



(51) International Patent Classification:

B25J 13/08 (2006.01)

B25J 19/02 (2006.01)

B25J 9/16 (2006.01)

(21) International Application Number:

PCT/KR2024/095301

(22) International Filing Date:

16 February 2024 (16.02.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/452,050

14 March 2023 (14.03.2023)

US

18/506,672

10 November 2023 (10.11.2023)

US

(71) Applicant: SAMSUNG ELECTRONICS CO., LTD.

[KR/KR]; 129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677 (KR).

(72) Inventors: KONAR, Abhisek; 1000 Sherbrooke St

W#1100, Montreal, Québec H3A 3G4 (CA). LIMYOY,

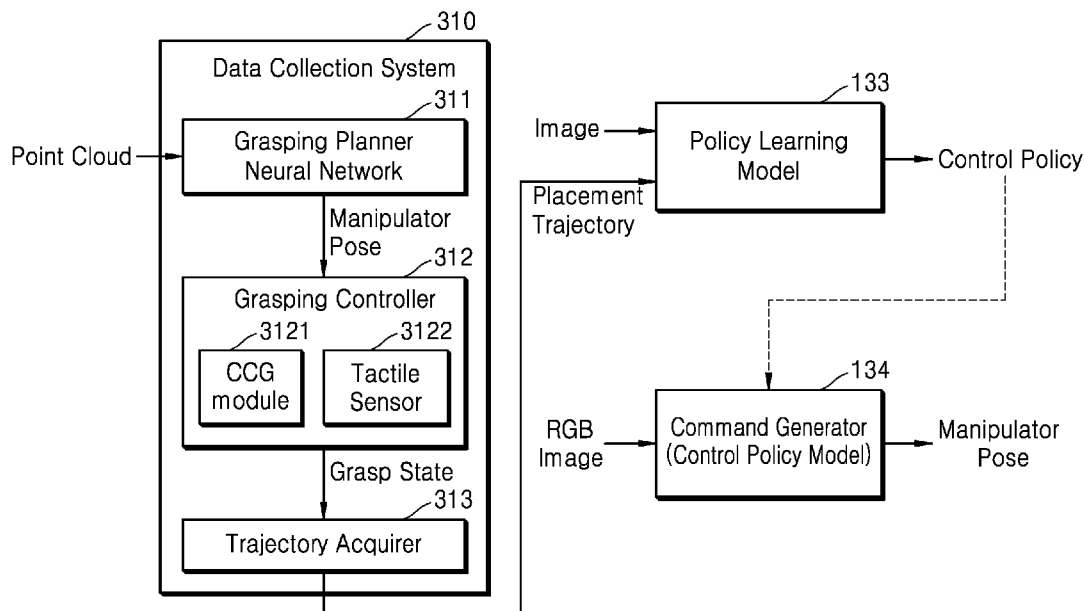
Oliver; 1000 Sherbrooke St W#1100, Montreal, Québec H3A 3G4 (CA). HOGAN, Francois; 1000 Sherbrooke St W#1100, Montreal, Québec H3A 3G4 (CA). DUDEK, Gregory Lewis; 1000 Sherbrooke St W#1100, Montreal, Québec H3A 3G4 (CA). ABLETT, Trevor; 1000 Sherbrooke St W#1100, Montreal, Québec H3A 3G4 (CA).

(74) Agent: Y.P.LEE, MOCK & PARTNERS; 12F Daelim Acrotel, 13 Eonju-ro 30-gil, Gangnam-gu, Seoul 06292 (KR).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH,

(54) Title: APPARATUS AND METHOD FOR CONTROLLING ROBOTIC MANIPULATORS



(57) Abstract: A method and an apparatus for controlling a robot including a manipulator is provided. The method may include: acquiring an image of a scene including a target object; inputting the image into a control policy model to obtain an object placement pose of the manipulator, as an output of the control policy model, wherein the control policy model is generated using data collected by a data collection system that is configured to acquire an object retrieval trajectory by observing a robot movement for object retrieval, and reverse the object retrieval trajectory into an object placement trajectory; and commanding the robot to position the manipulator according to the object placement pose, to place the target object at a designated location.

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, 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))*

Description

Title of Invention: APPARATUS AND METHOD FOR CONTROLLING ROBOTIC MANIPULATORS

Technical Field

- [1] The disclosure relates to an apparatus and a method for controlling robotic manipulators, and particularly, relates to an apparatus and a method for generating training data for an artificial intelligence (AI) model, and controlling the robotic manipulators using the AI model.

Background Art

- [2] Object placement is an important daily task for autonomous robotic manipulators. Examples of the object placement include household activities like setting tables or loading dishwashers. Robotic placement involves contact-rich interactions and typically requires vision-based feedback to account for the state of an in-hand object and the surroundings.
- [3] Developing effective object placement policies is intricate, and may involve solving numerous complex subtasks, including object tracking, scene understanding, motion planning, and control. Imitation learning (IL) presents an appealing, simple, and practical approach by learning end-to-end vision-based control policies from expert demonstrations. Given the intricacy of the real world, a demonstration dataset may need to cover diverse instances to encompass various factors such as backgrounds, lighting, and object types. A major bottleneck for imitation learning-based approaches is a substantial human time and effort necessary to collect expert demonstrations.

Disclosure of Invention

Solution to Problem

- [4] According to an embodiment of the disclosure, an electronic device for controlling a robot including a manipulator may include one or more processors, the one or more processors acquiring an image of a scene including a target object. The electronic device for controlling a robot including a manipulator may include one or more processors, the one or more processors inputting the image into a control policy model to obtain an object placement pose of the manipulator, as an output of the control policy model, wherein the control policy model is generated using data collected by a data collection system that is configured to acquire an object retrieval trajectory by observing a robot movement for object retrieval, and reverse the object retrieval trajectory into an object placement trajectory. The electronic device for controlling a robot including a manipulator may include one or more processors, the one or more processors commanding the robot to position the manipulator according to the object

placement pose, to place the target object at a designated location.

- [5] According to an embodiment of the disclosure, a method for controlling a robot including a manipulator may include acquiring an image of a scene including a target object. The method for controlling a robot including a manipulator may include inputting the image into a control policy model to obtain an object placement pose of the manipulator, as an output of the control policy model, wherein the control policy model is generated using data collected by a data collection system that is configured to acquire an object retrieval trajectory by observing a robot movement for object retrieval, and reverse the object retrieval trajectory into an object placement trajectory. The method for controlling a robot including a manipulator may include commanding the robot to position the manipulator according to the object placement pose, to place the target object at a designated location.

- [6] According to an aspect of the present disclosure, a computer-readable storage medium storing instruction is provided. The instructions, when executed by at least one processor, may cause the at least one processor to perform the method corresponding.

- [7] Additional aspects will be set forth in part in the description that follows and, in part, will be apparent from the description, or may be learned by practice of the presented embodiments of the disclosure.

Brief Description of Drawings

- [8] The above and other aspects, features, and aspects of embodiments of the disclosure will be more apparent from the following description taken in conjunction with the accompanying drawings, in which:

- [9] FIGS. 1A and 1B are block diagrams of a configuration of a system for controlling a robot including a manipulator, according to an embodiment of the present disclosure;

- [10] FIG. 2 illustrates a process of controlling a robot including a manipulator, according to an embodiment of the present disclosure;

- [11] FIG. 3 is a diagram illustrating a method for collecting data for training a policy learning model and performing an object placement task using a control policy model, according to an embodiment of the present disclosure;

- [12] FIG. 4 illustrates a method of picking and placing an object during a training phase and an interference phase according to an embodiment of the present disclosure;

- [13] FIG. 5 is a flowchart illustrating a method of obtaining a placement trajectory according to an embodiment of the present disclosure;

- [14] FIG. 6 illustrates a method of pruning candidate grasping pose according to an embodiment of the present disclosure;

- [15] FIG. 7 illustrates a method of grasping an object with complaint control according to an embodiment of the present disclosure;

- [16] FIGS. 8 and 9 illustrate a method of determining a grasp state is stable according to an embodiment of the present disclosure;
- [17] FIG. 10 illustrates a method of obtaining an object placement trajectory and reversing the object placement trajectory from the object retrieval trajectory according to an embodiment of the present disclosure;
- [18] FIG. 11 is a diagram of devices for generating a robot control policy and performing an object placement task using the robot control policy, according to an embodiment; and
- [19] FIG. 12 is a diagram of components of one or more electronic devices of FIG. 11 according to an embodiment.

