



(51) International Patent Classification:
B25J 9/16 (2006.01)

(21) International Application Number:
PCT/US2019/051177

(22) International Filing Date:
13 September 2019 (13.09.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/730,703 13 September 2018 (13.09.2018) US
62/730,933 13 September 2018 (13.09.2018) US
62/730,947 13 September 2018 (13.09.2018) US
62/730,934 13 September 2018 (13.09.2018) US
62/730,918 13 September 2018 (13.09.2018) US
62/731,398 14 September 2018 (14.09.2018) US

(71) Applicant: **THE CHARLES STARK DRAPER LABO-
RATORY, INC.** [US/US]; 555 Technology Square, Cam-
bridge, MA 02139 (US).

(72) Inventor; and

(71) Applicant: **TAYOUN, Anthony** [LB/US]; 15 Howard St.,
Unit 2, Cambridge, MA 02139 (US).

(72) Inventors: **JOHNSON, David, M.S.**; 21-1/2 Inman Street,
Apt. 1, Cambridge, MA 02139 (US). **WAGNER, Syler**; 94
Beacon St., Apt. 25, Somerville, MA 02143 (US). **LINES,
Steven**; 24 Williams Street, Brookline, MA 02446 (US).

(74) Agent: **MEAGHER, Timothy, J.** et al.; Hamilton, Brook,
Smith & Reynolds, P.C., 530 Virginia Rd., P.O. Box 9133,
Concord, MA 01742-9133 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

(54) Title: VOICE MODIFICATION TO ROBOT MOTION PLANS

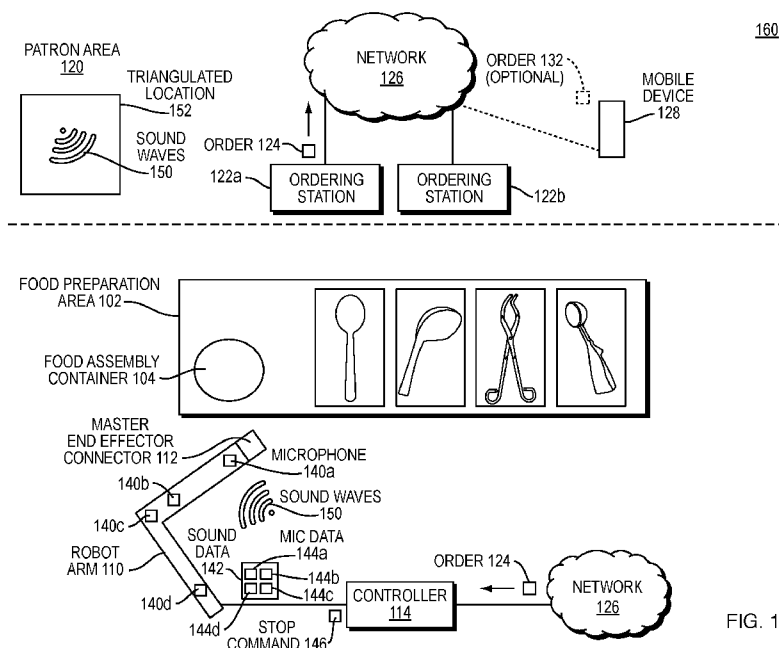


FIG. 1B

(57) Abstract: In an embodiment, a method during execution of a motion plan by a robotic arm includes determining a voice command from speech of a user said during the execution of the motion plan, determining a modification of the motion plan based on the voice command from the speech of the user, and executing the modification of the motion plan by the robotic arm.

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

Voice Modification To Robot Motion Plans

RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 62/730,703, filed on September 13, 2018, U.S. Provisional Application No. 62/730,947, filed on September 13, 2018, U.S. Provisional Application No. 62/730,933, filed on September 13, 2018, U.S. Provisional Application No. 62/730,918, filed on September 13, 2018, U.S. Provisional Application No. 62/730,934, filed on September 13, 2018, U.S. Provisional Application No. 62/731,398, filed on September 14, 2018.

[0002] This application is related to U.S. Patent Application titled “Manipulating Fracturable And Deformable Materials Using Articulated Manipulators”, Attorney Docket No. 5000.1049-001; U.S. Patent Application titled “Food-Safe, Washable, Thermally-Conductive Robot Cover”, Attorney Docket No. 5000.1050-000; U.S. Patent Application titled “Food-Safe, Washable Interface For Exchanging Tools”, Attorney Docket No. 5000.1051-000; U.S. Patent Application titled “An Adaptor for Food-Safe, Bin-Compatible, Washable, Tool-Changer Utensils”, Attorney Docket No. 5000.1052-001; U.S. Patent Application titled “Locating And Attaching Interchangeable Tools In-Situ”, Attorney Docket No. 5000.1053-001; U.S. Patent Application titled “Determining How To Assemble A Meal”, Attorney Docket No. 5000.1054-001; U.S. Patent Application titled “Controlling Robot Torque And Velocity Based On Context”, Attorney Docket No. 5000.1055-001; U.S. Patent Application titled “Stopping Robot Motion Based On Sound Cues”, Attorney Docket No. 5000.1056-000; U.S. Patent Application titled “Robot Interaction With Human Co-Workers”, Attorney Docket No. 5000.1057-001; and U.S. Patent Application titled “One-Click Robot Order”, Attorney Docket No. 5000.1059-000, all of the above U.S. Patent Applications having a first named inventor David M.S. Johnson and all being filed on the same day, September 13, 2019.

[0003] The entire teachings of the above applications are incorporated herein by reference.

BACKGROUND

[0004] Traditionally, the food industry employs human labor to manipulate ingredients with the purpose of either assembling a meal such as a salad or a bowl, or packing a box of

ingredients such as those used in grocery shopping, or preparing the raw ingredients. Robots have not yet been able to assemble complete meals from prepared ingredients in a food-service setting such as a restaurant, largely because the ingredients are arranged unpredictably and change shape in difficult-to-predict ways rendering traditional methods to move material ineffective without extensive modifications to existing kitchens. Additionally, traditional material handling methods are ill-suited to moving cooked foods without altering their texture and taste-profile. These difficulties arise because the friction, stiction, and viscosity of commonly consumed foods cause auger, conveyor, and suction mechanisms to become clogged and soiled, while these mechanisms simultaneously impart forces on the foodstuffs which alter their texture, consistency, and taste-profile in unappetizing ways.

SUMMARY

[0005] Traditionally, robots operate in constrained environments with previously known parameters, and without continuous and seamless interaction with humans. With the advancements of robotics in both hardware and software, robots are increasingly working alongside humans in unconstrained and unpredictable environments. However, there are currently no known methods for humans to interact with robots that are performing tasks by modifying those tasks using natural language in an efficient and reliable manner. Current methods to command robots require tactile input such as with a button or switch and often use additional displays to give the user additional information about the robot's internal state to aid in directing it to modify its current task. Using natural language and the context of the current robot motion allows the user to interact with the robot without requiring detailed knowledge of the robot's current internal state.

[0006] In embodiments, the below disclosure solves problems in relation to employing robotics in the quick service fast food restaurant environment.

[0007] In an embodiment, a method includes, during execution of an action of a group of actions by an autonomous system, determining a given action of the group of actions to modify based on user input received during the execution of the group of actions. The method further includes, during execution of the action of the group of actions by the autonomous system, modifying the given action of the group of actions based on the user input. The method further includes, during execution of the action of the group of actions by the autonomous system, executing, by the autonomous system, the group of actions modified based on the user input.

[0008] In an embodiment, modifying the given action further can include: (a) determining a new action as the given action, (b) modifying a quantity of material associated with the given action, (c) removing the given action from the plurality of actions, (d) changing a type of material associated with a given action, and (e) repeating a previous action of the plurality of actions. Modifying the given action can further include calculating a transition motion plan to align the autonomous system with a pose within (a) the new motion plan and (b) the previously unselected motion plan.

[0009] A person having ordinary skill in the art can recognize that an autonomous system can be a robot or robotic arm, as illustrated in Fig. 1 and described below.

[0010] In an embodiment, the user input is speech, a gesture as detected by a camera, or an input from a control.

[0011] In an embodiment, the group of actions is a sequence of ordered actions.

[0012] In an embodiment, the method further includes, during execution of the action of the group of actions by the autonomous system, notifying the user with information regarding the given action modified either through a visual display, an audio alert, or indication via motion of the autonomous system.

