detected in sensory input).

By way of an illustration of operation of a robotic device controller (e.g., 400 in FIG. 4A), sensory input (e.g., 402 comprising a sequence of video frames, inertial motion measurement, motor actuator feedback) may be analyzed using RKNN process in order to determine (predict) motor control signal (418).

Ex. 1004, 27:48-55, 28:51-56. Grotmol further explains that the inputs for the RKNN process may include specific pixels from individual video frames, and that these inputs (x) may be paired with motor control signals (y) produced at time (t):

Sensory input may comprise a plurality of features (e.g., representations of objects determined using video data). In some implementations, the RKNN process may comprise configuring a plurality of N KNN classifiers to process randomly selected subsets of features. For a given classifier C_i (e.g., 1408 in FIG. 14), a random subset x_i of features may be selected from a pool of potential features $x = \{x_1 \dots x_n\}$. As used herein, the term “classifier” may be used to describe a data processing operation configured to provide an output y (e.g., motor control command) based on analysis of a plurality of **inputs x_i (e.g., pixels of a digital image)**.

FIG. 15 illustrates an exemplary configuration for producing output y_p configured based on input x using $N=3$ classifiers and $k=4$ nearest neighbors. Tables 1500, 1530, 1560 in FIG. 15 may represent three classifier C_i instances corresponding to, e.g., index selection described by elements 1408, 1406, 1404. Individual rows 1502, 1504, 1506, 1508, 1510, 1512, 1514, 1516, 1532, 1534, 1536, 1538, 1540, 1542, 1544, 1546, 1562, 1564, 1566, 1568, 1570, 1572, 1574, 1576 may denote **training pairs (x,y) produced during training at time instances $t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_m$** .

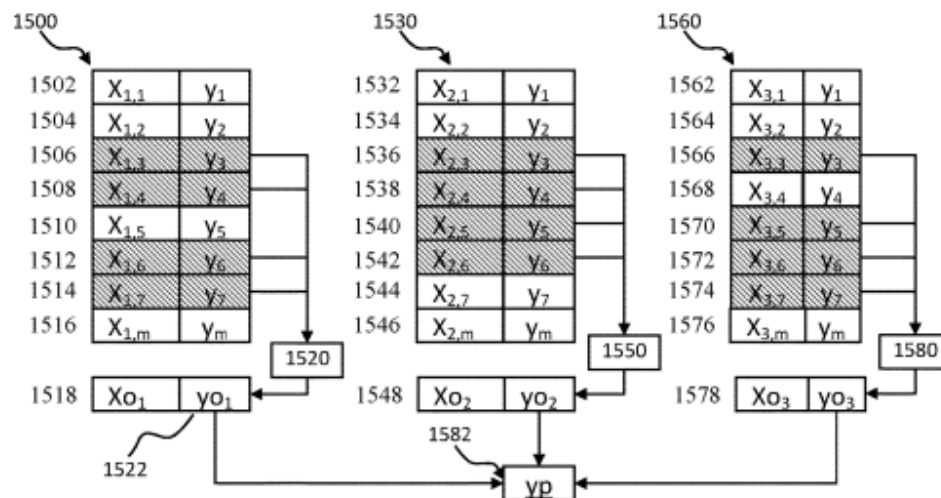


FIG. 15

Ex. 1004, 28:56-67, 29:37-46, Fig. 15. A POSITA would have understood these sections of Grotmol to explicitly disclose any number of circumstance representations x produced at time instances t_1 to t_m , including a third circumstance representation.

- e. [1.5] anticipating the first one or more instruction sets for operating the first device learned in the learning process based on at least partial match between the third circumstance representation and the first circumstance representation; and:

Limitation	Claimed Features
[1.5]	Anticipation of 1st Instructions by matching 1st CR and 3rd CR

78. Grotmol discloses anticipating motor commands (instruction sets for operating the first device) based on a partial match between the new image (third circumstance representation) and a stored image (first circumstance representation). Ex. 1004, 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”). A POSITA would have understood Grotmol’s disclosures for determining the best matching stored image by calculating a similarity measure between a current image and a stored image, and then determining whether that similarity measure breaches a threshold to correspond to the claimed “at least partial match between the third circumstance representation and the first circumstance representation.”

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. The method may comprise determining one or more potential control signals for the robot.

Ex. 1004, 2:30-38; *see also id.*, 39:13-41, 40:64-41:18. Grotmol clarifies that the term “feature” as used in the paragraph excerpted above can include values of image pixels or other image qualities. Ex. 1004, 28:19-24. (“[T]he 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 . . .”). Grotmol also clarifies that the term “training set” used in the paragraph excerpted above includes the images that a POSITA would have understood correspond to the claimed circumstance representation. Ex. 1004, 3:35-39 (“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”).

79. A POSITA would have appreciated that both Grotmol and the ’974 patent discuss pixel-wise difference analysis between two digital pictures to determine a “partial match.” ’974 patent, 66:24-28 (discussing “matching”), 73:21-54 (discussing object detection and recognition by “comparing regions of pixels”); Grotmol, 3:13-15 (“The first input type may comprise a digital image comprising a plurality of pixel values.”), 41:8-17 (“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.”).

- f. **[1.6] at least in response to the anticipating, executing the first one or more instruction sets for operating the first device learned in the learning process, wherein the first device or the second device autonomously performs one or more operations defined by the first**

one or more instruction sets for operating the first device learned in the learning process.

Limitation	Claimed Features
[1.6]	executing 1st instructions on 1st or 2nd device autonomously

80. As I discussed above for limitation [1.5], Grotmol discloses anticipating 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.

81. Grotmol discloses that after determination that the similarity measure between the new image and stored training image breaches a threshold, one or more potential control signals corresponding to the stored image obtained during training (the instruction sets correlated with the first circumstance representation) 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. 1004, 2:34-41, 3:12-13; *see also id.*, 43:5-9 (“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.”).

82. Grotmol discloses that the robot (first device) autonomously performs the operations defined by the first instruction set for operating the device. Ex. 1004, 34:56-62 (“In some implementations, the trainer may be provided with the control of the robot during training. Upon completion of the training, the robot may be configured to operate autonomously.”); 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.”), 43:5-9 (“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.”).

2. **Claim 2: The method of claim 1, wherein the knowledgebase further includes: a first knowledge cell and a second knowledge cell, and wherein the first knowledge cell includes the first correlation and the second knowledge cell includes the second correlation.**

Claim 2	Additional Features to Claim 1
	+ 1st and 2nd knowledge cells • 1st knowledge cell includes 1st correlation • 2nd knowledge cell includes 2nd correlation

83. Although Grotmol does not explicitly use the term “knowledge cell,” a POSITA would have understood Grotmol to disclose the claimed “knowledgebase” that comprises “a first knowledge cell” and a “second knowledge cell” for the reasons I set forth below.

84. The '974 patent describes that a “knowledge cell” may “compris[e] the first collection of object representations correlated with the first one or more instruction sets for operating the device.” '974 patent, 7:21-27.

Knowledge Cell 800 comprises the functionality for storing one or more Collections of Object Representations 525 correlated with any Instruction Sets 526 and/or Extra Info 527. Knowledge Cell 800 includes knowledge (i.e. unit of knowledge, etc.) of how Device 98 operated in a circumstance including objects with various properties.

Once created or generated, Knowledge Cells 800 can be used in/as neurons, nodes, vertices, or other elements in any of the data structures or arrangements (i.e. neural networks, graphs, sequences, etc.) used for storing the knowledge of Device's 98 operation in circumstances including objects with various properties, thereby facilitating learning functionalities herein.

Knowledge Cell 800 may include any data structure or arrangement that can facilitate such storing.

Id., 98:19-24, 98:24-31, 98:52-54.

85. The '974 patent further describes that “[k]nowledge cells may be connected, interrelated, or interlinked into knowledge structures using statistical,

artificial intelligence, machine learning, and/or other models or techniques.” ’974 patent, 153:66-154:2. “In some embodiments, Knowledge Structuring Unit 520 may correlate a Collection of Object Representations 525 with one or more temporally corresponding Instruction Sets 526 and/or Extra Info 527.” *Id.*, 99:11-14.

86. Grotmol describes storing sequences of images and the corresponding instruction sets (i.e., a sequence of knowledge cells). Ex. 1004, Abstract (“Motor commands associated with a given image may be stored.”); *see also id.*, 38:21-42. Grotmol further discloses that its training data is stored as a “sequence” and may be stored in “an ordered buffer in a non-volatile memory.” *Id.*, 38:39-42. An “ordered buffer” is a region of memory that is allocated for storing data in a specific order or sequence. Grotmol describes that the training buffer may be ordered according to when the images were acquired. *See* Ex. 1004, 39:65-40:4; *see also id.*, 39:17-19 (“Motor commands issued by a trainer corresponding to time instances when the images are acquired may be stored.”). A POSITA would have understood the training data stored in memory in a sequence corresponds to the claimed “knowledge cells.”

87. Furthermore, as I explained above with respect to limitations [1.2] and [1.3] of claim 1, Grotmol discloses that multiple images may be acquired during a training process, with images respectively stored in association with corresponding motor commands issued at the time the image was acquired. Thus, a POSITA would

have understood Grotmol to disclose first, second, third, etc. knowledge cells, where each knowledge cell includes the correlation.

3. **Claim 3: The method of claim 1, wherein the learning process includes: creating, inserting, deleting, modifying, or manipulating an element of the first correlation, or creating, inserting, deleting, modifying, or manipulating an element of the second correlation.**

Claim 3	Additional Features to Claim 1
	+ creation/deletion/modification of instruction or image while learning

88. As I described above with respect to claim 1, Grotmol discloses a learning process that includes operating a robot (first device) at least partially by a user. As part of this learning process, Grotmol discloses various manipulations of “element[s]”³ of training sets (first or second correlations).

89. For example, Grotmol discloses using a camera to take and store new images while following a trajectory and preprocessing operations that edit these images.

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.

³ I understand “an element” of the first or second correlations to encompass either a “circumstance representation” or “one or more instruction sets for operating a first device.” *See supra* limitation [1.2].

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. Additional 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.

Ex. 1004, Abstract, 39:13-22. A POSITA would have understood these disclosures to encompass *creation* or *modification* of the images corresponding to the first and second correlations.

90. Grotmol also discloses *insertion* of user specified instructions into the training sets that modify existing instruction sets. *Id.*, 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).”).

4. Claim 4

Claim 4	Additional Features to Claim 1
	+ 4th correlation (4th CR and 4th instructions) (user undefined) • 4th CR detected by 2nd device • 4th instructions for 2nd device

- a. **[4.1] The method of claim 1, wherein the learning process includes: generating or receiving the first circumstance representation, and generating or receiving the second circumstance representation; and,**

Limitation	Claimed Features
[4.1]	generating or receiving 1st CR & 2nd CR

91. Grotmol discloses that the learning process includes “generating or receiving the first circumstance representation, and generating or receiving the second circumstance representation.” As I discussed above for limitations [1.2] and [1.3], Grotmol discloses a learning process that results in a memory storing multiple correlations (training sets), where each correlation includes a “circumstance representation” (image of scene from a robot vehicle’s operation) correlated with an “instruction set[] for operation a first device” (motor control indication configured to cause the robot to execute an action). The images are generated by a camera during a training process and are received by the memory. Ex. 1004, 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.”); *see also id.*, 3:31-39, 8:66-9:3, 10:23-26, 38:36-58, 39:13-22.

- b. **[4.2] obtaining or receiving the first one or more instruction sets for operating the first device, and**

obtaining or receiving the second one or more instruction sets for operating the first device,

Limitation	Claimed Features
[4.2]	obtaining or receiving 1st instructions & 2nd instructions for 1st device

92. As I discussed above for limitations [1.2] and [1.3], Grotmol discloses a learning process in which multiple correlations (training sets) are stored in memory, where each correlation includes a “circumstance representation” (image of scene from a robot vehicle’s operation) correlated with an “instruction set[] for operation a first device” (motor control indication configured to cause the robot to execute an action). Grotmol further discloses that the motor commands are issued by a trainer and are received by the memory during the training process. Ex. 1004, 39:13-22 (“Motor commands issued by a trainer corresponding to time instances when the images are acquired may be stored”); *see also* Abstract, 3:36-39.

- c. **[4.3] and wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes: generating or receiving the fourth circumstance representation; and obtaining or**

receiving the fourth one or more instruction sets for operating the second device.

Limitation	Claimed Features
[4.3]	4th correlation for 2nd device (user undefined)

93. As I described above with respect to limitations [1.2] and [1.3], Grotmol discloses that multiple images may be acquired, with images respectively stored in association with corresponding motor commands issued at the time the image was acquired. Accordingly, a POSITA would have understood Grotmol to disclose first, second, third, fourth, etc. correlations, where each correlation includes a circumstance representation correlated with an instruction for operating the robotic device.

94. Grotmol also discloses that training sets and training configurations, which include pictures and associated instruction sets for operating the first device, can be stored or saved, and then loaded into the robot to provide the trained behavior Ex. 1004, 33:27-30 (“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.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that was operated during the learning. *Id.*, 33:30-32 (“[T]he trained configuration may be loaded into *one or more other robots* in order to provide learned behaviors”). A POSITA would

have understood this disclosure to mean that one robot can store the first and second correlations as well as a fourth correlation learned in another learning process involving the operation of a different robot—i.e., a circumstance detected by the sensor of a second robot correlated with instructions for operating that second robot.

5. **Claim 5: The method of claim 1, wherein the anticipating the first one or more instruction sets for operating the first device learned in the learning process based on the at least partial match between the third circumstance representation and the first circumstance representation includes:**

determining that a number of at least partially matching portions of the third circumstance representation and portions of the first circumstance representation exceeds a threshold number, or

determining that a percentage of at least partially matching portions of the third circumstance representation and portions of the first circumstance representation exceeds a threshold percentage.

Claim 5	Additional Features to Claim 1
	+ threshold number or % for partial match

95. As I described above with respect to limitation [1.5], Grotmol discloses anticipating an instruction set for operating the first device based on a partial match between a third circumstance representation and a first circumstance representation. *See, e.g.*, Ex. 1004, 43:5-9 (“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”); 33:59-62 (“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).” Grotmol discloses that determination of a control signal for the robot involves comparing a current image (i.e., third circumstance representation) with a stored image from the training buffer (i.e., first circumstance representation), computing a “similarity measure” for a given training set and another training set, and determining whether that similarity measure exceeds a threshold.

The one or more physical processors may configured to operate a random k-nearest neighbors learning process to determine a motor control indication by at least: determining a subset of features comprising the binary signal and a set of pixels randomly selected from the plurality of pixels; scaling individual pixels of the set of pixels by a scaling factor; scaling features of the subset by a scaling factor; comparing individual scaled features of the subset to corresponding features of individual ones of the training sets; based on the comparison, determining a **similarity measure** for a given training set, the similarity measure characterizing similarity between features of the subset and features of the given training set; based on an evaluation of the similarity measure, selecting one or more of the training sets; determining one or more potential control signals for the robot, individual ones of the one or more potential control signals being associated with a corresponding training set; and determining the control signal based on a transformation obtained from the one or more potential control signals.

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. The method may comprise determining one or more potential control signals for the robot.

Ex. 1004, 3:39-58, 4:30-39; *see also id.*, 39:23-41 (discussing selection of images from the buffer for comparison), 41:8-17 (discussing how a “similarity metric” can be computed between shifted or trimmed image frames), 43:5-9 (“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.”). A POSITA would have understood that a threshold in this context would be a threshold number, an example of which is a threshold percentage.

6. **Claim 6: The method of claim 1, wherein the at least the portion of the first correlation and the at least the portion of the second correlation are learned in the learning process while the user operates the first device, and wherein the one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process correspond to the user’s methodology of operating the first device learned in the learning process**

Claim 6	Additional Features to Claim 1
	+ user’s methodology of operation

96. As I described above with respect to limitations [1.2] and [1.3], Grotmol discloses that the first and second correlations are learned during a training process that involves a user operating the robotic vehicle. *See, e.g.*, Ex. 1004, 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. Motor commands associated with a given image may be stored.”); *see also* 3:27-39, 3:31-39, 8:66-9:3, 10:23-26, 38:36-58, 39:13-22.

97. The '974 patent does not define “user’s methodology” in any particular way, and does not explain the term or provide examples. Grotmol discloses training “based on” a user’s commands. Ex. 1004, 33:21-24 (“Training of the robot’s controller may be *based on* the user observing [the] robot’s actions and sending one or more *target control commands* to the robot via the training handset.”). Grotmol further explains that users can show the robot “how to perform [a] task,” in a particular way, and that these commands are learned and stored by the robot.

When a user shows the robot how to perform the task, he may do it in different ways on different occasions

In listen mode, the robot may be “listening” to or otherwise observing what the teacher teaches, and may not do anything on its own. However, the robot may learn an association.

In supervised learning, a user may train simple tasks by demonstrating example tasks to the system.

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).

In some implementations, the trainer may be provided with the control of the robot during training. Upon completion of the training, the robot may be configured to operate autonomously.

Ex. 1004, 12:16-17, 23:3-6, 24:19-20, 34:50-59. A POSITA would have understood these disclosures of learned user commands to correspond to a user's methodology for operating the robot.

7. Claim 7

Claim 7	Additional Features to Claim 1
	+ 4th correlation + 1st connection between 1st correlation and 2nd correlation + 2nd connection between 2nd correlation and 4th correlation

- a. [7.1] The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating: the first device, the second device, or a third device,

Limitation	Claimed Features
[7.1]	4th correlation (4th CR + 4th instructions for 1st , 2nd , or 3rd device)

98. As I described above with respect to claim 1, Grotmol discloses that multiple images may be acquired, with images respectively stored in association with corresponding motor commands issued at the time the image was acquired. Accordingly, A POSITA would have understood Grotmol to disclose first, second,

third, fourth, etc. correlations, where each correlation includes a circumstance representation correlated with an instruction for operating the robotic device.

99. Grotmol discloses that training sets and training configurations, which include pictures and associated instruction sets for operating a device, can be stored or saved, and then loaded into the robot to provide the trained behavior Ex. 1004, 33:27-32 (“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. In some implementations, the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that was operated during the learning. *Id.*, 33:30-32 (“[T]he trained configuration may be loaded to ***one or more other robots*** in order to provide learned behaviors”). Thus, a POSITA would have understood Grotmol to explicitly disclose a knowledgebase containing correlations that include instructions for operating second or third devices. For example, a POSITA would understand Grotmol to be disclosing that a first robot can have a memory that stores correlations learned during operation of that first robot, and the memory can also store correlations learned during operation of other robots.

- b. **[7.2] and wherein a first connection is generated to connect the first correlation with the second correlation, and wherein a second connection is**

generated to connect the second correlation with the fourth correlation.

Limitation	Claimed Features
[7.2]	• 1st connection between 1st correlation and 2nd correlation • 2nd connection between 2nd correlation and 4th correlation

100. The '974 patent describes that “[k]nowledge cells may be *connected*, interrelated, or interlinked into knowledge structures using statistical, artificial intelligence, machine learning, and/or other models or techniques.” '974 patent, 153:66-154:2. As I discussed above in relation to claim 2, the '974 patent discloses that these knowledge cells contain the claimed correlations. *See supra* claim 2.

101. In my opinion, Grotmol discloses these “connections” between the correlations in the knowledgebase. Grotmol describes storing sequences of images and the corresponding instruction sets, which a POSITA would have understood to correspond to a sequence of knowledge cells. Ex. 1004, Abstract (“Motor commands associated with a given image may be stored.”); 38:21-42 (“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...”). I understand that an ordered buffer is a region of memory that is allocated for storing data in a specific order or sequence. A POSITA would have understood Grotmol to describe that the training buffer may be ordered temporally,

according to when the images were acquired, due to its disclosures of images in the training buffer that precede and follow each other in time:

The given image may be compared to images of the primary particle set. In some implementations, the primary particle set may comprise a previously used image IO (e.g., 2000 in FIG. 20A), one or more (e.g., 2) images **preceding the IO image in time** in the training buffer, and one or more (e.g., 5) images **following the IO image in time** in the training buffer.

Motor commands issued by a trainer corresponding to time instances when the images are acquired may be stored.