Mode for the Invention

- [20] Example embodiments are described in greater detail below with reference to the accompanying drawings.
- [21] In the following description, like drawing reference numerals are used for like elements, even in different drawings. The matters defined in the description, such as detailed construction and elements, are provided to assist in a comprehensive understanding of the example embodiments. However, it is apparent that the example embodiments can be practiced without those specifically defined matters. Also, well-known functions or constructions are not described in detail since they would obscure the description with unnecessary detail.
- [22] Expressions such as "at least one of," when preceding a list of elements, modify the entire list of elements and do not modify the individual elements of the list. For example, the expression, "at least one of a, b, and c," should be understood as including only a, only b, only c, both a and b, both a and c, both b and c, all of a, b, and c, or any variations of the aforementioned examples.
- [23] While such terms as "first," "second," etc., may be used to describe various elements, such elements must not be limited to the above terms. The above terms may be used only to distinguish one element from another.
- [24] The term "module" or "component" is intended to be broadly construed as hardware, firmware, or a combination of hardware and software.
- [25] It will be apparent that systems and/or methods, described herein, may be implemented in different forms of hardware, firmware, or a combination of hardware and software. The actual specialized control hardware or software code used to implement these systems and/or methods is not limiting of the implementations. Thus, the operation and behavior of the systems and/or methods were described herein without reference to specific software code—it being understood that software and hardware may be designed to implement the systems and/or methods based on the description

herein.

- [26] Even though particular combinations of features are recited in the claims and/or disclosed in the specification, these combinations are not intended to limit the disclosure of possible implementations. In fact, many of these features may be combined in ways not specifically recited in the claims and/or disclosed in the specification. Although each dependent claim listed below may directly depend on only one claim, the disclosure of possible implementations includes each dependent claim in combination with every other claim in the claim set.
- [27] No element, act, or instruction used herein should be construed as critical or essential unless explicitly described as such. Also, as used herein, the articles "a" and "an" are intended to include one or more items, and may be used interchangeably with "one or more." Furthermore, as used herein, the term "set" is intended to include one or more items (e.g., related items, unrelated items, a combination of related and unrelated items, etc.), and may be used interchangeably with "one or more." Where only one item is intended, the term "one" or similar language is used. Also, as used herein, the terms "has," "have," "having," or the like are intended to be open-ended terms. Further, the phrase "based on" is intended to mean "based, at least in part, on" unless explicitly stated otherwise.
- [28] In an embodiment of the present disclosure, a learning to place by picking (LPP) method is utilized to automate a collection of training data for training an AI model configured to generate a robot arm trajectory to place an object and control a pose of the robot arm based on the generated robot arm trajectory. To autonomously gather placement demonstration data (also referred to as "placement training data" or "object placement trajectory data"), a self-supervised process may be used to obtain picking demonstration data and convert the picking demonstration data to placement demonstration data, based on a cyclical nature and inherent symmetry of the pick and place problem. A picking phase (i.e., grasping and retrieving) may offer training supervision or labels for a placement task, as picking and placing can be alternately performed by reversing a retrieval trajectory. A placement policy is generated based on this LPP approach, to be utilized beyond a training environment and in scenarios where objects are not initially positioned at their target locations.
- [29] The LPP method uses the concept of working backwards to the challenge of robotic object placement. An object is assumed to be in its intended place configuration during data collection. A robot manipulator may pick the object and reverse back to an initial location to generate data. The LPP method enables the robot manipulator to independently collect demonstration data within a larger scene containing multiple objects and varied placement positions.
- [30] The LPP method may use a self-supervised technique for collecting data in 6-degrees

of freedom (DOF) robotic object placement tasks. The data collection may be conducted based on tactile re-grasping and compliant control for grasping to achieve uninterrupted data collection without human intervention. A real-world experimental validation is conducted on the LPP method by using the collected data to train a vision-based control policy capable of placing multiple objects (e.g., plates) in an intended place (e.g., a dish rack).

[31] Various embodiments of the present disclosure will be described in detail with reference to the drawings below.

[32] FIGS. 1A-1B are block diagrams of a configuration of a system for controlling a robot including a manipulator, in accordance with an embodiment.

[33] As shown in FIGS. 1A and 1B, an apparatus 100 according to an embodiment may include a tool 110, a vision sensor 120, and a computer system 130. The computer system 130 may include an input/output interface 131, an image module 132, a policy learning model 133, and a command generator 134.

[34] In the apparatus 100, the image module 132 may be omitted when the policy learning model 133 and the command generator 134 use images directly received from the vision sensor 120. If the apparatus 100 does not support on-device learning, the apparatus 100 may omit the policy learning model 133, and the command generator 134 may use a pre-trained control policy model.

[35] The apparatus 100 may receive a task to be performed on a target object 102, and may estimate a movement path (i.e., trajectory) of the tool 110 to perform the task. For example, the apparatus 100 may train the policy learning model 133 that accepts as input an image of the target object 102 and a scene 104 into which the target object 102 is to be placed, and an object placement trajectory which is acquired by reversing an object retrieval trajectory, and generates as output a control policy for controlling a pose of the tool 110 to place the target object 102 at a location in the scene 104. The control policy may represent a neural network model configured to receive as input an image of the target object 102 and a scene 104, and generate pose information of the tool 110 for placing the target object 102 at a location in the scene 104. The pose information of the tool 110 may include rotational positions and transitional positions of the manipulator attached at an end of the tool 110 for placing the target object 102 at a location in the scene 104. The command generator 134 may use a neural network model (hereinafter "control policy model" or "control policy neural network"), which implements the control policy learned from the policy learning model 133, to provide a robot control command. In an embodiment, the control policy model may receive an image of the target object 102 and the scene 104 as input, and generate as output a robot control command including the pose information of the tool 110.

[36] Hereinafter, the elements of the apparatus 100 are described in further detail.

- [37] The tool 110 may be operated under the control of the computer system 130 to manipulate the target object 102. In an embodiment, the tool 110 may be a robot arm having a manipulator 112 positioned at one end thereof. The manipulator 112 may include a device such as an end-effector for interacting with the target object 102. Examples of the end-effector may include grippers, scoops, tweezers, force-torque sensors, material removal tools, welding torches, collision sensors, and tool changers, and the types of the end-effector are not limited thereto. Examples of the target object 102 to be manipulated by the tool 110 may include a hook, a cup, a container, a bag, and the like. For example, when a gripper of a robot arm is used as the manipulator 112, the command generator 134 may generate a robot control command to control the gripper to grasp, lift, and place the target object 102.
- [38] In an embodiment, the tool 110 may include one or more joints 111. For example, the joints 111 may include a joint 111a, a joint 111b, a joint 111c, a joint 111d, a joint 111e, a joint 111f, and a joint 111g, each of which may be located at different positions along the tool 110. In an embodiment, the tool 110 may move the manipulator 112 to a desired position in space by rotating, moving, or otherwise operating at least one of the joints 111 based on a robot control command. For example, based on the robot control command, the tool 110 may move and rotate the joints 111 to cause the manipulator 112 to be placed at a position that is suitable for placing the target object 102 in the scene 104.
- [39] The vision sensor 120 may include one or more cameras, and may be configured to capture images of at least one of the tool 110, the target object 102, and the scene 104 in which the target object 102 is to be placed. For example, in an embodiment, the vision sensor 120 may be attached to the robot arm such that the vision sensor 120 is located at a fixed position with respect to the manipulator 112, and therefore the vision sensor 120 may be moved by the robot arm along with the manipulator 112 according to the movements of the joints 111. The vision sensor 120 may be implemented as or include at least one of a red/green/blue (RGB) camera and an RGB depth (RGBD) camera, however an embodiment is not limited thereto.
- [40] The image module 132 may control the vision sensor 120 to obtain the images discussed above, and may provide the images to the policy learning model 133 and the command generator 133. In an embodiment, the image module 132 may perform processing on the images. For example, in an embodiment the image module 132 may obtain depth information based on multiple RGB images of the target object 102 or the scene 104, and may provide the depth information to the policy learning model 133 and the command generator 133. In an embodiment, when a data collection system 310 (see FIG. 3) is incorporated into the computer system 130, the image module 132 may generate a point cloud image of the target object 102 and the scene 104 and provide the

point cloud image to the data collection system 310. When the apparatus 100 does not support on-device learning, the apparatus 100 may omit the data collection system 310 and the policy learning model 133, and the command generator 133 may use a pre-trained control policy model.