[0013] In an embodiment, the given action of the group of actions has been previously executed by the autonomous system. The method further includes, during execution of the action of the group of actions by the autonomous system, generating one or more modified actions to undo the previously executed given action.

[0014] In an embodiment, a system includes a processor and a memory with computer code instructions stored thereon. The processor and the memory, with the computer code instructions, are configured to cause the system to, during execution of an action of a plurality of actions by an autonomous system, determine a given action of the plurality of actions to modify based on user input received during the execution of the plurality of actions. The instructions are further configured to, during execution of the action of the group of actions by the autonomous system, modify the given action of the plurality of actions based on user input. The instructions are further configured to, during execution of the action of the group of actions by the autonomous system, execute, by the autonomous system, the plurality of actions modified based on the user input.

[0015] In an embodiment, a method includes, during execution of an action of a plurality of actions by an autonomous system, assigning a confidence weight to the plurality of actions based on a received user input. The method further includes, during execution of an action of

a plurality of actions by an autonomous system, modifying a model or neural network that generates a second plurality of actions based on the confidence weight.

[0016] In an embodiment, the user input is positive, neutral, or negative feedback in relation to at least one of the plurality of actions.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] The foregoing will be apparent from the following more particular description of example embodiments, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating embodiments.

[0018] Fig. 1A is a block diagram illustrating an example embodiment of a quick service food environment 100 of embodiments of the present invention.

[0019] Fig. 1B is a block diagram illustrates an example embodiment of the present disclosure.

[0020] Fig. 2 is flow diagram illustrating an embodiment of a process of the present disclosure.

[0021] Fig. 3A is a state diagram illustrating an embodiment of robot decision making.

[0022] Fig. 3B is a flow diagram illustrating an example embodiment of a process employed by the present disclosure.

[0023] Fig. 4A is a state diagram illustrating an example embodiment of a robot decision making process according to the present disclosure.

[0024] Fig. 4B is a flow diagram illustrating an example embodiment of a process employed by the present disclosure.

[0025] Fig. 5 illustrates a computer network or similar digital processing environment in which embodiments of the present invention may be implemented.

[0026] Fig. 6 is a diagram of an example internal structure of a computer (e.g., client processor/device or server computers) in the computer system of Fig. 5.

DETAILED DESCRIPTION

[0027] A description of example embodiments follows.

[0028] Operating a robot in a food preparation environment, such as a quick service restaurant, can be challenging for several reasons. First, the end effectors (e.g., utensils), that the robot uses need to remain clean from contamination. Contamination can include allergens

(e.g., peanuts), dietary preferences (e.g., contamination from pork for a vegetarian or kosher customer), dirt/bacteria/viruses, or other non-ingestible materials (e.g., oil, plastic, or particles from the robot itself). Second, the robot should be operated within its design specifications, and not exposed to excessive temperatures or incompatible liquids, without sacrificing cleanliness. Third, the robot should be able to manipulate food stuffs, which are often fracturable and deformable materials, and further the robot must be able to measure an amount of material controlled by its utensil in order to dispense specific portions. Fourth, the robot should be able to automatically and seamlessly switch utensils (e.g., switch between a ladle and salad tongs). Fifth, the utensils should be adapted to be left in an assigned food container and interchanged with the robot as needed, *in situ*. Sixth, the interchangeable parts (e.g., utensils) should be washable and dishwasher safe. Seventh, the robot should be able to autonomously generate a task plan and motion plan(s) to assemble all ingredients in a recipe, and execute that plan. Eighth, the robot should be able to modify or stop a motion plan based on detected interference or voice commands to stop or modify the robot's plan. Ninth, the robot should be able to minimize the applied torque based on safety requirements or the task context or the task parameters (e.g., density and viscosity) of the material to be gathered. Tenth, the system should be able to receive an electronic order from a user, assemble the meal for the user, and place the meal for the user in a designated area for pickup automatically with minimal human involvement.

[0029] Fig. 1 is a block diagram illustrating an example embodiment of a quick service food environment 100 of embodiments of the present disclosure. The quick service food environment 100 includes a food preparation area 102 and a patron area 120.

[0030] The food preparation area 102 includes a plurality of ingredient containers 106a-d each having a particular foodstuff (e.g., lettuce, chicken, cheese, tortilla chips, guacamole, beans, rice, various sauces or dressings, etc.). Each ingredient container 106a-d stores *in situ* its corresponding ingredients. Utensils 108a-d may be stored *in situ* in the ingredient containers or in a stand-alone tool rack 109. The utensils 108a-d can be spoons, ladles, tongs, dishers (scoopers), spatulas, or other utensils. Each utensil 108a-e is configured to mate with and disconnect from a tool changer interface 112 of a robot arm 110. While the term utensil is used throughout this application, a person having ordinary skill in the art can recognize that the principles described in relation to utensils can apply in general to end effectors in other contexts (e.g., end effectors for moving fracturable or deformable materials in construction with an excavator or backhoe, etc.); and a robot arm can be replaced with any computer

controlled actuatable system which can interact with its environment to manipulate a deformable material. The robot arm 110 includes sensor elements/modules such as stereo vision systems (SVS), 3D vision sensors (e.g., Microsoft Kinect™ or an Intel RealSense™), LIDAR sensors, audio sensors (e.g., microphones), inertial sensors (e.g., internal motion unit (IMU), torque sensor, weight sensor, etc.) for sensing aspects of the environment, including pose (i.e., X, Y, Z coordinates and roll, pitch, and yaw angles) of tools for the robot to mate, shape and volume of foodstuffs in ingredient containers, shape and volume of foodstuffs deposited into food assembly container, moving or static obstacles in the environment, etc.

[0031] To initiate an order, a patron in the patron area 120 enters an order 124 in an ordering station 122a-b, which is forwarded to a network 126. Alternatively, a patron on a mobile device 128 can, within or outside of the patron area 120, generate an optional order 132. Regardless of the source of the order, the network 126 forwards the order to a controller 114 of the robot arm 110. The controller generates a task plan 130 for the robot arm 110 to execute.

[0032] The task plan 130 includes a list of motion plans 132a-d for the robot arm 110 to execute. Each motion plan 132a-d is a plan for the robot arm 110 to engage with a respective utensil 108a-e, gather ingredients from the respective ingredient container 106a-d, and empty the utensil 108a-e in an appropriate location of a food assembly container 104 for the patron, which can be a plate, bowl, or other container. The robot arm 110 then returns the utensil 108a-e to its respective ingredient container 106a-d, the tool rack 109, or other location as determined by the task plan 130 or motion plan 132a-d, and releases the utensil 108a-d. The robot arm executes each motion plan 132a-d in a specified order, causing the food to be assembled within the food assembly container 104 in a planned and aesthetic manner.

[0033] Within the above environment, various of the above described problems can be solved. The environment 100 illustrated by Fig. 1 can improve food service to patrons by assembling meals faster, more accurately, and more sanitarily than a human can assemble a meal. Some of the problems described above can be solved in accordance with the disclosure below.

[0034] Manipulation of deformable materials and movable objects can be challenging for autonomous systems. Automatically generated plans and motions may be sub-optimal, and therefore improvable as defects become apparent to human or other operators. Therefore, embodiments of this disclosure solve the problem of modifying the actions (e.g., a task plan or motion plan) of an autonomous system as the actions are in progress, without requiring a

full stop or restart, and without requiring an operator to intervene using a mechanical input device such as a button, keyboard, mouse, or motion tracker.

[0035] For example, when using a robot arm to scoop ice cream, the disher tool might not dig deep enough to extract a full scoop, or the viscosity of the ice cream may cause it to fall out of or otherwise escape the tool. In another example, when shoveling powdery snow, the snow may bounce out of a moving shovel or be blown by wind. In another example, when using a bulldozer to clear debris, the debris may be moved by other forces or is constrained by attachments, such as cabling or rebar. In another example, when manipulating food stuffs of variable consistencies, the food stuffs may not be manipulated as anticipated. In another example, when picking up moving objects, the movement of the object interacts or interferes with an end-effector of a robot.