Ex. 1004, 39:65-40:4, 39:17-19. A POSITA would also have understood that Grotmol discloses other ways to order the training buffer. For example, Grotmol describes “particle filtering,” images in the training buffer based on likely match to the path ahead, and ordering by random selection:

In some implementations, the search space may be narrowed using a form of particle filtering, where **the robot maintains a plurality of particles indicating the likely parts of the path.** That is, individual particle points at a particular image from the training buffer. As a new camera image arrives, the robot may search those images in the training buffer which are close to the particles. **Individual particles may be moved to a nearby location in the training buffer where the stored image matches closely with the newly arrived image.** Particles with poor match with the new image may be deleted. New particles may be created, either in the vicinity of the other particles, or from randomly sampled locations in the training buffer, shown in FIG. 22 ...

The particle characterized by the best match (e.g., comprising previously used image) may be referred to as the primary particle. In some implementations, the best match image may be complemented by one or more second best image matches, corresponding to secondary particles.

The given image may be compared to images of the primary particle set ... [or] may be compared to images of one or more secondary particle set(s)

In some implementations, the given image may be compared to images (e.g., 10-50) that may be **randomly selected from images in the training buffer**.

Ex. 1004, 39:42-40:19 Thus, a POSITA would have understood the training data stored in memory in a sequence to correspond to the claimed first and second “connections” between the first and second correlations, and the second and fourth correlations.

8. Claim 8

Claim 8	Additional Features to Claim 1
	modification of 1st instructions

- a. [8.1] The method of claim 1, further comprising:
modifying the first one or more instruction sets for
operating the first device learned in the learning
process or a copy of the first one or more instruction
sets for operating the first device learned in the
learning process,

Limitation [8.1]	Claimed Features
	modification of 1st instructions

102. Grotmol discloses that the robot may adjust its trajectory to compensate for, “mismatch[es]” in position or orientation, and further discloses modifying the motor command where a current image is shifted relative to a stored image:

Exemplary implementations of the methodology described herein may be applicable to controlling trajectory of a robotic device due to (i) **position mismatch** (e.g., the robot being located physically to the left of a target location); and/or (ii) due to **orientation mismatch** (e.g., the robot being in the same physical location while oriented towards the left of the target trajectory). To illustrate, assuming the camera faces straight ahead, the center of the image may be the spot which the robot is headed towards. Thus, if this spot is to the left of the spot where the robot is supposed to be headed towards (as defined by the camera image seen during training), then **the robot may need to adjust its heading rightwards**.

During operation when the robot may follow a target trajectory the shift amount determined using the image matching process may be close to 0 (this configuration may be referred to as “the robot stays on track”). In some implementations the shift amount may be utilized as an error metric by the control process of the robot. **The steering signal (which may be adjusted leftwards or rightwards) may be selected as the control variable for the process.** A negative feedback loop may be used in order to reduce the error metric to (and/or maintain at) a target level operation of the robot. In some implementations, the target error level may comprise zero displacement.

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.

Ex. 1004, 42:46-43:3, Abstract.

103. Grotmol further explains that motor commands corresponding to a given image and time may be combined with the output from a PID controller to “stabilize” the robot’s operation:

A PID controller may be used in order to reduce/maintain the error metric during operation of the robot. 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. Those motor commands may be combined with the output from the PID controller in order to **stabilize operation** of the robot.

Id., 43:4-11.

104. Accordingly, a POSITA would have understood Grotmol to disclose that the “first one or more instruction sets for operating the first device” learned in the learning process (i.e., motor command that is stored in association with image) are modified.

- b. **[8.2] and wherein the executing the first one or more instruction sets for operating the first device learned in the learning process includes executing the modified the first one or more instruction sets for operating the first device learned in the learning process or the modified the copy of the first one or more instruction sets for operating the first device learned in the learning process, and wherein the performing, by the first device or by the second device, the one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process includes performing, by the first device or by the second device, one or more operations defined by the modified the first one or more instruction sets for operating the first device learned in the learning process or by the modified the copy of the first one or more instruction sets for operating the first device learned in the learning process.**

	Claimed Features
--	-------------------------

Limitation [8.2]	Execution of modified 1st instructions
-----------------------------	---

105. As I discussed above for limitations [1.6] and [8.1], Grotmol discloses that motor commands at a given step may be obtained from the training buffer based on the best matching image(s), the motor commands may be modified, and the motor commands are executed to perform autonomous operation of the robotic vehicle. Accordingly, a POSITA would have understood Grotmol to disclose the executing and performing steps of limitation [8.2].

9. **Claim 9: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the first device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes operating the first device at least partially by the user.**

Claim 9	Additional Features to Claim 1
	+ 4th Correlation (including 4th CR detected by 1st device) learned by 1st user .

106. As I discussed above with respect to claim 1, Grotmol discloses 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 also explicitly discloses that multiple training processes can be conducted to learn multiple tasks of varying complexity, such as “fetch” and “grasping,” specifically noting that a user can omit training of “other behaviors” to focus on one task at a time:

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. To achieve an acceptable performance of a **complex task** by the system, the user may provide additional information to the system...

For example, **a task of playing fetch may be quite complex**: the robot may be supposed to find an object, grasp it, and bring it back to base. In response to the robot not “seeing” the target, the robot may perform a random exploration behavior avoiding obstacles. In response to the robot being close to the target, the robot may perform a positioning maneuver and grasp the target. If the target is in the robot's gripper, the robot may find its way to the base avoiding obstacles and then release the object. The whole task may be highly nonlinear and/or noisy...

According to exemplary implementations, **the user may train a component of the fetch behavior**. For example, the user may train how to search for a target and/or how to approach it while avoiding obstacles (e.g., “target approach”). The user may train the robot to perform a grasping behavior in various contexts (e.g., “grasping a target”). The user may assume the robot successfully grasped the target and train it to find its way to the base (“base approach”). 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).

Ex. 1004, 24:17-61. Grotmol also discloses creating discrete training “modules” for separate behaviors:

After is [sic] user is done training a particular behavior, he may create another module and activate it. Such creation is depicted in the figure below, in accordance with one or more implementations.

The system may be configured to learn **the second behavior** from the user. Figure above shows that the user may select a different set of input features and a different set of outputs for **“Behavior 2”**. The user may select a different type of the adaptive controller (“2”). For example, the user may control the robot using a gamepad (or other input device used to provide “user commands”) to turn and grasp the target (“Behavior 2”) using input from target tracker (“FE 1”) and state of the gripper (“FE 2”). “Adaptive controller” of type “2” (e.g., perceptron classifier) may learn to turn the robot appropriately and grasp the target by autonomously effectuating correct torques on the robot's wheel (“output 2”) and gripper (“output N”), while not controlling another wheel (“output 1”). “Sensory input” may include input from the cameras and encoders.

Id., 26:1-20. Thus, a POSITA would have understood Grotmol to disclose that the knowledgebase includes a fourth correlation, where the fourth correlation includes a circumstance representation detected by sensor of the device and at least a portion of the fourth correlation is learned in *another* learning process that includes operating the device by the user.

10. Claim 12

Claim 12	Additional Features to Claim 1
	+ CRs include objects detected at specific times

- a. **[12.1] The method of claim 1, wherein the first circumstance includes one or more objects detected at least in part by the one or more sensors of the first device at a first time or during a first time period, and**

wherein the second circumstance includes one or more objects detected at least in part by the one or more sensors of the first device at a second time or during a second time period,

Limitation [12.1]	Claimed Features
	1st CR includes objects detected at 1st time 2nd CR includes objects detected at 2nd time

107. Grotmol discloses that the claimed circumstance representations include objects detected by the sensors of the robot at different times:

In one or more implementations, images 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be **obtained at times $t_1 < t_2$, respectively**. The image 2000 may comprise **representations one or more objects**, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects. During motion of the vehicle 160 relative position and/or orientation of the camera and the visual scene may change. By way of an illustration, the vehicle 160 may execute a right turn. The image 2010 may be obtained subsequent to the turn execution. Representation of objects in the image 2010 may be shifted relative representations of the same objects in the frame obtained prior to the turn execution. Additional object representations (e.g., 2016 in FIG. 20A) may appear, while representations of some object that may be present in one image (e.g., 2000) may be absent from the subsequent image (e.g., 2010).

Ex. 1004, 38:52-39:1.

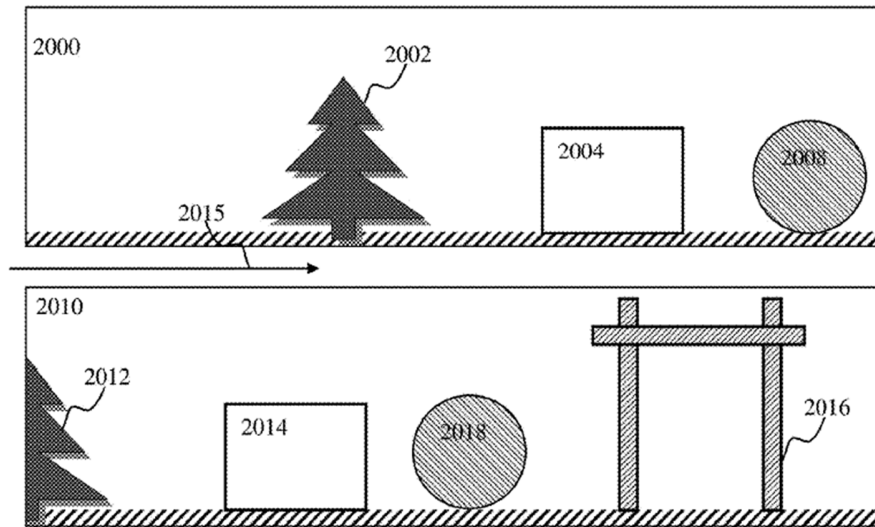


FIG. 20A
Ex. 1004, Fig. 20A

108. A POSITA would have appreciated the similarity between these disclosures and the examples of objects in pictures from the '974 patent, such as “person, cat, vehicle, building, street, tree, rock, etc.” '974 patent, 72:52-55. The '974 patent further describes various approaches to object recognition, such as “identifying and/or analyzing features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc.” *Id.*, 75:1-5.

109. 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. 1004, 8:57-9:13; *see, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or

classification may be implemented”). Grotmol, like the ’974 Patent, further includes edge detection. *Id.*, 8:60-63 (“[T]he term feature may refer to a representation of an object edge”).

110. Grotmol describes that the autonomous operation of a robot may be based in part on recognized objects and their characteristics. *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”).

- b. **[12.2] and wherein the third circumstance includes:
one or more objects detected at least in part by the
one or more sensors of the first device at a third time
or during a third time period, or one or more objects
detected at least in part by the one or more sensors of
the second device at a third time or during a third
time period.**

Limitation	Claimed Features
[12.2]	3rd CR includes objects detected at 3rd time

111. As I discussed above with respect to limitation [1.4] of claim 1, Grotmol discloses “generating or receiving a third circumstance representation, wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device.” A POSITA would have understood that in Grotmol, “the most recent, current” image

corresponds to the claimed “third circumstance representation” and it is compared to the stored images in the training buffer (the claimed first and second circumstance representations).

112. As I described above with respect to limitation [12.1], Grotmol further discloses that captured images include representations of objects and further discloses object detection by the sensor. Accordingly, a POSITA would have understood that in Grotmol, the third circumstance, like the first and second circumstances, comprises an image that includes one or more objects detected at a third time.

11. **Claim 13: The method of claim 1, wherein the first circumstance representation is a data structure including one or more data about the first circumstance of the first device, and wherein the second circumstance representation is a data structure including one or more data about the second circumstance of the first device, and wherein the third circumstance representation is a data structure including one or more data about: the third circumstance of the first device, or the third circumstance of the second device.**

Claim 13	Additional Features to Claim 1
	+ CRs = data structures

113. In my opinion, Grotmol discloses that the circumstance representations are data structures including data about the circumstances of the first device, within the meaning of the '974 patent.

114. The '974 patent lists various examples, of data structures “such as, for example, array, list, linked list, doubly linked list, queue, tree, heap, graph, map, grid, matrix, multi-dimensional matrix, table, database, database management system (DBMS), file, neural network, and/or any other type or form of a data structure including a custom one,” and states that “any variety of [these examples] can be used.” '974 patent, 65:45-51. As I discussed above with respect to limitation [1.2] of claim 1, Grotmol discloses circumstance representations comprising video frame instances. Grotmol discloses that a video frame “compris[es] a plurality of pixels.” Ex. 1004, 3:31-33. Thus, A POSITA would have understood the video frames disclosed by Grotmol to comprise “data structures” in the context of the '974 patent because they comprise at least an array, map, grid, or other form of pixels.

115. Further, A POSITA would also understand the entire video frame itself to comprise a “data structure” within the context of the '974 patent, because it is a file including image data. The '974 patent explains that a data structure can comprise a single file or field of data. '974 patent, 65:51-55 (“A data structure may include *one* or more fields or data fields that are part of or associated with the data structure. A field or data field may include *a data*, an object, a data structure, and/or any other

element or a reference/pointer thereto.”). Further, the’974 Patent states that data structures “can be stored in one or more memories, files, or other repositories,” (*id.*, 65:55-57) just as the video frame instances are stored in the “memory component” of Grotmol. Ex. 1004, 3:27-39.

116. Accordingly, a POSITA would have understood the video frame instances of Grotmol to include “one or more data” about the “circumstance[s] of the first device,” because these video frame instances are detected by a sensor attached to the first device, and capture images of the device’s surroundings for use in object detection and avoidance:

FIG. 1 depicts a mobile robotic apparatus that may be configured with an adaptive controller in accordance with one or more implementations of e.g., the learning apparatuses illustrated in FIGS. 4A-4B, *infra*. The robotic apparatus 160 may comprise a sensor component 166. The sensor component 166 may be characterized by an aperture/field of view 168 (e.g., an extent of the observable world that may be captured by the sensor at a given moment). The sensor component 166 may provide information associated with objects within the field-of-view 168. 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.

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 refreshed at, e.g., 25 Hz frame rate. The sensor output may be processed by a learning controller, e.g., as illustrated and described with respect to FIG. 4A.

In some implementations of robotic vehicle navigation, output of the sensor 166 in FIG. 1 may comprise representations of one or more objects (e.g., targets, and/or obstacles).

In one or more implementations, images 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects

Id., 8:57-9:3, 10:15-26, 38:52-58.

12. **Claim 14: The method of claim 1, wherein the first circumstance representation includes: one or more object representations, or one or more collections of object representations, and wherein the second circumstance representation includes: one or more object representations, or one or more collections of object representations, and wherein the third circumstance representation includes: one or more object representations, or one or more collections of object representations.**

Claim 14	Additional Features to Claim 1
	+ CRs include object representations

117. A POSITA would have understood Grotmol to disclose the claimed circumstance representations including object representations. As I discussed above in relation to limitations [1.2] and [1.4], Grotmol discloses first, second, and third circumstance representations representing first, second, and third circumstances detected at least in part by one or more sensors of the first device.

118. The '974 patent does not define “object representations” beyond its use in the phrase “collection of object representations,” which it defines as “one or more representations of objects detected by the sensor at a time.” '974 patent, 3:31-33. The “sensor” that detects the object representations in the '974 patent may include a “[c]amera... for capturing one or more pictures” or a “motion picture camera that can capture streams of pictures (i.e. motion pictures, videos, etc.).” *Id.*, 69:45-56.

119. The '974 patent also provides examples of objects represented in pictures such as “person, cat, vehicle, building, street, tree, rock, etc.” *Id.*, 72:52-55. The '974 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.*, 75:1-5.

120. A POSITA would have appreciated that Grotmol similarly discloses that images 2000 and 2010 may “comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively.” Ex. 1004, 38:54-56. These same 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 also id.*, 38:65-39:12 (referring to objects shown within images 2000 and 2010 as “object representations” and “representations”).

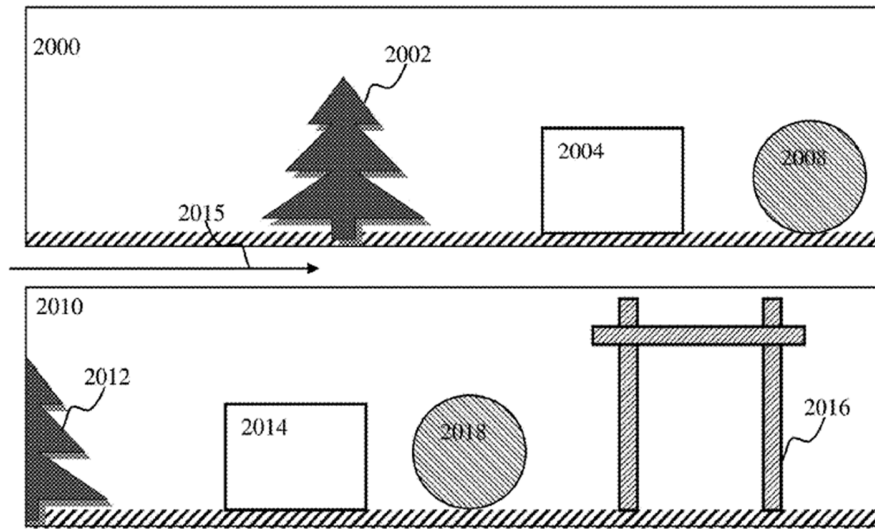


FIG. 20A
Ex. 1004, Fig. 20A

121. 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. 1004, 8:57-9:13. *See, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented”). Grotmol, like the ’974 Patent, further includes edge detection. *Id.*, 8:60-63 (“[T]he term feature may refer to a representation of an object edge”).

13. Claim 15

Claim 15	Additional Features to Claim 1
	+ 1st instructions temporally correspond to 1st CR + 1st CR includes object representations

- a. [15.1] The method of claim 1, wherein, to correlate the first circumstance representation with the first one or more instruction sets for operating the first device, a determination is made that the first one or more instruction sets for operating the first device temporally correspond to the first circumstance representation,

Limitation [15.1]	Claimed Features
	temporal correspondence between 1st instructions and 1st CR

122. As I discussed above with respect to limitation [1.2], the images in Grotmol are correlated with motor control indications issued contemporaneously. Grotmol describes training behaviors in a robotic device by “learning *associations* between sensory context . . . and a respective action . . . during training.” Ex. 1004, 33:59-62; *see also id.*, Abstract (“Motor commands *associated with* a given image may be stored.”). Grotmol further discloses that this association may be based on the timing of the images and the motor commands. *Id.*, 39:17-19 (“Motor commands issued by a trainer *corresponding to* time instances when the images are acquired may be stored.”). Accordingly, a POSITA would have understood Grotmol to

disclose that a determination is made that the first instruction set temporally corresponds to the first circumstance representation.

- b. **[15.2] and wherein the first circumstance representation includes: one or more object representations, or one or more collections of object representations.**

Limitation [15.2]	Claimed Features
	1st CR includes object representations

123. As I discussed above for claim 14, Grotmol discloses that the first circumstance representation includes one or more object representations, or one or more collections of object representations.

14. **Claim 16: The method of claim 1, wherein the anticipating the first one or more instruction sets for operating the first device learned in the learning process based on the at least partial match between the third circumstance representation and the first circumstance representation includes determining the first one or more instruction sets for operating the first device learned in the learning process based on the at least partial match between the third circumstance representation and the first circumstance representation.**

Claim 16	Additional Features to Claim 1
	+ determination of 1st instructions based on partially matching 1st CR and 3rd CR

124. As I discussed above for limitation [1.5], 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. 1004, 43:4-9. Grotmol discloses that the matching step may comprise calculating a similarity measure between a current image and a stored image, then determining whether that similarity measure breaches a threshold (i.e., “at least partial match between the third circumstance representation and the first circumstance representation”). Accordingly, a POSITA would have understood Grotmol to disclose determining the first instruction set for operating the device based on at least a partial match between third circumstance representation and the first circumstance representation.

15. Claim 17

Claim 17	Additional Features to Claim 1
	+ <i>various embodiments</i>

- a. **[17.1] The method of claim 1, wherein elements of the computing system are included in: a single device, or multiple devices,**

125. Grotmol discloses the method of claim 1 implemented using a “self-contained robotic apparatus,” *Id.*, 3:27-30 (“Another aspect of the disclosure relates to a self-contained robotic apparatus. The apparatus may comprise a platform, a first sensor, a second sensor, a memory component, and one or more physical

processors.”); *see supra* limitation [1.1]. Thus, a POSITA would have understood Grotmol to disclose that the elements of the computing system are included in a single device.

- b. [17.2] and wherein the one or more processor circuits include: one or more microcontrollers, one or more computing circuits, or one or more electronic circuits,**

126. A POSITA would have understood that Grotmol discloses one or more processor circuits including a microcontroller, computing circuits, or electronic circuits, for the same reasons I discussed above for the term “processor circuits” with respect to limitation [1.1].