- [41] The input/output interface 131 may enable communications between the tool 110, the vision sensor 120, and the computer system 130. The input/output interface 131 may include a transceiver and/or a separate receiver and transmitter that enables the computer system 130 to communicate with other devices, such as via a wired connection, a wireless connection, or a combination of wired and wireless connections. The input/output interface 131 may permit the computer system 130 to receive information from another device and/or provide information to another device. For example, the input/output interface 131 may include an Ethernet interface, an optical interface, a coaxial interface, an infrared interface, a radio frequency (RF) interface, a universal serial bus (USB) interface, a Wi-Fi interface, a cellular network interface, or the like. In an embodiment, the computer system 130 may be included in another element such as the tool 110 or the manipulator 112, or may be separate from and external to the other elements of the apparatus 100.
- [42] The image module 132, the policy learning model 133, and the command generator 134 may be implemented by at least one processor and at least one memory.
- [43] The processor may be implemented in hardware, firmware, or a combination of hardware and software. The processor may be a central processing unit (CPU), a graphics processing unit (GPU), an accelerated processing unit (APU), a micro-processor, a microcontroller, a digital signal processor (DSP), a field-programmable gate array (FPGA), an application-specific integrated circuit (ASIC), or another type of processing component. In an implementation, the processor may include one or more processors capable of being programmed to perform a function. The processor may access the memory and execute computer readable program instructions that are stored in the memory.
- [44] The memory may store information, data, an operating system, a plurality of program modules related to the operation and use of the apparatus 100. For example, the memory may include a hard disk (e.g., a magnetic disk, an optical disk, a magneto-optic disk, and/or a solid state disk), a compact disc (CD), a digital versatile disc (DVD), a universal serial bus (USB) flash drive, a Personal Computer Memory Card International Association (PCMCIA) card, a floppy disk, a cartridge, a magnetic tape, and/or another type of computer-readable medium, along with a corresponding drive. The memory may include program instructions and codes configured to be executed by the processor to perform the operations of the image module 132, the policy learning model 133, and the command generator 134.

- [45] FIG. 2 illustrates a process of controlling a robot including a manipulator, according to an embodiment of the present disclosure. In an embodiment, the operations illustrated in FIG. 2 may be performed using any element described herein, for example the apparatus 100 or any element thereof.
- [46] As shown in FIG. 2, the process 200 may include scanning the scene 104 in operation 210, and scanning the target object 102 in operation 220. In an embodiment, the scanning may refer to or include obtaining images using the vision sensor 120. For example, in operation 210 the vision sensor 120 may obtain one or more images of the scene 104 using an RGB camera, an RGBD camera, an IR camera, a LiDAR sensor, a stereo camera and the like, and in operation 220 the vision sensor 120 may obtain an image of the target object 102 using the RGB camera, the RGBD camera an IR camera, a LiDAR sensor, a stereo camera and the like. Operations 210 and 220 may occur as multiple steps, involving the capture of an image of the placement scene and the target object separately, or as a single step capturing an image of the placement scene including the target object.
- [47] As further shown in FIG. 2, in operation 230 the process 200 may include estimating a pose of the manipulator 112 for placing the target object 102 in the scene 104. The pose of the manipulator 112 may include rotational positions (e.g., yaw, pitch, and roll positions) and transitional positions (e.g., x, y, and z coordinates) of the manipulator 112. In an embodiment, the pose of the manipulator 112 may include a change in the rotation positions and the translation positions of the manipulator 112 from a current pose to a next pose.
- [48] In operation 230, the command generator 134 may determine the pose of the manipulator 112 by inputting the 2D or 3D images of the target object 102 and the scene 104 into the control policy model and obtaining pose information for the manipulator 112 as an output of the control policy model.
- [49] As further shown in FIG. 2, in operation 240 the process 200 may include placing the target object 102 at a location in the scene 104. For example, the command generator 134 may generate a robot control command based on the pose information, control the tool 110 to place the target object at the location in the scene 104 and release the target object 102 based on the robot control command.
- [50] FIG. 3 is a diagram illustrating a method for collecting data for training a policy learning model and performing an object placement task using a control policy model, according to an embodiment of the present disclosure. FIG. 4 illustrates a method of picking and placing an object during a training phase and an interference phase according to an embodiment of the present disclosure.
- [51] As shown in FIG. 3, a data collection system 310 may perform a self-supervised data collection method including a cycle with four main phases: (1) grasping planning, (2)

grasping, (3) retrieving and (4) placing by retrieving. The data collected by the data collection system 310 is used to train the policy learning model 133 during a training process. It is assumed that objects in a scene are initially in desired placement positions and orientations.

[52] (1) Grasping Planning

[53] During the grasping planning phase, the manipulator 112 positions itself at a pre-determined pose, ensuring the vision sensor 120 mounted on the robot (e.g., on the wrist of the robot manipulator) obtains an unobstructed view of objects of interest within the environment. The manipulator 112 may include any one or any combination of an arm, a joint, and an end effector which may include a gripper and/or fingers.

[54] A grasping planner neural network 311 may be used to generate L candidate grasping poses $T_g = \{T_{g,i}\}_{i=0}^L$, where $T_{g,i} \in SE(3)$ for all $i \in \{0, \dots, L\}$ based on a point cloud image of the target object 102 and the scene 104 that is input to the grasping planner neural network 311. For example, Contact-GraspNet may be used as the grasping planning neural network 311. Among an initial set of L candidate grasping poses output from the grasping planner neural network 311, grasping poses that do not belong to any objects of interest are pruned and filtered out, to obtain pruned candidate grasping poses, $\bar{T}_g = \{T_{g,i}\}_{i=0}^K$, where $K \leq L$ and $\bar{T}_g \subseteq T_g$. For example, the pruning may be performed based on an object class and/or an output of an obstacle detection module. After pruning, a single grasping pose $T_{g,k}$ is randomly selected from the pruned candidate grasping poses $\bar{T}_g$. In order to approach an object of interest without collision, a pre-grasping pose $T_{g,k}^{pre}$ may be determined by adding a translational offset to the grasping pose $T_{g,k}$.

[55] (2) Grasping

[56] During the grasping phase, a grasping controller 312 controls the manipulator 112 to position the manipulator 112 at the pre-grasping pose $T_{g,k}^{pre}$ and then at the grasping pose $T_{g,k}$ using linear interpolation or spherical linear interpolation. For example, the grasping controller 312 may use a compliant control for grasping (CCG) module 3121 and a tactile sensor 3122 including a tactile regrasping module to control the tool 100 to change the pose of the tool 100.

[57] The CCG module 3121 may use a Cartesian impedance controller to set compliance levels of the manipulator 112 at a plurality of different stages, and may maneuver the manipulator 112 between target poses according to the compliance levels. The term "compliance level" may refer to a degree to which the manipulator 112 is programmed or configured to be flexible or compliant in response to external forces and constraints. The compliance level may represent a translational stiffness level and a rotational

stiffness level of a manipulator controller (e.g., a manipulator actuator or the joints 111) of the tool 100, which determine how rigid or flexible the tool 100 is when interacting with the target object 102 in the environment. Higher compliance levels mean that the manipulator 112 is more flexible and can adapt better to external forces and constraints, while lower compliance levels make the manipulator 112 stiffer and less flexible. Compliance control may involve adjusting the stiffness of the joints 111 or the manipulator controller. Stiffness is a measure of how resistant the tool (e.g., robot) 100 is to deformation when subjected to external forces. Higher stiffness means less compliance, while lower stiffness means more compliance.