[0036] The approach of the present disclosure enables a robot to autonomously select and manipulate material or objects that may change in ways that may be difficult to predict by computational modeling or other autonomous prediction. Therefore, this disclosure enables these plans for selecting and manipulating material or objects to be subject to subsequent guidance and plan-modification via voice input during the manipulation. For example, when scooping ice cream, if self-adhesion begins pulling a partial scoop out of the right side of the disher tool, the scoop may be saved by altering the disher's trajectory so that it veers rightward. For a human scooping ice cream, it is easy to recognize this problem and remedy it as it occurs. In robotics, current state of the art does not solve this problem on a pure automation level. Therefore, human intervention via human speech can improve the robotic system by recognizing voice commands in real-time, such as the command “move a little to the right.” Therefore, voice commands can thus improve a robot system's performance in such tasks by providing an interface for input from a human for adjustments that a human can better determine.

[0037] Additionally, these verbal signals may be used to infer quality or corrective action which can serve as data for training a neural network. As an example, while the robot is performing an action, a user saying positive feedback like “yes” or “good” can train the neural network that the action being performed is desired given the order, the robot's environment, state of ingredients, etc. Conversely, a user saying negative feedback like “no,” “bad,” or “stop” can train the neural network that the action being performed is not desired given the order, the robot's environment, state of ingredients, etc.

[0038] Plans and trajectories for the automatic manipulation of materials and objects can be difficult to compute and costly to execute. Robotic actions based on heuristics and approximations may be improvable in real-time by way of recognized speech commands or commentary. Because materials and objects change during manipulation, actual shapes and other properties of the materials and objects may deviate from expectations. Before manipulation, the system stores a collection of possible motion plans with many variations and sequences of interchangeable components, as well as collection of recognizable speech commands and modifiers, as in verbs and adverbs. Mapping commands to plan components allows speech commands, when recognized by the system, to generate changes in the execution of motion and manipulation plans while in-progress. For example, a plan in progress may include torque limits and a trajectory including a sequence of points in a multidimensional space. This is further described in both U.S. Patent Application titled “Manipulating Fracturable And Deformable Materials Using Articulated Manipulators”, Attorney Docket No. 5000.1049-001 and U.S. Patent Application titled “Controlling Robot Torque And Velocity Based On Context”, Attorney Docket No. 5000.1055-001.

[0039] Upon recognizing the commands “turn right, softer”, or “softer, rightward”, the system decreases the torque limits and turns the current trajectory in the specified relative direction (e.g., by geometrics transformation, or by splicing into a different saved trajectory). In addition to specific commands, any other speech input may be recorded and used as input for more general purposes, such as a reinforcement learning system.

[0040] Based on the current state and executing plan, any recognized speech command is mapped to possible changes of state and plan. Current observed conditions and states may also inform the automatic speech recognition system. Recognized speech can be treated as an event and fed into a finite state machine (FSM) or other system controller. Some events may of course have no effect. In other words, the robotic system uses motion and physical feedback as inputs to determine the outside world and adjust in real time. However, in embodiments of the present disclosure, the system further includes voice or acoustic input in combination with other inputs.

[0041] Speech input may be processed in current time or later, offline, as potential input for automatic or other improvements to the system. For example, a comment such as “that's a good scoop” may be correlated with the measured result of the action and used as input for a machine learning system. To determine which speech may be used as command and/or comment input, the system may use some combination of wake-up word recognition,

command-prefix recognition, and speaker recognition (e.g., to identify an operator, as opposed to an on-looker).

[0042] In an embodiment, the disclosure employs a speech sensing and recognition apparatus (e.g., a microphone array connected to a computer or controller) in conjunction with a robot system (e.g., a jointed arm, end-effector, and connected cameras) which is able to manipulate the chosen materials or objects.

[0043] The robotic system (e.g., a controller) determines the current state of the world it can modify or change. The controller computes trajectories and other plan components for manipulating materials and objects in that world. The speech recognition system feeds additional events as input to the system controller, and the controller can modify the trajectories and other plan components (e.g., task plan, motion plan) based on all combined inputs. The robot executes the plan components, which the system controller monitors and corrects as they are executed. An operator voices commands and comments during execution (closed loop behavior) or commentary-only on the result of an executed plan (open loop). During execution, recognized commands become events fed into the controller, which may change plans and their executions in current time. Commentary may be used later to improve the system, as in becoming input to machine learning.

[0044] Several alternative methods can be used. For example, brute force with no operator intervention is an alternative. Modeling deformable materials and moveable objects requires a multi-dimensional state space and is subject to combinatorial explosion. Attempting to limit the size of the state space using heuristics and pruning is difficult, and may be intractable in some domains of interest, such as scooping ice cream or assembling salads.

[0045] In an alternate approach, an operator may intervene using a keyboard, pointer, or other mechanical input device. However, this means that operators must remain near the keyboard or other device, or use their hands. Automated speech recognition allows hands-free mobility.

[0046] In an alternate approach, the system can allow re-tries/repeated attempts after failed plan executions. Instead of fixing failed plans or handling unforeseen contingencies in real-time, plans may be executed to completion, upon which the system or operator evaluates the result as deficient, and the system executes another plan. This increases costs of both time and materials.

[0047] In a known system ASR is used to input commands to robots, but does not disclose changing complex plans during execution.

[0048] Typically, interactive robots execute simple commands without intervention, and robots that execute long-running plans, such as painting a car chassis, are not interactive.

[0049] Applicant's approach can use a combination of automatic speech recognition, search, neural nets, and model-based force control to accomplish the desired results. No previous approach combined these technologies to provide a better result of allowing real-time human input to be considered as an input to allow modification of a robot's path during operation.

[0050] Manipulation of deformable and granular materials using a robot can be difficult to plan, execute, and change while in progress. Failures may result in loss of material. Enabling operator intervention through a natural language like interface can thus save both time and materials.

[0051] This approach is meant to allow human experience to improve robot system performance both in real-time and in offline machine learning. Implementing a feedback loop between automatic speech recognition and semi-autonomous performance of complex tasks may support improvements in both, in these specific domains.

[0052] Fig. 1B is a block diagram 160 illustrates an example embodiment of the present disclosure. In a similar environment as Fig 1A, an array of several microphones 140a-d are mounted on the robot arm 110. The microphones 140a-d are configured to detect and record sound waves 142. As the microphones 140a-d record the sound data 142, it is reported to a controller 114. The sound data 142 can be organized into data from individual microphones as mic data 144a-d. The controller can process the sound data 142 and if a trigger word is detected (e.g., stop or ouch) then it can issue a stop command 146.

[0053] Fig. 1B further illustrates sound waves 150 beginning from the patron area 120. With the multiple microphones 140a-d, the controller 114 can determine a triangulated location 152 of the sound waves 150, and provide the sound waves 150 a lower weight or ignore the sound waves 150 entirely because the determined triangulated location 152 is within the patron area 120.

[0054] Fig. 2 is flow diagram 200 illustrating an embodiment of a process of the present disclosure. The process illustrated by Fig. 2 may be implemented by a processor executing computer program code in combination with the appropriate hardware (e.g., memory,

robotics, etc.). The computer program code may be stored on storage media, or may be transferred to a workstation over the Internet or some other type of network.

[0055] After initialization (200), the process connects sound capturing devices to the system (205). The sound capturing devices can, for example, be a microphone or an array of microphones. In embodiments, the array of microphones is used to detect the source (e.g., 2-dimensional or 3-dimensional location) of the sound.

[0056] The process then builds a keyword and command database (210). The keyword and command database is indexed and searchable by different parameters such as keywords, etc. The keyword and command database can either be built through a computer software that copies pre-defined keywords and sounds, by a live recording of a sound or keywords narrated by a human speaker, or through any other simulation of the keywords or sounds (e.g., a generation function training a neural network). In an embodiment, the process can dynamically update the keyword and command database based on self-generated feedback by the system or manually inputted feedback from a user to a particular recording.

[0057] The process then defines a set of rules and associates those rules with different keywords and commands (211). The rules can be pre-defined and copied via a computer software or can change dynamically based on user input. The rules can also dynamically change based on feedback captured by the overall system.

[0058] The process then connects the robot to the system (215). The robot may provide information to the overall system, such as motion data, image capture data, or other sensor output.

[0059] The robot then executes/performs the plan (216). The plan can be run by the same software described above or by an independent compatible software that is previously programmed.

