UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
62/798,978	01/30/2019		140	35447-42479/US		

CONFIRMATION NO. 6757

FILING RECEIPT

758
FENWICK & WEST LLP
SILICON VALLEY CENTER
801 CALIFORNIA STREET
MOUNTAIN VIEW, CA 94041

OC000000106011051

Date Mailed: 02/20/2019

Receipt is acknowledged of this provisional patent application. It will not be examined for patentability and will become abandoned not later than twelve months after its filing date. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Sam Anthony, Cambridge, MA;

Applicant(s)

Perceptive Automata, Inc., Cambridge, MA;

Power of Attorney: The patent practitioners associated with Customer Number 00758

Permission to Access Application via Priority Document Exchange: Yes

Permission to Access Search Results: Yes

Applicant may provide or rescind an authorization for access using Form PTO/SB/39 or Form PTO/SB/69 as appropriate.

If Required, Foreign Filing License Granted: 02/19/2019

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 62/798,978**

Projected Publication Date: None, application is not eligible for pre-grant publication

Non-Publication Request: No

Early Publication Request: No
**** SMALL ENTITY ****

Title

NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as

set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

Doc Code: PA..
Document Description: Power of Attorney

PTO/AIA/82A (07-13)
Approved for use through 01/31/2018. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FOR POWER OF ATTORNEY TO ONE OR MORE REGISTERED PRACTITIONERS

NOTE: This form is to be submitted with the Power of Attorney by Applicant form (PTO/AIA/82B) to identify the application to which the Power of Attorney is directed, in accordance with 37 CFR 1.5, unless the application number and filing date are identified in the Power of Attorney by Applicant form. If neither form PTO/AIA/82A nor form PTO/AIA82B identifies the application to which the Power of Attorney is directed, the Power of Attorney will not be recognized in the application.

Application Number	Not yet assigned
Filing Date	Herewith
First Named Inventor	Sam Anthony
Title	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES
Art Unit	n/a
Examiner Name	n/a
Attorney Docket Number	35447-42479/US

SIGNATURE of Applicant or Patent Practitioner

Signature	/Rajendra B. Panwar/	Date (Optional)	January 30, 2019
Name	Rajendra B. Panwar	Registration Number	63,165
Title (if Applicant is a juristic entity)	Attorney for Applicant		
Applicant Name (if Applicant is a juristic entity)	Perceptive Automata, Inc.		

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4(d) for signature requirements and certifications. If more than one applicant, use multiple forms.

☑ *Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Doc Code: PA..
Document Description: Power of Attorney

PTO/AIA/82B (07-13)
Approved for use through 01/31/2018. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

POWER OF ATTORNEY BY APPLICANT

I hereby revoke all previous powers of attorney given in the application identified in either the attached transmittal letter or the boxes below.

Application Number	Filing Date

(Note: The boxes above may be left blank if information is provided on form PTO/AIA/82A.)

☑ I hereby appoint the Patent Practitioner(s) associated with the following Customer Number as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above: 00758

OR

☐ I hereby appoint Practitioner(s) named in the attached list (form PTO/AIA/82C) as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the patent application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above. (Note: Complete form PTO/AIA/82C.)

Please recognize or change the correspondence address for the application identified in the attached transmittal letter or the boxes above to:

☑ The address associated with the above-mentioned Customer Number

OR

☐ The address associated with Customer Number:

OR

☐ Firm or Individual Name					
Address					
City		State		Zip	
Country					
Telephone		Email			

I am the Applicant (if the Applicant is a juristic entity, list the Applicant name in the box):

PERCEPTIVE AUTOMATA, INC.

☐ Inventor or Joint Inventor (title not required below)

☐ Legal Representative of a Deceased or Legally Incapacitated Inventor (title not required below)

☑ Assignee or Person to Whom the Inventor is Under an Obligation to Assign (provide signer's title if applicant is a juristic entity)

☐ Person Who Otherwise Shows Sufficient Proprietary Interest (e.g., a petition under 37 CFR 1.46(b)(2) was granted in the application or is concurrently being filed with this document) (provide signer's title if applicant is a juristic entity)

SIGNATURE of Applicant for Patent

The undersigned (whose title is supplied below) is authorized to act on behalf of the applicant (e.g., where the applicant is a juristic entity).

Signature		Date (Optional)	
Name	Kshitij "Sid" Misra		
Title	Chief Executive Officer		

NOTE: Signature - This form must be signed by the applicant in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. If more than one applicant, use multiple forms.

☑ Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 3.

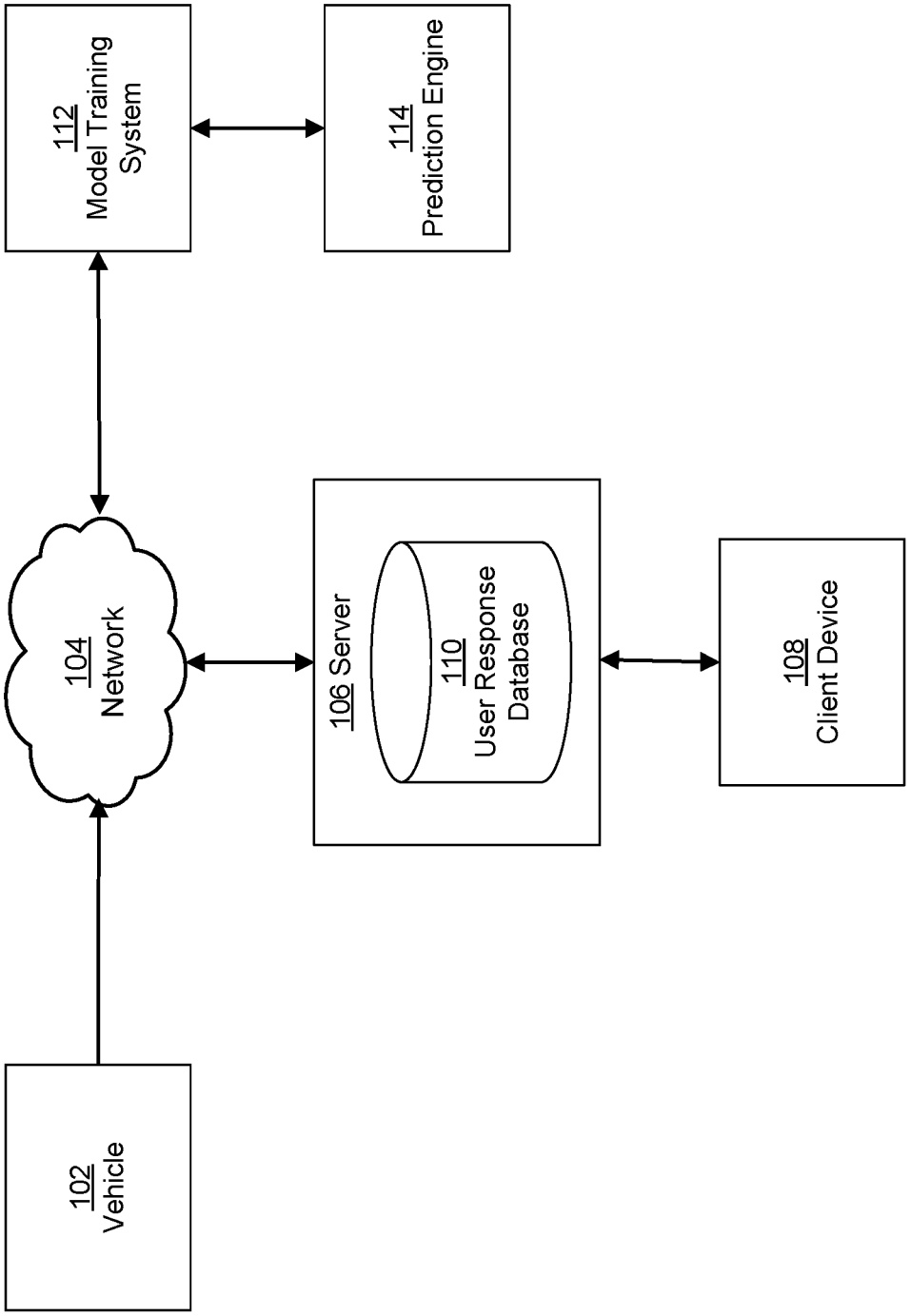

112
Model Training System
114
Prediction Engine
104
Network
106 Server
110
User Response Database
108
Client Device
102
Vehicle