- [58] Prior to engaging the manipulator 112 for grasping (e.g., prior to closing a gripper to perform a grasp), translational stiffness and rotational stiffness of the manipulator controller may be set to minimum values, to facilitate the adaptation of the manipulator 112 to properties and contact constraints of the target object 102 within the environment.
- [59] The tactile sensor 3122 operates under the assumption of grasp stability, implying that once the target object 102 is gripped, it remains securely in place without movement or slippage. To address this concern, the tactile sensor 3122 performs tactile sensing to proactively identify the stability of a grasp and initiates a regrasping action if instability is detected.
- [60] The tactile sensor 3122 may detect and measure physical interactions, such as pressure, force, or deformation, at contact points on a sensing surface of the tactile sensor 3122. When the tactile sensor 3122 captures an image of a contact area with the contact points, the tactile sensor 3122 may capture and record information regarding the distribution, locations, intensities, and properties of contact points across the sensing surface of the tactile sensor 3122.
- [61] The tactile sensor 3122 may include individual sensing elements or cells distributed across the sensing surface. The sensing elements are configured to respond to physical changes, such as pressure or deformation which occurs when the sensing elements are in contact with the target object 102.
- [62] When the target object 102 comes into contact with the sensing surface of the tactile sensor 3122, the individual sensing elements underneath the points of contact are activated and act collectively to capture an image of the contact area. Each sensing element corresponds to a specific point on the sensing surface, and a combined response of all the sensing elements provides a representation of the distribution of pressure or force across the contact area.
- [63] The grasping controller 312 determines whether a grasp state is stable based on a shape or a distribution pattern of the contact points acquired from the tactile sensor 3122. For example, the grasping controller 312 may identify a center position of the

distribution pattern of the contact points, and determine the grasp is stable when the identified center is located within a predetermined range of a stable position. The stable position may be pre-determined based on a shape of the target object. A method of determining grasp state stability and controlling the manipulator 112 to attain a stable grasp state will be discussed later in detail with reference to FIGS. 5-9.

[64] (3) Retrieving

[65] A trajectory acquirer 313 begins the retrieving phase when the grasp state is determined to be stable.

[66] Throughout the retrieval procedure, a manipulator pose and a corresponding timestamp are identified from a sequence of images and stored in a memory while the manipulator 112 moves along a retrieval trajectory. The retrieval trajectory is generated based on a grasping sequence, where the grasping controller 312 controls the manipulator 112 to position the manipulator 112 at the pre-grasping pose $T_{g,k}^{pre}$, in compliance with a rotational axis of the manipulator 112. Once at the pre-grasping pose $T_{g,k}^{pre}$, a clearance pose $T_{g,k}$ may be randomly selected, which is centered around a fixed pose above a target retrieval place (e.g., a dish rack). Given a fixed clearance pose, random translations that are sampled from $\mathcal{N}(\mathbf{0}, \sigma_{tr}^2)$ are added. At the end of the retrieval phase, a dense retrieval trajectory τ_r^{expert} of length M from the grasping pose $T_{g,k}$ to the clearance pose $T_{c,k}$ is acquired via the pre-grasping pose $T_{g,k}^{pre}$ as follows:

[67] $\tau_r^{expert} = ((T_0, t_0), \dots, (T_M, t_M))$ Equation (1)

[68] Where $T_0 = T_{g,k}$, and $T_M = T_{c,k}$. $t_i \in \mathbb{R}_{\geq 0}$ for all $i \in \{0, \dots, M\}$ are respective timestamps of the retrieval trajectory starting with $t_0 = 0$. To match a desired control frequency of a manipulator control policy, states are sampled Δt apart in order to generate an sparse retrieval trajectory $\bar{\tau}_r^{expert}$. A total of $\bar{M} + 1$ states are extracted as follows, where $\bar{M} = \frac{t_M - t_0}{\Delta t}$:

[69] $\bar{\tau}_r^{expert} = (T_{\Delta t_0}, \dots, T_{\Delta t_{\bar{M}+1}})$, where

[70] $T_{\Delta t_i} = T_{n^*}$, with $n^* = \arg \min_{n \in \{0, \dots, M\}} |i \cdot \Delta t - t_n|$ Equation (2)

[71] In Equation (2), a nearest pose is identified from the dense retrieval trajectory τ_r^{expert} at every Δt interval.

[72] (4) Placing by Retrieving

[73] During the placing by retrieving phase, the trajectory acquirer 313 reverses the sparse retrieval trajectory $\bar{\tau}_r^{expert}$ into a placement trajectory $\bar{\tau}_p^{expert}$ expressed as global poses.

[74] $\bar{\tau}_p^{expert} = (T_{\Delta t_{\bar{M}+1}}, \dots, T_{\Delta t_0})$ Equation (3)

[75] The trajectory acquirer 313 converts the global poses in placement trajectory $\bar{\tau}_p^{expert}$ into relative poses in order to use state differences as expert actions:

[76] $T_{\Delta cmd,i} = T_{\Delta t_{\bar{M}-i+1}}^{-1} T_{\Delta t_{\bar{M}-i}}$ Equation (4)

[77] Where $i \in \{0, \dots, \bar{M}\}$, and $T_{\Delta cmd,i} \in SE(3)$ is a relative change in the manipulator pose. Six dimensional vectors $[t, \theta]^T$ may be used to represent poses, where $t \in \mathbb{R}^3$ is a translation, and $\theta \in \mathbb{R}^3$ is a rotation vector. The rotation vector is defined as $\theta = \theta e$, where $\theta \in \mathbb{R}^1$ represents a rotational angle and $e \in \mathbb{R}^3$ represents rotational axes (i.e., x, y, and z rotational axes). The expert action is defined as $a_i = (a_{\Delta cmd,i}, a_{gripper,i})$, where $a_{\Delta cmd,i} = [t_i, \theta_i]^T$ is a six dimensional vector representation of the pose $T_{\Delta cmd,i}$ and $a_{gripper,i} \in \{0, 1\}$ is a binary variable describing the gripper command (i.e., open or close). Robustness of the placement trajectory may be increased via noise removal, which is described below.

[78] To gather an expert data episode, a series of downsampled placement commands, denoted as $\{a_i\}_{i=0}^{\bar{M}}$, are executed, and sequences of observation and actions are stored as an expert trajectory. The sequence of observations may be acquired by capturing RGB images using the vision sensor 120 affixed to the tool 100, where the tool observation space $o_i \in \mathbb{R}^{\omega \times h \times 3}$. During the object placement phase, the tool 100 maintains a gripper of the manipulator 120 in a closed position, meaning that the gripper action is defined as $a_{gripper,i} = 1$ for all $i \in \{0, \dots, \bar{M}\}$. At the end of the placement trajectory, the tool 100 is commanded to retain the manipulator 120 in the current pose while opening the gripper for N discrete time steps (i.e., $a_{\Delta cmd,j} = [0, 0, 0, 0, 0, 0]^T$ and $a_{gripper,j} = 0$ for all $j \in \{\bar{M} + 1, \dots, \bar{M} + N\}$). For example, N may be set to 5 to give the robot an adequate amount of time to open the gripper and release the target object 102. A single expert demonstration or trajectory is defined as a set of training tuples:

[79] $\tau_r^{expert} = \{o_t, a_t, o_{t+1}\}_{t=0}^T$ Equation (5)

[80] Where a total length of $\tau = \bar{M} + N$.