[0060] The process then monitors sounds (220). The process checks whether a keyword and/or phrase and/or command has been detected by matching measured information with the pre-built database (225). The checking/determining may involve receiving words and phrases from a system that processes this data, the processing including removing noise or performing other transformations to the data. If no match occurs, then the monitoring process continues (220). However, if a match is detected, the system executes and processes the rule or rules (230). These rules can be executed by the robot, such as to change the robot's plan (216). Optionally, the rules can be dynamically updated based on the fact that this rule or

rules have been executed (not shown). Optionally, the database of keywords and commands can be updated based on the fact that this rule or rules have been executed (not shown).

[0061] Fig. 3A is a state diagram 300 illustrating an embodiment of robot decision making. Traditionally, upon a request for modification, the plan is cancelled and begun again from the beginning. The state begins with creating a plan 302 for the robot. After the plan is created, the robot executes the action 304. However, upon new inputs, the action can be reset back to the planning stage 302. When the actions are completed, the robot enters a serve state 306, which resets the state diagram back to the planning state 302 for a new order.

[0062] Fig. 3B is a flow diagram illustrating an example embodiment of a process employed by the present disclosure. The process first creates a plan (352). The process then executes the plan (354) while monitoring sounds and sensory input (356). If the process determines that the monitored sounds and sensory inputs require modification to the plan (358), the system returns to create the plan again (352). Otherwise, the system finishes the plan (360), and continues to create another plan for a new order, if necessary (352).

[0063] Fig. 4A is a state diagram 400 illustrating an example embodiment of a robot decision making process according to the present disclosure. After providing an initial plan 402, during execution of the plan 404 and upon additional input, such as voice modification, the system transforms the plan, or interpolates a new plan 408. Such a modification is possible in food service and certain robotics applications because there is enough time for a human to notice a change, and further ability for the system to act in time. For example, intelligent voice assistants are bots that cannot be modified. Other robotic tasks are things humans cannot understand as they are computed, such as sensor fusion. However, ice cream scooping, or other food applications, are slow enough that a human can intervene and positively affect the outcome of the scooping.

[0064] Upon voice input, there are two possibilities for changing the plan. First, the system can transform the plan, or the system can interpolate to a new plan. If transforming the plan, the system can calculate, as the current plan continues, a new plan from the current position, velocity, torque, state of the food container, and amount of food in the container. Alternatively, the system could transition to a previously calculated motion plan that had not been selected. For example, an unselected motion plan may have moved the disher to the right by a number of degrees. That unselected plan can be selected, and a transition motion plan can be calculated to move the tool in position to execute that plan.

[0065] Fig. 4B is a flow diagram 400 illustrating an example embodiment of a process employed by the present disclosure. The process first creates a plan (452). The process then executes the plan (454) while monitoring sounds and sensory input (456). In this embodiment, if the process determines that the monitored sounds and sensory inputs require modification to the plan (458), the system modifies the plan (460) instead of creating a new one, and then executes the modified plan (454). Otherwise, the system finishes the plan (462), and continues to create another plan for a new order, if necessary (452).

[0066] Fig. 5 illustrates a computer network or similar digital processing environment in which embodiments of the present invention may be implemented.

[0067] Client computer(s)/devices 50 and server computer(s) 60 provide processing, storage, and input/output devices executing application programs and the like. The client computer(s)/devices 50 can also be linked through communications network 70 to other computing devices, including other client devices/processes 50 and server computer(s) 60. The communications network 70 can be part of a remote access network, a global network (e.g., the Internet), a worldwide collection of computers, local area or wide area networks, and gateways that currently use respective protocols (TCP/IP, Bluetooth®, etc.) to communicate with one another. Other electronic device/computer network architectures are suitable.

[0068] Fig. 6 is a diagram of an example internal structure of a computer (e.g., client processor/device 50 or server computers 60) in the computer system of Fig. 5. Each computer 50, 60 contains a system bus 79, where a bus is a set of hardware lines used for data transfer among the components of a computer or processing system. The system bus 79 is essentially a shared conduit that connects different elements of a computer system (e.g., processor, disk storage, memory, input/output ports, network ports, etc.) that enables the transfer of information between the elements. Attached to the system bus 79 is an I/O device interface 82 for connecting various input and output devices (e.g., keyboard, mouse, displays, printers, speakers, etc.) to the computer 50, 60. A network interface 86 allows the computer to connect to various other devices attached to a network (e.g., network 70 of Fig. 5). Memory 90 provides volatile storage for computer software instructions 92 and data 94 used to implement an embodiment of the present invention (e.g., plan creation module, plan modification module, and environment monitoring module code detailed above). Disk storage 95 provides non-volatile storage for computer software instructions 92 and data 94

used to implement an embodiment of the present invention. A central processor unit 84 is also attached to the system bus 79 and provides for the execution of computer instructions.

[0069] In one embodiment, the processor routines 92 and data 94 are a computer program product (generally referenced 92), including a non-transitory computer-readable medium (e.g., a removable storage medium such as one or more DVD-ROM's, CD-ROM's, diskettes, tapes, etc.) that provides at least a portion of the software instructions for the invention system. The computer program product 92 can be installed by any suitable software installation procedure, as is well known in the art. In another embodiment, at least a portion of the software instructions may also be downloaded over a cable communication and/or wireless connection. In other embodiments, the invention programs are a computer program propagated signal product embodied on a propagated signal on a propagation medium (e.g., a radio wave, an infrared wave, a laser wave, a sound wave, or an electrical wave propagated over a global network such as the Internet, or other network(s)). Such carrier medium or signals may be employed to provide at least a portion of the software instructions for the present invention routines/program 92.

[0070] The teachings of all patents, published applications and references cited herein are incorporated by reference in their entirety.

[0071] While example embodiments have been particularly shown and described, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the embodiments encompassed by the appended claims.

CLAIMS

What is claimed is:

1. A method comprising:
 - during execution of an action of a plurality of actions by an autonomous system:
 - determining a given action of the plurality of actions to modify based on user input received during the execution of the plurality of actions;
 - modifying the given action of the plurality of actions based on the user input; and
 - executing, by the autonomous system, the plurality of actions modified based on the user input.
2. The method of Claim 1, wherein modifying the given action further includes one or more of: determining a new action as the given action, modifying a quantity of material associated with the given action, removing the given action from the plurality of actions, changing a type of material associated with a given action, and repeating a previous action of the plurality of actions.
3. The method of Claim 2, wherein modifying the given action further includes calculating a transition motion plan to align the autonomous system with a pose within one or more of: the new motion plan and the previously unselected motion plan.
4. The method of Claim 1, wherein the user input is at least one of speech, a gesture as detected by a camera, and using a control.
5. The method of Claim 1, wherein the plurality of actions is a sequence of ordered actions.
6. The method of Claim 1, further comprising notifying the user with information regarding the given action modified either through a visual display, an audio alert, or indication via motion of the autonomous system.
7. The method of Claim 1, wherein,

the given action of the plurality of actions has been previously executed by the autonomous system; and further comprising:

generating at least one modified action to undo the previously executed given action.

8. A system comprising:

a processor; and

a memory with computer code instructions stored thereon, the processor and the memory, with the computer code instructions, being configured to cause the system to:

during execution of an action of a plurality of actions by an autonomous system:

determine a given action of the plurality of actions to modify based on user input received during the execution of the plurality of actions;

modify the given action of the plurality of actions based on user input;

and

execute, by the autonomous system, the plurality of actions modified based on the user input.

9. The system of Claim 8, wherein modifying the given action further includes one or more of:

determining a new action as the given action, modifying a quantity of material associated with the given action, removing the given action from the plurality of actions, changing a type of material associated with a given action, and repeating a previous action of the plurality of actions.

10. The system of Claim 9, wherein modifying the given action further includes calculating a transition motion plan to align the autonomous with a pose within one or more of: the new motion plan and the previously unselected motion plan.

11. The system of Claim 8, wherein the user input is at least one of speech, a gesture, and using a control.

12. The system of Claim 8, wherein the plurality of actions is a sequence of ordered actions.

13. The system of Claim 8, wherein the computer code further causes the processor to:
notify the user with information regarding the given action modified either through a visual display, an audio alert, or indication via motion of the autonomous system.
14. The system of Claim 8, wherein the given action of the plurality of actions has been previously executed by the autonomous system; and
wherein the computer code is further configured to generate at least one modified action to undo the previously executed given action.
15. A method comprising:
during execution of an action of a plurality of actions by an autonomous system:
assigning a confidence weight to the plurality of actions based on a received user input; and
modifying a model or neural network that generates a second plurality of actions based on the confidence weight.
16. The method of Claim 15, wherein the user input is positive, neutral, or negative feedback in relation to at least one of the plurality of actions.