FIG. 1A

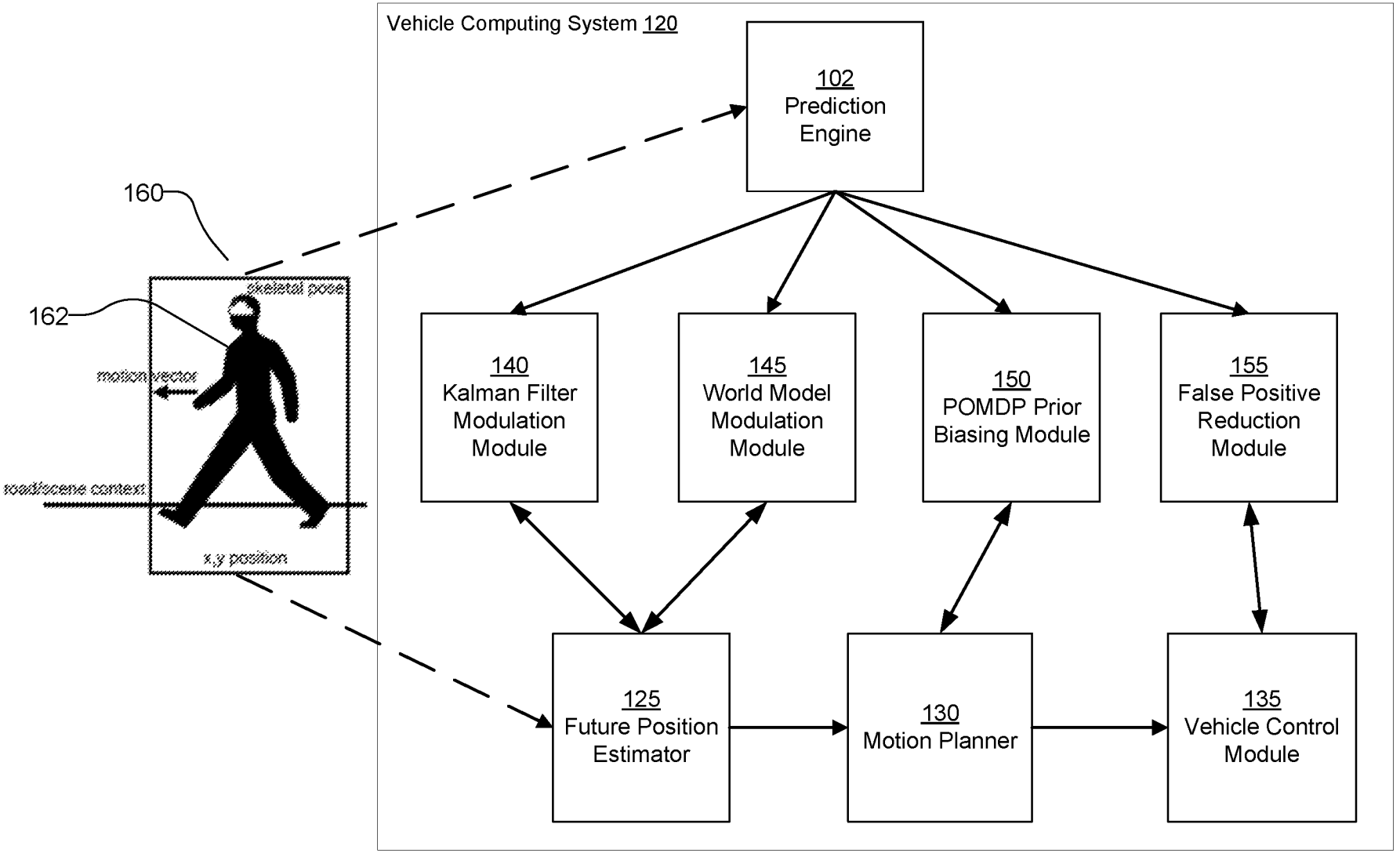


FIG. 1B

TESLA EXHIBIT 1007, Page 7

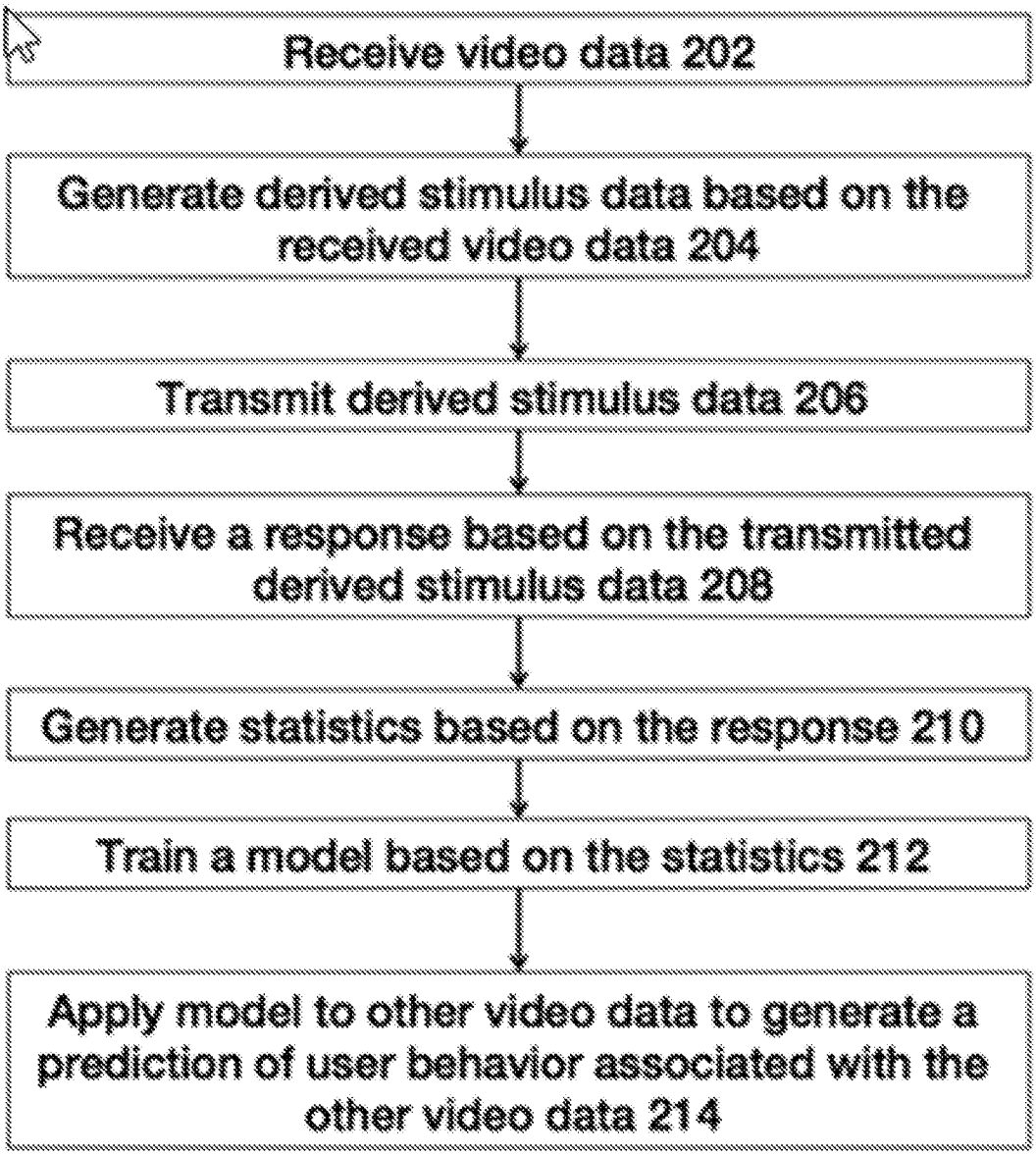

Receive video data 202
Generate derived stimulus data based on the received video data 204
Transmit derived stimulus data 206
Receive a response based on the transmitted derived stimulus data 208
Generate statistics based on the response 210
Train a model based on the statistics 212
Apply model to other video data to generate a prediction of user behavior associated with the other video data 214

FIG. 2A

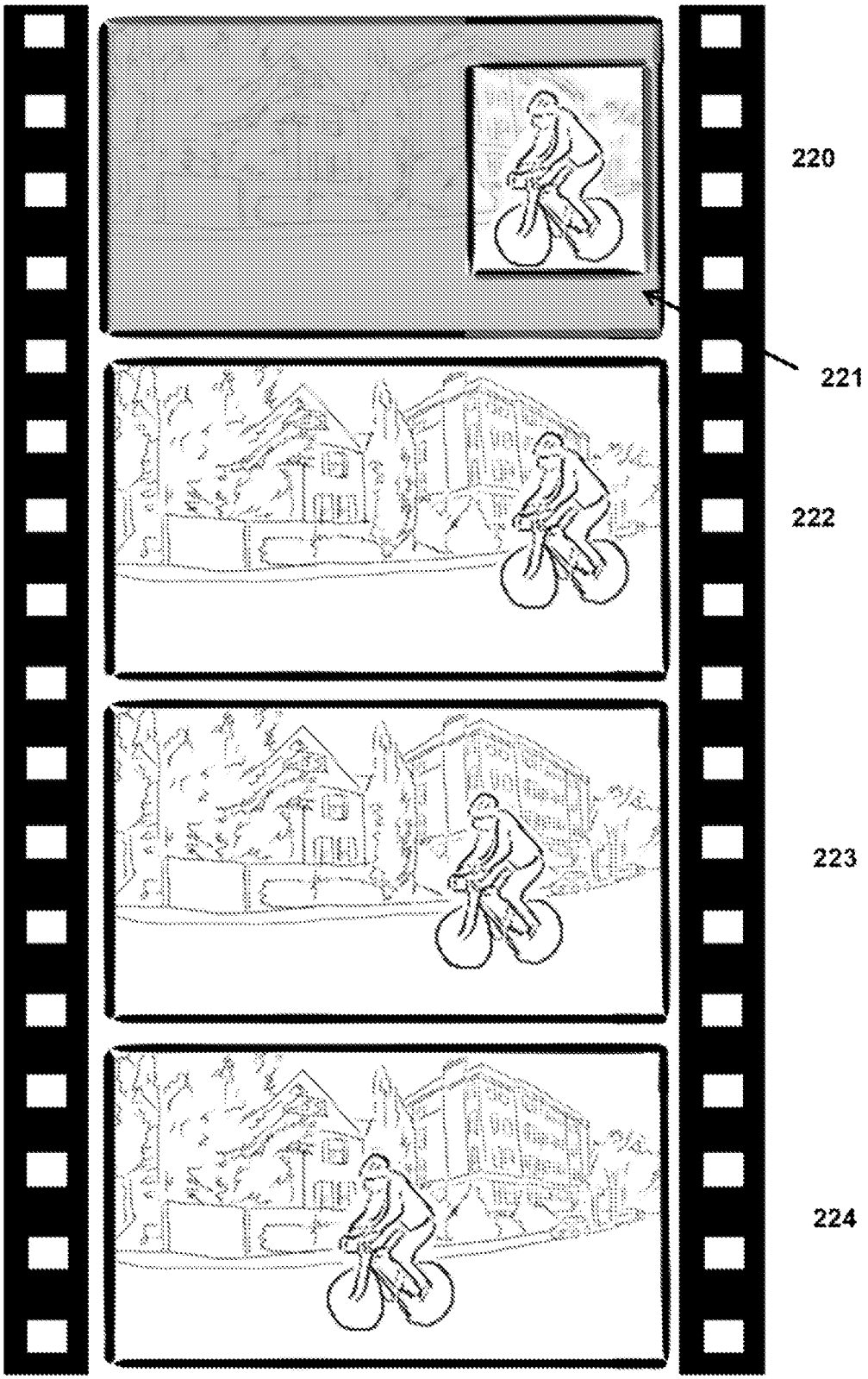


FIG. 2B

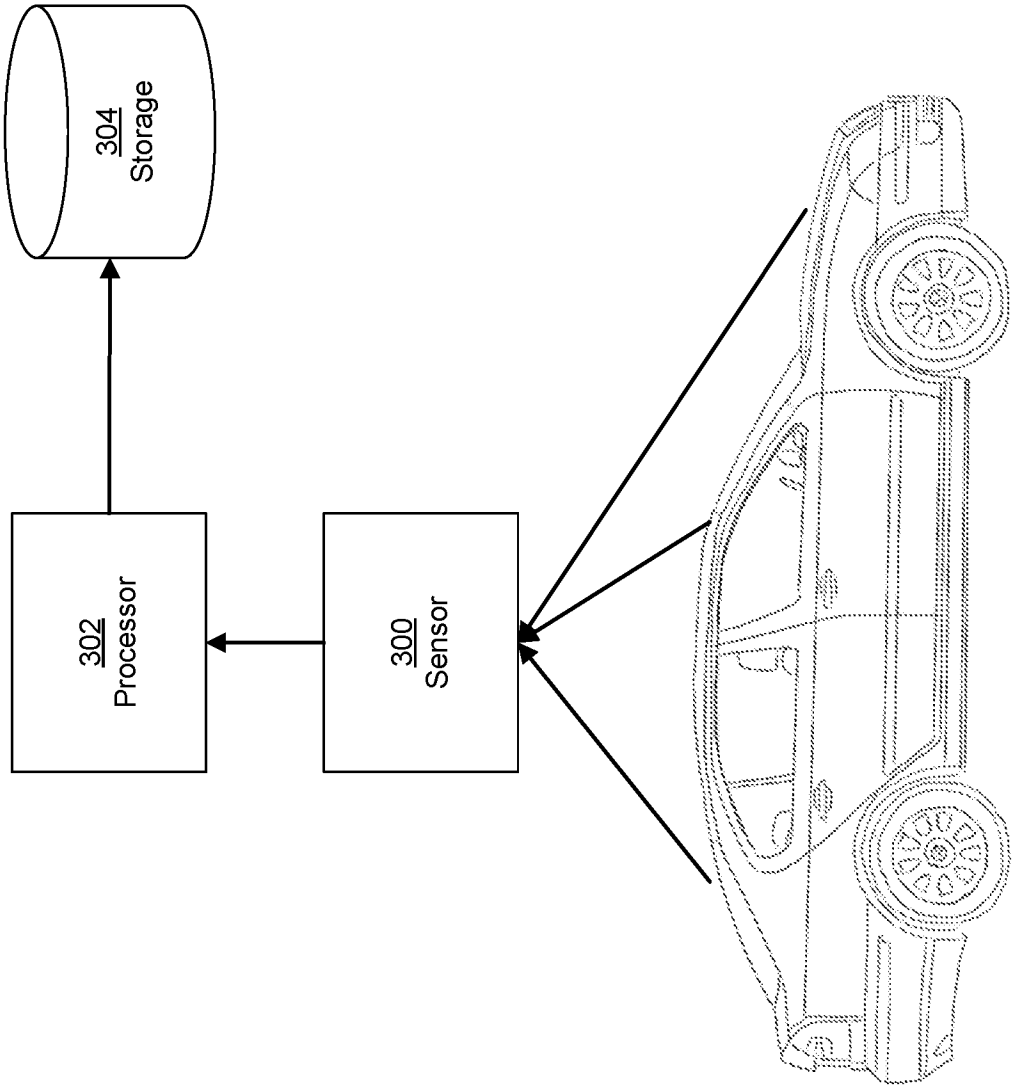


FIG. 3

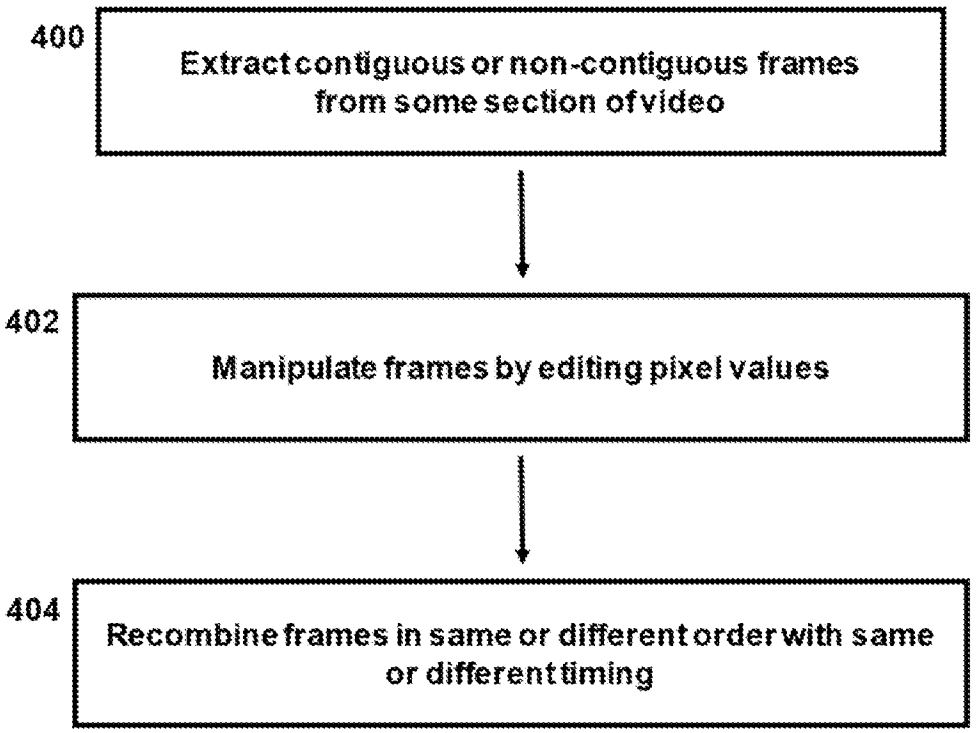

400
Extract contiguous or non-contiguous frames from some section of video
402
Manipulate frames by editing pixel values
404
Recombine frames in same or different order with same or different timing