- c. [17.3] and wherein the memory includes: a volatile memory, or a non-volatile memory,**

127. Grotmol discloses that its “memory component” may comprise a “non-volatile memory.” *Id.*, 38:39-42 (“In some implementations, the images may be stored in an ordered buffer in a non-volatile memory.”); *see supra* limitation [1.2].

- d. [17.4] and wherein the first device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine,**

128. Grotmol discloses a robot (“self-contained robotic apparatus”) that may include a vehicle. *Id.*, 3:27-30; *id.*, 38:36-39 (“In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1)”); *see supra* claim 1.

- e. **[17.5] and wherein the second device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine,**

129. Grotmol discloses a second device comprising a robot. Grotmol discloses that training sets and training configurations can be stored or saved, and then loaded into the robot to provide the trained behavior. *Id.*, 33:27-32 (“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. In some implementations, the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that did the learning. *Id.*, 33:30-32 (“the trained configuration may be loaded into ***one or more other robots*** in order to provide learned behaviors”). Thus, a POSITA would have understood Grotmol to disclose 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 that is loaded into a second one or more other robots to “provide learned behaviors.”

- f. **[17.6] and wherein an instruction set of the first one or more instruction sets for operating the first device includes at least one selected from the group comprising: only one instruction, a plurality of instructions, one or more inputs, one or more commands, one or more computer commands, one or more keywords, one or more symbols, one or more operators, one or more variables, one or more values,**

one or more objects or object references, one or more data structures or data structure references, one or more functions or function references, one or more parameters, one or more signals, one or more characters, one or more digits, one or more user operating directions, one or more user directions, one or more user inputs, one or more representations of one or more user actions, one or more representations of one or more user clicks, one or more binary bits, one or more assembly language commands, one or more states, one or more codes, one or more data, and one or more information,

130. As I discussed above with respect to limitation [1.2], Grotmol discloses “motor control indication[s]” and “motor commands” that a POSITA would have understood to correspond to the claimed first instruction set(s) for operating the first device, including, for example, one or more instructions, one or more commands, one or more computer commands, and/or one or more information:

The memory component may be configured to store training sets. 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.

In listen mode, the robot may be “listening” to or otherwise observing what the teacher teaches, and may not do anything on its own. However, the robot may learn an association. But if the teacher stops sending a command, the robot may stay still and wait for the next command. In some implementations, teaching the robot may begin with the listen mode. Once enough examples have been provided and the robot has learned something, the override mode may be used.

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. Additional 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.

A PID controller may be used in order to reduce/maintain the error metric during operation of the robot. 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. Those motor commands may be combined with the output from the PID controller in order to stabilize operation of the robot.

Id., 3:35-39, 23:3-11, 39:13-22, 43:4-11.

- g. **[17.7] and wherein the one or more sensors of the first device include at least one selected from the group comprising: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, and one or more apparatuses for detecting objects or object properties, and wherein the one or more sensors of the second device include at least one selected from the group comprising: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars,**

and one or more apparatuses for detecting objects or object properties,

131. As I discussed above for limitation [1.4] and claim 14, Grotmol discloses that a robotic vehicle includes a sensor comprising a camera for detecting objects or object properties:

FIG. 1 depicts a mobile robotic apparatus that may be configured with an adaptive controller in accordance with one or more implementations of e.g., the learning apparatuses illustrated in FIGS. 4A-4B, *infra*. The robotic apparatus 160 may comprise a sensor component 166. The sensor component 166 may be characterized by an aperture/field of view 168 (e.g., an extent of the observable world that may be captured by the sensor at a given moment). The sensor component 166 may provide information associated with objects within the field-of-view 168. 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.

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 refreshed at, e.g., 25 Hz frame rate. The sensor output may be processed by a learning controller, e.g., as illustrated and described with respect to FIG. 4A.

In some implementations of robotic vehicle navigation, output of the sensor 166 in FIG. 1 may comprise representations of one or more objects (e.g., targets, and/or obstacles).

One or more implementations described herein may provide a mechanism for enabling a robot to learn navigating a target trajectory while reducing deviation from a target path. In some

implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG. 1) **comprising a camera** (e.g., 166). 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).

FIG. 20A illustrates exemplary images for use with training of path navigation. Images 2000, 2010 may be obtained with a **variety of sensors, e.g., a video camera, IR camera, RADAR, SONAR, LiDAR, ultrasonic imager, and/or other device**. In some implementations, the camera may be spanning the entire horizontal view field of 360 degrees. A 360 degree horizontal view field may provide additional information about robot's environment and enable cyclical shift of images.

In one or more implementations, images 2000, 2010 may be obtained with **a camera 166 mounted on a robotic vehicle 160** of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.

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. Additional 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.

Id., 8:57-9:3, 10:15-26, 38:33-58, 39:13-22.

132. Furthermore, as I discussed above for limitation [4.3], Grotmol discloses a second robotic vehicle that includes a camera for detecting objects or object properties:

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. In some implementations, **the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.** Subsequent to loading of the saved configuration, the controller learning process may match process configuration being present during saving of the configuration.

Id., 33:21-35.

- h. **[17.8] and wherein the at least the portion of the first correlation includes: one portion of the first correlation, multiple portions of the first correlation, all portions of the first correlation, or the entire first correlation, and wherein the at least the portion of the second correlation includes: one portion of the second correlation, multiple portions of the second correlation, all portions of the second correlation, or the entire second correlation,**

133. As I discussed above for limitations [1.2] and [1.3], Grotmol discloses that first and second correlations are learned during a learning process. Accordingly, a POSITA would have understood Grotmol to disclose that at least the portion of the first correlation includes the entire first correlation, and wherein the at least the portion of the second correlation includes the entire second correlation.

- i. [17.9] and wherein an object of the first circumstance is the same as an object of the third circumstance, or multiple objects of the first circumstance are the same as multiple objects of the third circumstance, or all objects of the first circumstance are the same as all objects of the third circumstance, or all objects of the first circumstance are different than all objects of the third circumstance.

134. As I discussed above with respect to limitation [1.5] and claim 12, Grotmol discloses comparing first circumstance and third circumstance to determine whether there is at least a partial match, and further discloses that the first and third circumstance include one or more objects, as well as object detection. A POSITA would have understood that the object(s) in the first circumstance and the third circumstance can be the same or different. For example, Grotmol shows in Fig. 20A that image 2010, taken subsequent to image 2000, contains the same tree, rectangle, and ball objects as the previous image, but also contains a new object, 2016:

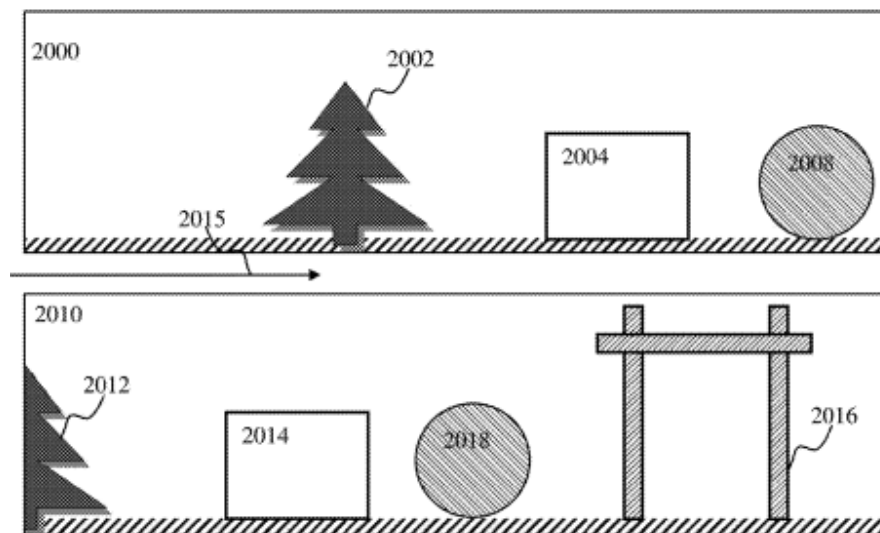


FIG. 20A

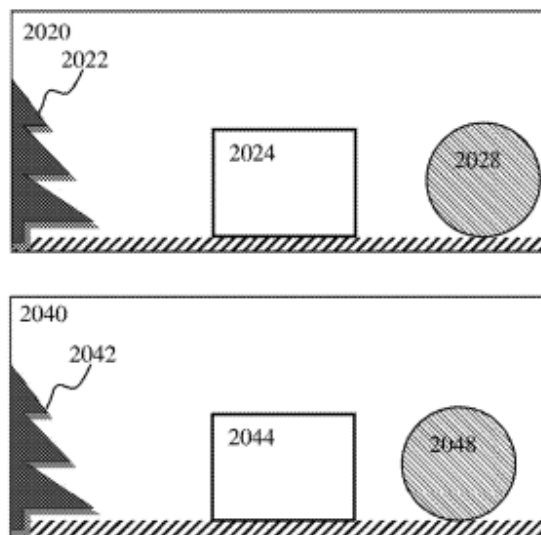


FIG. 20B

In one or more implementations, images 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects. During motion of the vehicle 160 relative position and/or orientation of the camera and the visual scene may change. By way of an illustration, the vehicle 160 may execute a right turn. The image 2010 may be obtained subsequent to the turn execution. Representation of objects in the image 2010 may be shifted relative representations of the same objects in the frame obtained prior to the turn execution. **Additional object representations (e.g., 2016 in FIG. 20A) may appear, while representations of some object that may be present in one image (e.g., 2000) may be absent from the subsequent image (e.g., 2010).** Comparing image 2010 to image 2000, representations of the tree 2012 the rectangle, 2014, and the ball 2018 are shifted (horizontally) within the frame 2010 relative representations, 2002, 2004, 2008 in the frame 2010. The amount of the shift may be configured based on relative motion between the images (e.g., motion speed and time interval $t_2 - t_1$ between the images). It may be further noticed from comparing images 2000 and 2010 that the ball 2018 in image 2010 may be disposed

closer to the rectangle 2014 compared to distance between respective object representations 2008, 2004 in image 2000.

Id., Fig. 20A, Fig. 20B, 38:52-39:12. Accordingly, a POSITA would have understood that the images may have one object, multiple or all objects that are the same, or have all different objects.

16. Claim 18

135. Claim 18 is nearly identical to claim 1 but has a different preamble. My analysis for claim 1 applies to claim 18 as shown in the chart below.

Claim 18	Disclosure of Grotmol
[18.1] One or more non-transitory machine readable media storing machine readable code that when executed by one or more processor circuits causes the one or more processor circuits to perform at least:	Grotmol discloses a “non-transitory computer-readable storage medium having instructions embodied thereon, the instructions being executable by a processor” Ex. 1004, claim 17. As I discussed above with respect to limitation [1.1], A POSITA would have understood Grotmol’s “processor” to comprise one or more processor circuits.

Claim 18	Disclosure of Grotmol
[18.2] accessing a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device and a second correlation including a second circumstance representation correlated with a second one or more instruction sets for operating the first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device and the second circumstance representation represents a second circumstance detected at least in part by the one or more sensors of the first device,	<i>See supra</i> limitation [1.2].
[18.3] and wherein at least a portion of the first correlation and at least a portion of the second correlation are learned in a learning process that includes operating the first device at least partially by a user;	<i>See supra</i> limitation [1.3].
[18.4] generating or receiving a third circumstance representation, wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device or at least in part by one or more sensors of a second device;	<i>See supra</i> limitation [1.4].

Claim 18	Disclosure of Grotmol
[18.5] anticipating the first one or more instruction sets for operating the first device learned in the learning process based on at least partial match between the third circumstance representation and the first circumstance representation; and	<i>See supra</i> limitation [1.5].
[18.6] at least in response to the anticipating, causing the first device or the second device to perform one or more operations defined by the first one or more instruction sets for operating the first device learned in the learning process at least by causing the one or more processor circuits or another one or more processor circuits to execute the first one or more instruction sets for operating the first device learned in the learning process.	<i>See supra</i> limitation [1.6].

17. Claim 19

136. Independent claim 19 is a system claim that corresponds to method claim 1. As I explained in the claim construction section, a corresponding structure for the “means for generating or receiving,” “means for anticipating,” and the “means for executing” is “one or more processor circuits.” Claim 1 likewise requires that the generating/receiving, anticipating, and executing steps are performed by one or more processor circuits. Accordingly, my analysis for claim 1 applies to claim 19 as shown in the chart below. In my opinion, Grotmol anticipates claim 19.

Claim 19	Disclosure of Grotmol
[19.1] A system comprising:	To the extent the preamble is limiting, Grotmol discloses a system implemented as a “self-contained robot apparatus.” Ex. 1004, 3:27-30; <i>see also</i> Abstract (“An apparatus and methods for training and/or operating a robotic device to follow a trajectory.”); 10:45-48.
[19.2] a memory that stores at least a knowledgebase that includes: a first correlation including a first circumstance representation correlated with a first one or more instruction sets for operating a first device and a second correlation including a second circumstance representation correlated with a second one or more instruction sets for operating the first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device and the second circumstance representation represents a second circumstance detected at least in part by the one or more sensors of the first device,	<i>See supra</i> limitation [1.2]
[19.3] and wherein at least a portion of the first correlation and at least a portion of the second correlation are learned in a learning process that includes operating the first device at least partially by a user;	<i>See supra</i> limitation [1.3]

Claim 19	Disclosure of Grotmol
[19.4] means for generating or receiving a third circumstance representation, wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device or at least in part by one or more sensors of a second device;	<i>See supra</i> limitations [1.1] and [1.4]
[19.5] means for anticipating the first one or more instruction sets for operating the first device learned in the learning process based on at least partial match between the third circumstance representation and the first circumstance representation; and	<i>See supra</i> limitations [1.1] and [1.5]
[19.6] means for executing, at least in response to the anticipating, the first one or more instruction sets for operating the first device learned in the learning process, wherein the first device or the second device autonomously.	<i>See supra</i> limitations [1.1] and [1.6]

18. Claim 20

137. Claim 20 depends from claim 19 and requires that the “means for generating or receiving,” the “means for anticipating,” and the “means for executing” include “one or more processor circuits.” As I explained above for claim 19, it is my opinion that a corresponding structure for the “means for anticipating,” and the “means for executing” is “one or more processor circuits.” Accordingly, my analysis of why claim 20 is anticipated by Grotmol is the same as for claim 19.

C. Ground 2: Claims 10 and 11 Are Obvious Based on Grotmol in View of the Knowledge of a POSITA

138. As set forth below, claims 10 and 11 of the '974 patent are rendered obvious by Grotmol in view of the knowledge of a POSITA.

1. **Claim 10: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the first device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes operating the first device at least partially by another user.**

Claim 10	Additional Features to Claim 1
	+ 4th correlation (4th CR and 4th instructions) learned with 2nd user • 4th CR detected by 1st device • 4th instructions for 1st device

139. As I discussed above in Ground 1, Grotmol discloses that a human user can operate the robot during a training process for learning correlations.

In listen mode, the robot may be “listening” to or otherwise observing what the teacher teaches, and may not do anything on its own. However, the robot may learn an association.

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. Additional 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.

Ex. 1004, 23:3-6, 39:13-22. Grotmol further discloses that multiple learning processes can be done to train multiple behaviors, including “simple tasks” and “complex tasks.” *Id.*, 24:17-67 (detailing methodologies for training simple and complex tasks), 26:1-20 (detailing process for learning a “second behavior from the user”). Grotmol does not place any limit on the number of users that can train the multiple behaviors, and a POSITA would have understood that different users can train different behaviors, particularly given that Grotmol describes the training of both simple and complex tasks. A POSITA would have understood that different users can possess different skills and knowledge relating to the functionality and operation of a robot. Therefore, it would have been obvious to a POSITA that different trainers could be used to train behaviors involving different levels of complexity in order to take advantage of the varying skills and knowledge different trainers possess. In my opinion, there is no reason a POSITA would have understood Grotmol’s disclosure to be limited to a single user training all possible tasks, especially due to the inefficiency this limitation would impose on learning across all interoperable robots.

2. **Claim 11: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes operating the second device at least partially by another user.**

Claim 11	Additional Features to Claim 1
	+ 4th correlation (4th CR and 4th instructions) learned with 2nd user • 4th CR detected by 2nd device • 4th instructions for 2nd device

140. As I discussed above for claim 10, Grotmol discloses that multiple learning processes can be done to train multiple behaviors, and that it would be obvious that the multiple learning processes could be done by different users operating the robot. A POSITA would also understand that Grotmol discloses that a robot can be loaded with a first trained configuration learned during operation of that robot by a user, and the robot can also be loaded with a trained configuration learned during operation of a different robot:

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. In some implementations, **the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.** Subsequent to loading of the saved configuration, the controller learning process may match process configuration being present during saving of the configuration.

Id., 33:24-35. For the same reasons discussed with respect to claim 10, it is my opinion that it would be obvious to a POSITA that a different user could operate the different robot to learn the trained configuration that is then loaded into the first robot.

D. Ground 3: Claims 4 and 11 Are Obvious Based on Grotmol in View of Hickman

141. As set forth in detail below, claims 4 and 11 of the '974 patent are separately rendered obvious by Grotmol in view of Hickman.

1. Motivation to Combine Grotmol and Hickman

142. A POSITA would have been motivated to combine the autonomous robotic system of Grotmol with the shared robot knowledge base of Hickman, which would allow for simultaneous learning across multiple devices with multiple users, increasing efficiency and conserving local system resources. As detailed in Section VI.A above, Grotmol teaches a training methodology for autonomous navigation by one or more robots. *See* Section VI.A. A POSITA would have looked to expand upon the functionality of the robots and training methods of Grotmol by using the teachings of Hickman, which, as discussed in Section VI.B above, discloses a shared robot knowledge base that is updated based on robot experiences, allowing any

particular robot to benefit from the prior experiences of any other connected robot.
See Section VI.B.

143. Grotmol discloses that training sets and training configurations, which include pictures and associated instruction sets for operating the robot, can be stored or saved, and then loaded into the robot to provide the trained behavior. *Id.*, 33:27-32 (“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. In some implementations, the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that did the learning. *Id.*, 33:30-32 (“the trained configuration may be loaded into ***one or more other robots*** in order to provide learned behaviors”).

144. Grotmol does not go into further detail regarding this sharing of learning behaviors. Thus, a POSITA would seek out other references, such as Hickman, for disclosures related to sharing knowledge between robots. A POSITA would have specifically looked to Hickman because, like Grotmol, it is directed to the training and autonomous operation of robots. Hickman is also cited on the face of Grotmol as a cited U.S. patent document. *Id.*, p. 2. A POSITA would certainly consider the references listed on the face of Grotmol when looking to expand upon its teachings.

145. In addition, Grotmol and Hickman both address similar robotic training applications, such as learning how to “fetch” based on human input or commands:

Another example application may be called “Fetch.” The “Fetch” application may include robot instructions for playing fetch with a dog... When the robot executes the “Fetch” application, the person may adjust or further configure certain aspects of the individual tasks via a robot user interface on a smartphone, tablet computer, etc. to “fine tune” the application. For example, the person may increase or decrease the force with which the robot throws the ball or vary the praise given to the dog upon returning with the ball, e.g., “Good dog!” Creating the “Fetch” application in this manner is akin to “teaching” the robot a new skill.

After creating the “Fetch” application, the person may upload the application (or perhaps instruct the robot to upload the application) to the application knowledge base 311. Once uploaded to the application knowledge base 311, other robots can download and/or execute the “Fetch” application. Thus, the skill (i.e., playing fetch with a dog) learned by a single robot can be shared with all other robots via the application knowledge base 311 so that each robot does not have to learn the skill on its own.

Ex. 1005, 23:49-24:7.

For example, a task of playing fetch may be quite complex: the robot may be supposed to find an object, grasp it, and bring it back to base. In response to the robot not “seeing” the target, the robot may perform a random exploration behavior avoiding obstacles. In response to the robot being close to the target, the robot may perform a positioning maneuver and grasp the target. If the target is in the robot's gripper, the robot may find its way to the base avoiding obstacles and then release the object. The whole task may be highly nonlinear and/or noisy, so a very powerful classifier may infer the right sequence of actions from just the examples and be resistant to noise. The classifier may need to determine which particular input is important for a particular task. For example, a state of the gripper may not be

relevant to target search behavior. Additional training may be needed to show the predictor that the search behavior should be performed independently of the state of the gripper.