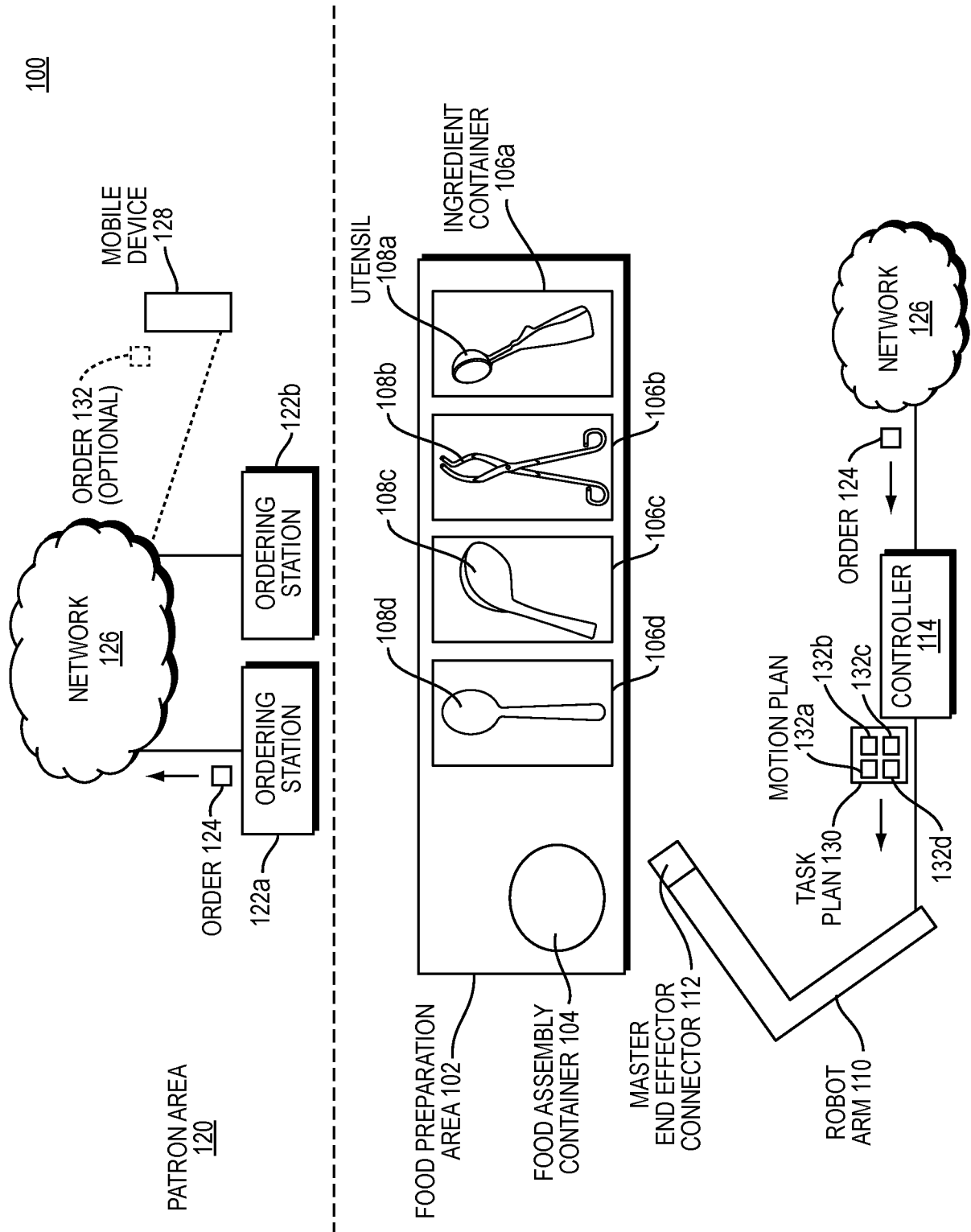
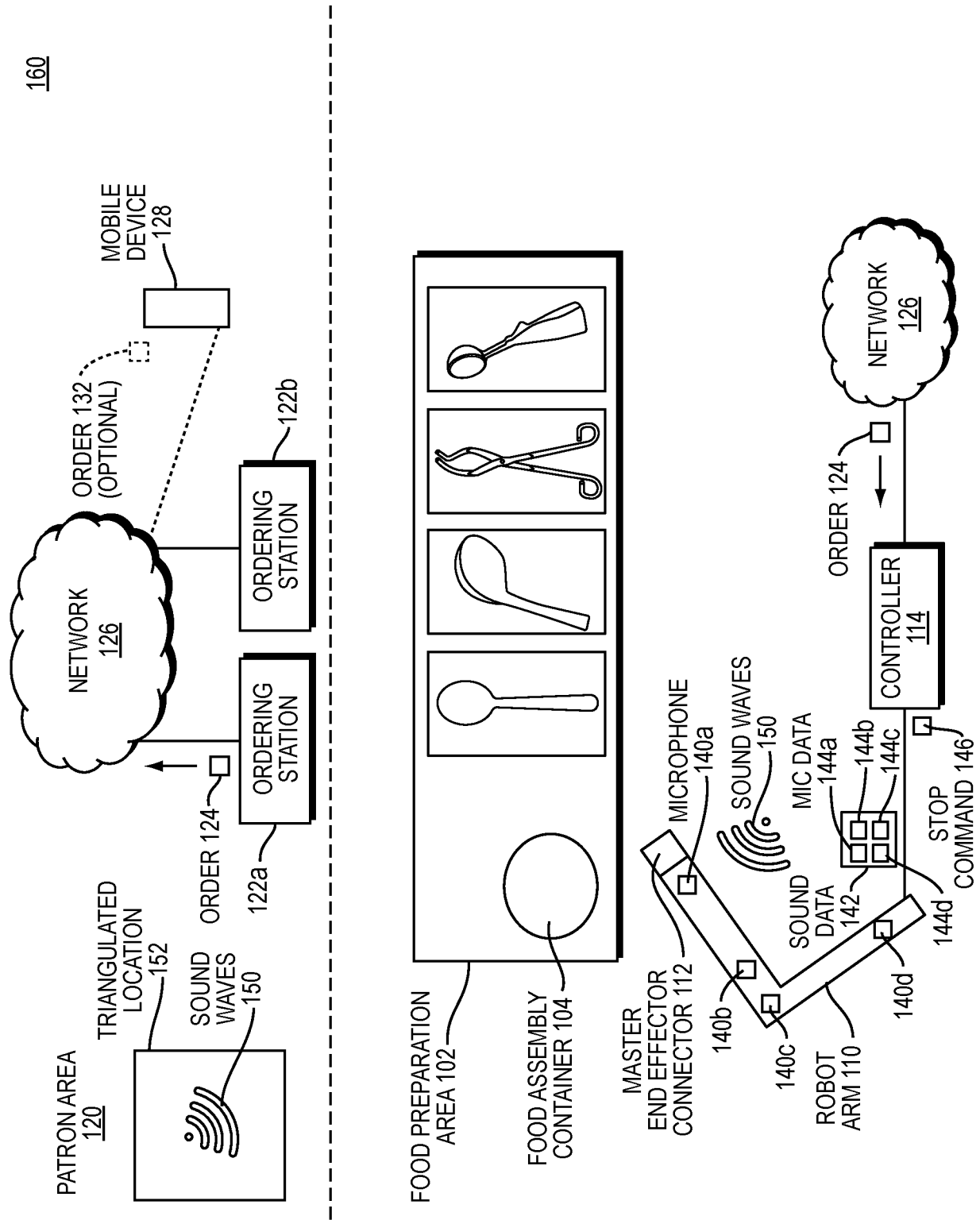