FIG. 4

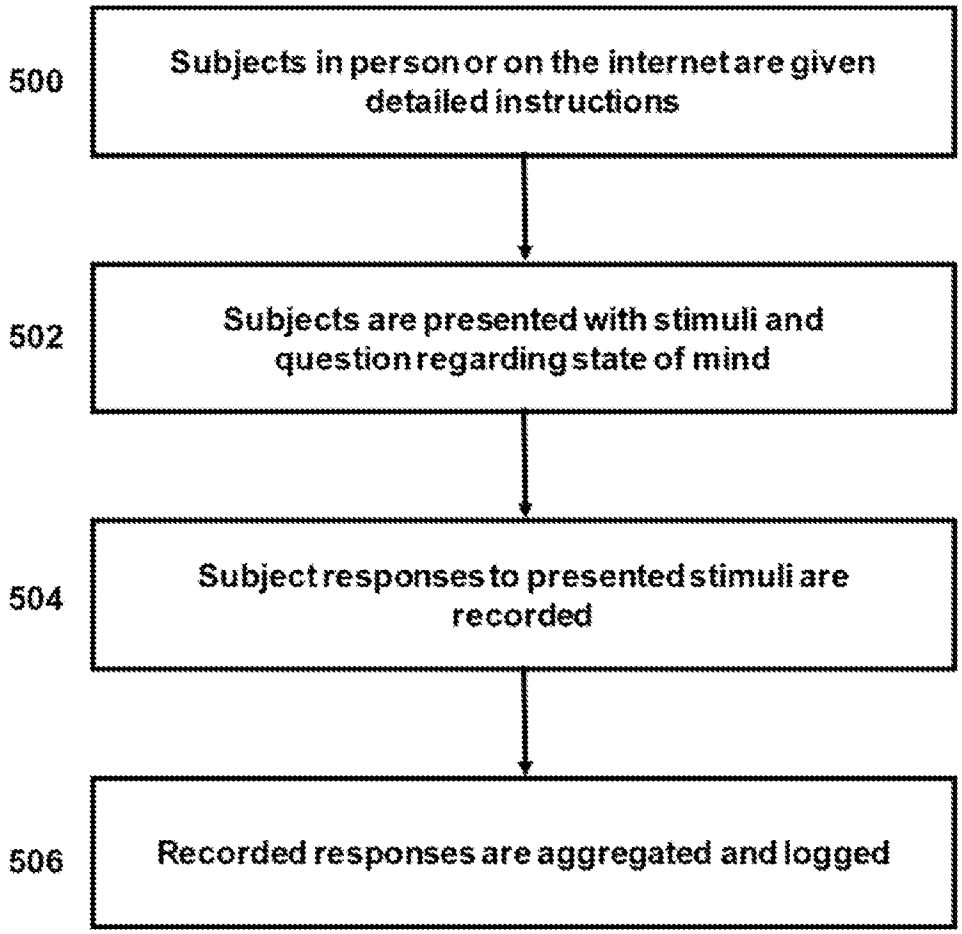

500
Subjects in person or on the internet are given detailed instructions
502
Subjects are presented with stimuli and question regarding state of mind
504
Subject responses to presented stimuli are recorded
506
Recorded responses are aggregated and logged

FIG. 5

600 602 604 606

subject_id	derived_stim	response	response_time
2329	/ims/seg114/vid44m22s23-28	3	440
2329	/ims/seg3932/vid329m24s01-16	1	600
2330	/ims/seg114/vid44m22s23-28	1	280
2330	/ims/seg118/vid51m02s44-50	4	500

FIG. 6

TESLA EXHIBIT 1007, Page 13

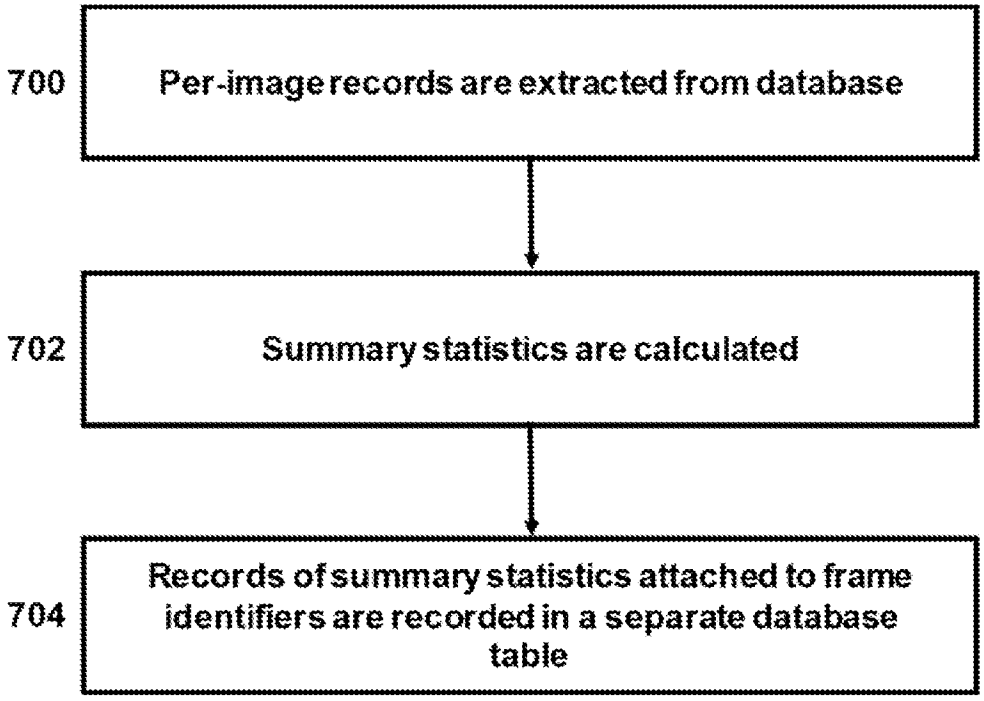

700
Per-image records are extracted from database
702
Summary statistics are calculated
704
Records of summary statistics attached to frame identifiers are recorded in a separate database table

FIG. 7

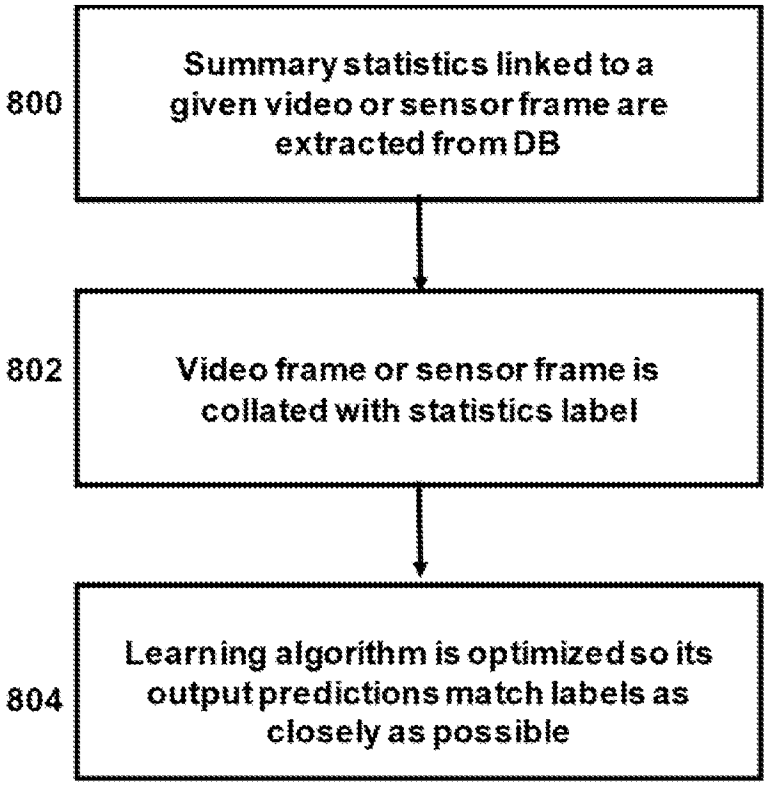

800
Summary statistics linked to a given video or sensor frame are extracted from DB
802
Video frame or sensor frame is collated with statistics label
804
Learning algorithm is optimized so its output predictions match labels as closely as possible

FIG. 8

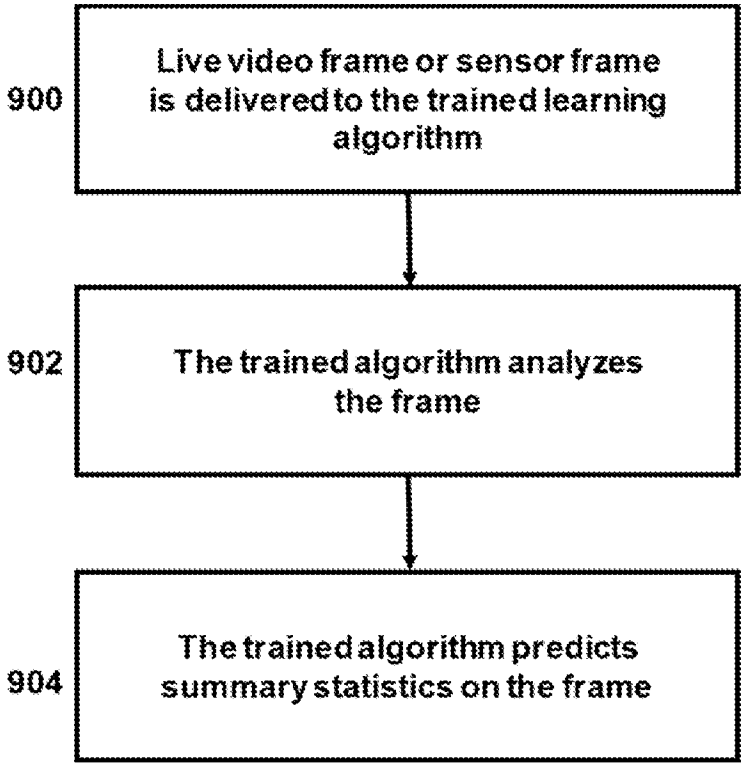

900
Live video frame or sensor frame is delivered to the trained learning algorithm
902
The trained algorithm analyzes the frame
904
The trained algorithm predicts summary statistics on the frame