According to the hierarchy depicted in FIG. 9, the “Fetch switcher” may be a high level behavior that activates one or more low level behavior. Such activation may depend on the context. The user may train the “Fetch switcher” using supervised learning. For example, the user may demonstrate to the switcher which behavior it should select in a particular situation. The user may use a user interface for such demonstrations, according to some implementations. In some implementations, if the target is far or not visible, the user may send a “select target approach” command. If the target is close, the user may send a “select grasping a target” command.

Ex. 1004, 24:31-48, 25:10-21; *see also id.*, 27:28-40, Fig. 9.

146. Grotmol states that the fetch task can be “quite complex.” *Id.*, 24:31-33. Thus, a POSITA would consider references, such as Hickman, that explicitly teach ways to “optimize” such complex training:

Also, as described above, some applications in the application knowledge base 311 may build upon data from other knowledge bases. For example, the “Fetch” application may rely on a “grasp” task from the task knowledge base 309 and a “ball” object from the general object knowledge base 308. As individual robots become better at applying the “grasp” task to the “ball” object through either trial and error and/or human “coaching,” the individual robots can update data in the shared robot knowledge base 306 based on their individual experiences. As a result, **the cloud computing system 304 can optimize the “Fetch” application based on the collective experiences of many different robots.** Thus, each time a robot connects to the cloud computing system 304 to execute the “Fetch” application from the application knowledge base 311, that robot

gains the benefit of the collective experiences of many other robots that have executed the “Fetch” application in the past.

Ex. 1005, 24:8-25.

147. More generally, a POSITA would have understood that pairing Hickman’s shared robot knowledge base with Grotmol’s autonomous robotic system would facilitate wider and faster learning among all robots connected to the knowledgebase and conserve local system resources. Further, A POSITA would expect these advantages to scale proportionally as the number of robots connected to the shared knowledgebase increases. Hickman explains as follows:

The advantages of a shared robot knowledge base 306 multiply as robots are deployed around the world. If information about a particular object is “public” and therefore accessible by all robots, then information that a single robot learns about that object and uploads to the shared robot knowledge base 306 may be accessible by other robots all around the world. In operation, many millions or even many billions of robots may ultimately access the shared robot knowledge base 306. As a result, the shared robot knowledge base 306 enables robots to benefit from the collective information learned by many other robots.

[D]ividing application execution and storage between client devices and the cloud computing system allows more processing to be performed by the cloud computing system, thereby taking advantage of the cloud computing system's processing power and capability, for example.

In many situation, cloud computing may free users from certain hardware and software installation and maintenance tasks

through the use of simplified hardware on the user's computer. Also, because the user's computer can access a vast network of computing resources (e.g., processors, disk drives, etc.), the user is not limited just to the computing and storage power of his or her local computer. Additionally, the sharing of computing resources across many users may reduce computing costs to individuals. For example, multiple computers connected to the cloud may be able to share the same pool of computing power, applications, and files.

Id., 11:31-41, 5:42-46, 1:38-48.

148. A POSITA familiar with Grotmol during or prior to 2016 would have readily appreciated the advantages of combination with Hickman's shared robot knowledgebase. For example, in a 2014 Science Article, the CEO of Brain Corporation, the assignee of Grotmol, discussed the "profound" benefits of shared robot learning:

Efficient chips can run complex—and normally power-hungry—algorithms that enable robots to learn. The Internet, in turn, will allow far-flung robots to share those lessons and skills. A robot that learns how to pick a strawberry without crushing it, for instance, could uplink that skill for the benefit of its kind around the globe. That means neuromorphic computing could offer robots something far more profound than enhanced perceptual skills. When humans collectively pass along life lessons, we call that culture.

Ex. 1006 at 183.

149. In my opinion, a POSITA would have found the resulting combination of Grotmol's autonomous robotic system and Hickman's shared robot knowledgebase to be expected and predictable. Further, it is my opinion that a

POSITA would have had a reasonable expectation of success in combining the teachings of Grotmol and Hickman, which describe clearly complementary systems. Grotmol expressly contemplates the sharing of trained configurations among robots, and Hickman's expansion upon this idea does not interfere with Grotmol's autonomous robotic system.

2. **Claim 4: The method of claim 1, wherein the learning process includes: generating or receiving the first circumstance representation, and generating or receiving the second circumstance representation; and obtaining or receiving the first one or more instruction sets for operating the first device, and obtaining or receiving the second one or more instruction sets for operating the first device, and wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that includes: generating or receiving the fourth circumstance representation; and obtaining or receiving the fourth one or more instruction sets for operating the second device.**

Claim 4	Additional Features to Claim 1
	+ 4th correlation (4th CR and 4th instructions) (user undefined) • 4th CR detected by 2nd device • 4th instructions for 2nd device

150. As I discussed above with respect to Ground 1, Grotmol discloses that multiple images may be acquired, with images respectively stored in association with corresponding motor commands issued at the time the image was acquired:

Another aspect of the disclosure relates to a self-contained robotic apparatus. The apparatus may comprise a platform, a first sensor, a second sensor, a memory component, and one or more physical processors. The platform may comprise a motor. The first sensor component may be configured to provide a signal conveying a video frame comprising a plurality of pixels. The second sensor component may be configured to provide a binary sensor signal characterized by one of two states. The memory component may be configured to store training sets. 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.

In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG. 1) comprising a camera (e.g., 166). The robot may comprise a controller configured to store a sequence of images obtained during training phase.

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. In some implementations, the matching process may comprise comparing the given image to every image in the training buffer.

Ex. 1004, 3:27-39; 38:36-39; 39:23-28; *see supra* limitations [1.2], [1.4].

Accordingly, a POSITA would have understood Grotmol to disclose first, second, third, fourth, etc. correlations, where each correlation includes a circumstance representation correlated with an instruction for operating the robotic device.

151. The combination of Grotmol and Hickman renders obvious the additional requirement in claim 4 that the knowledgebase includes a fourth correlation, which includes a circumstance representation detected by sensors of a second device correlated with an instruction set for operating the second device.

152. Grotmol discloses that training sets and training configurations, which include pictures and associated instruction sets for operating the first device, can be stored or saved, and then loaded into the robot to provide the trained behavior. *Id.*, 33:27-32 (“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. In some implementations, the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that did the learning. *Id.*, 33:30-32 (“the trained configuration may be loaded into ***one or more other robots*** in order to provide learned behaviors”). Thus, A POSITA would have understood Grotmol to disclose the memory component of a first robot that can store the claim first and second correlations as well as a fourth correlation learned from operating a second robot.

153. Grotmol discloses, for example, situations in which the user may wish to train the robot to “find an object, grasp it, and bring it back” (i.e., fetch). *Id.*, 24:31-25:5. Grotmol describes that this “task of playing fetch may be quite complex” and

outlines numerous user training steps that would be needed in order to teach this skill to the robot. *Id.* In my opinion, it would be desirable to implement a system wherein learned robot behaviors could be uploaded to a shared knowledge base for use in numerous other robots around the world, such as the system described in Hickman.

154. Hickman discloses “a shared robot knowledge bases for use with cloud computing systems” that “collects data from a robot about an object the robot has encountered in its environment, and stores the received data in the shared robot knowledge base ... sends instructions for interacting with an object to a robot ... sends instructions to a robot for executing an application based on information stored in the shared robot knowledge base.”Ex. 1005, Abstract. Hickman notes that the “information in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other robots.” *Id.* Hickman discloses that these prior “robot experiences” may include human coaching:

By analyzing past experiences from many different robots, the cloud computing system 304 may be configured to synthesize certain rules for robots to follow when executing the “Groceries” application. For example, some rules that could be synthesized or generalized from a large number of robot experiences might include (i) unidentifiable frozen items should be stored in a freezer, (ii) identifiable cold items should be stored in a refrigerator, (iii) certain types of vegetables should be stored in a certain refrigerator drawer, (iv) items should not be placed on top of bread in a pantry, etc. In this manner, **robot experiences (including “coaching” from humans)** can be used to improve the way that robots execute the “Groceries” application.

Id., 23:36-48. A POSITA would have appreciated the similarity between this human coaching and the human training described in Grotmol. For example, similar to the “fetch” behavior described in Grotmol, Hickman describes a human coach training a robot how to properly grasp a cup, uploading and storing this information on a shared knowledge base, and then having another robot obtain these instructions from the memory and executing the instructions:

In the example shown in FIG. 4, the robot 404 may be performing an application that includes moving the cup 405 from one location to another location...

In this example interaction, the grasping force specified in the instructions was too high, and as a result, the robot crushed the cup and caused a mess. When an accident happens, **a human can “coach” the robot on how much force to use** by, for example, manually controlling the robot's hand to grasp a new cup. The robot can capture the grasping force that it used to successfully grasp the cup under the manual control of the human, and then **send feedback 411 to the cloud processing system 402 with the modified grasping force**. The cloud processing system 402 can then update 412 task and object data in the shared robot knowledge base 403 to improve how the “grasp” task is applied this particular “cup” object.

Later, a second robot 413 may encounter an object 414 while performing the same (or a similar) application that requires grasping the same type of cup ...

After determining the object is a cup 414 and obtaining information about the cup 414, the cloud computing system 401 can send object data 418 associated with the cup 414 to the robot 413 in response to the identification query 415 received from the robot 413. The object data 418 may include both the identity of the cup 414 and instructions for interacting with the cup 414 based on the application or task that the robot 413 has been instructed to execute or perform...

Importantly, the cloud computing system 401 updated the amount of force to use when grasping the particular type of cup that the robot 413 must now grasp (i.e., cup 414) based on the feedback 411 previously received from robot 404 regarding the same type of cup. As a result, when the robot 413 applies the “grasp” task to the type of “cup” object here (i.e., when robot 413 grasps cup 414), the robot 413 successfully grasps the cup without crushing the cup 419. Thus, **the second robot 413 has in effect learned from the experience of the earlier robot 404.**

Id., 26:29-27:21.

155. The combination of Grotmol and Hickman discloses a shared knowledgebase that includes multiple training sets (i.e., circumstance representations correlated with instruction sets) learned during the operation of multiple devices (i.e., first robot, second robot, etc.). A POSITA would understand that this shared knowledgebase would include training sets learned during operation of a first robot (i.e., first and second circumstance representations correlated with first and second instruction sets for operating a first device) as well as trainings sets learned during operation of another robot (i.e., fourth circumstance representation correlated with instruction set for operating a second device).

3. **Claim 11: The method of claim 1, wherein the knowledgebase further includes a fourth correlation including a fourth circumstance representation correlated with a fourth one or more instruction sets for operating the second device, and wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by the one or more sensors of the second device, and wherein at least a portion of the fourth correlation is learned in another learning process that**

includes operating the second device at least partially by another user.

Claim 11	Additional Features to Claim 1
	+ 4th correlation (4th CR and 4th instructions) learned with 2nd user • 4th CR detected by 2nd device • 4th instructions for 2nd device

156. As I discussed above in Ground 2, Grotmol discloses that multiple learning processes can be performed to train multiple behaviors. As I discussed above with respect to claim 4 in Ground 1, Grotmol discloses that a single robot can learn from the training of multiple robots. As I discussed above with respect to claim 4 in Ground 3, Hickman discloses that humans can coach the various robots connected to its shared knowledgebase. Ex. 1005, 23:36-48; 26:21-27:21. As also detailed above, it would have been obvious to combine Hickman's shared robot knowledgebase with Grotmol's autonomous robot system such that the shared knowledge base includes the training of multiple robots by multiple users.

157. The combination of Grotmol and Hickman discloses a shared knowledgebase which contains any number of correlations, consisting of circumstance representations and associated instruction sets, detected and learned on any number of devices being trained by any number of users. A POSITA would understand that the shared knowledgebase would include training sets learned during

operation of a first robot by a user (i.e., first and second circumstance representations correlated with first and second instruction sets for operating a first device) as well as trainings sets learned during operation of another robot by another user (i.e., fourth circumstance representation correlated with instruction set for operating a second device).

IX. The '344 Patent

A. The Claims

158. I understand that the petition challenges claims 1-20 of the '344 patent. Independent claim 1 of the '344 patent is directed to a method in which a device anticipates and autonomously executes instruction(s) corresponding to a newly detected circumstance based on comparison to a previously detected and stored circumstance. As I explain in this Declaration, a POSITA, in and prior to 2016, would have understood that there was nothing inventive about this combination of features, which were well known in the art at the time. Claim 1 is reproduced below:

1. A system comprising:
one or more processors configured to perform at least:
(a) accessing a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device, and wherein at least a portion of the first

circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;

(b) generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;

(c) anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation; and

(d) at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.

159. Independent claim 19 is a method claim that is nearly identical to system claim 1.

160. Independent claim 20 is a system claim that is nearly identical to system claim 1, except that claim 20 requires a “means for processing,” whereas claim 1 requires “one or more processors.”

161. Dependent claims 2 and 3 depend from claim 1, claims 4-13, 16-18 depend from claim 3, and claims 14 and 15 depend from claim 13. Each of these dependent claims can be summarized as follows in the table below, based on the additional steps or features that they require. As I further explain in this Declaration, the steps and features added by these dependent claims were also well-known at or prior to 2016, and a POSITA would have understood that there was nothing inventive about the combination of steps and features recited in these claims.

Dependent Claim	Additional Features
Claim 2	In addition to the limitations of claim 1, claim 2 further requires the first instruction sets to include information about the state of the first device. + State information of 1st device
Claim 3	In addition to the limitations of claim 1, claim 3 further requires the 1st and 2nd circumstance representations to include one or more object representations. + 1st CR and 2nd CR include object representations
Claim 4	In addition to the limitations of claim 3, claim 4 further limits the system of claim 3 to only the first device. detection of 2nd CR and performance of 1st instructions on 1st Device

Dependent Claim	Additional Features
Claim 5	In addition to the limitations of claim 3, claim 5 further requires the second circumstance representation to be detected by the second device, and the second device to perform the operations defined by the first instruction sets. + detection of 2nd CR by 2nd device + performance of 1st instructions on 2nd Device
Claim 6	In addition to the limitations of claim 3, claim 6 further requires a server to receive the first circumstance representation or instructions from the first device, and then send what was received to the second device. + server receives 1st CR or 1st instructions from 1st device + server sends 1st CR or 1st instructions to 2nd device

Dependent Claim	Additional Features
Claim 7	In addition to the limitations of claim 3, claim 7 further requires modification of the first instruction set(s) for operating the first device, and performance of those instruction(s) on either the first or second device. + modification of 1st instructions + performance of modified 1st instructions on 1st device or 2nd device
Claim 8	In addition to the limitations of claim 3, claim 8 further requires modification of the first or second circumstance representations, and then comparison between the two with one or both modified. + modification of 1st CR or 2nd CR + comparison of 1st CR to 2nd CR, with one or both modified
Claim 9	In addition to the limitations of claim 3, claim 9 further requires the knowledgebase to include a third circumstance representation detected on the first device correlated with a second instruction set for the first device learned with a first user. + 3rd CR (detected by 1st device) correlated with 2nd instructions (for 1st device) learned with 1st user

Dependent Claim	Additional Features
Claim 10	In addition to the limitations of claim 3, claim 10 further requires the knowledgebase to include a third circumstance representation detected on a third device correlated with second instruction set for the first device learned with a second user. + 3rd CR (detected by 1st Device) correlated with 2nd instructions (for 1st Device) learned with 2nd user
Claim 11	In addition to the limitations of claim 3, claim 11 further requires the knowledgebase to include a third circumstance representation detected on a third device correlated with second instruction set for the third device learned with a first or second user. + 3rd CR (detected by 3rd Device) correlated with 2nd instructions (for 3rd Device) learned with 1st user or 2nd user

Dependent Claim	Additional Features
Claim 12	In addition to the limitations of claim 3, claim 12 further requires the knowledgebase to include a third circumstance representation detected on a third device correlated with second instruction set for the third device and a fourth circumstance representation detected on a fourth device. The second instructions for the third device must be anticipated by a match between the fourth and third circumstance representations, and the fourth device must autonomously perform the second instructions for the third device. + 3rd CR (detected by 3rd Device) correlated with 2nd Instructions (for 3rd Device) + 4th CR (detected by 4th Device) + anticipation of 2nd Instructions by matching 3rd CR & 4th CR + autonomous performance of 2nd instructions on 4th Device
Claim 13	In addition to the limitations of claim 3, claim 13 further requires the learning of either the first circumstance representation or first instruction sets, and the anticipation of the first instruction sets to be performed by an “artificial intelligence system.” + artificial intelligence system

Dependent Claim	Additional Features
Claim 14	In addition to the limitations of claim 13, claim 14 further requires the artificial intelligence system to include object representations as inputs and instruction sets as outputs. + input object representations to AI system and output instructions
Claim 15	In addition to the limitations of claim 13, claim 15 further requires the artificial intelligence system to include either a knowledgebase or a neural network. + AI system includes knowledgebase or neural network
Claim 16	In addition to the limitations of claim 3, claim 16 further requires the first circumstance representation to be connected to the 1st instruction sets. + connection between 1st CR and 1st Instructions
Claim 17	In addition to the limitations of claim 3, claim 17 further requires the anticipation of the first instruction sets to be based on a partial match between the first and second object representations. + matching between 1st object representation and 2nd object representation

Dependent Claim	Additional Features
Claim 18	In addition to the limitations of claim 3, claim 18 further requires all of the structures listed in claim 1 to include one of various listed embodiments. + various embodiments

B. Ground 1: Claims 1-5, 7-9, and 13-20 Are Anticipated by Grotmol

162. As set forth below, Grotmol discloses each limitation of claims 1-5, 7-9, and 13-20.

1. Claim 1

a. [1.1] A system comprising:

Limitation	Claimed Features
[1.1]	system

163. Grotmol discloses a system implemented as a “self-contained robot apparatus”:

Another aspect of the disclosure relates to a self-contained robotic apparatus. The apparatus may comprise a platform, a first sensor, a second sensor, a memory component, and one or more physical processors.

An apparatus and methods for training and/or operating a robotic device to follow a trajectory.

In some implementations, a user (e.g., human operator may train the system to Switch between tasks based on the sensory context. FIG. 2 shows a diagram of the **system**, according to some implementations.

Ex.1004, 3:27-30, Abstract, 10:45-48.

b. [1.2] one or more processors configured to perform at least:

Limitation	Claimed Features
[1.2]	processor circuit(s)

164. Grotmol discloses that its robot may contain “one or more physical processors.” *Id.*, 3:72-30. Although Grotmol does not specifically use the term “processor circuits,” a POSITA would have understood the “one or more processors” of Grotmol to include “one or more processor circuits.” The ’344 patent only defines “processor circuits” as “includ[ing] a logic circuit,” which itself “may include a microcontroller.” ’344 patent, 2:34-36. Grotmol defines “processor” as follows:

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.

Ex. 10004, 8:5-15. Given this definition, a POSITA would have understood Grotmol's use of the term "processor" to include processor circuits that include logic circuits that may include a microcontroller.

- c. **[1.3] accessing a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device;**

Limitation	Claimed Features
[1.3]	access memory including 1st CR correlated with and 1st instructions for 1st Device • CR = image/video frame

165. Grotmol discloses accessing a memory that stores multiple training sets, each including an image of a scene from a robot vehicle's operation (circumstance representation) correlated with a motor control indication configured to cause the robot to execute an action (instruction set):

The **memory component** may be configured to store training sets. 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.

A robotic vehicle may utilize a camera and stores the sequence of images of a visual scene when following a trajectory during

training in an ordered buffer. **Motor commands associated with a given image may be stored.**

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.

In some implementations of robotic vehicle navigation, **output of the sensor 166 in FIG. 1 may comprise representations** of one or more objects (e.g., targets, and/or obstacles).

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. Additional 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.

Id., 3:35-39, Abstract, 8:66-9:3, 10:23-26, 39:13-22. Grotmol also teaches that one embodiment for this method is autonomous vehicle navigation. *Id.*, 38:36-58.

166. Grotmol discloses that the stored images are “associated with” or “correspond to” contemporaneously issued motor control indications. *Id.*, 33:59-62 (“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”), Abstract

(“Motor commands *associated with* a given image may be stored.”). A POSITA would have understood the associations between the motor commands and images in Grotmol to correspond to the claimed “correlat[ion].” Grotmol further explains that the associated motor commands may correspond to the time when the images are captured. *Id.*, 39:17-19 (“Motor commands issued by a trainer *corresponding to* time instances when the images are acquired may be stored.”), 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.”).