[81] In order to increase the coverage of the demonstrations and train a more robust control policy, perturbations are introduced to a predetermined percentage (e.g., 75%) of poses through the application of noise. For example, isotropic Gaussian noise is injected into the demonstration data to achieve persistent excitation, which is a condition where the training data is informative enough to learn a model that is robust to compounding errors during deployment. In detail, a translation of a post t is perturbed as follows:

[82] $t_{perturbed} = t + \delta t$ Equation (6)

[83] Where $\delta t \in \mathbb{R}^3$ is isotropic Gaussian noise and each dimension of δt is sampled from $\mathcal{N}(0, 0.5 \text{ cm})$. Similarly, the rotation is perturbed as follows:

[84] $\theta_{perturbed} = (\theta + \delta\theta)(e + \delta e)$ Equation (7)

[85] Where $\delta e \in \mathbb{R}^3$ is isotropic Gaussian noise, and each dimension of δe is sampled from $\mathcal{N}(0, \sigma_e)$, and $\delta\theta$ is sampled from $\mathcal{N}(0, \sigma_\theta)$.

[86] Referring to FIGS. 3 and 4, during a training phase, the policy learning model 133 may receive as input a placement trajectory and an image of the target object 102 and the scene 104, and generates as output a control policy model for controlling a pose of the tool 100.

[87] The control policy model $\pi_\phi(a_t | o_t)$ is parameterized by ϕ , and represents a neural network model trained to generate a tool control policy based on data collected via the LPP method. Behavioral cloning may be used to train the neural network model for generating the control policy, using a likelihood-based loss or objective function. Behavioral cloning may refer to a technique where the neural network model is trained to replicate the behavior of a human or expert by learning from their demonstrated actions. It is a form of supervised learning where the neural network model learns to map input observations to corresponding output actions, mimicking the behavior exhibited in the training data. Given a dataset of N_{train} expert demonstration trajectories, a loss function $\mathcal{L}$ is calculated as follows:

[88]
$$\mathcal{L} = \frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \sum_{t=0}^T -\log \pi_\phi(a_t | o_t)$$
 Equation (8)

[89] The neural network model for generating the control policy may include one or more convolutional neural networks (CNN), such as, but not limited to, ResNET, followed by a multilayer perceptron (MLP) layer that maps an output of the CNN into parameters of the distribution of control actions. The distribution of control actions is represented by a unimodal Gaussian and/or a multimodal mixture of Gaussians.

[90] Referring to FIGS. 3 and 4, in an inference phase, the command generator 134 may use the control policy model to obtain a manipulator pose for placing the target object 102 at a location in the scene 104. The control policy model may be trained by the policy learning model 133, and may receive as input an RGB image or an RGBD image of the target object 102 and the scene 104 and provide as output the manipulator pose for placing the target object 102.

[91] FIG. 5 is a flowchart illustrating a method of obtaining an object placement trajectory according to an embodiment of the present disclosure.

[92] In operation 510, an object grasping plan is carried out to obtain candidate grasping

poses for grasping the target object 102 using the grasping planner neural network 311 such as Contac-GraspNet.

- [93] In operation 520, among an initial set of candidate grasping poses output from the grasping planner neural network 311, grasping poses that do not belong to a class of the target object 102 may be pruned and filtered out, to obtain pruned candidate grasping poses. For example, when the target object 102 is a plate, pruning may be performed based on the object class "plate" to remove candidate poses for grasping objects (e.g., cups) which are not classified as plates. After pruning, a single grasping pose may be randomly selected from the pruned candidate grasping poses.
- [94] In operation 530, the manipulator 112 is controlled to grasp the target object 102 with compliant control using the compliant control for grasping (CCG) module 3121. Prior to engaging a grasping action, the CCG module 3121 may set translational stiffness and rotational stiffness of the manipulator controller (e.g., the joints 111 or the actuator of the manipulator 112) to minimum values, to facilitate the adaptation of the manipulator 112 to properties and contact constraints of the target object 102 in the scene 104.
- [95] In operation 540, while the manipulator 112 is grasping the target object 102, the tactile sensor 3122 mounted on the manipulator 112 may capture an image of contact points within a contact area between a sensing surface of the tactile sensor 3122 and the target object 102. The tactile sensor 3122 may include individual sensing elements or cells distributed across the sensing surface. The sensing elements may be configured to respond to physical changes, such as pressure or deformation which occurs when the sensing elements are in contact with the target object 102. When the target object 102 comes into contact with the sensing surface of the tactile sensor 3122, the individual sensing elements underneath the points of contact may be activated and act collectively to capture the image of the contact area. Each sensing element may correspond to a specific point on the sensing surface, and a combined response of all the sensing elements may provide a representation of the distribution of pressure or force across the contact area.
- [96] In operation 540, the grasping controller 312 may determine whether the grasp is stable or not based on a shape or a distribution pattern of the contact points. For example, the grasping controller 312 may identify a center position of the distribution pattern of the contact points, and determine the grasp is stable when the identified center is located within a predetermined range of a stable position. The stable position may be predetermined based on a shape of the target object.
- [97] If the grasping controller 312 determines the grasp is unstable, the method proceeds to operation 550 to control the manipulator 112 to re-grasp the target object 102 so that the center position of the distribution pattern of the contact points falls within the pre-

determined range of the stable position. The grasping controller 312 may determine a position vector for changing the center position to the predetermined stable position, and generates a robot control command based on the position vector to enable the manipulator 112 to re-grasp the target object 102 based on the robot control command.

- [98] If the grasp is determined to be stable, the method proceeds to operation 560 to acquire an object retrieval trajectory based on the stable grasp position of the manipulator 112 and then to acquire an object placement trajectory by reversing the object retrieval trajectory.
- [99] Operations 520, 530, and 550 of FIG. 5 will be further discussed with references to FIGS. 6-9 below.
- [100] FIG. 6 illustrates a method of pruning candidate grasping pose according to an embodiment of the present disclosure. The method shown in FIG. 6 may correspond to operation 520 of FIG. 5.
- [101] In operation 521, the vision sensor 120 may obtain an image of the scene 104 including the target object 102, and the image module 132 may generate a point cloud image of the scene 104 via 3D image reconstruction.
- [102] In operation 522, the grasping planner neural network 311 may obtain as input the point cloud image, and generate as output an initial set of candidate grasping poses.
- [103] In operation 523, a class of the target object 102 may be identified, and from the initial set of candidate grasping poses, grasping poses that do not correspond to objects of the target class may be removed. In an embodiment, a set of pruned candidate grasping poses may be obtained.
- [104] FIG. 7 illustrates a method of grasping an object with compliant control according to an embodiment of the present disclosure. The method shown in FIG. 7 may correspond to operation 530 of FIG. 5.
- [105] In operation 531, the manipulator 112 may be commanded to move to a grasping position. The grasping position is selected from the pruned candidate grasping poses.
- [106] In operation 532, prior to taking a grasping action, translational stiffness and rotational stiffness of the manipulator controller (e.g., the joints 111 or the actuator of the manipulator 112) may be set to predetermined minimum values. In an embodiment, setting translational stiffness and rotational stiffness of the manipulator controller to predetermined minimum values may facilitate the adaptation of the manipulator 112 to properties and contact constraints of the target object 102 in the scene 104. In an embodiment, the compliant control for grasping (CCG) module 3121 may set translational stiffness and rotational stiffness of the manipulator controller to predetermined minimum values
- [107] In operation 533, a gripper mounted at the manipulator 112 may be closed to securely grasp the target object 102, with the translational stiffness and rotational

stiffness of the manipulator controller set to the minimum values.

[108] In operation 534, following the grasping action, the translational stiffness and rotational stiffness may be set to a predetermined normal value.

[109] FIGS. 8 and 9 illustrate a method of determining a grasp state is stable according to an embodiment of the present disclosure. The method shown in FIG. 8 may correspond to operation 540 of FIG. 5.

[110] Referring to FIG. 8, in operation 541, the grasping controller 312 may read a tactile image captured by the tactile sensor 3122 which is mounted on the manipulator 112.

[111] In operation 542, the grasping controller 312 may compute a contact area between the manipulator 112 and the target object 102 based on the tactile image that shows contact points between a sensing surface of the tactile sensor 3122 and the target object 102.