FIG. 9

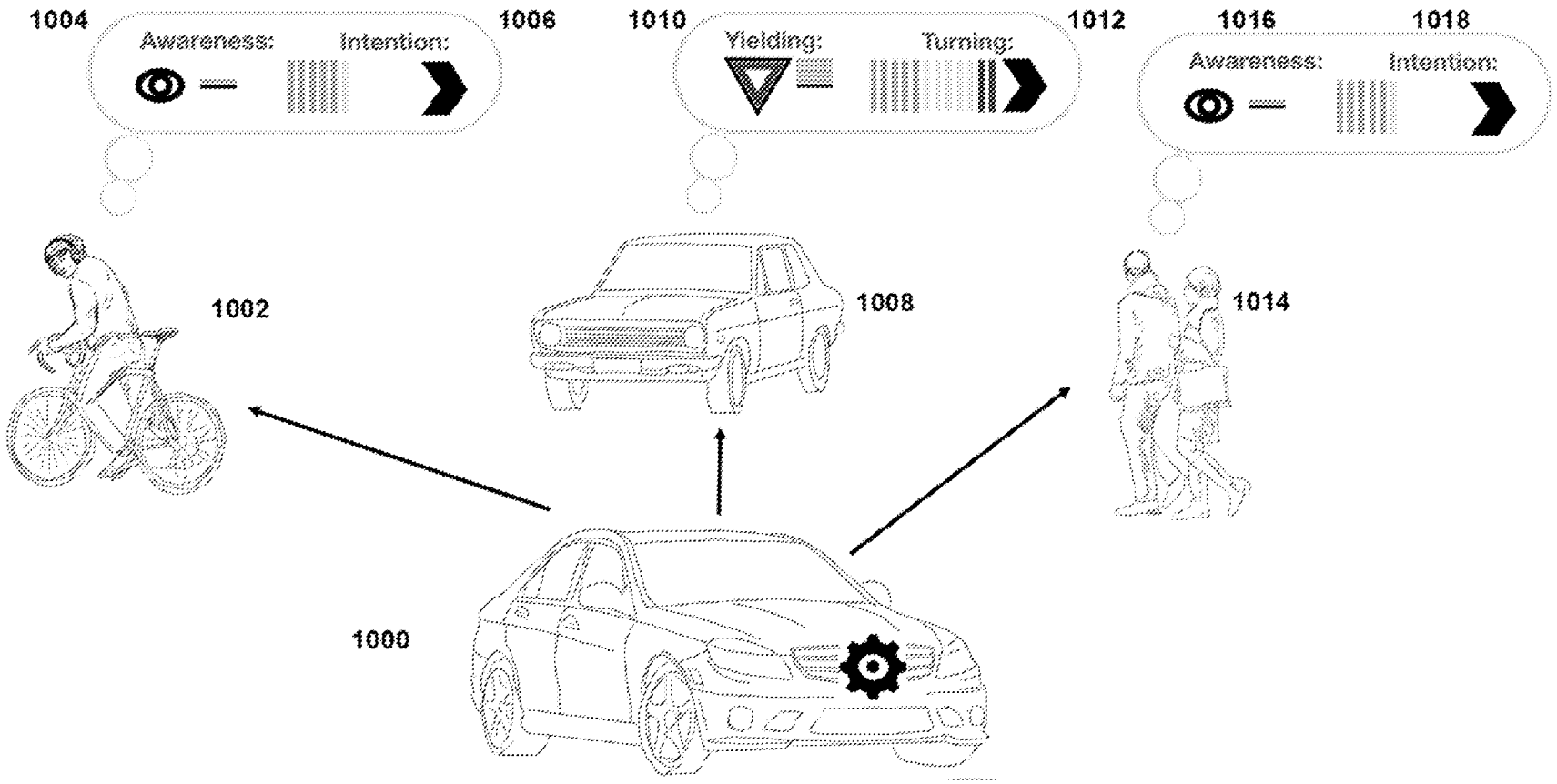


FIG. 10

TESLA EXHIBIT 1007, Page 17

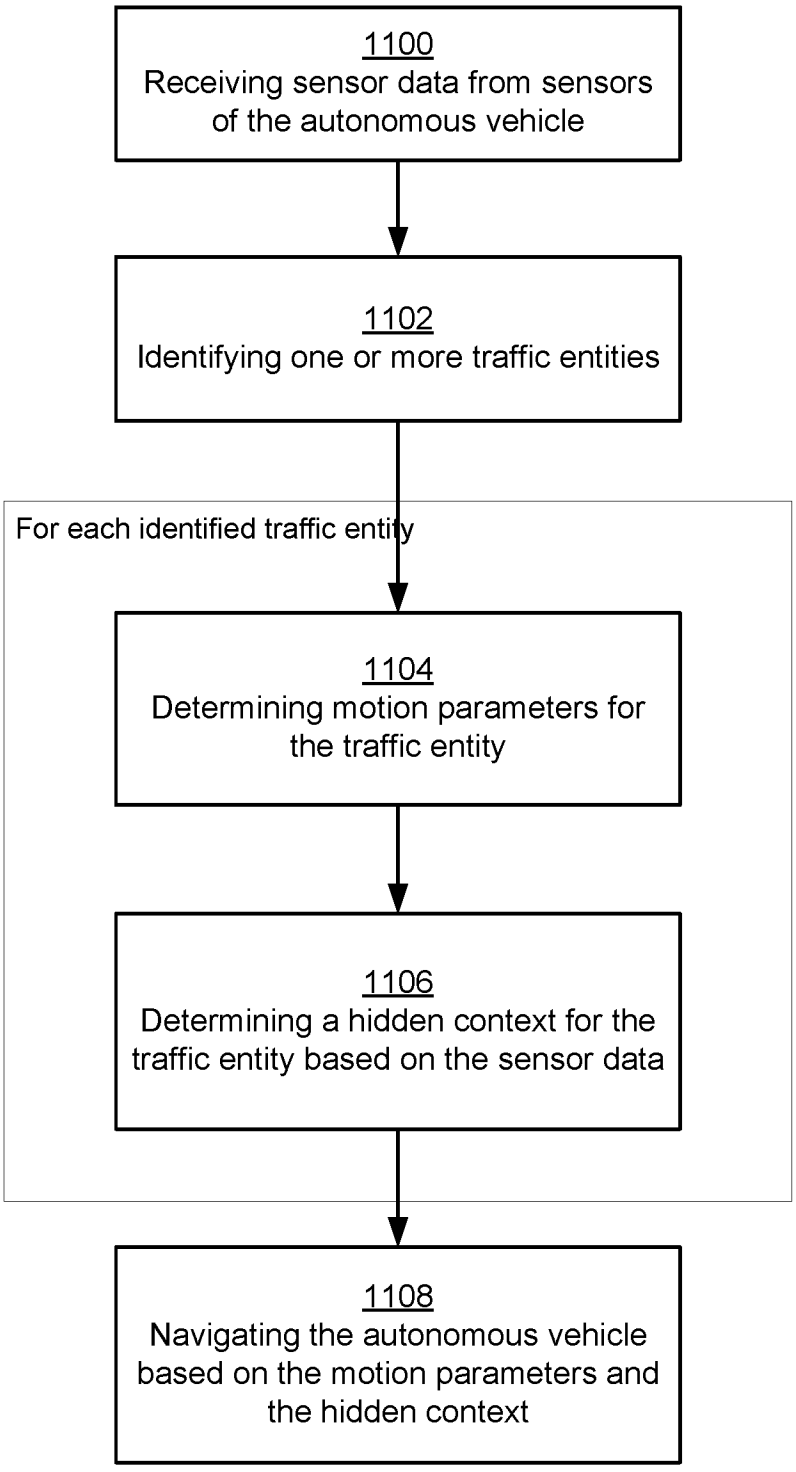

1100
Receiving sensor data from sensors of the autonomous vehicle
1102
Identifying one or more traffic entities
For each identified traffic entity
1104
Determining motion parameters for the traffic entity
1106
Determining a hidden context for the traffic entity based on the sensor data
1108
Navigating the autonomous vehicle based on the motion parameters and the hidden context

FIG. 11

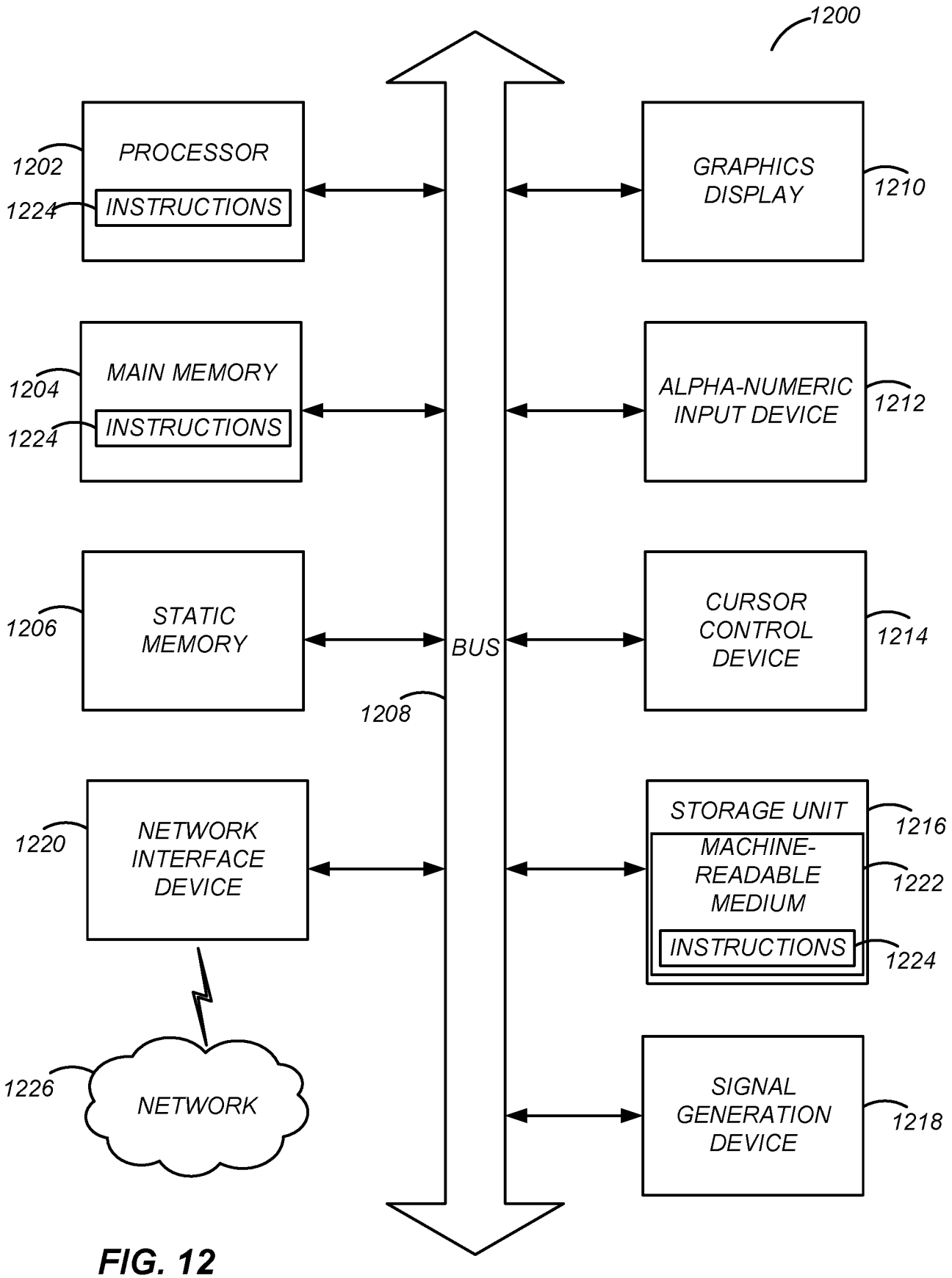

1200
PROCESSOR
1202
INSTRUCTIONS
1224
MAIN MEMORY
1204
INSTRUCTIONS
1224
STATIC MEMORY
1206
BUS
1208
NETWORK INTERFACE DEVICE
1220
NETWORK
1226
GRAPHICS DISPLAY
1210
ALPHA-NUMERIC INPUT DEVICE
1212
CURSOR CONTROL DEVICE
1214
STORAGE UNIT
1216
MACHINE-READABLE MEDIUM
1222
INSTRUCTIONS
1224
SIGNAL GENERATION DEVICE
1218

FIG. 12

Electronic Patent Application Fee Transmittal

Application Number:	
Filing Date:	
Title of Invention:	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES
First Named Inventor/Applicant Name:	Sam Anthony
Filer:	Rajendra Bhagwatisingh Panwar
Attorney Docket Number:	35447-42479/US

Filed as Small Entity

Filing Fees for Provisional

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
PROVISIONAL APPLICATION FILING FEE	2005	1	140	140
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**140**

Electronic Acknowledgement Receipt

EFS ID:	35012725
Application Number:	62798978
International Application Number:	
Confirmation Number:	6757
Title of Invention:	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES
First Named Inventor/Applicant Name:	Sam Anthony
Customer Number:	758
Filer:	Rajendra Bhagwatisingh Panwar
Filer Authorized By:	
Attorney Docket Number:	35447-42479/US
Receipt Date:	30-JAN-2019
Filing Date:	
Time Stamp:	19:58:59
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	$140
RAM confirmation Number	013119INTEFSW00017020192555
Deposit Account	192555
Authorized User	Rajendra Panwar

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.16 (National application filing, search, and examination fees)

37 CFR 1.17 (Patent application and reexamination processing fees)

37 CFR 1.19 (Document supply fees)

37 CFR 1.20 (Post Issuance fees)

37 CFR 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	42479_ADS.PDF	1255912 69d7b428c403d52e5dcfba398703a72995b8dfed	no	8
Warnings:					
Information:					
2	Power of Attorney	42479_POA.PDF	746167 068aaf0e8717bfe8dc8d5cd49404ca1ccc8b4731	no	2
Warnings:					
Information:					
3	Specification	42479_SPEC.pdf	175004 3ff6dc791fa81db95d1bd4c60d5846bbfc3bdf0e	no	36
Warnings:					
Information:					
4	Drawings-only black and white line drawings	42479_FIGS.pdf	1415860 483b7d28e267681b8268e4c3b736729765bf377f	no	14
Warnings:					
Information:					
5	Fee Worksheet (SB06)	fee-info.pdf	30046 fbae4897a443d4dd7e96dd5eb89580ac88d3f3a6	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	3622989		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1 — Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Sam		Anthony	

Residence Information (Select One) ● US Residency Non US Residency Active US Military Service

City	State/Province	Country of Residence
Cambridge	MA	US

Mailing Address of Inventor:

Address 1	Perceptive Automata, Inc.
Address 2	10 Ware Street

City	State/Province	Postal Code	Country
Cambridge	MA	02138	US

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

☐ **An Address is being provided for the correspondence Information of this application.**

Customer Number	00758		
Email Address	ptoc@fenwick.com	Add Email	Remove Email

Application Information:

Title of the Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES		
Attorney Docket Number	35447-42479/US	**Small Entity Status Claimed**	☒
Application Type	Provisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	14	**Suggested Figure for Publication (if any)**	

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	● Customer Number	US Patent Practitioner	○ Limited Recognition (37 CFR 11.9)
Customer Number	00758		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.
When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button. Add

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX)[i] the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Remove

Application Number	Country[i]	Filing Date (YYYY-MM-DD)	Access Code[i] (if applicable)

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button. Add

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant 1 — Remove

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

Clear

● Assignee	Legal Representative under 35 U.S.C. 117	Joint Inventor
Person to whom the inventor is obligated to assign.	Person who shows sufficient proprietary interest	

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here. ☒

Organization Name	Perceptive Automata, Inc.		
Mailing Address Information For Applicant:			
Address 1	10 Ware Street		
Address 2			
City	Cambridge	**State/Province**	MA
Country	US	Postal Code	02138
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button. Add

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

Assignee 1

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

Remove

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

Prefix	**Given Name**	Middle Name	**Family Name**	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1			
Address 2			
City		**State/Province**	
Country i		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button. Add

Signature:

Remove

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Rajendra Panwar/			Date (YYYY-MM-DD)	2019-01-30
First Name	Rajendra B.	Last Name	Panwar	Registration Number	63,165

Additional Signature may be generated within this form by selecting the Add button. Add

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	35447-42479/US
	Application Number	
Title of Invention	NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES	

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1 The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3 A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent CooperationTreaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

NAVIGATING AUTONOMOUS VEHICLES BASED ON HIDDEN CONTEXT ASSOCIATED WITH TRAFFIC ENTITIES

Inventor

Sam Anthony

TECHNICAL FIELD

[0001] The present disclosure relates generally navigating an autonomous vehicle through traffic and more specifically to routing of an autonomous vehicle through traffic based on prediction of hidden context associated with traffic entities.