167. As an initial matter, a POSITA would have understood the “circumstance representations” of the ’344 patent to comprise digital images or instances of a video frame. The ’344 patent specification does not define the term “circumstance representation,” but instead defines a “collection of object representations,” as “one or more representations of objects detected by the sensor at a time.” ’344 patent, 3:42-44. The ’344 patent further explains that the “sensor” that detects the object representations may include a “[c]amera... for capturing one or more pictures” or a “motion picture camera that can capture streams of pictures (i.e. motion pictures, videos, etc.).” *Id.*, 69:51-70:32. Consistent with the ’344 patent’s description, Grotmol discloses that images from the robot’s sensors represent circumstances detected by a sensor of the first device. Specifically,

Grotmol discloses the video frame instances are detected by a “sensor component” of the robotic vehicle. Ex. 1004, 3:31-33 (“The first sensor component may be configured to provide a signal conveying a video frame comprising a plurality of pixels.”), *see also* Abstract (“A robotic vehicle may utilize a camera and stores the sequence of images of a visual scene when following a trajectory during training in an ordered buffer.”).

168. A POSITA would further understand the “circumstance representations” of the ’344 patent to comprise the images from Grotmol due to additional similarities. Both Grotmol and the ’344 patent discuss CCD image sensors that capture pictures in the form of color encoded pixels. ’344 patent, 70:7-24; Ex. 1004, 9:8-22. Grotmol further discloses that these pixels are used for object (or circumstance) recognition. Ex. 1004, 8:66-9:3 (“In one or more implementations, such as object recognition, and/or obstacle avoidance, the output provided by the sensor component may comprise a stream of pixel values associated with one or more digital images.”).

169. A POSITA would also understand the “instruction sets” of the ’344 patent to correspond to the motor commands of Grotmol, due to the way the term is broadly described in the specification. The ’344 patent discloses instruction sets of any programming language or construct that can be executed on a processor or logic

circuit, used by an actuator to operate a device, and can be received by any element of the system.

In some embodiments, an instruction set (i.e. Instruction Set 526, etc.) may be used or executed by a processor (i.e. Processor 11, etc.) in operating a device... or executed by a logic circuit (i.e. Logic Circuit 250, etc.) in operating a device... [or] used by an actuator in operating a device. For example, such instruction set may be or include one or more inputs into an actuator... In further aspects, an instruction set can be received from a processor, application program, logic circuit, and/or other processing element... In further designs, an instruction set can be received from memory (i.e. Memory 12, etc.), hard drive, or any other storage element or repository. In further designs, an instruction set can be received over a network such as Internet, local area network, wireless network, and/or other network... In general, **an instruction set can be received by any element of the system.**

Instruction Set 526 may include one or more instructions or commands of a high-level programming language such as Java or SQL, a low-level language such as assembly or machine language, an intermediate language or construct such as bytecode, and/or **any other language or construct.**”

’344 patent, 152:5-45, 93:34-39.

170. Grotmol shows in Figure 8, copied below, that the “memory” component is connected to the “sensory” functionality through the “processing” functionality, consistent with claim 1’s requirements that the images stored in the memory are detected by the robot’s sensors:

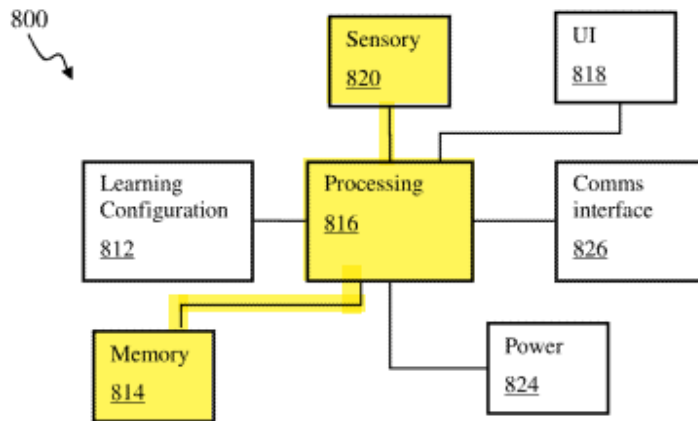


FIG. 8

Ex. 1004, Fig. 8 (annotated).

171. Grotmol further discloses accessing the training sets stored in the memory component to perform autonomous operation. *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.”); *see also 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.”).

172. Additionally, a POSITA would have understood the memory component of Grotmol to comprise a “knowledgebase” within the meaning of the ’344 patent. The ’344 patent states that a knowledgebase “may be or include *any data structure or arrangement* capable of storing the knowledge of a device’s

operation in circumstances including objects with various properties.” ’344 patent, 102:15-18.

- d. [1.4] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;

Limitation	Claimed Features
[1.4]	portion of 1st CR or 1st Instructions learned with user operation

173. Grotmol discloses that a human⁴ can operate the robot during a training process and “[t]rained 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.” Ex. 1004, 33:59-62. In both the “listen mode” and “override mode,” Grotmol teaches that the robot will listen to commands sent by a user “and try to remember it for next time the same situation occurs.” *Id.*, 23:3-21; *see also id.*, 24:56-61 (“While training a particular behavior, the *user* may concentrate on a

⁴ The challenged claims of the ’344 patent recite “user(s),” which in my analysis I have considered to include humans. However, the ’344 patent does not limit the term “user” to only humans. ’344 patent, 68:29-31 (“User 50 (also referred to simply as user or other suitable name or reference) comprises a human user or non-human user.”).

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).”). 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**”).

- e. **[1.5] generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;**

Limitation	Claimed Features
[1.5]	generation or receipt of 2nd CR detected by 1st or 2nd device

174. Grotmol discloses generating or receiving a new image (second circumstance representation) detected by the robot’s sensors in addition to the first stored circumstance representation:

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

In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1) comprising a camera (e.g., 166). The robot may comprise a controller configured to store a *sequence of images* obtained during training phase.

Ex. 1004, 39:23-26, 38:36-39. A POSITA would have understood this “most, recent current” image to correspond to the claimed “second circumstance representation.”

175. Grotmol further discloses that training sets can be saved and loaded onto other robots (i.e., robots other than the one that was operated as part of the training). *Id.*, 33:27-32 (“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. In some implementations, *the trained configuration may be loaded to one or more other robots* in order to provide learned behaviors.”). Thus, a POSITA would have understood Grotmol to disclose that the most recent image can be detected by a camera of the first robot that was used during the training mode or by the camera of a second robot that has been loaded with a training set.

176. Grotmol further discloses that a sequence of video frames may be analyzed using a random k-nearest neighbors (RKNN) process in order to determine or (predict) motor control signals:

In some implementations, e.g., Such as described above in connection with FIGS. 4 and/or 10A-10B a random k-nearest neighbors (RKNN) approach may be used for associating sensory context with one or motor actions. In some implementations, the RKNN methodology may comprise online learning of predicting output y (e.g., motor commands) based on the input (e.g., plurality of features detected in sensory input).

By way of an illustration of operation of a robotic device controller (e.g., 400 in FIG. 4A), sensory input (e.g., 402 comprising a sequence of video frames, inertial motion measurement, motor actuator feedback) may be analyzed using RKNN process in order to determine (predict) motor control signal (418).

Id., 27:48-55, 28:51-56. Grotmol further explains that the inputs for the RKNN process may include specific pixels from individual video frames, and that these inputs (x) may be paired with motor control signals (y) produced at time (t):

Sensory input may comprise a plurality of features (e.g., representations of objects determined using video data). In some implementations, the RKNN process may comprise configuring a plurality of N KNN classifiers to process randomly selected subsets of features. For a given classifier C_i (e.g., 1408 in FIG. 14), a random subset x_i of features may be selected from a pool of potential features $x = \{x_1 \dots x_n\}$. As used herein, the term “classifier” may be used to describe a data processing operation configured to provide an output y (e.g., motor control command) based on analysis of a plurality of **inputs x_i (e.g., pixels of a digital image)**.

FIG. 15 illustrates an exemplary configuration for producing output y_p configured based on input x using $N=3$ classifiers and $k=4$ nearest neighbors. Tables 1500, 1530, 1560 in FIG. 15 may

represent three classifier C_i instances corresponding to, e.g., index selection described by elements 1408, 1406, 1404. Individual rows 1502, 1504, 1506, 1508, 1510, 1512, 1514, 1516, 1532, 1534, 1536, 1538, 1540, 1542, 1544, 1546, 1562, 1564, 1566, 1570, 1572, 1574, 1576 may denote **training pairs (x,y) produced during training at time instances $t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_m$.**

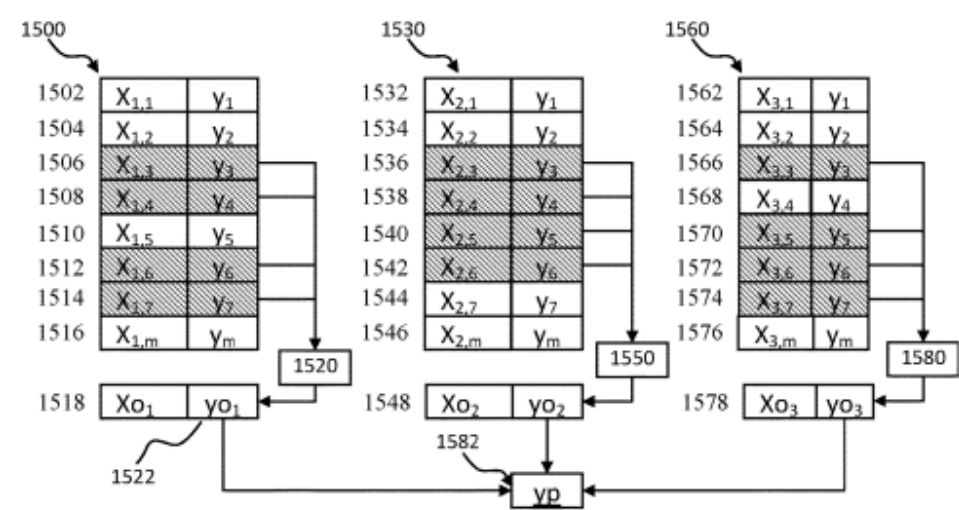


FIG. 15

Id., 28:56-67, 29:37-44, Fig. 15. A POSITA would have understood these sections of Grotmol to explicitly disclose any number of circumstance representations x produced at time instances t_1 to t_m , including a second circumstance representation.

- f. [1.6] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation;

Limitation	Claimed Features
[1.6]	Anticipation of 1st instructions by matching 1st CR and 2nd CR

177. Grotmol discloses anticipating motor commands (instruction sets for operating the first device) based on a partial match between the new image (second circumstance representation) and a stored image (first circumstance representation). *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”). A POSITA would have understood Grotmol’s disclosures for determining the best matching stored image by calculating a similarity measure between a current image and a stored image and then determining whether that similarity measure breaches a threshold to correspond to the claimed “at least partial match between the third circumstance representation and the first circumstance representation.”

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. The method may comprise determining one or more potential control signals for the robot.

Id., 2:29-38; *see also id.*, 39:13-41, 40:64-41:17. Grotmol clarifies that the term “feature” as used in the paragraph excerpted above can include values of image pixels or other image qualities. *Id.*, 28:19-24. (“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”). Grotmol also clarifies that the term “training set” used in the paragraph excerpted above includes the images that a POSITA would have understood correspond to the claimed circumstance representation. *Id.*, 3:36-39 (“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”).

178. A POSITA would have appreciated that both Grotmol and the ’974 patent discuss pixel-wise difference analysis between two digital pictures to determine a “partial match.” ’344 patent, 66:34-38 (discussing “matching”), 73:28-61 (discussing object detection and recognition by “comparing regions of pixels”); Ex. 1004, 3:14-16 (“The first input type may comprise a digital image comprising a plurality of pixel values.”), 41:8-17 (“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.”).

- g. **[1.7] and at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.**

	Claimed Features
--	-------------------------

Limitation [1.7]	executing 1st instructions on 1st or 2nd device autonomously
-----------------------------------	---

179. As I discussed above for limitation [1.6], Grotmol discloses anticipating the learned instruction sets for operating the robot 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.

180. Grotmol discloses that after determination that the similarity measure between the new image and stored training image breaches a threshold, one or more potential control signals corresponding to the stored image obtained during training (the instruction sets correlated with the first circumstance representation) 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**.

Id., 2:34-41, 3:12-13; *see also* Grotmol, 43:5-9 (“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.”).

181. Grotmol discloses that the robot (first device) autonomously performs the operations defined by the first instruction set for operating the device. *Id.*, 34:56-59 (“In some implementations, the trainer may be provided with the control of the robot during training. Upon completion of the training, the robot may be configured to operate autonomously.”); *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.”), 43:5-9 (“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.”).

182. Grotmol also discloses that another robot (second device) can be loaded with a training set and be autonomously operated using that training set. *Id.*, 33:21-32 (“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. In some implementations, ***the trained configuration may be loaded to one or more other robots*** in order to provide learned

behaviors.”). A POSITA would understand that where a second robot is loaded with a training set, the same process disclosed in Grotmol of matching a current / new image with a stored image to determine an associated motor command to execute would be performed by the second robot. *Id.*, Abstract (“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.”).

2. **Claim 2: The system of claim 1, wherein the first one or more instruction sets for operating the first device include one or more information about one or more states of: the first device, or a portion of the first device.**

Claim 2	Additional Features to Claim 1
	+ State information of 1st device

183. Grotmol discloses training sets, (instruction sets) that contain information about the state of the robot (first device). As I discussed above for limitation [1.3], Grotmol discloses accessing 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.” *Id.*, 3:36-39. 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.” *Id.*, 39:19-22. This “additional sensory information” corresponds to the claimed “information about one or more states” of the first device.

3. **Claim 3: The system of claim 1, wherein the first circumstance representation includes a first one or more object representations, and wherein the second circumstance representation includes a second one or more object representations.**

Claim 3	Additional Features to Claim 1
	+ 1st CR and 2nd CR include object representations

184. A POSITA would have understood Grotmol to disclose the claimed circumstance representations including object representations. As I discussed above for limitations [1.3] and [1.4], Grotmol discloses the claimed first and second circumstance representations detected at least in part by one or more sensors of the first device or second device.

185. The ’344 patent does not define “object representations” beyond its use in the phrase “collection of object representations,” which it defines as “one or more representations of objects detected by the sensor at a time.” ’344 patent, 3:42-44. The “sensor” that detects the object representations may include a “[c]amera... for capturing one or more pictures” or a “motion picture camera that can capture streams of pictures (i.e. motion pictures, videos, etc.).” *Id.*, 69:51-63.

186. The '344 patent provides examples of detected objects such as “person, cat, vehicle, building, street, tree, rock, etc.” '344 patent, 72:59-64. The '344 patent further describes techniques for recognizing these objects within pictures, such as through the identification of edges, corners, and relative position. *Id.*, 75:8-12 (“In some aspects, picture recognition techniques and/or tools involve identifying and/or analyzing features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc., which may then be used to search for pictures with matching features.”).

187. A POSITA would have appreciated the similarity of Grotmol's disclosures that images taken of the devices' visual field at particular time instances contain object representations:

[I]mages 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. **The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.**

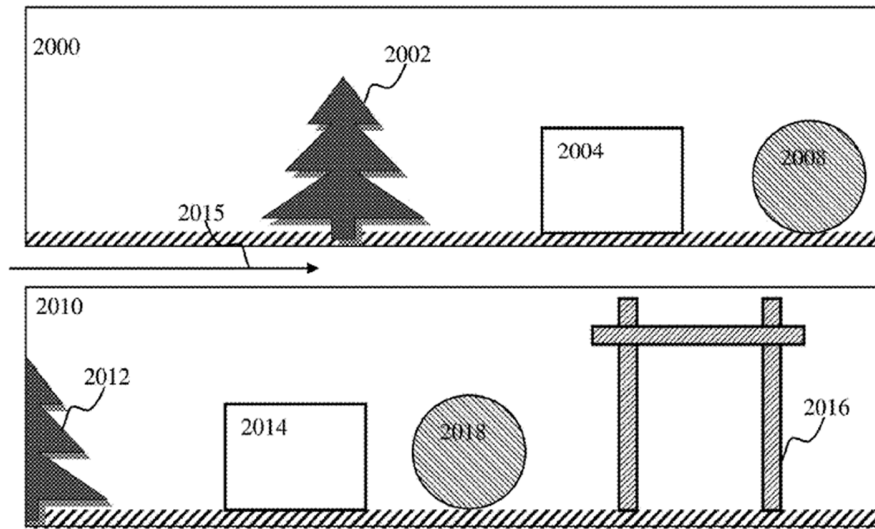


FIG. 20A

Ex. 1004, 38:52-58, Fig. 20A.

188. 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.” *Id.*, 8:57-9:3. *See, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented . . .”).

4. **Claim 4: The system of claim 3, wherein the second circumstance representation represents the second circumstance detected at least in part by the one or more sensors of the first device, and wherein the first device autonomously performs the one or more operations defined**

by the first one or more instruction sets for operating the first device.

Claim 4	Additional Features to Claim 3
	+ detection of 2nd CR and performance of 1st instructions on 1st Device

189. As I discussed above with respect to limitations [1.5] and [1.7] of claim 1, Grotmol discloses that a second image, which a POSITA would have understood to correspond to the claimed second circumstance representation, is detected by a camera of the robot used for training (i.e., first device), and the same robot autonomously performs the operations defined by the instructions stored in the training set.

5. **Claim 5: The system of claim 3, wherein the second circumstance representation represents the second circumstance detected at least in part by the one or more sensors of the second device, and wherein the second device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.**

Claim 5	Additional Features to Claim 3
	+ detection of 2nd CR by 2nd device + performance of 1st instructions on 2nd Device

190. As described above with respect to limitations [1.5] and [1.7] of claim 1, Grotmol discloses that training sets (including images corresponding to instructions at given times) from operating one robot can be loaded into other robots,

so that the other robots can perform the trained behaviors autonomously. Thus, a POSITA would have understood Grotmol to disclose a second robot, which is loaded with training sets from a first robot, detects a second circumstance representation with a camera, and autonomously performs operations defined by instructions for operating the first robot.

6. Claim 7

Claim 7	Additional Features to Claim 3
	+ modification of 1st instructions + performance of modified 1st instructions on 1st device or 2nd device

- a. [7.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: **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,**

Limitation [7.1]	Claimed Features
	modification of 1st instructions

191. Grotmol discloses that the robot may adjust its trajectory to compensate for, “mismatch[es]” in position or orientation, and further discloses modifying the motor command where a current image is shifted relative to a stored image:

Exemplary implementations of the methodology described herein may be applicable to controlling trajectory of a robotic device due to (i) **position mismatch** (e.g., the robot being located physically to the left of a target location); and/or (ii) due to

orientation mismatch (e.g., the robot being in the same physical location while oriented towards the left of the target trajectory). To illustrate, assuming the camera faces straight ahead, the center of the image may be the spot which the robot is headed towards. Thus, if this spot is to the left of the spot where the robot is supposed to be headed towards (as defined by the camera image seen during training), then **the robot may need to adjust its heading rightwards**.

During operation when the robot may follow a target trajectory the shift amount determined using the image matching process may be close to 0 (this configuration may be referred to as “the robot stays on track”). In some implementations the shift amount may be utilized as an error metric by the control process of the robot. **The steering signal (which may be adjusted leftwards or rightwards) may be selected as the control variable for the process.** A negative feedback loop may be used in order to reduce the error metric to (and/or maintain at) a target level operation of the robot. In some implementations, the target error level may comprise zero displacement.

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.

Id., 42:46-43:3; Abstract.

192. Grotmol further explains that motor commands corresponding to a given image and time may be combined with the output from a PID controller to “stabilize” the robot’s operation:

A PID controller may be used in order to reduce/maintain the error metric during operation of the robot. 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. Those motor commands may be combined with the output from the PID controller in order to **stabilize operation** of the robot.

Id., 43:4-11.

193. Accordingly, a POSITA would have understood Grotmol to disclose that the “first one or more instruction sets for operating the first device” learned in the learning process (i.e., motor command that is stored in association with image) are modified.

- b. **[7.2] and wherein the anticipating the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation includes: anticipating the modified the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, or anticipating the modified the copy of the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation, and**

Limitation	Claimed Features
[7.2]	Anticipation of modified 1st instructions

194. As I discussed above for limitation [1.6], Grotmol discloses that an instruction set is selected based on partially matching a stored image (first circumstance representation) with the current, most recent image (second

circumstance representation). *Id.*, 43:5-9; 2:29-39; 39:13-41; 40:64-41:18; 28:19-24; 41:8-17; 34:50-55; 42:46-43:3; 43:5-11. Further, as I discussed above for limitation [7.1], Grotmol discloses modifying instruction sets for operating the first device. Thus, a POSITA would have understood Grotmol to disclose that where a partial match occurs between a stored image and a new image, and the instruction set(s) associated with the stored image are modified, the anticipated instruction would be the modified instruction.