[112] In operation 543, the grasping controller 312 may align a predetermined shaped region (e.g., an eclipse region) with the contact area including the contact points, so that the predetermined shaped region encompasses all the contact points as shown in FIG. 9 image 920.

[113] Referring to FIGS. 8 and 9, in operation 544, the grasping controller 312 may determine whether the predetermined shaped region indicates that the current grasp is a stable grasp. The grasping controller 312 identifies a center point of the predetermined shaped region and measures a distance between the center point and a stable position. The stable position is predetermined based on a shape of the target object 102.

[114] When the distance between the center point and the stable position exceeds a predetermined range, the grasping controller 312 may identify the current grasp as unstable in operation 545 and proceeds to adjust the grasp pose of the manipulator 112. The grasp pose may be adjusted to bring the center position closer to the stable position, ensuring the center position falls within the predetermined range. For example, referring to FIG. 9, the image 921 and the image 922 shows an example that the distance between the center point and the stable position exceeds the predetermined shape region, so that the grasping controller 312 may identify the current grasp as unstable in operation 545.

[115] When the distance between the center point and the stable position is within the predetermined range, the grasping controller 312 determines that the current grasp is stable in operation 546. When the current grasp is stable, the grasping controller 312 commands the manipulator 112 to execute an object retrieval action while maintaining the current grasp state. For example, referring to FIG. 9, the image 923 shows an example that the distance between the center point and the stable position is within the predetermined shape region, so that the grasping controller 312 may determine the current grasp is stable in operation 546. The data collection system 310 may record an

object retrieval trajectory while the manipulator 112 is executing the object retrieval action, and derive an object placement trajectory by reversing the object retrieval trajectory.

[116] FIG. 10 illustrates a method of obtaining an object placement trajectory and reversing the object placement trajectory from the object retrieval trajectory according to an embodiment of the present disclosure. The method shown in FIG. 10 may correspond to operation 560 shown in FIG. 5.

[117] In operation 561, a retrieval pose is determined based on a retrieval trajectory that extends from the grasping pose to a clearance pose (e.g., a standby pose of the manipulator 112 before initiating a retrieval task or a placement task).

[118] In operation 562, the manipulator 112 is set to the retrieval pose while recording the retrieval pose as an object placement trajectory. The retrieval pose and a corresponding timestamp are stored while the manipulator 112 moves along the retrieval trajectory.

[119] In operation 563, the object placement trajectory is downsampled to match a control policy frequency of the control policy model for performing an object placement task. For example, if the original object placement trajectory includes 24 retrieval poses spaced at 0.2 second intervals, downsampling is performed on the object placement trajectory so that the downsampled object placement trajectory includes 12 retrieval poses at 0.4 second intervals, to match the control policy frequency.

[120] In operation 564, the downsampled object placement trajectory is reversed into an object placement trajectory, for example, using Equations (3)-(5).

[121] In operation 565, the object placement task is executed based on the object placement trajectory, and a result of the object placement task is stored. The result of the object placement task is used as training data for the policy learning model 133 as shown in FIG. 3.

[122] FIG. 11 is a diagram of devices for generating a robot control policy and performing an object placement task using the robot control policy according to an embodiment. FIG. 11 includes a user device 610, a server 620, and a communication network 630. The user device 610 and the server 620 may interconnect via the communication network 630. The communication network 630 may be, such as, but not limited to, wired connections, wireless connections, or a combination of wired and wireless connections.

[123] The user device 610 may include one or more devices (e.g., a processor 611 and a data storage 612) configured to retrieve an image corresponding to a search query. For example, the user device 610 may include a computing device (e.g., a desktop computer, a laptop computer, a tablet computer, a handheld computer, a smart speaker, a server, etc.), a mobile phone (e.g., a smart phone, a radiotelephone, etc.), a camera device, a wearable device (e.g., a pair of smart glasses, a smart watch, etc.), a home

appliance (e.g., a robot vacuum cleaner, a smart refrigerator, etc.), or a similar device. The data storage 612 of the user device 610 may include at least one of the data collection system 310, the policy learning model 133, and the control policy model generated from the policy learning model 133. When the policy learning model 133 is stored in the server 602 instead of being stored in the user device 610, the server 602 may train the policy learning model 133 and provides the user device 610 with the pre-trained policy learning model 133.

- [124] The server 620 may include one or more devices (e.g., a processor 621 and a data storage 622) configured to train the control policy model. The data storage 622 of the server 620 may include at least one of the data collection system 310, the policy learning model 133, and the control policy model generated from the policy learning model 133.
- [125] The communication network 630 includes one or more wired and/or wireless networks. For example, network 1300 may include a cellular network, a public land mobile network (PLMN), a local area network (LAN), a wide area network (WAN), a metropolitan area network (MAN), a telephone network (e.g., the Public Switched Telephone Network (PSTN)), a private network, an ad hoc network, an intranet, the Internet, a fiber optic-based network, or the like, and/or a combination of these or other types of networks.
- [126] The number and arrangement of devices and networks shown in FIG. 11 are provided as an example. In practice, there may be additional devices and/or networks, fewer devices and/or networks, different devices and/or networks, or differently arranged devices and/or networks than those shown in FIG. 11. In an embodiment, two or more devices shown in FIG. 11 may be implemented within a single device, or a single device shown in FIG. 11 may be implemented as multiple, distributed devices. In an embodiment, a set of devices (e.g., one or more devices) may perform one or more functions described as being performed by another set of devices.
- [127] FIG. 12 is a diagram of components of one or more electronic devices of FIG. 11 according to an embodiment. An electronic device 1000 in FIG. 12 may correspond to the user device 610 and/or the server 620.
- [128] FIG. 12 is for illustration only, and an embodiment of the electronic device 1000 could be used without departing from the scope of this disclosure. For example, the electronic device 1000 may correspond to a client device or a server.
- [129] The electronic device 1000 may include a bus 1010, a processor 1020, a memory 1030, an interface 1040, and a display 1050.
- [130] The bus 1010 may include a circuit for connecting the components 1020 to 1050 with one another. The bus 1010 may function as a communication system for transferring data between the components 1020 to 1050 or between electronic devices. The bus

1010 may include one or more components that may permit communication among the set of components of the electronic device 1000. For example, the bus 1010 may be a communication bus, a cross-over bar, a network, or the like. Although the bus 1010 is depicted as a single line in FIG. 12, the bus 1010 may be implemented using multiple (e.g., two or more) connections between the set of components of the electronic device 1000. The present disclosure is not limited in this regard.

[131] The processor 1020 may include one or more of a central processing unit (CPU), a graphics processor unit (GPU), an accelerated processing unit (APU), a many integrated core (MIC), a field-programmable gate array (FPGA), or a digital signal processor (DSP). The processor 1020 is able to perform control of any one or any combination of the other components of the electronic device 1000, and/or perform an operation or data processing relating to communication. For example, the processor 1020 may correspond to the computer system 130 illustrated in FIG. 1B, and/or may include at least one of the data collection system 310, the policy learning model 133, and the control policy model generated from the policy learning model 133. The processor 1020 executes one or more programs stored in the memory 1030.

[132] The memory 1030 may include a volatile and/or non-volatile memory. In an embodiment, the memory 1030 may include volatile memory such as, but not limited to, random access memory (RAM), static RAM (SRAM), dynamic RAM (DRAM), and the like. In an embodiment, the memory 1030 may include non-volatile memory such as, but not limited to, read only memory (ROM), electrically erasable programmable ROM (EEPROM), NAND flash memory, phase-change RAM (PRAM), magnetic RAM (MRAM), resistive RAM (RRAM), ferroelectric RAM (FRAM), magnetic memory, optical memory, and the like. However, the present disclosure is not limited in this regard, and the memory 1030 may include other types of dynamic and/or static memory storage. In an embodiment, the memory 1030 may store information and/or instructions for use (e.g., execution) by the processor 1020. The memory 1030 stores information, such as one or more of commands, data, programs (one or more instructions), applications 1034, etc., which are related to at least one other component of the electronic device 1000 and for driving and controlling the electronic device 1000. For example, commands and/or data may formulate an operating system (OS) 1032. Information stored in the memory 1030 may be executed by the processor 1020. In an embodiment, the memory 1030 may store at least one of the data collection system 310, the policy learning model 133, and the control policy model generated from the policy learning model 133.