FIG. 1A



3/9

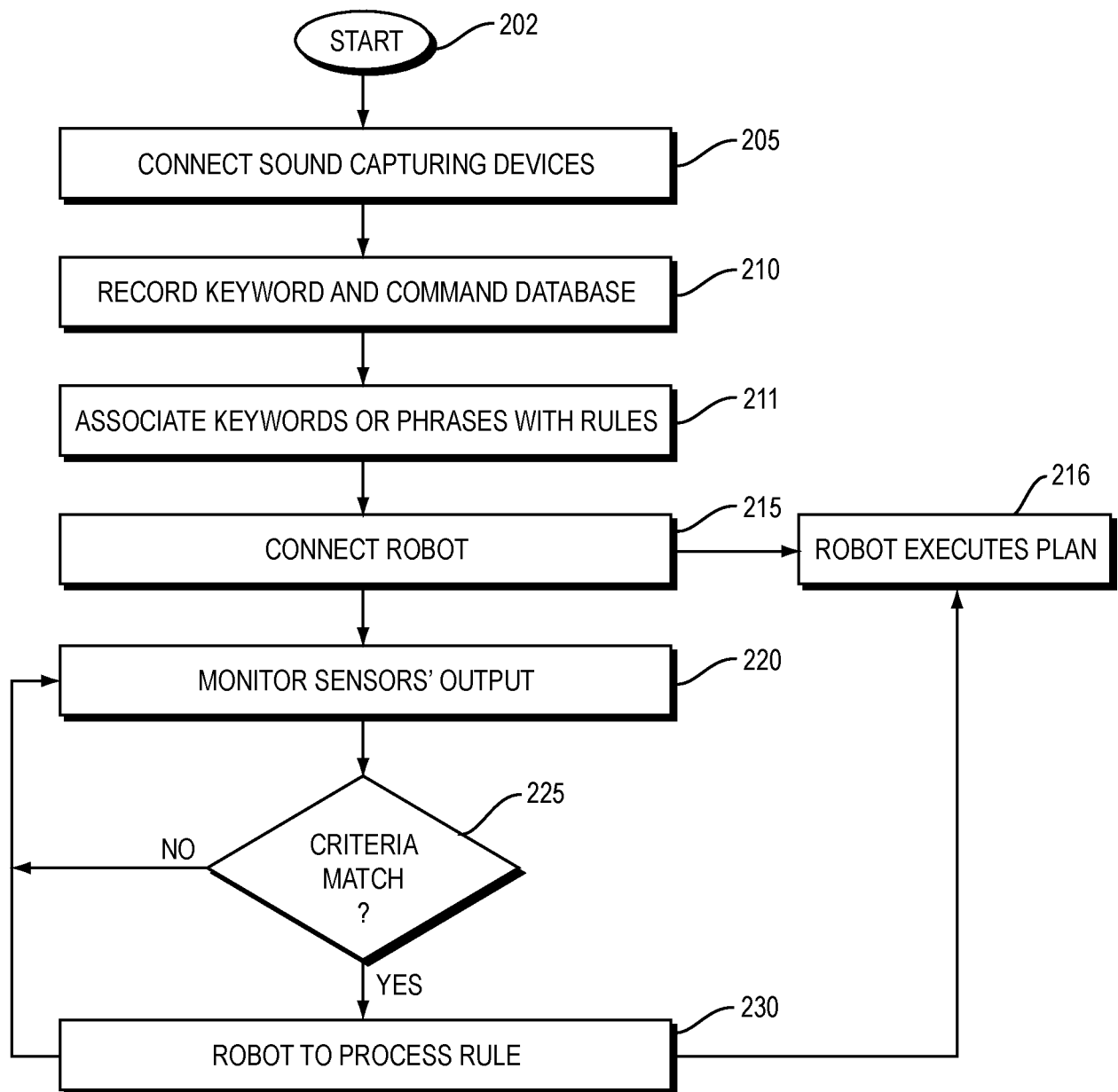
200

FIG. 2

300

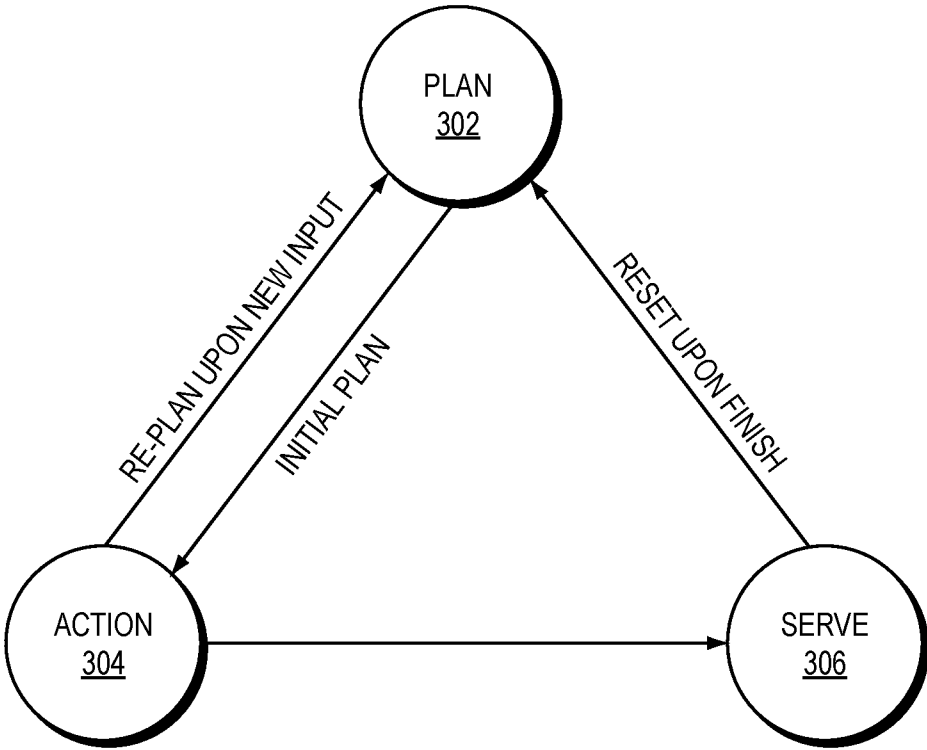


FIG. 3A

5/9

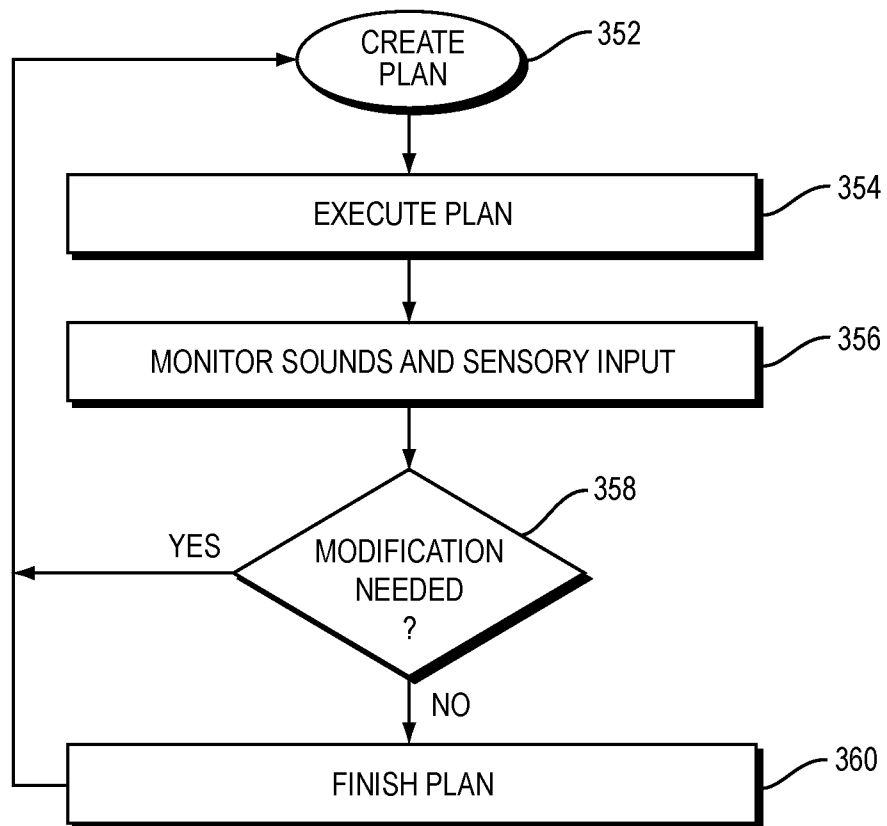
350

FIG. 3B

400

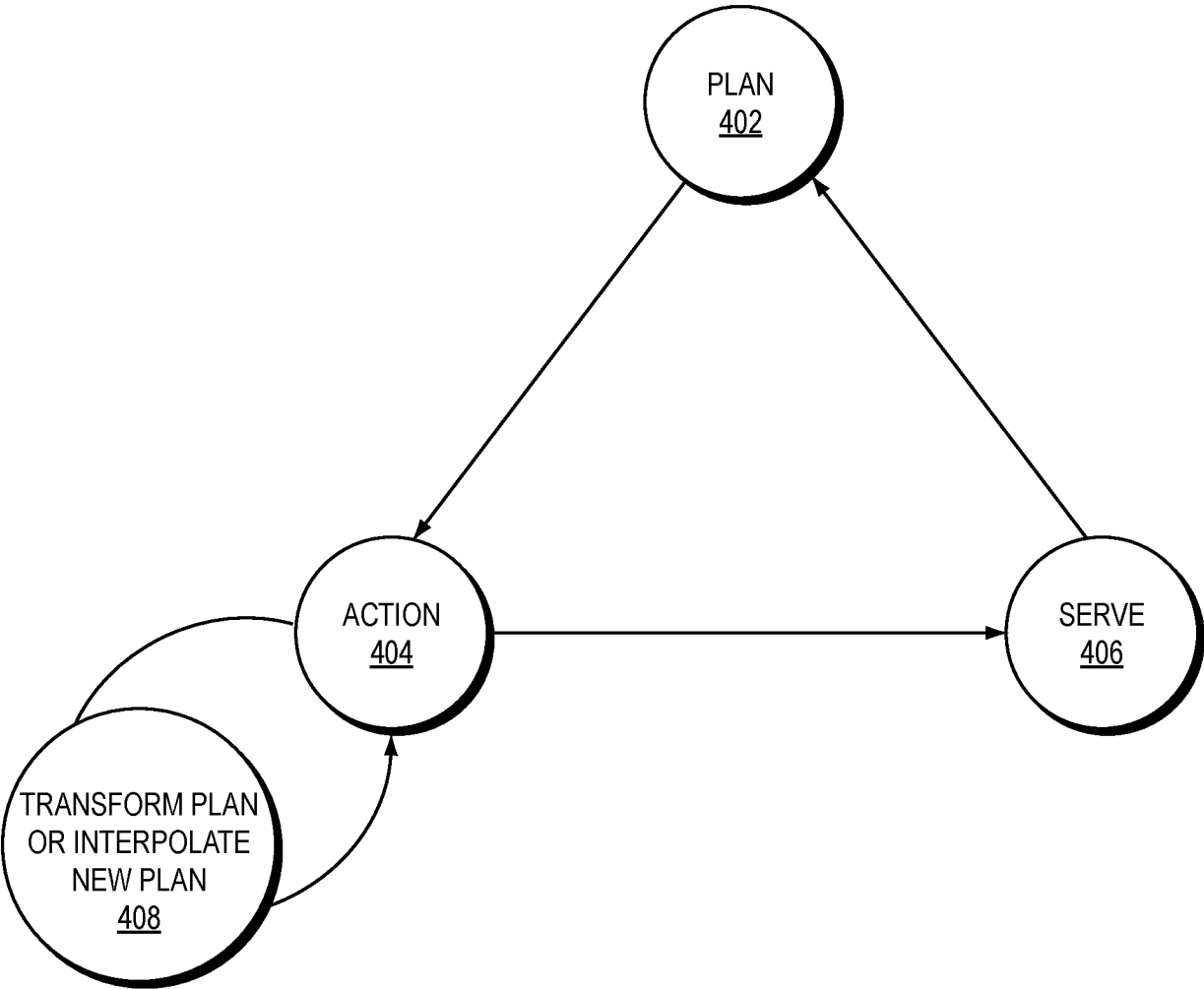


FIG. 4A

7/9

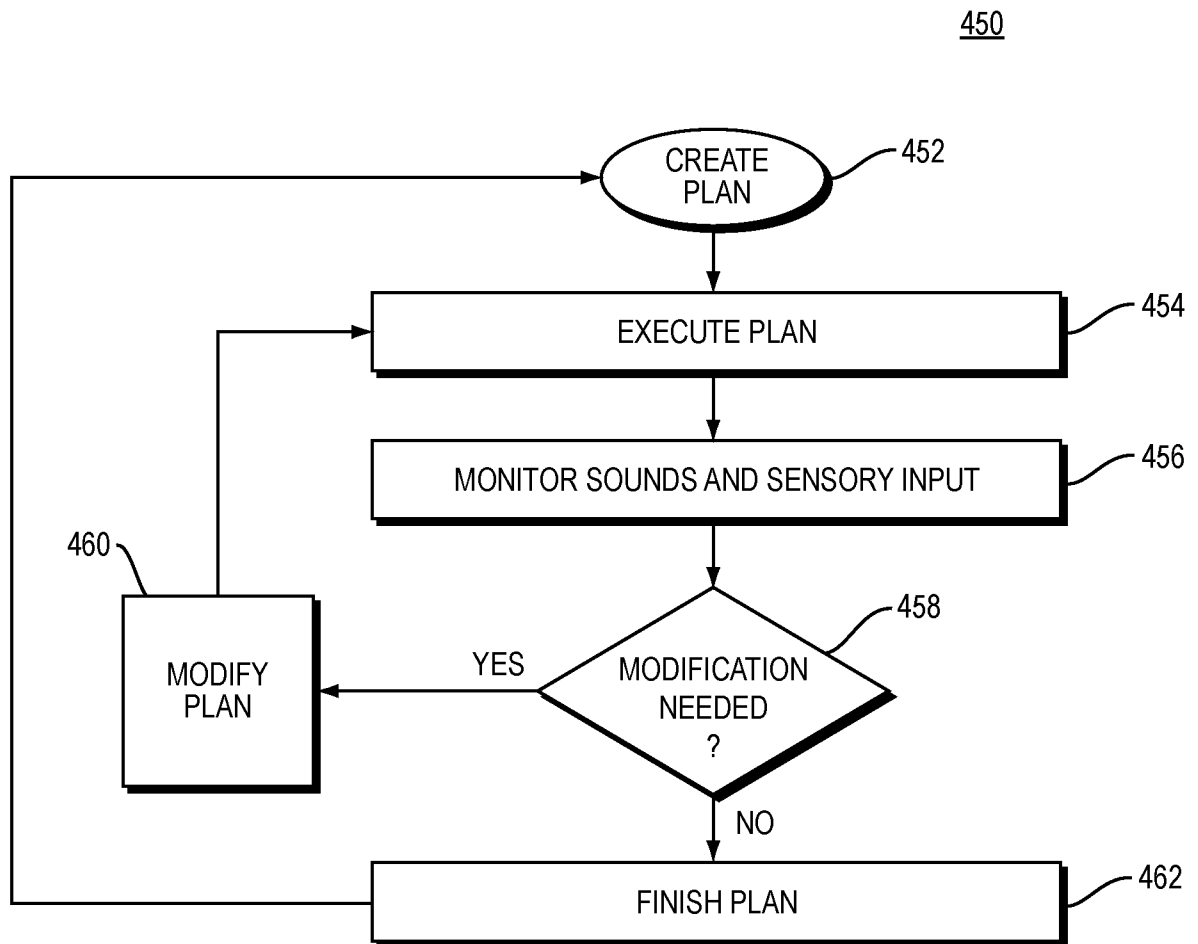


FIG. 4B

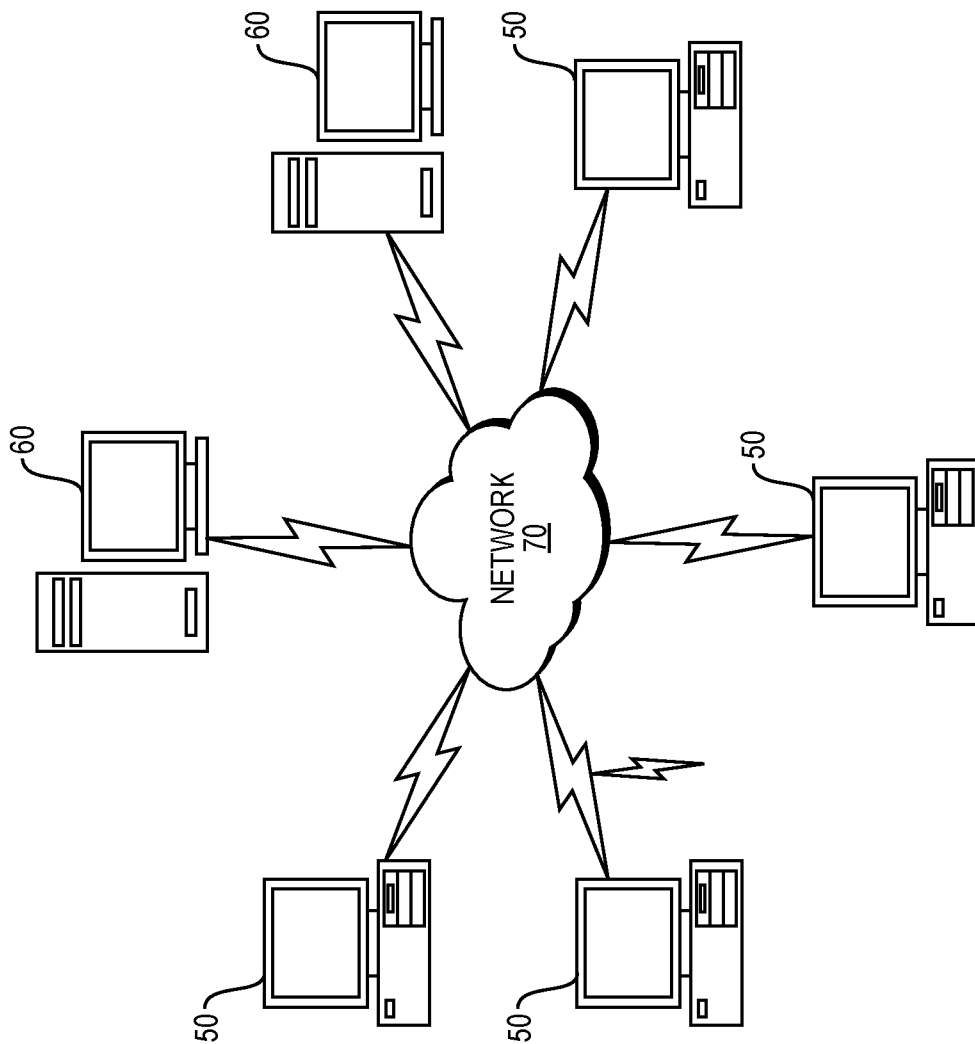


FIG. 5

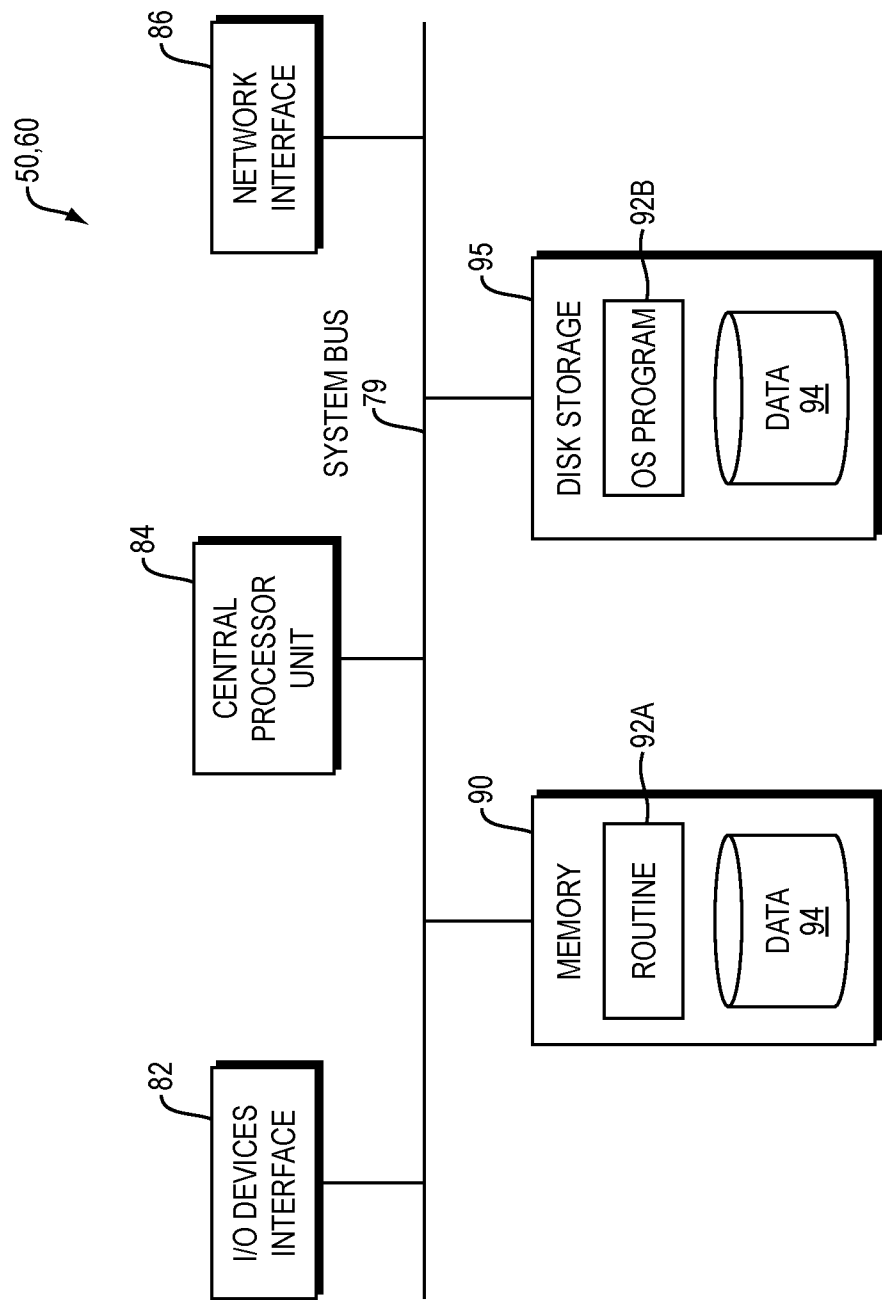


FIG. 6

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2019/051177

A. CLASSIFICATION OF SUBJECT MATTER
INV. B25J9/16
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G05B B25J

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DE 38 23 102 A1 (SIEMENS AG [DE]) 11 January 1990 (1990-01-11) column 1, line 66 - column 2, line 4 column 3, line 21 - line 21 -----	1-16