- c. **[7.3] wherein the executing the first one or more instruction sets for operating the first device includes: executing the modified the first one or more instruction sets for operating the first device, or executing the modified the copy of the first one or more instruction sets for operating the first device, and wherein the autonomously performing, by the first device or by the second device, the one or more operations defined by the first one or more instruction sets for operating the first device includes: autonomously performing, by the first device or by the second device, one or more operations defined by the modified the first one or more instruction sets for operating the first device, or autonomously performing, by the first device or by the second device, one or more operations defined by the modified the copy of the first one or more instruction sets for operating the first device.**

Limitation	Claimed Features
[7.3]	Execution of modified 1st instructions on first device or second device autonomously

195. As I discussed above for limitations [1.6], [7.2], and [7.3], Grotmol discloses that an instruction set is selected based on partially matching a stored image (first circumstance representation) with the current, most recent image (second circumstance representation), and that where a partial match occurs with a stored image that is associated with an instruction, and the instruction has been modified, a POSITA would have understood that the anticipated instruction would be the modified instruction. Furthermore, as I discussed above for limitation [1.7], Grotmol discloses executing the anticipated instruction set, and the first device or the second device autonomously performs operations defined by the instruction set. In light of these disclosures, a POSITA would have understood that where the anticipated instruction is modified, the modified instruction would be executed and the first device or second device would perform the operations defined by the modified instruction set.

7. Claim 8

Claim 8	Additional Features to Claim 3
	+ modification of 1st CR or 2nd CR + comparison of 1st CR to 2nd CR, with one or both modified

- a. **[8.1] The system of claim 3, wherein the one or more processors are further configured to perform at least: modifying at least one of: the first circumstance representation, a copy of the first circumstance representation, the second circumstance**

representation, or a copy of the second circumstance representation,

Limitation	Claimed Features
[8.1]	modification of 1st CR or 2nd CR

196. Grotmol discloses modifying captured images (circumstance representations) in a variety of ways, such as transformation, scaling, normalization, and/or equalization.

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

In order to facilitate contributions from different types of signals for determining a distance measure between features in a metric space (e.g., Euclidian distance), data from highly redundant input (e.g., the RGB pixel values) may be **normalized**. Various other distance measures (metrics) may be utilized, e.g., Mahalanobis, Manhattan, Hamming, Chebyshev, Minkowski, and/or other metrics.

In some implementations, **the normalization may comprise shifting and/or scaling input features to a given value range** (e.g., A1=64 to A2=196 for an 8-bit pixel value, 0 to 1 range, and/or other range). In one or more implementations, the normalization may be configured based on determining an on-line estimate of the mean and standard deviation of feature values to obtain z-score for individual feature (pixel). In one such implementation, for a given pixel (e.g., pixel at location (i1,i2)) a pair of values may be stored in history memory: one for the pixel mean and another for the pixel standard deviation. In some implementations, one or more parameters related to history of the input (e.g., pixel statistics) may be computed over a given interval, and/or the total duration of training. In one or more implementations, the learnings process may be configured to enable a user to reset contents of the parameter (e.g., pixel statistics).

Equalization may be applied so that contribution of individual sensory modality on expected distances may be comparable relative contribution from another modality data. In some implementations, the equalization may comprise determining an on-line estimate of the mean and standard deviation of individual features; using the on-line estimates to calculate a normalizing constant C_s for individual sensory modality s such that the expected Euclidean distance between two samples, measured only using the features in modality s is 1.0. Weighting data of a given modality (to further reduce the mean squared error) as training parameters that may be optimized during training.

The one or more physical processors may configured to operate a random k-nearest neighbors learning process to determine a motor control indication by at least: ... **scaling individual pixels** of the set of pixels by a scaling factor; **scaling features of the subset by a scaling factor**; **comparing individual scaled features of the subset** to corresponding features of individual ones of the training sets[.]

In some implementations, data for one or more inputs may be scaled by a parameter NF , where NF is configured based on the overall number of features of a given feature types (i.e., the number of pixels in a subset t). In some implementations, the scaling parameter may be selected from the range between $\sqrt{NF}$ and $10 \times NF$.

Id., 4:18-22, 30:47-31:2, 31:61-32:5, 3:39-48; 31:3-8. Thus, a POSITA would have understood Grotmol to disclose a system with processors configured to modify the first and second circumstance representations.

- b. **[8.2] and wherein the at least partial match between the second circumstance representation and the first circumstance representation includes: (i) at least partial match between the modified the second circumstance representation and the first**

circumstance representation, (ii) at least partial match between the modified the copy of the second circumstance representation and the first circumstance representation, (iii) at least partial match between the second circumstance representation and the modified the first circumstance representation, (iv) at least partial match between the second circumstance representation and the modified the copy of the first circumstance representation, (v) at least partial match between the modified the second circumstance representation and the modified the first circumstance representation, (vi) at least partial match between the modified the copy of the second circumstance representation and the modified the copy of the first circumstance representation, (vii) at least partial match between the modified the second circumstance representation and the modified the copy of the first circumstance representation, or (viii) at least partial match between the modified the copy of the second circumstance representation and the modified the first circumstance representation.

Limitation	Claimed Features
[8.2]	comparison of 1st CR to 2nd CR, with one or both modified

197. As I discussed above with respect to limitations [1.6] and [7.1], Grotmol discloses that an anticipated instruction set is selected based on partially matching a stored image (first circumstance representation) with the current, most recent image (second circumstance representation), and that these stored images can be modified. A POSITA would have understood that where either the current image or the stored image have been modified, that modified image (circumstance

representation) would then be used to determine whether there is a partial match. Thus, in my opinion, Grotmol discloses a partial match between (i) a modified second circumstance representation and a first circumstance representation, (iii) a second circumstance representation and a modified first circumstance representation, (v) a modified second circumstance representation and a modified first circumstance representation.

8. **Claim 9: The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating the first device, and wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the third circumstance representation or at least a portion of the second one or more instruction sets for operating the first device is learned in another learning process that includes operating the first device at least partially by the user.**

Claim 9	Additional Features to Claim 3
	+ 3rd CR (detected by 1st device) correlated with 2nd instructions (for 1st device) learned with 1st user

198. As I discussed above with respect to claim 1, Grotmol discloses 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 also explicitly discloses that multiple training processes can be conducted to learn multiple tasks of varying complexity, such as “fetch” and “grasping,”

specifically noting that a user can omit training of “other behaviors” to focus on one task at a time:

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. To achieve an acceptable performance of a **complex task** by the system, the user may provide additional information to the system...

For example, **a task of playing fetch may be quite complex**: the robot may be supposed to find an object, grasp it, and bring it back to base. In response to the robot not “seeing” the target, the robot may perform a random exploration behavior avoiding obstacles. In response to the robot being close to the target, the robot may perform a positioning maneuver and grasp the target. If the target is in the robot's gripper, the robot may find its way to the base avoiding obstacles and then release the object. The whole task may be highly nonlinear and/or noisy...

According to exemplary implementations, **the user may train a component of the fetch behavior**. For example, the user may train how to search for a target and/or how to approach it while avoiding obstacles (e.g., “target approach”). The user may train the robot to perform a grasping behavior in various contexts (e.g., “grasping a target”). The user may assume the robot successfully grasped the target and train it to find its way to the base (“base approach”). 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).

Id., 24:17-61. Grotmol also discloses that a single user can create discrete training “modules” for separate behaviors:

After is [sic] user is done training a particular behavior, he may create another module and activate it. Such creation is

depicted in the figure below, in accordance with one or more implementations.

The system may be configured to learn **the second behavior** from the user. Figure above shows that the user may select a different set of input features and a different set of outputs for “**Behavior 2**”. The user may select a different type of the adaptive controller (“2”). For example, the user may control the robot using a gamepad (or other input device used to provide “user commands”) to turn and grasp the target (“Behavior 2”) using input from target tracker (“FE 1”) and state of the gripper (“FE 2”). “Adaptive controller” of type “2” (e.g., perceptron classifier) may learn to turn the robot appropriately and grasp the target by autonomously effectuating correct torques on the robot's wheel (“output 2”) and gripper (“output N”), while not controlling another wheel (“output 1”). “Sensory input” may include input from the cameras and encoders.

Id., 26:1-20. Thus, a POSITA would have understood Grotmol to disclose that the knowledgebase includes a first circumstance representation and additional circumstance representations (including the claimed third circumstance representation), where the additional circumstance representation is detected by a sensor of the first device and is learned in *another* learning process that includes operating the device by the same user.

9. **Claim 13: The system of claim 3, wherein the system further comprising: an artificial intelligence system, and wherein the learning the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device includes learning, at least in part by the artificial intelligence system, the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device, and wherein the anticipating the first one or more instruction**

sets for operating the first device includes anticipating, at least in part by the artificial intelligence system, the first one or more instruction sets for operating the first device.

Claim 13	Additional Features to Claim 3
	+ Artificial Intelligence System

199. Claim 13 requires the use of an “artificial intelligence system to the learn the first circumstance representation or the first instruction set, and to anticipate the first instruction set, which as I discussed above in claim 1, involves matching a stored image with a newly acquired image.

200. The '344 patent gives a broad description of the term “artificial intelligence unit” as a generic piece of hardware or software:

In certain embodiments, the artificial intelligence unit is coupled to the sensor. In further embodiments, the artificial intelligence unit is coupled to the memory unit. In further embodiments, the artificial intelligence unit is part of, operating on, or coupled to the processor circuit. In further embodiments, the system further comprises: an additional processor circuit, wherein the artificial intelligence unit is part of, operating on, or coupled to the additional processor circuit. In further embodiments, the artificial intelligence unit is part of, operating on, or coupled to a remote computing device or a remote computing system, the remote computing device or the remote computing system coupled to the processor circuit via a network or an interface. In further embodiments, the artificial intelligence unit includes a circuit, a computing apparatus, or a computing system attachable to the processor circuit. In further embodiments, the artificial intelligence unit includes a circuit, a computing apparatus, or a computing system attachable to the device. In further embodiments, the artificial intelligence unit is attachable to an application for operating the device, the application running on

the processor circuit. In further embodiments, the artificial intelligence unit includes a circuit, a computing apparatus, or a computing system built into the processor circuit. In further embodiments, the artificial intelligence unit includes a circuit, a computing apparatus, or a computing system built into the device. In further embodiments, the artificial intelligence unit is built into an application for operating the device, the application running on the processor circuit. In further embodiments, the artificial intelligence unit is provided as a feature of the processor circuit. In further embodiments, the artificial intelligence unit is provided as a feature of an application running on the processor circuit. In further embodiments, the artificial intelligence unit is provided as a feature of the device. In further embodiments, the artificial intelligence unit is further configured to: take control from, share control with, or release control to the processor circuit. In further embodiments, the artificial intelligence unit is further configured to: take control from, share control with, or release control to an application or an object of the application, the application running on the processor circuit.

'344 patent, 3:1-41. The '344 patent does not limit the artificial intelligence system to a specific model or technique, specifically stating that both supervised and unsupervised learning can be utilized, and going as far as to say that "any model and/or technique can be utilized to facilitate the functionalities described herein":

Referring to FIG. 17, the disclosed **artificially intelligent devices, systems, and methods for learning** and/or using a device's circumstances for autonomous device operation may include various artificial intelligence models and/or techniques. 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, **supervised learning, unsupervised learning**, neural networks (i.e. convolutional neural network, recurrent neural network, deep neural network, etc.), search-based, logic and/or fuzzy logic-based, optimization-

based, tree/graph/other data structure-based, hierarchical, symbolic and/or sub-symbolic, evolutionary, genetic, multi-agent, deterministic, probabilistic, statistical, **and/or other models and/or techniques.**

Id., 103:33-48.

201. When anticipating instruction sets based on a partial match, the '344 patent states that “any level of similarity, however high or low, may be utilized” and “any one or more properties (i.e., existence, type, identity, distance, bearing/angle, location, shape/size, activity, etc.) of one or more objects can be utilized for determining similarity or match of circumstances.” *Id.*, 110:28-44, 111:10-15. Given these broad disclosures, in my opinion, a POSITA would not consider the '344 patent's instruction anticipation process to require any specific artificial intelligence model or technique.

202. As I discussed above for limitation [1.3], Grotmol discloses a system in which one or more processors facilitate the learning of associations between circumstance representations detected by cameras on the robotic device and instruction sets for operating the device. Like the '344 patent, Grotmol states that either supervised or unsupervised learning may be utilized. Ex. 1004, 16:8-12 (“the 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 one and/or multiple tasks.”), 24:17-20 (“Apparatus and methods for hierarchical learning are disclosed herein, in accordance with one

or more implements. In *supervised learning*, a user may train simple tasks by demonstrating example tasks to the system.”). And as I discussed above for limitation [1.6], Grotmol discloses a system in which processors perform the operation of anticipating an instruction by calculating a similarity measure between a current image and a stored image, then determining whether that similarity measure breaches a threshold. Thus, in my opinion, a POSITA would have understood Grotmol to disclose an artificial intelligence system that performs that the learning and anticipating steps.

10. Claim 14

Claim 14	Additional Features to Claim 13
	+ input object representations to AI system and output instructions

- a. [14.1] The system of claim 13, wherein the artificial intelligence system includes: one or more inputs for receiving one or more object representations, and one or more outputs for providing one or more instruction sets,

Limitation [14.1]	Claimed Features
	input object representations to AI system and output instructions

203. As described above for limitations [1.3], [1.6], claim 3, and claim 13, Grotmol discloses an artificial intelligence system that utilizes one or more processors to learn correlations between images including object representations and

instruction sets, and to anticipate instructions for operating a robot based on a partial match between a current image and a stored image. As shown in Figure 8 of Grotmol below, the processor is connected to the “sensor” functionality to receive input images containing object representations:

[T]he output provided by the sensor component 166 may comprise a stream of pixel values associated with one or more digital images.

In one or more implementations, images 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.

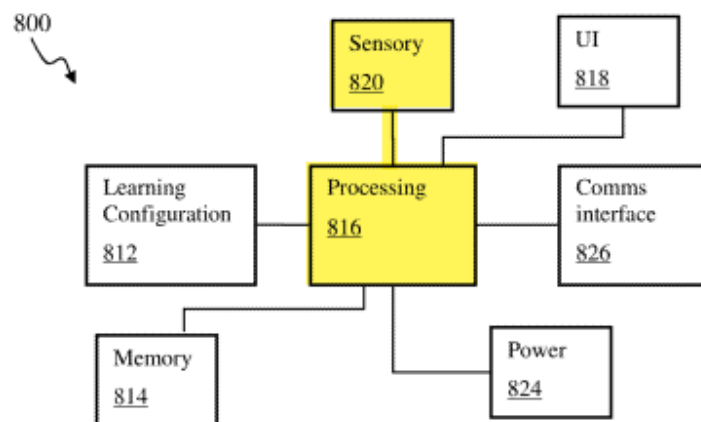


FIG. 8

Id., 9:1-3, 38:52-58, Fig. 8 (annotated).

204. Grotmol also discloses that motor commands associated with images are outputs for operation of the robot:

A PID controller may be used in order to reduce/maintain the error metric during operation of the robot. 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. **Those motor commands may be combined with the output from the PID controller in order to stabilize operation of the robot.**

Id., 43:4-11. Further, Grotmol describes an “adaptive control system,” shown in Fig. 4A below, that facilitates a supervised and/or reinforcement learning process, by using motor commands as “control signal 420,” which a POSITA would have understood to be the output of the control system:

FIG. 4A illustrates an implementation of adaptive control system 400. The adaptive control system 400 of FIG. 4A may comprise a corrector 412, an adaptive predictor 422, and a combiner 414 cooperating to control a robotic platform 430. **The learning process of the adaptive predictor 422 may comprise a supervised learning process, a reinforcement learning process, and/or a combination thereof.** The corrector 412, the predictor 422 and the combiner 414 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.

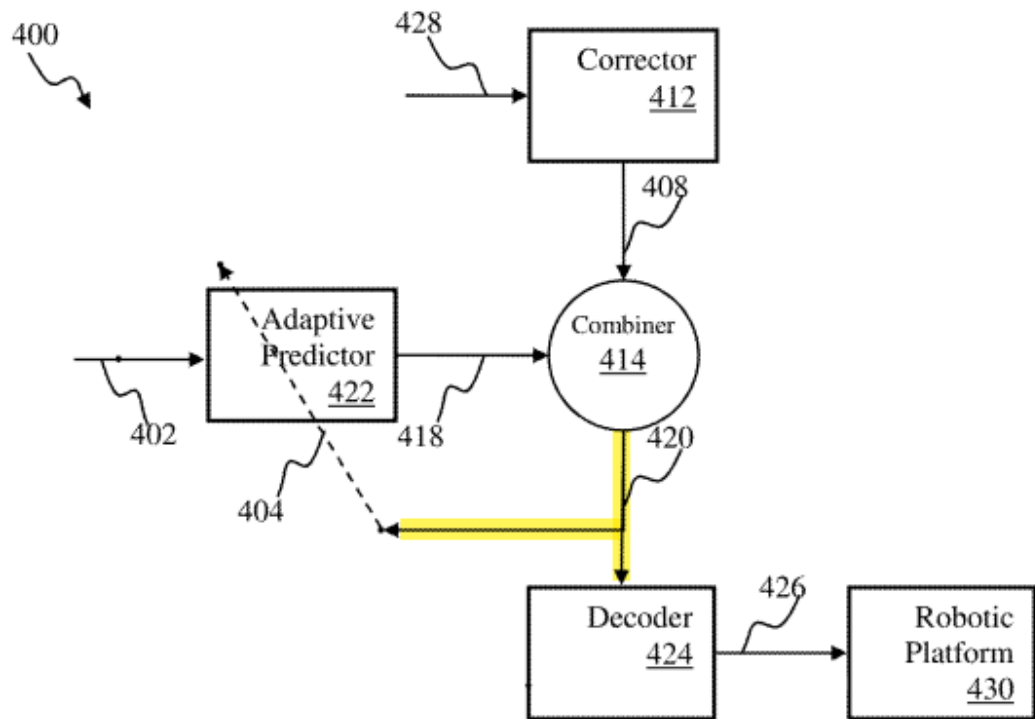


FIG. 4A

Id., 15:19-33, Fig. 4A (annotated). Thus, a POSITA would have understood Grotmol to disclose inputs for receiving object representations and outputs for providing instruction sets.

- b. [14.2] and wherein the learning, at least in part by the artificial intelligence system, the at least the portion of the first circumstance representation or the at least the portion of the first one or more instruction sets for operating the first device includes applying the first one or more object representations to the one or more inputs of the artificial intelligence system,

Limitation	Claimed Features
[14.2]	Applying 1st object representations to input of AI system to learn 1st CR or 1st instructions

205. As I discussed above with respect to claim 3 and limitation [14.1], Grotmol discloses that the learning process includes recognizing objects within a given image captured by a camera, and further that the processor includes an input for receiving images having one or more object representations. Thus, Grotmol discloses that a learning process includes applying one or more object representations to the input of an artificial intelligence system.

- c. **[14.3] and wherein the anticipating, at least in part by the artificial intelligence system, the first one or more instruction sets for operating the first device includes: applying the second one or more object representations to the one or more inputs of the artificial intelligence system; and receiving the first one or more instruction sets for operating the first device from the one or more outputs of the artificial intelligence system.**

Limitation [14.3]	Claimed Features
	applying 2nd object representations to input of AI system and receiving 1st instructions as output

206. As I discussed above with respect to limitations [1.3] and [1.6], Grotmol discloses a processor that performs the step of anticipating an instruction for operating a first device by comparing a new image (second one or more object representations) with a stored picture at least by matching portions of the pictures and determining whether those matching portions exceed a threshold number to output a control signal. Thus, A POSITA would have understood Grotmol to disclose

applying a new image (second one or more object representations) to an input of an artificial intelligence system running on the processor, and receiving from an output of an artificial intelligence system running on the processor a motor command (instruction set) for operating the first device.

11. Claim 15: The system of claim 13, wherein the artificial intelligence system includes at least one of: the knowledgebase, or a neural network.