INTRODUCTION

[0002] An autonomous vehicle uses different types of sensors to receive input describing the surroundings (or environment) of the autonomous vehicle while driving through traffic. For example, an autonomous vehicle may perceive the surroundings using camera images and lidar scans. The autonomous vehicle determines whether an object in the surroundings is stationary, for example, buildings or trees or the object is non-stationary, for example, a pedestrian, a vehicle, and so on. The autonomous vehicle system predicts the motion of non-stationary objects to make sure that the autonomous vehicle is able to navigate through non-stationary obstacles in the traffic.

[0003] Conventional systems predict motion of non-stationary objects using kinematics. For example, autonomous vehicles may rely on methods that make decisions on how to control the vehicles by predicting "motion vectors" of people near the vehicles. This is accomplished by collecting data of a person's current and past movements, determining a motion vector of the person at a current time based on these movements, and extrapolating a future motion vector representing the person's predicted motion at a future time based on the current motion vector. However, these techniques fail to predict motion of certain non-

stationary objects for example, pedestrians, bicyclists, and so on. For example, if the autonomous vehicle detects a pedestrian standing in a street corner, the motion of the pedestrian does not help predict whether the pedestrian will cross the street or whether the pedestrian will remain standing on a street corner. Similarly, if the autonomous vehicle detects a bicyclist in a lane, the current motion of the bicycle does not help the autonomous vehicle predict whether the bicycle will change lanes. Failure of autonomous vehicles fail to accurately predict motion of non-stationary traffic objects results in unnatural movement of the autonomous vehicle, for example, as a result of the autonomous vehicle suddenly stopping due to a pedestrian moving in the road or the autonomous vehicle continuing to wait for a person to cross a street even if the person never intends to cross the street.

BRIEF DESCRIPTION OF FIGURES

[0004] Various objectives, features, and advantages of the disclosed subject matter can be more fully appreciated with reference to the following detailed description of the disclosed subject matter when considered in connection with the following drawings, in which like reference numerals identify like elements.

[0005] FIG. 1A is a system diagram of a networked system for predicting human behavior according to some embodiments of the invention.

[0006] FIG. 1B is the system architecture of a vehicle computing system that routes an autonomous vehicle based on prediction of hidden context associated with traffic objects according to some embodiments of the invention.

[0007] FIG. 2A is a flow chart showing a process of predicting human behavior, according to some embodiments of the invention.

[0008] FIG. 2B shows an example of a derived stimulus, according to some embodiments of the invention.

TESLA EXHIBIT 1007, Page 34

[0009] FIG. 3 is a system diagram showing a sensor system associated with a vehicle, according to some embodiments of the invention.

[0010] FIG. 4 shows a process of generating derived stimuli from raw camera or sensor data in the vehicle, according to some embodiments of the invention.

[0011] FIG. 5 is a flowchart showing a process of collecting predictions and other information from human observers based on derived stimuli, according to some embodiments of the invention.

[0012] FIG. 6 shows a data structure associated with tracking video frame data, according to some embodiments of the invention.

[0013] FIG. 7 is a flowchart showing a process for producing summary statistics of a video frame or derived stimulus according to some embodiments of the invention.

[0014] FIG. 8 is a flowchart showing a process of training a learning algorithm using summary statistics, according to some embodiments of the invention.

[0015] FIG. 9 is a flowchart showing a process of predicting the state of mind of road users using a trained learning algorithm, according to some embodiments of the invention.

[0016] FIG. 10 is a diagram showing an example of an application of a context user prediction process in an automobile context, according to some embodiments of the invention.

[0017] FIG. 11 represents a flowchart illustrating the process of navigating the autonomous vehicle based on hidden context, according to an embodiment.

[0018] FIG. 12 is a block diagram illustrating components of an example machine able to read instructions from a machine-readable medium and execute them in a processor (or controller).

DETAILED DESCRIPTION

[0019] Embodiments of the invention predict hidden context associated with traffic entities that determines behavior of these traffic entities in the traffic. A traffic entity represents an object in traffic, for example, a pedestrian, a bicycle, a vehicle, a delivery robot, and so on. Hidden context includes factors that affect the behavior of such traffic entities, for example, a state of mind of a pedestrian, a degree of awareness of the existence of the autonomous vehicle in the vicinity (for example, whether a bicyclist is aware of the existence of the autonomous vehicle in the proximity of the bicyclist), and so on. The system uses the hidden context to predict behavior of people near a vehicle in a way that more closely resembles how human drivers would judge the behavior.

[0020] In one embodiment, a group of users (or human observers) view sample images of people (such as pedestrians) near streets and/or vehicles and indicate or are measured for their understanding of how they believe the people will behave. These indicators or measurements are then used as a component for training a machine learning based model that predicts how people will behave in a real-world context. In other words, after being trained based on the reactions of human observers to sample images in a training environment, the machine learning based model predicts behavior of traffic entities in a real-world environment, for example, actual pedestrian behavior in a real-world environment.

[0021] A non-stationary object may also be referred to as a movable object. An object in the traffic may also be referred to as an entity. Systems for predicting human interactions with vehicles are disclosed in U.S. Patent Application Number 15/830,549, filed on December 4, 2017 which is incorporated herein by reference in its entirety.

System Environment

[0022] FIG. 1 is a system diagram of a networked system for predicting human behavior according to some embodiments of the invention. FIG. 1 shows a vehicle 102, a network

104, a server 106, a user response database 110, a user terminal 108, a model training system 112 and a prediction engine 114.

[0023] The vehicle 102 can be any type of manual or motorized vehicle such as a car, bus, train, scooter, or bicycle. As described in more detail below, the vehicle 102 can include sensors for monitoring the environment surrounding the vehicle. In one implementation, the sensors can include a camera affixed to any portion of the vehicle for capturing a video of people near the vehicle.

[0024] The network 104 can be any wired and/or wireless network capable of receiving sensor data collected by the vehicle 102 and distributing it to the server 106, the model training system 112, and, through the model training system 112, the prediction engine 114.

[0025] The server 106 can be any type of computer system capable of (1) hosting information (such as image, video and text information) and delivering it to a user terminal (such as client device 108), (2) recording responses of multiple users (or human observers) to the information, and (3) delivering such information and accompanying responses (such as responses input via client device 108) back to the network 104.

[0026] The user response database 110 can be any type of database or data storage system capable of storing the image, video, and text information and associated user responses and subsequently recalling them in response to a query.

[0027] The model training system 112 can be implemented in any type of computing system. In one embodiment, the system 112 receives the image, video, and/or text information and accompanying, or linked, user responses from the database 110 over the network 104. In some embodiments, the text segments are discrete values or free text responses. The model training system 112 can use images, video segments and text segments as training examples to train an algorithm, and can create labels from the accompanying user responses based on the trained algorithm. These labels indicate how the algorithm predicts

the behavior of the people in the associated image, video, and/or text segments. After the labels are created, the model training system 112 can transmit them to the prediction engine 144.

[0028] The prediction engine 114 can be implemented in any computing system. In an illustrative example, the prediction engine 114 includes process that executes a model that has been trained by the model training system 112. This process estimates a label for a new (*e.g.,* an actual "real-world") image, video, and/or text segment based on the labels and associated image, video, and/or text segments that it received from the model training system 112. In some embodiments, this label comprises aggregate or summary information about the responses of a large number of users (or human observers) presented with similar image, video, or text segments while the algorithm was being trained.

[0029] FIG. 1B is the system architecture of a vehicle computing system that routes an autonomous vehicle based on prediction of hidden context associated with traffic objects according to some embodiments of the invention. The vehicle computing system 120 comprises the prediction engine 102, a future position estimator 125, a motion planner 130, a vehicle control module 135, a Kalman filter modulation module 140, a world model modulation module 145, a POMDP (partially observable Markov decision process) prior biasing module 150, and a false positive reduction module 155. Other embodiments may include more or fewer modules than those shown in FIG. 1B. Actions performed by a particular module as indicated herein may be performed by other modules than those indicated herein.

[0030] The sensors of an autonomous vehicle capture sensor data 160 representing a scene describing the traffic surrounding the autonomous vehicle. Examples of sensors used by an autonomous vehicle include cameras, lidars, GNSS (global navigation satellite system

such as a global positioning system, or GPS), IMU (inertial measurement unit), and so on. Examples of sensor data includes camera images and lidar scans.

[0031] The traffic includes one or more traffic entities, for example, a pedestrian 162. The vehicle computing system 120 analyzes the sensor data 160 and identifies various traffic entities in the scene, for example, pedestrians, bicyclists, other vehicles, and so on. The vehicle computing system 120 determines various parameters associated with the traffic entity, for example, the location (represented as x and y coordinates), a motion vector describing the movement of the traffic entity, and so on.

[0032] The future position estimator 125 estimates the future position of a traffic entity. The motion planner 130 determines a plan for the motion of the autonomous vehicle. The vehicle control module 135 sends signals to the vehicle controls (for example, accelerator, brakes, steering, emergency braking system, and so on) to control the movement of the autonomous vehicle. In an embodiment, the future position estimates for a traffic entity determined by the future position estimator 125 based on sensor data 160 are provided as input to the motion planner 130. The motion planner 130 determines a plan for navigating the autonomous vehicle through traffic, and provides a description of the plan to the vehicle control module 135. The vehicle control module 135 generates signals for providing to the vehicle controls. For example, the vehicle control module 135 may send control signals to an emergency braking system to stop the vehicle suddenly while driving, the vehicle control module 135 may send control signals to the accelerator to increase or decrease the speed of the vehicle, or the vehicle control module 135 may send control signals to the steering of the autonomous vehicle to change the direction in which the autonomous vehicle is moving.

[0033] FIG. 2A is a flow chart showing a process of predicting human behavior, according to some embodiments of the invention. In one implementation, video or other data is captured by a camera or sensor mounted on the vehicle 102. The camera or other sensor

can be mounted in a fixed or temporary manner to the vehicle 102. Of course, the camera does not need to be mounted to an automobile, and could be mounted to another type of vehicle, such as a bicycle. As the vehicle travels along various streets, the camera or sensor captures still and/or moving images (or other sensor data) of pedestrians, bicycles, automobiles, *etc.* moving or being stationary on or near the streets. In step 202, this video or other data captured by the camera or other sensor is transmitted from the vehicle 102, over the network 104, and to the server 106 where it is stored.

[0034] Then, in step 204, video frames or segments are extracted from the stored video or other data and are used to create stimulus data including derived stimulus (or stimuli). In one implementation, the derived stimulus corresponds to a scene in which one or more humans are conducting activities (*e.g.,* standing, walking, driving, riding a bicycle, *etc.*) beside or on a street and/or near a vehicle. As explained in more detail below for example in step 214 and in the text accompanying FIG. 9, as part of the training process for the prediction algorithm, human observers view the derived stimulus and predict how they believe the humans shown in the derived stimulus will act. In yet a further implementation, after the video frames or segments are extracted from the stored data, the derived stimulus is generated by manipulating the pixels or equivalent array data acquired from the camera or other sensor in step 204, producing a new data file that conveys a portion of the information from the original video with certain aspects highlighted or obscured, as described below in FIG 4.