X	US 5 136 223 A (KARAKAMA TATSUO [JP] ET AL) 4 August 1992 (1992-08-04) column 3, line 24 - line 66 -----	1-16
X	EP 1 145 804 A1 (MATSUSHITA ELECTRIC IND CO LTD [JP]) 17 October 2001 (2001-10-17) paragraph [0008] - paragraph [0017] paragraph [0031] - paragraph [0056] -----	1-16
X	US 2018/056520 A1 (OZAKI SHUNICHI [JP] ET AL) 1 March 2018 (2018-03-01) paragraph [0045] - paragraph [0071] -----	1-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

20 December 2019

Date of mailing of the international search report

09/01/2020

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Hauser, Leon

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2019/051177

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
DE 3823102	A1	11-01-1990	NONE

US 5136223	A	04-08-1992	DE 68926292 D1 23-05-1996
			DE 68926292 T2 19-09-1996
			EP 0400154 A1 05-12-1990
			JP H02160487 A 20-06-1990
			US 5136223 A 04-08-1992
			WO 9006836 A1 28-06-1990

EP 1145804	A1	17-10-2001	EP 1145804 A1 17-10-2001
			JP 4122652 B2 23-07-2008
			JP 2001088071 A 03-04-2001
			US 6522949 B1 18-02-2003
			WO 0123149 A1 05-04-2001

US 2018056520	A1	01-03-2018	CN 107767760 A 06-03-2018
			DE 102017007729 A1 29-03-2018
			JP 6517762 B2 22-05-2019
			JP 2018030185 A 01-03-2018
			US 2018056520 A1 01-03-2018