Claim 15	Additional Features to Claim 13
	+ AI system includes knowledgebase or neural network

207. The '344 patent describes that an artificial intelligence unit may be coupled to a memory unit that stores a knowledgebase. '344 patent, 3:1-3 ("In certain embodiments, the artificial intelligence unit is coupled to the sensor. In further embodiments, the artificial intelligence unit is coupled to the memory unit."), 16:39-40 ("The knowledgebase may be stored in the memory unit."). As I discussed above for claim 13, Grotmol discloses an artificial intelligence system running on a processor that performs the learning and anticipating steps, and as I further discussed for limitations [1.3] and [1.4], Grotmol discloses the claimed knowledgebase stored in a memory unit that is coupled to the processor. Thus, a POSITA would have understood Grotmol to disclose an artificial intelligence system that includes a knowledgebase.

12. **Claim 16: The system of claim 3, wherein the first circumstance representation correlated with the first one or more instruction sets for operating the first device include at least a portion of the first one or more object representations connected, using at least one or more connections, with at least a portion of the first one or more instruction sets for operating the first device.**

Claim 16	Additional Features to Claim 3
	+ connection between 1st CR and 1st Instructions

208. As I discussed above with respect to claim 3, Grotmol discloses storing object representations and instruction sets as trained behaviors, which may be configured based on associations between the presence of an object and the action taken during training to avoid that object:

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.

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.

Ex. 1004, 33:24-26, 33:59-62. In my opinion, a POSITA would have understood the association between an object representation and an action to be a “connection” between the two. At least the data structures described in Grotmol linking object representations and instruction sets together in memory, for example, would comprise “one or more connections” as recited in the claim. *Id.*, 3:35-39 (“The

memory component may be configured to store training sets. *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.”), Abstract (“Motor commands associated with a given image may be stored.”), 39:13-19 Thus, Grotmol discloses a system where a first circumstance representation correlated with a first instruction includes a connection between an object representation and an instruction.

13. **Claim 17: The system of claim 3, wherein the anticipating the first one or more instruction sets for operating the first device based on the at least partial match between the second circumstance representation and the first circumstance representation includes determining the first one or more instruction sets for operating the first device based on at least partial match between the second one or more object representations and the first one or more object representations.**

Claim 17	Additional Features to Claim 3
	+ matching between 1st object representation and 2nd object representation

209. As I discussed above with respect to limitation [1.6] and claim 3, Grotmol discloses anticipating an instruction set based on a partial match between the first and second circumstance representation, and further that the first and second circumstance representations include object representations. Grotmol further

discloses determining a partial match between object representations from two images, one before and one after a right turn:

In one or more implementations, images 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. **The image 2000 may comprise representations one or more objects**, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects. During motion of the vehicle 160 relative position and/or orientation of the camera and the visual scene may change. By way of an illustration, the vehicle 160 may execute a right turn. The image 2010 may be obtained subsequent to the turn execution. Representation of objects in the image 2010 may be shifted relative representations of the same objects in the frame obtained prior to the turn execution. Additional object representations (e.g., 2016 in FIG. 20A) may appear, while representations of some object that may be present in one image (e.g., 2000) may be absent from the subsequent image (e.g., 2010). **Comparing image 2010 to image 2000, representations of the tree 2012 the rectangle, 2014, and the ball 2018 are shifted** (horizontally) within the frame 2010 relative representations, 2002, 2004, 2008 in the frame 2010. The amount of the shift may be configured based on relative motion between the images (e.g., motion speed and time interval $t_2 - t_1$ between the images). **It may be further noticed from comparing images 2000 and 2010 that the ball 2018 in image 2010 may be disposed closer to the rectangle 2014 compared to distance between respective object representations 2008, 2004 in image 2000.**

Id., 38:52-39:12. Thus, a POSITA would have understood that Grotmol discloses anticipating instructions sets based on a partial match between first and second object representations.

14. Claim 18

Claim 18	Additional Features to Claim 3
	+ various embodiments

- a. **[18.1] The system of claim 3, wherein at least some elements of the system are included in: a single device, or multiple devices, and**

210. Grotmol discloses the method of claim 1 implemented using a “self-contained robotic apparatus,” *Id.*, 3:27-30 (“Another aspect of the disclosure relates to a self-contained robotic apparatus. The apparatus may comprise a platform, a first sensor, a second sensor, a memory component, and one or more physical processors.”); *see supra* limitation [1.1]. Thus, a POSITA would have understood Grotmol to disclose that the elements of the computing system are included in a single device.

- b. **[18.2] wherein the one or more processors include: one or more microcontrollers, one or more computing devices, or one or more electronic devices, and**

211. A POSITA would have understood that Grotmol discloses one or more processor circuits including a microcontroller, computing device, or electronic device for the same reasons I discussed above for the term “processor circuits” with respect to limitation [1.2].

- c. **[18.3] wherein the memory includes: a non-transitory machine readable medium, a volatile memory, a non-**

**volatile memory, a storage device, or a storage system,
and**

212. Grotmol explicitly discloses that its “memory component” may comprise a “non-volatile memory.” *Id.*, 38:39-42 (“In some implementations, the images may be stored in an ordered buffer in a non-volatile memory.”); *see supra* limitation [1.3].

- d. [18.4] wherein the one or more sensors of the first device include at least one of: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, or one or more apparatuses for detecting objects or object properties, and wherein the one or more sensors of the second device include at least one of: one or more cameras, one or more microphones, one or more radars, one or more lidars, one or more sonars, or one or more apparatuses for detecting objects or object properties, and**

213. As described above for limitation [1.3] and claim 3, Grotmol discloses that the robot includes a sensor comprising a camera for detecting objects or object properties:

FIG. 1 depicts a mobile robotic apparatus that may be configured with an adaptive controller in accordance with one or more implementations of e.g., the learning apparatuses illustrated in FIGS. 4A-4B, *infra*. The robotic apparatus 160 may comprise a sensor component 166. The sensor component 166 may be characterized by an aperture/field of view 168 (e.g., an extent of the observable world that may be captured by the sensor at a given moment). The sensor component 166 may provide information associated with objects within the field-of-view 168. 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.

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 refreshed at, e.g., 25 Hz frame rate. The sensor output may be processed by a learning controller, e.g., as illustrated and described with respect to FIG. 4A.

In some implementations of robotic vehicle navigation, output of the sensor 166 in FIG. 1 may comprise representations of one or more objects (e.g., targets, and/or obstacles).

One or more implementations described herein may provide a mechanism for enabling a robot to learn navigating a target trajectory while reducing deviation from a target path. In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG. 1) **comprising a camera** (e.g., 166). 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).

FIG. 20A illustrates exemplary images for use with training of path navigation. Images 2000, 2010 may be obtained with **a variety of sensors, e.g., a video camera, IR camera, RADAR, SONAR, LiDAR, ultrasonic imager, and/or other device**. In some implementations, the camera may be spanning the entire horizontal view field of 360 degrees. A 360 degree horizontal view field may provide additional information about robot's environment and enable cyclical shift of images.

In one or more implementations, images 2000, 2010 may be obtained with **a camera 166 mounted on a robotic vehicle 160** of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. The

image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.

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. Additional 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.

Id., 8:57-9:3, 10:15-26, 38:33-58, 39:13-22.

214. Furthermore, as described above for claim 5, Grotmol discloses a second robot that includes a camera for detecting objects or object properties:

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. In some implementations, **the trained configuration may be loaded to one or more other robots in order to provide learned behaviors**. Subsequent to loading of the saved configuration, the controller learning process may match process configuration being present during saving of the configuration.

Id., 33:21-35.

- e. **[18.5] wherein the first device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine, and**

215. Grotmol discloses a robot (“self-contained robotic apparatus”) that may include a vehicle. *Id.*, 3:27-30; *id.*, 38:36-39 (“In some implementations, the robot may comprise a robotic vehicle (e.g., 160 in FIG.1)”); *see supra* claim 1.

- f. **[18.6] wherein the second device includes: a robot, a vehicle, an appliance, an electronic device, or a mechanical machine, and**

216. Grotmol discloses a second device comprising a robot. Grotmol discloses that training sets and training configurations can be stored or saved, and then loaded into the robot to provide the trained behavior. *Id.*, 33:27-32 (“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. In some implementations, the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that did the learning. *Id.*, 33:30-32 (“the trained configuration may be loaded into ***one or more other robots*** in order to provide learned behaviors”). Thus, a POSITA would have understood Grotmol to disclose 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 that is loaded into a second one or more other robots to “provide learned behaviors.”

- g. [18.7] wherein the first circumstance representation is a data structure that includes one or more data about the first circumstance, and wherein the second circumstance representation is a data structure that includes one or more data about the second circumstance,**

217. In my opinion, Grotmol discloses that the circumstance representations are data structures including data about the circumstances of the first device, within the meaning of the '344 patent.

218. The '344 patent lists various examples, of data structures “such as, for example, array, list, linked list, doubly linked list, queue, tree, heap, graph, map, grid, matrix, multi-dimensional matrix, table, database, database management system (DBMS), file, neural network, and/or any other type or form of a data structure including a custom one,” and states that “any variety of [these examples] can be used.” '344 patent, 65:55-63. As I discussed above with respect to limitation [1.3] of claim 1, Grotmol discloses circumstance representations comprising video frame instances. Grotmol discloses that a video frame “compris[es] a plurality of pixels.” Ex. 1004, 3:31-33. Thus, A POSITA would have understood the video frames disclosed by Grotmol to comprise “data structures” in the context of the '344 patent because they comprise at least an array, map, grid, or other form of pixels.

219. Further, A POSITA would also understand the entire video frame itself to comprise a “data structure” within the context of the ’974 patent, because it is a file including image data. The ’344 patent explains that a data structure can comprise a single file or field of data. ’344 patent, 65:61-65 (“A data structure may include *one* or more fields or data fields that are part of or associated with the data structure. A field or data field may include *a data*, an object, a data structure, and/or any other element or a reference/pointer thereto.”). Further, the ’344 Patent states that data structures “can be stored in one or more memories, files, or other repositories,” (*id.*, 65:65-67) just as the video frame instances are stored in the “memory component” of Grotmol. Ex. 1004, 3:27-39.

220. Accordingly, a POSITA would have understood the video frame instances of Grotmol to include “one or more data” about the “circumstance[s] of the first device,” because these video frame instances are detected by a sensor attached to the first device, and capture images of the device’s surroundings for use in object detection and avoidance:

FIG. 1 depicts a mobile robotic apparatus that may be configured with an adaptive controller in accordance with one or more implementations of e.g., the learning apparatuses illustrated in FIGS. 4A-4B, *infra*. The robotic apparatus 160 may comprise a sensor component 166. The sensor component 166 may be characterized by an aperture/field of view 168 (e.g., an extent of the observable world that may be captured by the sensor at a given moment). The sensor component 166 may provide information associated with objects within the field-of-view 168. 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.

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 refreshed at, e.g., 25 Hz frame rate. The sensor output may be processed by a learning controller, e.g., as illustrated and described with respect to FIG. 4A.

In some implementations of robotic vehicle navigation, output of the sensor 166 in FIG. 1 may comprise representations of one or more objects (e.g., targets, and/or obstacles).

In one or more implementations, images 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.

Id., 8:57-9:3, 10:15-26, 38:52-58.

- h. [18.8] and wherein the first circumstance includes one or more objects detected at least in part by the one or more sensors of the first device at a first time or during a first time period, and wherein the second circumstance includes: one or more objects detected at least in part by the one or more sensors of the first device at a second time or during a second time period, or one or more objects detected at least in part by the one or more sensors of the second device at a second time or during a second time period, and**

221. As I discussed above for limitations [1.3] and [1.5], Grotmol discloses a knowledgebase including a first circumstance representation detected by a first

device during a training session (i.e., a first time period), and second circumstance representation detected by a first or second device in the context of autonomous operation (i.e., a second time period). As I also discussed above for claim 3, Grotmol discloses that the first and second circumstance representations include detected objects.

222. The '344 patent provides examples of detected objects such as “person, cat, vehicle, building, street, tree, rock, etc.” '344 patent, 72:59-64. The '344 patent further describes techniques for recognizing these objects within pictures, such as through the identification of edges, corners, and relative position. *Id.*, 75:8-12 (“In some aspects, picture recognition techniques and/or tools involve identifying and/or analyzing features such as lines, edges, ridges, corners, blobs, regions, and/or their relative positions, sizes, shapes, etc., which may then be used to search for pictures with matching features.”).

223. A POSITA would have appreciated the similarity of Grotmol's disclosures that images taken of the devices' visual field at particular time instances contain detected objects:

[I]mages 2000, 2010 may be obtained with a camera 166 mounted on a robotic vehicle 160 of FIG. 1. Images 2000, 2010 may comprise representations of visual field that may be obtained at times $t_1 < t_2$, respectively. **The image 2000 may comprise representations one or more objects, e.g., a tree 2002, a rectangle, a ball 2008, and/or other objects.**

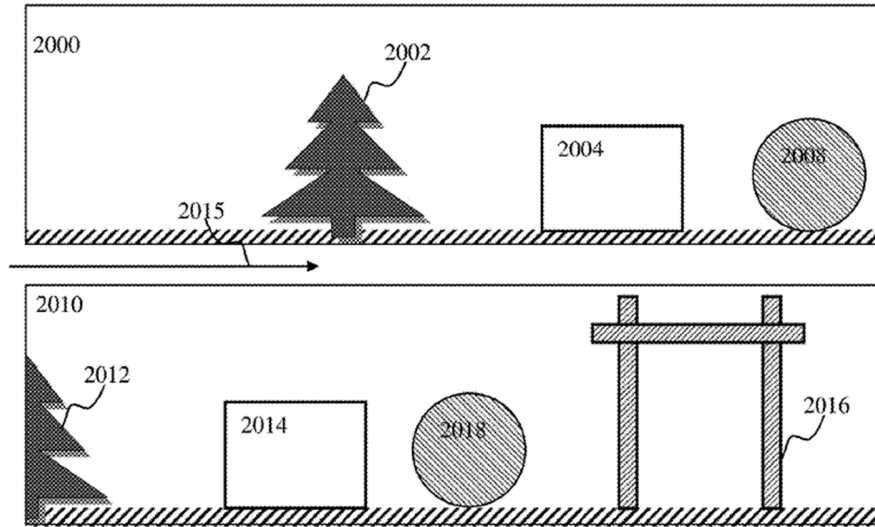


FIG. 20A

Ex. 1004, 38:52-58, Fig. 20A.

224. 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.” *Id.*, 8:57-9:13. *See, e.g., id.*, 9:41-52 (“In one or more implementations, object recognition and/or classification may be implemented”). Grotmol, like the ’344 patent, further includes edge detection. *Id.*, 8:60-63 (“[T]he term feature may refer to a representation of an object edge”).

225. Grotmol describes that the autonomous operation of a robot may be based in part on recognized objects and their characteristics. *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”).

- i. **[18.9] 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, multiple instructions, one or more inputs, one or more commands, one or more keywords, one or more symbols, one or more operators, one or more variables, one or more parameters, one or more characters, one or more numbers, one or more values, one or more signals, one or more binary bits, one or more functions, one or more function references, one or more objects, one or more object references, one or more data structures, one or more data structure references, one or more states, one or more representations of one or more states, one or more representations of one or more user inputs, one or more codes, one or more data, or one or more information, and**

226. As I discussed above for limitation [1.3], 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, one or more computer commands, and/or one or more information. *See id.*, 3:35-39; 34:3-11; 39:13-22; 43:4-11.

- j. **[18.10] wherein the first one or more object representations include: one object representation, ... and wherein the second one or more object representations include: one object representation, ... and wherein the first one or more object representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the second one or more object**

representations include: ... one or more digital pictures that depict one or more objects, ... and wherein the at least the portion of the first circumstance representation includes: at least a portion of one object representation ... and

227. As I discussed above for limitation [1.3] and claim 3, Grotmol discloses a first circumstance representation includes an object representation, and a second circumstance representation includes an object representation, and also discloses the first and second circumstance representations include digital pictures that depict one or more objects.

- k. [18.11] wherein the at least the portion of the first one or more instruction sets for operating the first device includes: at least a portion of one instruction set of the first one or more instruction sets for operating the first device ...**

228. As I discussed above for limitations [1.3] and [1.4], Grotmol discloses motor commands (i.e., at least a portion of the first one or more instruction sets for operating the first device) are learned in a learning process that includes operating the first device at least partially by a user. Accordingly, Grotmol discloses wherein the at least the portion of the first one or more instruction sets for the operating the first device includes at least a portion of one instruction of the first one or more instruction sets for operating the first device.

- l. [18.12] and wherein: an object of the first circumstance is the same as an object of the second circumstance, multiple objects of the first circumstance are the same as multiple objects of the**

second circumstance, all objects of the first circumstance are the same as all objects of the second circumstance, or all objects of the first circumstance are different than all objects of the second circumstance.

229. As described above with respect to limitations [1.6] and [18.8], Grotmol discloses comparing a first circumstance and second circumstance to determine whether there is at least a partial match, and further discloses that the first and second circumstance include one or more objects, as well as object detection. The object(s) in the first circumstance and the second circumstance can be the same or different. *See, e.g., id.*, 38:52-39:12, Fig. 20A, Fig. 20B. Grotmol further discloses that “the similarity [between the images] 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.” The images may have one object, multiple or all objects that are the same, or have all different objects.

15. Claim 19

230. Independent claim 19 is a method claim that corresponds to system claim 1. My analysis above explaining how Grotmol discloses the limitations of claim 1 applies to the limitations of claim 19 as set forth in the chart below.

Claim 19	Corresponding Analysis for Claim 1
[19.1] A method implemented using a computing system that includes one or more processors, the method comprising:	<i>See</i> analysis for limitations [1.1] and [1.2] above.

Claim 19	Corresponding Analysis for Claim 1
[19.2] accessing a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device,	<i>See analysis for limitation [1.3] above.</i>
[19.3] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user;	<i>See analysis for limitation [1.4] above.</i>
[19.4] generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;	<i>See analysis for limitation [1.5] above.</i>
[19.5] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation; and	<i>See analysis for limitation [1.6] above.</i>

Claim 19	Corresponding Analysis for Claim 1
[19.6] at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.	<i>See analysis for limitation [1.7] above.</i>

16. Claim 20

231. Independent claim 20 is a system claim that corresponds to claim 1, except that claim 20 requires “means for processing” whereas claim 1 requires “one or more processors.” As I explained in the claim construction section, a corresponding structure for “means for processing” is “one or more processor circuits.” For the remainder of claim 20, my analysis above explaining how Grotmol discloses the limitations of claim 1 applies to the limitations of claim 20 as set forth in the chart below:

Claim 20	Corresponding Analysis for Claim 1
[20.1] A system comprising:	<i>See analysis for limitation [1.1] above.</i>
[20.2] means for processing; and	<i>See Section VII.B and limitation [1.2] above.</i>

Claim 20	Corresponding Analysis for Claim 1
[20.3] a memory that stores at least a knowledgebase that includes a first circumstance representation correlated with a first one or more instruction sets for operating a first device, wherein the first circumstance representation represents a first circumstance detected at least in part by one or more sensors of the first device,	<i>See analysis for limitations [1.3] above.</i>
[20.4] and wherein at least a portion of the first circumstance representation or at least a portion of the first one or more instruction sets for operating the first device is learned in a learning process that includes operating the first device at least partially by a user, and	<i>See analysis for limitations [1.4] above.</i>
[20.5] wherein the means for processing is configured to perform at least: generating or receiving a second circumstance representation, wherein the second circumstance representation represents a second circumstance detected at least in part by: the one or more sensors of the first device, or one or more sensors of a second device;	<i>See analysis for limitation [1.5] above.</i>

Claim 20	Corresponding Analysis for Claim 1
[20.6] anticipating the first one or more instruction sets for operating the first device based on at least partial match between the second circumstance representation and the first circumstance representation; and	<i>See analysis for limitation [1.6] above.</i>
[20.7] at least in response to the anticipating, executing the first one or more instruction sets for operating the first device, wherein the first device or the second device autonomously performs one or more operations defined by the first one or more instruction sets for operating the first device.	<i>See analysis for limitation [1.7] above.</i>

C. Ground 2: Claims 10 and 11 Are Obvious Based on Grotmol in View of the Knowledge of a POSITA

232. As set forth below, claims 10 and 11 of the '344 patent are rendered obvious by Grotmol in view of the knowledge of a POSITA.

1. **Claim 10: The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating the first device, and wherein the third circumstance representation represents a third circumstance detected at least in part by the one or more sensors of the first device, and wherein at least a portion of the third circumstance representation or at least a portion of the second one or more instruction sets for operating the**

first device is learned in another learning process that includes operating the first device at least partially by another user.