[0035] FIG. 2B shows an example of a derived stimulus, according to some embodiments of the invention. As shown, three non-consecutive frames 220, 222, 223, and 224, are extracted from a short segment of video. A copy of the first frame is made 220. On the copy of the first frame 220, a box 221 is drawn around the figure of interest, highlighting that figure. The copy of the first frame 220 and the three extracted frames 222, 223, and 224 are

then recombined into a short video segment that loops. This encourages human observers to focus on a figure of interest in the video segment.

[0036] In step 206, the derived stimulus is transmitted from the server 106 and displayed to a large number of users (or human observers) on the user terminal 108 (or multiple terminals 108). The terminal(s) 108 prompt the human observers to predict how the people shown in the derived stimulus will act, and upon viewing the displayed stimulus, the observers input their responses corresponding to their predictions. For example, in the derived stimulus shown in FIG. 2B, the human observers may predict that the bicyclist will continue riding to the left in the figure. In other derived stimuli, the responses may predict whether a first person in the stimulus will cross the street, another person will remain standing on a street corner, and yet another person will change lanes on his or her bicycle. In an illustrative embodiment, the human observers may make a continuous or ordinal judgment about the state of mind or the predicted behavior of the people shown in the derived stimulus and record that judgment. For example, the human observers may select an appropriate icon displayed on the terminal(s) 108 by clicking a mouse or by pressing a key to indicate their judgment or prediction. The judgment or prediction may correspond to the human observers' assessment of the state of mind of the person in the derived stimulus or other awareness or intention that would be relevant to a hypothetical driver who sees the person in the derived stimulus while driving. In step 208, the derived stimulus and associated human observer responses are transmitted from the terminal(s) 108 to the server 106 and recorded in the user response database 110.

[0037] In step 210, summary statistics are generated based on the user responses. For example, the statistics may characterize the aggregate responses of multiple human observers to a particular derived stimulus. For instance, if the derived stimulus shows a pedestrian walking on a sidewalk towards an intersection, the response can be categorized in terms of

how many human observers believe that the pedestrian will stop upon reaching the intersection, continue walking straight across the intersection, turn a corner and continue walking along the sidewalk without crossing the intersection, *etc.* These summary statistics can characterize the human observer responses in terms of certain parameters associated with the statistics, such as a content of a response, a time associated with entering a response, and a position of an eye of a human observer associated with the response. The parameters can also be associated with a (1) central tendency, variance, skew, kurtosis, scale, or histogram. For example, the amount of time users took to input their responses can be characterized in terms of central tendency, variance, skew, kurtosis, scale, histogram. Also, the statistics can include a parameter that additionally or alternatively characterizes the movement of the human observers' eyes relative to a display when making the judgments in terms of central tendency, variance, skew, kurtosis, scale, histogram or two-dimensional distribution. In one embodiment, the statistics are stored in the user response database 110 with an index that identifies the raw video or sensor data from which the derived stimulus was generated. In a further embodiment, the statistics stored in the database 110 cover a large set of images of people on or near roads and are categorized in a number of different categories, such as pedestrian, driver, motorcyclist, bicyclist, scooter driver, self-balancing scooter rider, unicyclist, motorized wheelchair user, skateboarder, or others. Moreover, the statistics are respectively stored along with, or linked to, the images of the derived stimuli corresponding to the statistics.

[0038] In step 212, the stored statistics and corresponding images (*e.g.,* the video frames or segments that were extracted from the video or other data (captured from the camera or sensor of the vehicle 202)) are sent over the network 104 to the model training system 112 and used to train a prediction algorithm. For example, the collection of images and statistics can be used to train a supervised learning algorithm, which can comprise a random forest

TESLA EXHIBIT 1007, Page 42

regressor, a support vector regressor, a simple neural network, a deep convolutional neural network, a recurrent neural network, a long-short-term memory **(LSTM)** neural network with linear or nonlinear kernels that are two dimensional or three dimensional, or any other supervised learning algorithm which is able to take a collection of data labeled with continuous values and adapt its architecture in terms of weights, structure or other characteristics to minimize the deviation between its predicted label on a novel stimulus and the actual label collected on that stimulus using the same method as was used on the set of stimuli used to train that network. The model is given data which comprises some subset of the pixel data from the video frames that the summary statistics were generated from. In one implementation, this subset includes the pixel data contained in a box (such as the box 221 shown in FIG. 2B) drawn to contain the boundaries of the person, cyclist, motorist and vehicle, or other road user, including their mode of conveyance. In some other implementations, it also includes the entire pixel data from the rest of the image. In one of those implementations, that pixel data is selected according to criteria such as the salience of those features in terms of contrast, lighting, presence of edges, or color. In an additional implementation, the features can include descriptive meta-data about the images such as the dimensions and location of the bounding box, the shape of the bounding box or the change in size or position of the bounding box from one frame to the next.

[0039] In step 214, the prediction engine 114 uses the trained model from the model training system 112 to predict the actual, "real-world" or "live data" behavior of people on or near a road. In one embodiment, the prediction engine 114 receives "live data" that matches the format of the data used to train the trained model. For example, if the trained model was trained based on video data received from a camera on the vehicle 102, the "live data" that is input to the algorithm likewise is video data from the same or similar type camera. On the other hand, if the model was trained based on another type of sensor data received from

another type of sensor on the vehicle 102, the "live data" that is input to the prediction engine 114 likewise is the other type of data from the same or similar sensor.

[0040] The trained model or algorithm makes a prediction of what a pedestrian or other person shown in the "live data" would do based on the summary statistics and/or training labels of one or more derived stimulus. The accuracy of the model is determined by having it make predictions of novel derived stimuli that were not part of the training images previously mentioned but which do have human ratings attached to them, such that the summary statistics on the novel images can be generated using the same method as was used to generate the summary statistics for the training data, but where the correlation between summary statistics and image data was not part of the model training process. The predictions produced by the trained model comprise a set of predictions of the state of mind of road users that can then be used to improve the performance of autonomous vehicles, robots, virtual agents, trucks, bicycles, or other systems that operate on roadways by allowing them to make judgments about the future behavior of road users based on their state of mind.

[0041] FIG. 3 is a system diagram showing a sensor system associated with a vehicle, according to some embodiments of the invention. FIG. 3 shows a vehicle 306 with arrows pointing to the locations of its sensors 300, a local processor and storage 302, and remote storage 304.

[0042] Data is collected from cameras or other sensors 300 including solid state Lidar, rotating Lidar, medium range radar, or others mounted on the car in either a fixed or temporary capacity and oriented such that they capture images of the road ahead, behind, and/or to the side of the car. In some embodiments, the sensor data is recorded on a physical storage medium (not shown) such as a compact flash drive, hard drive, solid state drive or dedicated data logger. In some embodiments, the sensors 300 and storage media are managed by the processor 302.

[0043] The sensor data can be transferred from the in-car data storage medium and processor 302 to another storage medium 304 which could include cloud-based, desktop, or hosted server storage products. In some embodiments, the sensor data can be stored as video, video segments, or video frames.

[0044] In some embodiments, data in the remote storage 304 also includes database tables associated with the sensor data. When sensor data is received, a row can be added to a database table that records information about the sensor data that was recorded, including where it was recorded, by whom, on what date, how long the segment is, where the physical files can be found either on the internet or on local storage, what the resolution of the sensor data is, what type of sensor it was recorded on, the position of the sensor, and other characteristics.

[0045] FIG. 4 shows a process of generating derived stimuli from raw camera or sensor data in the vehicle, according to some embodiments of the invention.

[0046] Referring to step 400, video or data frames are selected according to a method which selects either random or specific frames from video or sensor data segments. As described above, sensor data can include video segments or specific frames. These frames can either be contiguous or non-contiguous, and can be in the original order, in a permuted order, in reverse order, or in random order. Some of the frames can be repeated once or more than once. The frames are selected so that when human observers are later presented with reordered versions of the sequences of frames, they perceive either the actual motion of humans in the scene that was captured in the original video, or motion that is a product of the ordering of the frames but which diverges via reordering from the motion that was originally captured. The temporal information that the human observers looking at the sets of images or derived stimulus are able to use to answer questions about the sets of images may be

therefore limited or expanded to include specific temporal cues that were or were not in the original sequence.

[0047] Referring to step 402, some of the frames can be manipulated. The frames can be manipulated by adjusting pixel values. These manipulations can include blurring, the addition or one or more occluding bars, bands, or shapes, sharpening, the removal of color information, the manipulation of color information, the drawing of non-occluding or highlighting shapes on the image, other manipulations, or a combination of the manipulations listed here, or a combination of the manipulations listed here with other manipulations, or other manipulations of the pixels not listed combined with each other. The manipulations serve the purpose of highlighting, occluding or degrading portions of the image, so that when the images are shown to the human observers, they are directed to people or specific portions of the image when predicting what the people in the images will do. For example, using the highlighting described above, a certain pedestrian in a scene can be isolated such that a human observer's feedback can be more reliably associated with the pedestrian.

[0048] Referring to step 404, frames can be recombined to form a derived stimulus. In some embodiments, if there is only one frame that frame comprises the derived stimulus. If there is more than one frame those frames may then be recombined. The recombination into a temporal segment can involve the frames having the same display time and interval as in the original video clip or could involve faster, slower or variable display times. The timing of the recombination is chosen in order to shape the perceived temporal duration of the activities visible in the original video segment so the judgments the observers make will rely on temporal information that is or is not identical to that featured in the original image.

[0049] FIG. 5 is a flowchart showing a process of collecting predictions and other information from human observers based on derived stimuli, according to some embodiments of the invention. In step 500, human observers are given detailed instructions about how to

answer questions about derived stimuli. In step 502, those observers are presented with derived stimuli and asked to answer questions about them. In step 504, the observers respond to the stimuli and those responses are recorded. In step 506, the recorded responses are aggregated and logged in a database. For example, the responses may be sent to and logged in the user response database 110 (FIG. 1).

[0050] Referring to step 500, in one example, some number of human observers (two or more) are recruited to participate on one or several crowdsourcing websites, such as Amazon's Mechanical Turk or at a physical location provided with a display. The observers are given detailed written and pictorial instructions explaining the task that they are about to complete. These instructions give examples of situations that might be depicted in the derived stimuli, and the kinds of responses that would be appropriate for those situations. For example, if human observers are asked to judge the intention of a pedestrian to walk in front of a car, the instructions may be: "In this test, we want you to pretend you're driving a car. You'll see road scenes with cyclists highlighted, and you'll have to decide what they're planning to do. Try to answer as quickly as you can. If a box contains more than one cyclist, try to judge if any of them intends to ride into the path of your car. Some of the images might be small, or dark, but just do your best."

[0051] Referring to step 502, the human observers may be shown a display which includes the derived stimulus. The display also includes a mechanism for making a judgment about the stimulus. The mechanism for making the judgment can be a continuous indicator such as a ribbon on which the observer could drag a control to a certain point. The mechanism can also be an ordinal measure such as a Likert scale where the observer can make a judgment about a degree of certainty of the judgment. The mechanism can also be a control that the human observer drags with their mouse to draw a trajectory onscreen indicating a judgment. The mechanism can also be a text entry field where the observer types

a description of their judgment. The judgment that the human observer makes is an evaluation of the state of mind of a road user depicted in the derived stimulus. The evaluation can be of the intention, awareness, personality, state of consciousness, level of tiredness, aggressiveness, enthusiasm, thoughtfulness or another characteristic of the internal mental state of the pictured road user. If the ratings collected are on an ordinal scale they can describe the characteristic using language of probability, such as "the other driver may be attentive" or "the other driver" is definitely attentive" or "the other driver is definitely not attentive". The ratings of large numbers of human observers are collected. Summary statistics are generated based on the responses of all of the observers who looked at an image. Individual variability in responses to a given stimulus can be characterized in the information given by the observers to the learning algorithm. The summary statistics might include unweighted information from all observers, or might exclude observers based on extrinsic or intrinsic criteria such as the time it took an observer to respond, the geographical location of an observer, the observer's self-reported driving experience, or the observer's reliability in making ratings of a set of other images.

[0052] Referring to step 504, the explicit response of the observer is recorded as well as implicit data. The implicit data can include how long the subject took to respond, if they hesitated in their motions, if they deleted keystrokes, if they moved the mouse anywhere other than the location corresponding to the response they eventually chose, where their eyes moved, or other implicit measures.

[0053] Referring to step 506, the responses are aggregated and recorded in a data structure, such as the user response database 110 (FIG. 1). This data structure is then sent as a text field to a networked computer system running database software and logged in a database.

TESLA EXHIBIT 1007, Page 48

[0054] FIG. 6 shows a data structure (*e.g.,* table) associated with tracking video frame data according to some embodiments of the invention. Column 600 shows an example of a unique identifier for each human observer who rated an image. Column 602 shows a unique identifier for a given derived stimulus. Column 604 shows an example of a response from an observer recorded in the database. Column 606 shows an example of additional information - in this case response time — recorded by the database.