[133] The applications 1034 may include the above-discussed embodiments. These functions can be performed by a single application or by multiple applications that each carry out one or more of these functions. For example, the applications 1034 may

include an artificial intelligence (AI) model for performing operations 200 and 500 illustrated in FIGS. 2 and 5.

- [134] The display 1050 may include, for example, a liquid crystal display (LCD), a light emitting diode (LED) display, an organic light emitting diode (OLED) display, a quantum-dot light emitting diode (QLED) display, a microelectromechanical systems (MEMS) display, or an electronic paper display. The display 1050 can also be a depth-aware display, such as a multi-focal display. The display 1050 may be able to present, for example, various contents, such as text, images, videos, icons, and symbols.
- [135] The interface 1040 may include input/output (I/O) interface 1042, communication interface 1044, and/or one or more sensors 1046. The I/O interface 1042 serves as an interface that can, for example, transfer commands and/or data between a user and/or other external devices and other component(s) of the electronic device 1000.
- [136] The communication interface 1044 may enable communication between the electronic device 1000 and other external devices, via a wired connection, a wireless connection, or a combination of wired and wireless connections. The communication interface 1044 may permit the electronic device 1000 to receive information from another device and/or provide information to another device. For example, the communication interface 1044 may include an Ethernet interface, an optical interface, a coaxial interface, an infrared interface, a radio frequency (RF) interface, a universal serial bus (USB) interface, a Wi-Fi interface, a cellular network interface, or the like. The communication interface 1044 may receive images from an external device, such as a server.
- [137] The sensor(s) 1046 of the interface 1040 can meter a physical quantity or detect an activation state of the electronic device 1000 and convert metered or detected information into an electrical signal. For example, the sensor(s) 1046 can include one or more cameras or other imaging sensors such as the vision sensor 120 for capturing images of scenes, and the tactile sensor 3122 to detect contact points between the manipulator 112 and the target object 102. The sensor(s) 1046 can also include any one or any combination of a microphone, a keyboard, a mouse, and one or more buttons for touch input. The sensor(s) 1046 can further include an inertial measurement unit. In addition, the sensor(s) 1046 can include a control circuit for controlling at least one of the sensors included herein. Any of these sensor(s) 1046 can be located within or coupled to the electronic device 1000. The sensor(s) 1046 may receive a text and/or a voice signal that contains one or more queries.
- [138] The policy learning model and the control policy model may be written as computer-executable programs or instructions that may be stored in a medium.
- [139] The medium may continuously store the computer-executable programs or instructions, or temporarily store the computer-executable programs or instructions for

execution or downloading. Also, the medium may be any one of various recording media or storage media in which a single piece or plurality of pieces of hardware are combined, and the medium is not limited to a medium directly connected to electronic device 100, but may be distributed on a network. Examples of the medium include magnetic media, such as a hard disk, a floppy disk, and a magnetic tape, optical recording media, such as CD-ROM and DVD, USB flash drive, PCMCIA card, magneto-optical media such as a floptical disk, and ROM, RAM, and a flash memory, which are configured to store program instructions. Other examples of the medium may include recording media and storage media managed by application stores distributing applications or by websites, servers, and the like supplying or distributing other various types of software.

[140] The data collection system, the policy learning model and the control policy model may be provided in a form of downloadable software. A computer program product may include a product (for example, a downloadable application) in a form of a software program electronically distributed through a manufacturer or an electronic market. For electronic distribution, at least a part of the software program may be stored in a storage medium or may be temporarily generated. In an embodiment, the storage medium may be a server or a storage medium of server.

[141] The foregoing disclosure provides illustration and description, but is not intended to be exhaustive or to limit the implementation to the precise form disclosed. Modifications and variations are possible in light of the above disclosure or may be acquired from practice of the implementation.

[142] It will be apparent that systems and/or methods, described herein, may be implemented in different forms of hardware, firmware, or a combination of hardware and software. The actual specialized control hardware or software code used to implement these systems and/or methods is not limiting of the implementations. Thus, the operation and behavior of the systems and/or methods were described herein without reference to specific software code—it being understood that software and hardware may be designed to implement the systems and/or methods based on the description herein.

[143] Even though particular combinations of features are recited in the claims and/or disclosed in the specification, these combinations are not intended to limit the disclosure of possible implementations. In fact, many of these features may be combined in ways not specifically recited in the claims and/or disclosed in the specification. Although each dependent claim listed below may directly depend on only one claim, the disclosure of possible implementations includes each dependent claim in combination with every other claim in the claim set.

[144] A model related to the neural networks described above may be implemented via a

software module. When the model is implemented via a software module (for example, a program module including instructions), the model may be stored in a computer-readable recording medium.

[145] Also, either one or both of the policy learning model and the control policy model may be a part of the electronic device 1000 described above by being integrated in a form of a hardware chip. For example, the model may be manufactured in a form of a dedicated hardware chip for artificial intelligence, or may be manufactured as a part of an existing general-purpose processor (for example, a CPU or application processor) or a graphic-dedicated processor (for example a GPU).

[146] Also, the model may be provided in a form of downloadable software. A computer program product may include a product (for example, a downloadable application) in a form of a software program electronically distributed through a manufacturer or an electronic market. For electronic distribution, at least a part of the software program may be stored in a storage medium or may be temporarily generated. In this case, the storage medium may be a server of the manufacturer or electronic market, or a storage medium of a relay server.

[147] While the embodiments of the disclosure have been described with reference to the figures, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope as defined by the following claims.

[148] According to an embodiment of the disclosure, an electronic device may include a vision sensor configured to capture the image of the scene including the target object.

[149] According to an embodiment of the disclosure, the electronic device, wherein the object placement pose of the manipulator that is output from the control policy model may include information of rotational positions comprising a yaw angle, a pitch angle, and a roll angle of the manipulator. According to an embodiment of the disclosure, the electronic device, wherein the object placement pose of the manipulator that is output from the control policy model may include information of transitional positions comprising an x-direction position, a y-direction position, and a z-direction position of the manipulator.

[150] According to an embodiment of the disclosure, the electronic device may include the data collection system that is configured to obtain a first set of manipulator grasping poses via a grasping planning neural network. The electronic device may include the data collection system that is configured to obtain a second set of manipulator grasping poses by pruning grasping poses that do not belong to an object class of interest, from the first set of grasping poses. The electronic device may include the data collection system that is configured to determine a grasping pose from the second set of manipulator grasping poses. The electronic device may include the data collection system

that is configured to obtain the object retrieval trajectory that extends from the determined grasping pose to a clearance pose.

- [151] According to an embodiment of the disclosure, the electronic device may include the data collection system including a tactile sensor mounted on the manipulator and is configured to identify a center position of a plurality of contact points in a contact area between the tactile sensor and the target object. The electronic device may include the data collection system including a tactile sensor mounted on the manipulator and is configured to determine a position vector for changing the center position of the plurality of contact points to a predetermined stable position. The electronic device may include the data collection system including a tactile sensor mounted on the manipulator and is configured to output a robot control command based on the position vector.
- [152] According to an embodiment of the disclosure, the electronic device may include the data collection system, which is configured to command the robot to re-grasp the target object based on the robot control command including the position vector to acquire the object retrieval trajectory.
- [153] According to an embodiment of the disclosure, the electronic device may include the data collection system, which is configured to determine a retrieval pose of the manipulator based on a movement path extending from a grasping pose to a clearance pose of the manipulator. The electronic device may include the data collection system, which is configured to position the manipulator according to the retrieval pose and store the retrieval pose to generate the object retrieval trajectory based on the retrieval pose. The electronic device may include the data collection system, which is configured to downsample the object retrieval trajectory to match a predetermined control policy frequency. The electronic device may include the data collection system, which is configured to reverse the downsampled object retrieval trajectory into the object placement trajectory. The electronic device may include the data collection system, which is configured to generate the control policy model by training a policy learning model based on the object placement trajectory and the image of the scene including the target object.
- [154] According to an embodiment of the disclosure, the electronic device may include the data collection system, which is configured to command the robot to position the manipulator at a grasping position. The electronic device may include the data collection system, which is configured to change a rotational stiffness level and a translation stiffness level of the manipulator from original values to predetermined minimum values, command the robot to close a gripper of the manipulator. The electronic device may include the data collection system, which is configured to acquire the object retrieval trajectory while the rotational stiffness level and the translation stiffness level

are set to the predetermined minimum values.