Claim 10	Additional Features to Claim 3
	+ 3rd CR (detected by 1st Device) correlated with 2nd instructions (for 1st Device) learned with 2nd user

233. As I discussed above in Ground 1, Grotmol discloses that a human user can operate the robot during a training process for learning correlations between circumstance representations and instruction sets:

In listen mode, the robot may be “listening” to or otherwise observing what the teacher teaches, and may not do anything on its own. However, the robot may learn an association.

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. Additional 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.

Ex. 1004, 23:3-6, 39:13-22. Grotmol further discloses that multiple learning processes can be done to train multiple behaviors, including “simple tasks” and “complex tasks.” *Id.*, 24:17-67 (detailing methodologies for training simple and complex tasks); 26:1-20 (detailing process for learning a “second behavior from the

user”). Grotmol does not place any limit on the number of users that can train the multiple behaviors, and a POSITA would have understood that different users can train different behaviors, particularly given that Grotmol describes the training of both simple and complex tasks. A POSITA would have understood that different users can possess different skills and knowledge relating to the functionality and operation of a robot. Therefore, it would have been obvious to a POSITA that different trainers could be used to train behaviors involving different levels of complexity in order to take advantage of the varying skills and knowledge different trainers possess. In my opinion, there is no reason a POSITA would have understood Grotmol’s disclosure to be limited to a single user training all possible tasks, especially due to the inefficiency this limitation would impose on learning across all interoperable robots.

2. **Claim 11: The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating a third device, and wherein the third circumstance representation represents a third circumstance detected at least in part by one or more sensors of the third device, and wherein at least a portion of the third circumstance representation or at least a portion of the second one or more instruction sets for operating the third device is learned in another learning process that**

includes operating the third device at least partially by: the user, or another user.

Claim 11	Additional Features to Claim 3
	+ 3rd CR (detected by 3rd Device) correlated with 2nd instructions (for 3rd Device) learned with 1st user or 2nd user

234. Grotmol in view of the knowledge of a POSITA renders obvious claim 11, which requires: (i) learning a correlation between a circumstance representation detected by a sensor of another (third) device and an instruction set for operating the third device; and (ii) the correlation is learned during an additional learning process that involves operating the other (third) device by the same user or another user.

235. As I discussed above with respect to limitation 1.5, Grotmol discloses that a robot can receive images correlated with instructions sets (training sets) from other robots. Ex. 1004, 33:27-32 (“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. In some implementations, *the trained configuration may be loaded to one or more other robots* in order to provide learned behaviors.”). A POSITA would understand “one *or more*” other robots to encompass a third device. Thus, a POSITA would have understood Grotmol to disclose that a third circumstance representation can be detected by a camera of a third robot trained

by a user to learn instruction sets correlated with that third circumstance representation.

236. Further, as described above for claim 10, Grotmol discloses that multiple learning processes can be done to train multiple behaviors, and that it would be obvious that the multiple learning processes could be done by different users operating a robot. For the same reasons discussed with respect to claim 10 for a single robot, it would be obvious to a POSITA that the training process involving operation of the first robot (resulting in the first trained configuration), and the training process involving operation of the third robot (resulting in another trained configuration) could be performed by different users.

D. Ground 3: Claims 6, 11, and 12 Are Obvious Based on Grotmol in View of Hickman

237. As set forth in detail below, claims 6, 11, and 12 of the '344 patent are separately rendered obvious by Grotmol in view of Hickman.

1. Motivation to Combine Grotmol and Hickman

238. For the same reasons as I discussed above in Section VIII.D.1 above, a POSITA would have been motivated to combine the autonomous robotic system of Grotmol with the shared robot knowledge base of Hickman, which would allow for learning across multiple devices with multiple users, increasing efficiency and conserving local system resources.

2. **Claim 6: The system of claim 3, wherein the system further comprising: a server that receives from the first device at least one of: the first circumstance representation, or the first one or more instruction sets for operating the first device, and wherein the second device receives from the server at least one of: the first circumstance representation, or the first one or more instruction sets for operating the first device; and wherein the second device autonomously performs the one or more operations defined by the first one or more instruction sets for operating the first device.**

Claim 6	Additional Features to Claim 3
	+ server receives 1st CR or 1st instructions from 1st device + server sends 1st CR or 1st instructions to 2nd device + 2nd device autonomously performs 1st instructions

239. As described above in Ground 1, Grotmol discloses: (1) that circumstance representations are detected by a sensor of a first device and are correlated with an instruction set for operating a first device; and (2) that training sets may be loaded into the robot used for training or may be loaded to another robot. The combination of Grotmol and Hickman renders obvious the limitations that a server receives a first circumstance representation or first one or more instruction sets for operating a first device, and a second device receiving from the server a first circumstance representation or first one or more instruction sets for operating the first device.

240. Grotmol discloses that training sets and training configurations, which include pictures and associated instruction sets for operating the first device, can be

stored or saved, and then loaded into the robot to provide the trained behavior Ex. 1004, 33:27-30 (“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.”). Grotmol explains that the saved training configuration can also be loaded into robots other than the one that did the learning. *Id.*, 33:30-32 (“[T]he trained configuration may be loaded to *one or more other robots* in order to provide learned behaviors”). Thus, a POSITA would have understood Grotmol to disclose both the storing of a circumstance representation or first one or more instruction sets for operating a first device, and the receipt of the first circumstance representation or the first one or more instruction sets for operating the first device by a second device.

241. Grotmol discloses, for example, situations in which the user may wish to train the robot to “find an object, grasp it, and bring it back” (i.e., fetch). Ex. 1004, 24:31-25:5. Grotmol describes that this “task of playing fetch may be quite complex” and outlines numerous user training steps that would be needed in order to teach this skill to the robot. *Id.* In my opinion, it would be desirable to implement a system wherein learned robot behaviors could be uploaded to a shared knowledge base for use in numerous other robots around the world, such as the system described in Hickman.

242. Hickman discloses a “ shared robot knowledge bases for use with cloud computing systems” that “collects data from a robot about an object the robot has encountered in its environment, and stores the received data in the shared robot knowledge base ... sends instructions for interacting with an object to a robot ... sends instructions to a robot for executing an application based on information stored in the shared robot knowledge base.” Ex. 1005, Abstract. Hickman discloses using a cloud computing system, where “[t]he ‘cloud’ may refer to a service or a group of service accessible over a network (e.g., the Internet) by clients, *server devices*, and cloud computing systems”*Id.*, 5:20-. Moreover, “[t]he cloud computing system may be configured to (i) send and receive data and queries to and from multiple robots, (ii) send data retrieved from the shared robot knowledge base to robots, and (iii) update data in the shared robot knowledge base based on data received from robots.” *Id.*, 5:7-11; *see also* Fig. 6 (below).

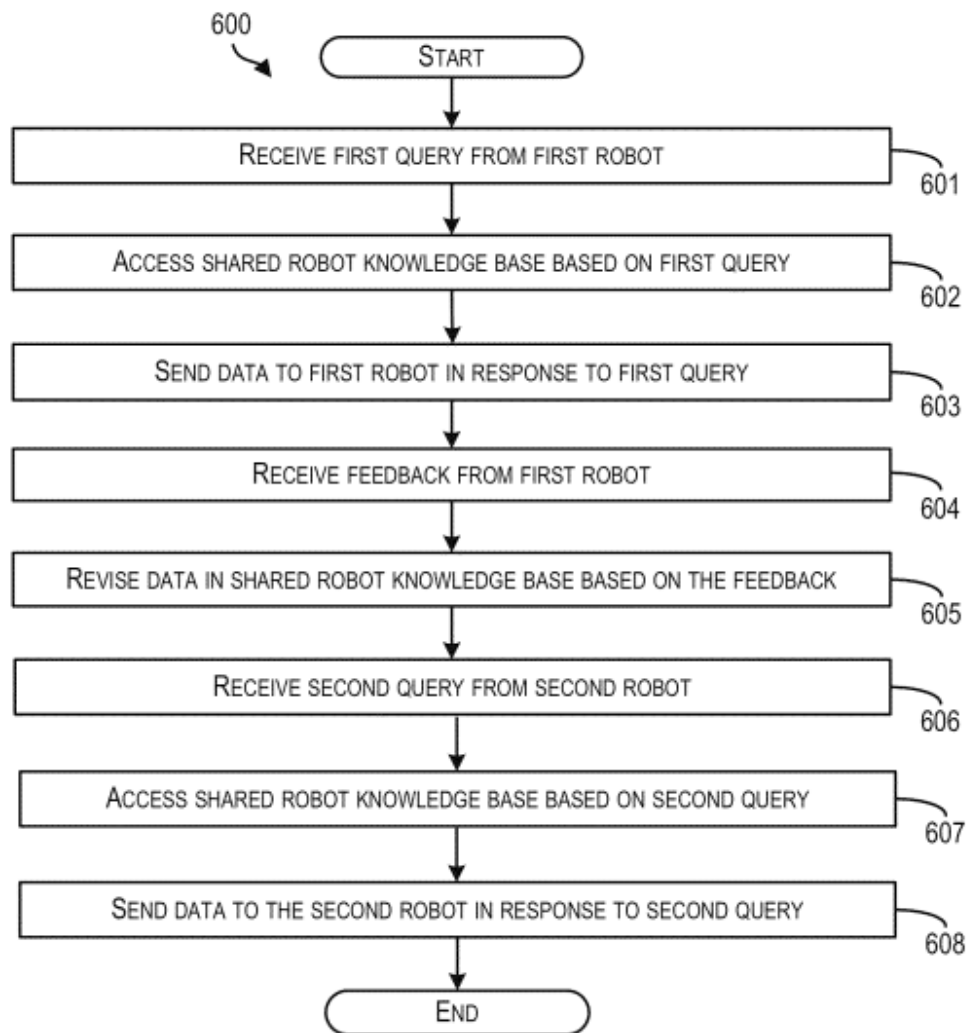


FIGURE 6

243. Hickman notes that the robot shared knowledge base using the cloud computing system can update based on individual robot experiences. *Id.*, Abstract (“[I]nformation in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other

robots.”). For example, Hickman describes a robot that is trained to properly grasp a cup uploading and storing the information it learned onto the shared knowledge base, and then having another robot obtain these instructions from the memory and executing the instructions:

In the example shown in FIG. 4, the robot 404 may be performing an application that includes moving the cup 405 from one location to another location ...

In this example interaction, the grasping force specified in the instructions was too high, and as a result, the robot crushed the cup and caused a mess. When an accident happens, **a human can “coach” the robot on how much force to use** by, for example, manually controlling the robot's hand to grasp a new cup. The robot can capture the grasping force that it used to successfully grasp the cup under the manual control of the human, and then **send feedback 411 to the cloud processing system 402 with the modified grasping force**. The cloud processing system 402 can then update 412 task and object data in the shared robot knowledge base 403 to improve how the “grasp” task is applied this particular “cup” object.

Later, a second robot 413 may encounter an object 414 while performing the same (or a similar) application that requires grasping the same type of cup ...

After determining the object is a cup 414 and obtaining information about the cup 414, the cloud computing system 401 can send object data 418 associated with the cup 414 to the robot 413 in response to the identification query 415 received from the robot 413. The object data 418 may include both the identity of the cup 414 and instructions for interacting with the cup 414 based on the application or task that the robot 413 has been instructed to execute or perform ...

Importantly, the cloud computing system 401 updated the amount of force to use when grasping the particular type of cup that the robot 413 must now grasp (i.e., cup 414) based on the

feedback 411 previously received from robot 404 regarding the same type of cup. As a result, when the robot 413 applies the “grasp” task to the type of “cup” object here (i.e., when robot 413 grasps cup 414), the robot 413 successfully grasps the cup without crushing the cup 419. Thus, **the second robot 413 has in effect learned from the experience of the earlier robot 404.**

Ex. 1005, 26:29-27:21. Hickman teaches that when a robot queries the cloud processing engine, “the cloud computing system sends data associated with the object of the second query to the second robot in response to the second query received from the second robot at block 606. The data may include ... instructions for interacting with the object of the second query.” *Id.*, 29:33-41. Accordingly, a POSITA would have understood Hickman to disclose a second device autonomously performing the one or more operations defined by the first one or more instruction sets.

244. Thus, A POSITA would have understood that in the combination of Grotmol and Hickman, the training configuration learned by operation of a robot described in Grotmol would be uploaded to the server containing the “shared robot knowledge base” described in Hickman, and a second robot would receive the circumstance representations and instructions stored in the knowledge base in order to autonomously perform the learned behaviors.

3. Claim 11

Claim 11	Additional Features to Claim 3
	+ 3rd CR (detected by 3rd Device) correlated with 2nd instructions (for 3rd Device) learned with 1st user or 2nd user

245. As I discussed above in Ground 1 for limitation [1.5] and claim 5, in Ground 2, and in claim 6, Grotmol discloses that multiple learning processes can be performed to train multiple behaviors and that a single robot can learn from the training of multiple robots. In my opinion, combining Hickman’s shared robot knowledgebase with these disclosures from Grotmol renders obvious that the knowledgebase of claim 1 contains training configurations from two different learning processes, each involving a different user operating a different robot.

246. Hickman discloses a “shared robot knowledge bases for use with cloud computing systems” that “collects data from a robot about an object the robot has encountered in its environment, and stores the received data in the shared robot knowledge base ... sends instructions for interacting with an object to a robot ... sends instructions to a robot for executing an application based on information stored in the shared robot knowledge base.” Ex. 1005, Abstract. Hickman further teaches that the “information in the shared robot knowledge bases is updated based on robot experiences so that any particular robot may benefit from prior experiences of other

robots.” *Id.* Hickman discloses that these prior “robot experiences” may include human coaching:

By analyzing past experiences from many different robots, the cloud computing system 304 may be configured to synthesize certain rules for robots to follow when executing the “Groceries” application. For example, some rules that could be synthesized or generalized from a large number of robot experiences might include (i) unidentifiable frozen items should be stored in a freezer, (ii) identifiable cold items should be stored in a refrigerator, (iii) certain types of vegetables should be stored in a certain refrigerator drawer, (iv) items should not be placed on top of bread in a pantry, etc. In this manner, **robot experiences (including “coaching” from humans)** can be used to improve the way that robots execute the “Groceries” application.

Ex. 1005, 23:36-48. A POSITA would have appreciated the similarity between this human coaching and the human training described in Grotmol. For example, similar to the “fetch” behavior described in Grotmol, Hickman describes a human coach training a robot how to properly grasp a cup, uploading and storing this information on a shared knowledge base, and then having another robot obtain these instructions from the memory and executing the instructions:

In the example shown in FIG. 4, the robot 404 may be performing an application that includes moving the cup 405 from one location to another location ...

In this example interaction, the grasping force specified in the instructions was too high, and as a result, the robot crushed the cup and caused a mess. When an accident happens, **a human can “coach” the robot on how much force to use** by, for example, manually controlling the robot's hand to grasp a new cup. The robot can capture the grasping force that it used to successfully grasp the cup under the manual control of the human, and then

send feedback 411 to the cloud processing system 402 with the modified grasping force. The cloud processing system 402 can then update 412 task and object data in the shared robot knowledge base 403 to improve how the “grasp” task is applied this particular “cup” object.

Later, a second robot 413 may encounter an object 414 while performing the same (or a similar) application that requires grasping the same type of cup ...

After determining the object is a cup 414 and obtaining information about the cup 414, the cloud computing system 401 can send object data 418 associated with the cup 414 to the robot 413 in response to the identification query 415 received from the robot 413. The object data 418 may include both the identity of the cup 414 and instructions for interacting with the cup 414 based on the application or task that the robot 413 has been instructed to execute or perform ...

Importantly, the cloud computing system 401 updated the amount of force to use when grasping the particular type of cup that the robot 413 must now grasp (i.e., cup 414) based on the feedback 411 previously received from robot 404 regarding the same type of cup. As a result, when the robot 413 applies the “grasp” task to the type of “cup” object here (i.e., when robot 413 grasps cup 414), the robot 413 successfully grasps the cup without crushing the cup 419. Thus, **the second robot 413 has in effect learned from the experience of the earlier robot 404.**

Ex. 1005, 26:29-27:21.

247. In my opinion, it would have been obvious to combine Hickman’s shared robot knowledge base with Grotmol’s autonomous robot system such that the shared knowledgebase includes the training of multiple robots by multiple users. The combination discloses a knowledgebase that includes a circumstance representation correlated with instructions for operating a robot (i.e., claimed first circumstance

correlated with first instruction set) learned during a training using the robot, and also a circumstance representation correlated with instructions for operating a different robot (i.e., claimed third circumstance representation correlated with second instruction set) learned during a training using a different robot.

4. Claim 12

Claim 12	Additional Features to Claim 3
	+ 3rd CR (detected by 3rd Device) correlated with 2nd Instructions (for 3rd Device) + 4th CR (detected by 4th Device) + anticipation of 2nd Instructions by matching 3rd CR & 4th CR + autonomous performance of 2nd instructions on 4th Device

- a. [12.1] The system of claim 3, wherein the knowledgebase further includes a third circumstance representation correlated with a second one or more instruction sets for operating a third device, and wherein the third circumstance representation represents a third circumstance detected at least in part by one or more sensors of the third device, and wherein the one or more processors are further configured to perform at least:

Limitation	Claimed Features
[12.1]	3rd CR (detected by 3rd Device) correlated with 2nd Instructions (for 3rd Device)

248. Limitation [12.1] includes the same limitations regarding the knowledgebase including an additional circumstance representation correlated with

an instruction set as claim 11. Accordingly, my opinions as to why Grotmol in combination with Hickman renders obvious limitation [12.1] are the same as the reasons I discussed above for claim 11.

- b. **[12.2] generating or receiving a fourth circumstance representation, wherein the fourth circumstance representation represents a fourth circumstance detected at least in part by one or more sensors of a fourth device; anticipating the second one or more instruction sets for operating the third device based on at least partial match between the fourth circumstance representation and the third circumstance representation; and at least in response to the anticipating the second one or more instruction sets for operating the third device, executing the second one or more instruction sets for operating the third device, wherein the fourth device autonomously performs one or more operations defined by the second one or more instruction sets for operating the third device;**

Limitation	Claimed Features
[12.2]	• 4th CR (detected by 4th Device) • anticipation of 2nd Instructions by matching 3rd CR & 4th CR • autonomous performance of 2nd instructions on 4th Device

249. As I discussed above for claim 11 and limitation [12.1], a POSITA would have understood the combination of Grotmol and Hickman to render obvious a knowledgebase including the third circumstance representation correlated with the second instruction set. Grotmol and Hickman further disclose that the

knowledgebase can be shared to perform autonomous operation of any number of other devices, with no explicitly defined limit. Ex. 1004, 33:27-32; Ex. 1005, 11:31-41. Grotmol teaches that the autonomous operation of a robot is performed by acquiring an image (circumstance representation) during operation of the robot, comparing the newly acquired image to images already stored in the knowledgebase to determine a partial match, and then executing an operation defined by the motor command (instruction sets) correlated with the matched image. Ex. 1004, Abstract, 39:13-41.

250. Thus, when combined with the shared robot knowledgebase of Hickman, the combination discloses the generation and storage of a plurality of circumstance representations correlated with corresponding instructions, each learned during a training process involving different robots and then uploaded to the shared knowledgebase to allow loading of the learned training sets into all other robots. To illustrate within the language of the '344 patent, the combination's shared knowledgebase would naturally encompass a first circumstance representation detected by a first device correlated with a first instruction set for operating the first device, and a third circumstance representation detected by a third device correlated with a second instruction set for operating the third device. A fourth device connected to the shared robot knowledgebase could then be loaded with the training sets from both the first and third devices before encountering a new, fourth

circumstance representation detected by its sensors. If that fourth circumstance representation partially matches the third circumstance representation, then the fourth device could anticipate the second instruction set learned on the third device, execute those instructions, and autonomously perform the operations defined by those instructions. In effect, any additional device could compare newly acquired images with previously acquired images from any other robot, and possibly make use of the learned behaviors corresponding to those stored images in the event of a match.

251. Thus, in my opinion, the combination of Grotmol and Hickman renders obvious anticipating an instruction for operating a fourth device based on a partial match between a circumstance representation detected by sensors of a fourth device and a third circumstance representation stored in the knowledgebase, and then executing operations defined by an instruction set correlated with the third circumstance representation.

X. Declaration

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

Dated: 6/9/23

By: 
Jason Janet