[0055] In reference to column 600, each observer who is shown the stimuli is given a unique identifier so that information about their response may be analyzed in the context of their responses to other images.

[0056] In reference to column 602, each derived stimulus that is rated by a human observer is listed with a filename and path that uniquely identifies that derived stimulus, and also indicates the original video frames from which that derived stimulus was generated.

[0057] In reference to column 604, for each stimulus rated by each human observer, a response is recorded that could be a continuous, discrete, or ordinal value. This value may refer to the probability of the pictured human road user has a given state of mind — e.g. that a pedestrian is likely to cross the street or that an oncoming vehicle is unlikely to be willing to yield to the vehicle containing the sensor if the vehicle containing the sensor needs to tum. In some embodiments, a higher ordinal value (e.g., the ordinal 4 as shown in FIG. 6) indicates that a human observer believes that there is a higher probability that the pictured human road user has a given state of mind or will perform a particular action. On the other hand, a lower ordinal value (e.g., the ordinal 1 as shown in FIG. 6) indicates that the human observer believes that there is a lower probability that the pictured human road user has the state of mind or will perform the particular action. On the other hand, in some embodiments, a lower ordinal value can indicate a higher probability of an action, and a higher ordinal value can indicate a lower probability of an action.

[0058] In reference to column 606, an amount of time associated with a subject responding to the derived stimulus is also recorded. In some embodiments, this time is associated with the overall reliability of the human observer's rating. For example, a response associated with a lower response time may be weighted higher and a response associated with a slower response time may be weighted lower.

[0059] FIG. 7 is a flowchart showing a process for producing summary statistics of a video frame or derived stimulus according to some embodiments of the invention. In step 700, per-image results are extracted from the database which stores responses from a human observer. In step 702, summary statistics are calculated using those per-image results. In step 704, records of the summary statistics attached to frames from the raw car data (not the derived stimuli) are recorded in a separate table.

[0060] Records of all of the responses by human observers on a given task for a given frame are extracted from the database in step 700. In one example, all of the responses of a given type are aggregated into a single data structure.

[0061] This data structure is queried to produce summary statistics in step 702. These summary statistics could include measurements of the central tendency of the distribution of scores like the mean, median, or mode. They could include measurements of the heterogeneity of the scores like variance, standard deviation, skew, kurtosis, heteroskedasticity, multimodality, or uniformness. They could also include summary statistics like those above calculated from the implicit measurements of the responses listed above.

[0062] The calculated summary statistics are recorded in a database table in step 704 attached to a unique identifier which links them to the video frame or sensor data frame associated with the responses from which they were calculated.

[0063] FIG. 8 is a flowchart showing a process of training a learning algorithm using summary statistics, according to some embodiments of the invention. For example, in one embodiment, the process may train an algorithm used in the model training system 112 in FIG. 1. In step 800, summary statistics linked to a given video or sensor frame are extracted from the database 110 in FIG. 1. Subsequently, in step 802, the video or sensor frame may be collated with the statistics label for the frame. The collated frame is provided to the learning algorithm in the model training system 112 in step 804, and this process yields a final learning algorithm in step 806.

[0064] In some embodiments, between one and three partitions are generated from the database records in step 800. These partitions each comprise a list of images in the set of images that have been manipulated to create derived stimuli and rated by human observers. At least one of these partitions is made to comprise a set of training data. An additional partition may be created to use as a set of validation data, a second list of images in the set of images that have been manipulated to create derived stimuli and rated by human images, but where the second validation set is not shown to the model in the learning phase, but is instead preserved to be used later for evaluation.

[0065] The list of images in the training data partition is collated with some or all of the summary statistics that have been generated from the human observer responses to those images and stored in the database in step 802, such that the summary statistics are associated with the appropriate image in the list which comprises the training partition.

[0066] Each combination of image and summary statistics comprises a single training sample for the learning algorithm in step 804. The algorithm may be any type of supervised learning algorithm capable of predicting a continuous label for a two or three dimensional input, including but not limited to a random forest regressor, a support vector regressor, a simple neural network, a deep convolutional neural network, a recurrent neural network, a

long-short-term memory (LSTM) neural network with linear or nonlinear kernels that are two dimensional or three dimensional.

[0067] The learning algorithm is optimized by a process of progressively adjusting the parameters of that algorithm in response to the characteristics of the images and summary statistics given to it in the training phase to minimize the error in its predictions of the summary statistics for the training images in step 804. In one embodiment of the model training system 112, the algorithm can be a deep neural network. In this embodiment the parameters are the weights attached to the connections between the artificial neurons comprising the network. Pixel data from an image in a training set collated with human observer summary statistics in step 802 can serve as an input to the network. This input can be transformed according to a mathematical function by each of the artificial neurons, and then the transformed information can be transmitted from that artificial neuron to other artificial neurons in the neural network. The transmission between the first artificial neuron and the subsequent neurons can be modified by the weight parameters discussed above. In this embodiment, the neural network can be organized hierarchically such that the value of each input pixel can be transformed by independent layers (e.g., 10 to 20 layers) of artificial neurons, where the inputs for neurons at a given layer come from the previous layer, and all of the outputs for a neuron (and their associated weight parameters) go to the subsequent layer. At the end of the sequence of layers, in this embodiment, the network can produce numbers that are intended to match the human summary statistics given at the input. The difference between the numbers that the network output and the human summary statistics provided at the input comprises an error signal. An algorithm (e.g., back-propagation) can be used to assign a small portion of the responsibility for the error to each of the weight parameters in the network. The weight parameters can then be adjusted such that their estimated contribution to the overall error is reduced. This process can be repeated for each

image (or for each combination of pixel data and human observer summary statistics) in the training set collected via step 802. At the end of this process the model is "trained", which in some embodiments, means that the difference between the summary statistics output by the neural network and the summary statistics calculated from the responses of the human observers in step 506 is minimized.

[0068] FIG 9. is a flowchart showing a process of predicting the state of mind of road users using a trained learning algorithm, according to some embodiments of the invention. In step 900, the training algorithm receives a "real world" or "live data" video or sensor frame. Then in step 902, the trained algorithm analyzes the frame, thus enabling the algorithm in step 904 to output a prediction of summary statistics on the frame.

[0069] The "real world" or "live data" video or other sensor frames from a car-mounted sensor are delivered to the trained learning algorithm in step 900. These frames have the same resolution, color depth and file format as the frames used to train the algorithm. These frames are delivered as individual frames or as sequences according to the format used to train the original algorithm.

[0070] Each of these frames is analyzed by being passed through the trained model in step 902. In one embodiment, the data from the frame that was passed through the model would comprise the pixel data from a camera. This data would be transformed by an artificial neural network that had been trained according to step 804. At the final stage of the processing in the artificial network, it would produce an output. This output is the model output in step 904.

[0071] The model outputs a number or set of numbers that comprise the predicted summary statistics for the "real world" or "live data" image in step 904. The predicted summary statistics are the model's best estimation of what the summary statistics would be on the image if the image had human annotations collected. The prediction is generated

automatically by passing the sensor data through the model, where the information is transformed by the internal mechanisms of the model according to the parameters that were set in the training process shown in step 804. Because these summary statistics characterize the distribution of human responses that predict the state of mind of a road user pictured in the stimulus, the predicted statistics are therefore a prediction of the aggregate judgment of human observers of the state of mind of the pictured road user and thus an indirect prediction of the actual state of mind of the road user.

[0072] FIG. 10 is a diagram showing an example of an application of a context user prediction process in an automobile context, according to some embodiments of the invention. In this example intention 1006 1018 means that the road user 1002 1014 has the goal of moving into the path of the vehicle 1000 before the vehicle 1000 (on which the system is mounted) reaches their position. Awareness 1004 1016 in this example means that the road user 1002 1014 understands that the vehicle on which the system is mounted 1000 is present in their vicinity. In this example, when cyclist 1002 rides into the field of view of a camera mounted on vehicle 1000, the pixel data of the camera image of the cyclist is fed to a trained algorithm as described above in step 900. The trained algorithm analyzes the image as described above in step 902. The trained algorithm would predict summary statistics as in step 904. These summary statistics are an estimate of what the summary statistics would be for a collection of human observers who were shown a derived stimulus of the camera data as in step 504. The estimates summary statistics are therefore the system's best answer to the question "does this cyclist intend to enter the path of the vehicle." The vehicle is therefore able to make a guess 1006 about the intention of the cyclist that is closely matched to the guess that a human driver would make in that same situation. In this example, the intention of the cyclist 1006 is relatively high, as indicated by the number of horizontal bars in the display. The system installed on an automobile 1000 also makes predictions about the

awareness 1004 of cyclists of the vehicle 1000, by the same method described for intention. It also makes predictions about the willingness of an automobile 1008 to yield 1010 or its desire to turn across the system-containing vehicle's path 1012 by the same method described above. In the case of the automobile the questions that human subjects answered that would be predicted by the algorithm are "would the vehicle be willing to yield" 1010 and "does the vehicle wish to turn across your path" 1012. It also makes predictions about the desire of pedestrians 1014 to cross in front of the vehicle 1018, and whether those pedestrians are aware of the vehicle 1016, by the same method described above.

[0073] The models described above can be implemented as a real-time module that makes predictions of road user behavior based on input from cameras or other sensors installed on a car 1000. In the case of an autonomous car, these predictions can be used to make inferences about the intent of road users such as cyclists 1002, other motorists 1008, and pedestrians 1014 to cross into the path of the car, as well as whether the road users are aware of the car and its future path. They can also be used to predict whether other road users would be surprised, welcoming, or aggressively unwelcoming if the car were to engage in maneuvers which would take it into the path of another road user (e.g., would an oncoming car yield if the car implementing the systems and methods described herein were to turn left).

Determining a Stable Signal for Predicting Behavior of Traffic Entities

[0074] A pedestrian consumes information about the world that the pedestrian is walking in. The sensory apparatus of a traffic entity, for example, a pedestrian ingests many gigabytes of information about the world. This information is converted into a vast range of inferential judgments. Those judgments are used to update and execute moment-by-moment motion plans that propel the pedestrian through the traffic, and allow it to reach the pedestrian's target.

[0075] A human brain is able to negotiate complex situations with many contingent factors. This flexibility and capacity that makes human behavior very difficult to model using a computer. Even given a consistent goal structure the complexity of predicting human inference scales with the complexity of the inputs to the human perceptual system. In a complex environments, such as a pedestrian navigating a busy intersection, the number of perceptual inputs with high valence to moment-by-moment motion planning is extremely large. At any time the pedestrians actions at the next moment are influenced by many hundreds of different aspects of the scene. These can include moving agents, contextual cues, distractions, fixed obstacles, newly noticed scene aspects, and the internal process of reflection. Because of this, predicting the behavior of a pedestrian at any time that is not extremely close to the present is intractable. Even with a complete world model, the number of possible actions at, say, one second from the present moment are subject to a combinatorial explosion.

[0076] Embodiments of the invention extract a stable signal of potential behavior of a traffic entity, for example, pedestrian that is not subject to the combinatorial explosion of inputs. In one embodiment, this stable signal is used to modulate the performance envelope of a vehicle under full or partial automatic control. In another embodiment, this stable signal is used to inform the human operator of a vehicle that the behavior of upcoming road users may require their attention.

[0077] In an embodiment, determining the stable signal comprises generating a computational model of the mental state of the pedestrian. Humans observe other humans and with a high degree of reliability, make estimates of their internal mental states as they apply to behavior. Embodiments of the invention train a model by leveraging indirect but highly reliable measurements of human theory of mind reasoning. This allows the model to make predictions about the internal mental states of others.

[0078] Embodiments of the invention predict an internal mental state of a pedestrian as a stable signal. The internal mental state of a pedestrian comprises a goal-oriented planning and evaluation system with short and long-term memory. A pedestrian on the sidewalk has come to occupy the sidewalk in pursuance of a specific goal (for example, to get to the other side). This goal has a duration that is (in general) far longer than the amount of time in which a computing system of an automatic vehicle would have to consider that pedestrian's state of mind. If a pedestrian has the goal of eventually reaching a destination on the other side of the street down which the vehicle in question is traveling, that goal will obtain until they have reached that destination. This goal will not generally be changed by transient conditions like the appearance of a "do not walk" sign. However, the likelihood that the pedestrian steps into the street at any given moment is very contingent on that transient condition. Embodiments of the invention predict estimates of state of mind, which aim to predict the internal goal and belief structure of the pedestrian, are therefore not subject to the combinatorial explosion of contingent factors.

Overall Process of Navigating an Autonomous Vehicle through Traffic

[0079] FIG. 11 represents a flowchart illustrating the process of navigating the autonomous vehicle based on hidden context, according to an embodiment. The steps illustrated in the flowchart may be performed in an order different from that illustrated in FIG. 11. For example, certain steps may be performed in parallel. The steps may be performed by modules other than those indicated herein.

[0080] The vehicle computing system 120 receives 1100 sensor data from sensors of the autonomous vehicle. For example, the vehicle computing system 120 may receive lidar scans from lidars and camera images from cameras mounted on the autonomous vehicle. In an embodiment, the vehicle computing system 120 builds a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data. The point cloud

representation includes coordinates of points surrounding the vehicle, for example, three dimensional points and parameters describing each point, for example, the color, intensity, and so on. The vehicle computing system 120 identifies 1102 one or more traffic entities based on the sensor data, for example, pedestrians, bicyclists, or other vehicles driving in the traffic. The traffic entities represent non-stationary objects in the surroundings of the autonomous vehicle.

[0081] The vehicle computing system 120 repeats the following steps 1104 and 1106 for each identified traffic entity. The vehicle computing system 120 determines 1104 motion parameters for the traffic entity, for example, speed and direction of movement of the traffic entity. The vehicle computing system 120 also determines 1106 a hidden context associated with the traffic entity using the prediction engine 102. The hidden context may represent a state of mind of a person represented by the traffic entity. For example, the hidden context may represent a near term goal of the person represented by the traffic entity, for example, indicating that the person is likely to cross the street, or indicating that the person is likely to pick up an object (e.g., a wallet) dropped on the street but stay on that side of the street, or any other task that the person is likely to perform within a threshold time interval. The hidden context may represent a degree of awareness of the person about the autonomous vehicle, for example, whether a bicyclist driving in front of the autonomous vehicle is likely to be aware that the autonomous vehicle is behind the bicycle.

[0082] In an embodiment, the vehicle computing system 120 represents the hidden context as a vector of values, each value representing a parameter, for example, a likelihood that a person represented by a traffic entity is going to cross the street in front of the autonomous vehicle, a degree of awareness of the presence of autonomous vehicle in the mind of a person represented by a traffic entity, and so on.

[0083] The vehicle computing system 120 navigates 1108 the autonomous vehicle based on the motion parameters as well as the hidden context. For example, the vehicle computing system 120 may determine a safe distance from the traffic entity that the autonomous vehicle should maintain based on the motion parameters of the traffic entity. The vehicle computing system 120 modulates the safe distance based on the hidden context. The vehicle computing system 120 may adjust the safe distance based on whether the near-term goal of the person indicating that the person intends to reach a location in the direction of the movement of the traffic entity or in a different direction.

[0084] For example, based on the motion parameters, the vehicle computing system 120 may determine that the autonomous vehicle can drive within X meters of the traffic entity. However, the hidden context indicates that the person represented by the traffic entity intends to cross the street in a direction different from the direction indicated by the motion parameters. In this situation, the vehicle computing system 120 adjusts the safe distance such that the autonomous vehicle is able to drive closer to the traffic entity than the distance X. On the other hand, if the hidden context indicates that the person represented by the traffic entity intends to cross the street in a direction same as the direction indicated by the motion parameters, the vehicle computing system 120 adjusts the safe distance such that the autonomous vehicle maintains a distance greater than X from the traffic entity.

Navigating Autonomous Vehicle Based on Hidden Context

[0085] Embodiments of the vehicle computing system 120 predict hidden context representing intentions and future plans of an agent (e.g., a pedestrian or a bicyclist) and use the hidden context for navigating the autonomous vehicle, for example, by adjusting the path planning of the autonomous robot based on the hidden context. For example, the vehicle computing system 120 may improve the path planning by taking a machine learning based model that predicts the level of human uncertainty about the future actions of pedestrians and

cyclists and uses that as an input into the autonomous vehicle's motion planner. The training vector further includes information about the ground truth of the world obtained from more computer vision models. The vehicle computing system 120 may use the output of the prediction engine 102 to generate a probabilistic map of the risk of encountering an obstacle given different possible motion vectors at the next time step. Alternatively, the vehicle computing system 120 may use the output of the prediction engine 102 to a motion plan which incorporates the probabilistic uncertainty of the human assessment.

[0086] In an embodiment, the prediction engine 102 determines a metric representing a degree of uncertainty in human assessment of the near-term goal of a pedestrian or any user representing a traffic entity. The specific form of the representation of uncertainty is a model output that comes in the form of a probability distribution, capturing the expected distributional characteristics of human assessments of the goal and awareness state of road users. The model output may comprise the central tendency, i.e., the mean likelihood that a person will act in a certain way and one or more parameters including the variance, kurtosis, skew, heteroskedasticity, and multimodality of the predicted human distribution. These summary statistics carry information about the level of human uncertainty.

Computing Machine Architecture

[0100] FIG. 12 is a block diagram illustrating components of an example machine able to read instructions from a machine-readable medium and execute them in a processor (or controller). Specifically, FIG. 12 shows a diagrammatic representation of a machine in the example form of a computer system 1200 within which instructions 1224 (e.g., software) for causing the machine to perform any one or more of the methodologies discussed herein may be executed. In alternative embodiments, the machine operates as a standalone device or may be connected (e.g., networked) to other machines. In a networked deployment, the machine

TESLA EXHIBIT 1007, Page 60

may operate in the capacity of a server machine or a client machine in a server-client network environment, or as a peer machine in a peer-to-peer (or distributed) network environment.

[0101] The machine may be a server computer, a client computer, a personal computer (PC), a tablet PC, a set-top box (STB), a personal digital assistant (PDA), a cellular telephone, a smartphone, a web appliance, a network router, switch or bridge, or any machine capable of executing instructions 1224 (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute instructions 1224 to perform any one or more of the methodologies discussed herein.

[0102] The example computer system 1200 includes a processor 1202 (e.g., a central processing unit (CPU), a graphics processing unit (GPU), a digital signal processor (DSP), one or more application specific integrated circuits (ASICs), one or more radio-frequency integrated circuits (RFICs), or any combination of these), a main memory 1204, and a static memory 1206, which are configured to communicate with each other via a bus 1208. The computer system 1200 may further include graphics display unit 1210 (e.g., a plasma display panel (PDP), a liquid crystal display (LCD), a projector, or a cathode ray tube (CRT)). The computer system 1200 may also include alphanumeric input device 1212 (e.g., a keyboard), a cursor control device 1214 (e.g., a mouse, a trackball, a joystick, a motion sensor, or other pointing instrument), a storage unit 1216, a signal generation device 1218 (e.g., a speaker), and a network interface device 1220, which also are configured to communicate via the bus 1208.

[0103] The storage unit 1216 includes a machine-readable medium 1222 on which is stored instructions 1224 (e.g., software) embodying any one or more of the methodologies or functions described herein. The instructions 1224 (e.g., software) may also reside,

completely or at least partially, within the main memory 1204 or within the processor 1202 (e.g., within a processor's cache memory) during execution thereof by the computer system 1200, the main memory 1204 and the processor 1202 also constituting machine-readable media. The instructions 1224 (e.g., software) may be transmitted or received over a network 1226 via the network interface device 1220.

[0104] While machine-readable medium 1222 is shown in an example embodiment to be a single medium, the term "machine-readable medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, or associated caches and servers) able to store instructions (e.g., instructions 1224). The term "machine-readable medium" shall also be taken to include any medium that is capable of storing instructions (e.g., instructions 1224) for execution by the machine and that cause the machine to perform any one or more of the methodologies disclosed herein. The term "machine-readable medium" includes, but not be limited to, data repositories in the form of solid-state memories, optical media, and magnetic media.

Additional Considerations

[0087] Although embodiments disclosed describe techniques for navigating autonomous vehicles, the techniques disclosed are applicable to any mobile apparatus, for example, a robot, a delivery vehicle, a drone, and so on.

[0088] The subject matter described herein can be implemented in digital electronic circuitry, or in computer software, firmware, or hardware, including the structural means disclosed in this specification and structural equivalents thereof, or in combinations of them. The subject matter described herein can be implemented as one or more computer program products, such as one or more computer programs tangibly embodied in an information carrier (e.g., in a machine readable storage device) or in a propagated signal, for execution by, or to control the operation of, data processing apparatus (e.g., a programmable processor,

a computer, or multiple computers). A computer program (also known as a program, software, software application, or code) can be written in any form of programming language, including compiled or interpreted languages, and it can be deployed in any form, including as a stand-alone program or as a module, component, subroutine, or other unit suitable for use in a computing environment. A computer program does not necessarily correspond to a file. A program can be stored in a portion of a file that holds other programs or data, in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers at one site or distributed across multiple sites and interconnected by a communication network.

[0089] The processes and logic flows described in this specification, including the method steps of the subject matter described herein, can be performed by one or more programmable processors executing one or more computer programs to perform functions of the subject matter described herein by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus of the subject matter described herein can be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application specific integrated circuit).

[0090] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processor of any kind of digital computer. Generally, a processor will receive instructions and data from a read only memory or a random access memory or both. The essential elements of a computer are a processor for executing instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto optical disks, or optical disks. Information carriers

suitable for embodying computer program instructions and data include all forms of nonvolatile memory, including by way of example semiconductor memory devices, (e.g., EPROM, EEPROM, and flash memory devices); magnetic disks, (e.g., internal hard disks or removable disks); magneto optical disks; and optical disks (e.g., CD and DVD disks). The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

[0091] To provide for interaction with a user, the subject matter described herein can be implemented on a computer having a display device, e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor, for displaying information to the user and a keyboard and a pointing device, (e.g., a mouse or a trackball), by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well. For example, feedback provided to the user can be any form of sensory feedback, (e.g., visual feedback, auditory feedback, or tactile feedback), and input from the user can be received in any form, including acoustic, speech, or tactile input.

[0092] The subject matter described herein can be implemented in a computing system that includes a back end component (e.g., a data server), a middleware component (e.g., an application server), or a front end component (e.g., a client computer having a graphical user interface or a web browser through which a user can interact with an implementation of the subject matter described herein), or any combination of such back end, middleware, and front end components. The components of the system can be interconnected by any form or medium of digital data communication, e.g., a communication network. Examples of communication networks include a local area network ("LAN") and a wide area network ("WAN"), e.g., the Internet.

[0093] It is to be understood that the disclosed subject matter is not limited in its application to the details of construction and to the arrangements of the components set forth

in the following description or illustrated in the drawings. The disclosed subject matter is capable of other embodiments and of being practiced and carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein are for the purpose of description and should not be regarded as limiting.

[0094] As such, those skilled in the art will appreciate that the conception, upon which this disclosure is based, may readily be utilized as a basis for the designing of other structures, methods, and systems for carrying out the several purposes of the disclosed subject matter. It is important, therefore, that the claims be regarded as including such equivalent constructions insofar as they do not depart from the spirit and scope of the disclosed subject matter.

[0095] Although the disclosed subject matter has been described and illustrated in the foregoing exemplary embodiments, it is understood that the present disclosure has been made only by way of example, and that numerous changes in the details of implementation of the disclosed subject matter may be made without departing from the spirit and scope of the disclosed subject matter, which is limited only by the claims which follow.

Claimable material includes the following but is not limited to (human-derived estimates of state of mind for navigating autonomous vehicles):

1. A method comprising:

receiving, by an autonomous vehicle, sensor data from sensors mounted on the autonomous vehicle;

identifying, one or more traffic entities based on the sensor data, the traffic entities representing non-stationary objects in traffic in which the autonomous vehicle is driving;

for each of the one or more traffic entities:

determining one or more motion parameters describing movement of the traffic entity; and

determining a hidden context based on a machine learning based model, wherein the machine learning based model is trained based on feedback received from users responsive to presenting images or videos showing traffic scenarios to the users; and

navigating the autonomous vehicle based on the motion parameters and the hidden context of each of the traffic entities.

2. The method as in any of the preceding claims, wherein the hidden context represents a state of mind of a user represented by the traffic entity.

3. The method as in any of the preceding claims, wherein the hidden context represents a task that a user represented by the traffic entity is planning on accomplishing.

4. The method as in any of the preceding claims, wherein the hidden context represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity.

5. The method as in any of the preceding claims, wherein the hidden context represents a goal of a user represented by the traffic entity, wherein the user expects to achieve the goal within a threshold time interval.

6. The method as in any of the preceding claims, wherein navigating the autonomous vehicle comprises:

generating signals for controlling the autonomous vehicle based on the motion parameters and the hidden context of each of the traffic entities; and

sending the generated signals to controls of the autonomous vehicle.

7. The method as in any of the preceding claims, wherein the sensor data represents one or more images captured by a camera mounted on the autonomous vehicle.

8. The method as in any of the preceding claims, wherein the sensor data represents a lidar scan captured by a lidar mounted on the autonomous vehicle.

9. The method as in any of the preceding claims, wherein navigating the autonomous vehicle comprises ensuring that the autonomous vehicle does not collide with any of the traffic entities.

10. A non-transitory computer readable storage medium storing instructions for performing the steps of methods of claims 1-9.

11. A computer system comprising:

one or more computer processors; and

the non-transitory computer readable storage medium of claim 10.