- [155] According to an embodiment of the disclosure, the electronic device may include the control policy model including one or more convolutional neural networks (CNNs). The electronic device may include the control policy model including a multilayer perceptron (MLP) layer that maps an output of the CNNs into parameters of a distribution of robot control actions.
- [156] According to an embodiment of the disclosure, a method, wherein the object placement pose of the manipulator may include information of rotational positions comprising a yaw angle, a pitch angle, and a roll angle of the manipulator. The method, wherein the object placement pose of the manipulator may include information of transitional positions comprising an x-direction position, a y-direction position, and a z-direction position of the manipulator.
- [157] According to an embodiment of the disclosure, the method may include acquiring the object retrieval trajectory by obtaining a first set of manipulator grasping poses via a grasping planning neural network. The method may include acquiring the object retrieval trajectory by obtaining a second set of manipulator grasping poses by pruning grasping poses that do not belong to an object class of interest, from the first set of grasping poses. The method may include acquiring the object retrieval trajectory by determining a grasping pose from the second set of manipulator grasping poses. The method may include acquiring the object retrieval trajectory by acquiring the object retrieval trajectory that extends from the determined grasping pose to a clearance pose.
- [158] According to an embodiment of the disclosure, the method, wherein a tactile sensor mounted on the manipulator, and the method may include acquiring the object retrieval trajectory by identifying a center position of a plurality of contact points in a contact area between the tactile sensor and the target object. The method, wherein a tactile sensor mounted on the manipulator, and the method may include acquiring the object retrieval trajectory by determining a position vector for changing the center position of the plurality of contact points to a predetermined stable position. The method, wherein a tactile sensor mounted on the manipulator, and the method may include acquiring the object retrieval trajectory by outputting a robot control command based on the position vector to place the manipulator at the predetermined stable position. The method, wherein a tactile sensor mounted on the manipulator, and the method may include acquiring the object retrieval trajectory by acquiring the object retrieval trajectory by executing an object retrieval task from the predetermined stable position.
- [159] According to an embodiment of the disclosure, the method may include acquiring the object placement trajectory by downsampling the object retrieval trajectory. The method may include acquiring the object placement trajectory by downsampling the object retrieval trajectory to match a predetermined control policy frequency. The

method may include acquiring the object placement trajectory by reversing the downsampled object retrieval trajectory into the object placement trajectory. The method may include acquiring the object placement trajectory by generating the control policy model by training a policy learning model based on the object placement trajectory and the image of the scene including the target object.

[160] According to an embodiment of the disclosure, the method may include acquiring the object placement trajectory by commanding the robot to position the manipulator at a grasping position. The method may include acquiring the object placement trajectory by changing a rotational stiffness level and a translation stiffness level of the manipulator from original values to predetermined minimum values. The method may include acquiring the object placement trajectory by commanding the robot to close a gripper of the manipulator. The method may include acquiring the object placement trajectory by acquiring the object retrieval trajectory while the rotational stiffness level and the translation stiffness level are set to the predetermined minimum values.

[161] According to an embodiment of the disclosure, the method may include the policy model including one or more convolutional neural networks (CNNs). The method may include a multilayer perceptron (MLP) layer that maps an output of the CNNs into parameters of a distribution of robot control actions.

Claims

- [Claim 1] An electronic device for controlling a robot including a manipulator, the electronic device comprising:
one or more processors configured to:
acquire an image of a scene including a target object;
input the image into a control policy model to obtain an object placement pose of the manipulator, as an output of the control policy model, wherein the control policy model is generated using data collected by a data collection system that is configured to acquire an object retrieval trajectory by observing a robot movement for object retrieval, and reverse the object retrieval trajectory into an object placement trajectory; and
command the robot to position the manipulator according to the object placement pose, to place the target object at a designated location.
- [Claim 2] The electronic device of claim 1, further comprising a vision sensor configured to capture the image of the scene including the target object.
- [Claim 3] The electronic device any one of claims 1 to 2, wherein the object placement pose of the manipulator that is output from the control policy model comprises:
information of rotational positions comprising a yaw angle, a pitch angle, and a roll angle of the manipulator; and
information of transitional positions comprising an x-direction position, a y-direction position, and a z-direction position of the manipulator.
- [Claim 4] The electronic device any one of claims 1 to 3, wherein the data collection system is configured to:
obtain a first set of manipulator grasping poses via a grasping planning neural network;
obtain a second set of manipulator grasping poses by pruning grasping poses that do not belong to an object class of interest, from the first set of grasping poses;
determine a grasping pose from the second set of manipulator grasping poses; and
obtain the object retrieval trajectory that extends from the determined grasping pose to a clearance pose.
- [Claim 5] The electronic device any one of claims 1 to 4, wherein the data collection system comprises a tactile sensor mounted on the manipulator and is configured to:

identify a center position of a plurality of contact points in a contact area between the tactile sensor and the target object;
determine a position vector for changing the center position of the plurality of contact points to a predetermined stable position; and
output a robot control command based on the position vector.

[Claim 6] The electronic device any one of claims 1 to 5, wherein the data collection system is configured to command the robot to re-grasp the target object based on the robot control command including the position vector to acquire the object retrieval trajectory.

[Claim 7] The electronic device any one of claims 1 to 6, wherein the data collection system is further configured to:
determine a retrieval pose of the manipulator based on a movement path extending from a grasping pose to a clearance pose of the manipulator;
position the manipulator according to the retrieval pose and store the retrieval pose to generate the object retrieval trajectory based on the retrieval pose;
downsample the object retrieval trajectory to match a predetermined control policy frequency;
reverse the downsampled object retrieval trajectory into the object placement trajectory; and
generate the control policy model by training a policy learning model based on the object placement trajectory and the image of the scene including the target object.

[Claim 8] The electronic device any one of claims 1 to 7, wherein the data collection system is configured to command the robot to position the manipulator at a grasping position,
change a rotational stiffness level and a translation stiffness level of the manipulator from original values to predetermined minimum values,
command the robot to close a gripper of the manipulator, and
acquire the object retrieval trajectory while the rotational stiffness level and the translation stiffness level are set to the predetermined minimum values.

[Claim 9] A method for controlling a robot including a manipulator, the method comprising:
acquiring an image of a scene including a target object;
inputting the image into a control policy model to obtain an object placement pose of the manipulator, as an output of the control policy

model, wherein the control policy model is generated using data collected by a data collection system that is configured to acquire an object retrieval trajectory by observing a robot movement for object retrieval, and reverse the object retrieval trajectory into an object placement trajectory; and
commanding the robot to position the manipulator according to the object placement pose, to place the target object at a designated location.

[Claim 10]

The method of claim 9, wherein the object placement pose of the manipulator comprises:

information of rotational positions comprising a yaw angle, a pitch angle, and a roll angle of the manipulator; and

information of transitional positions comprising an x-direction position, a y-direction position, and a z-direction position of the manipulator.

[Claim 11]

The method any one of claims 9 to 10, further comprising acquiring the object retrieval trajectory by:

obtaining a first set of manipulator grasping poses via a grasping planning neural network;

obtaining a second set of manipulator grasping poses by pruning grasping poses that do not belong to an object class of interest, from the first set of grasping poses;

determining a grasping pose from the second set of manipulator grasping poses; and

acquiring the object retrieval trajectory that extends from the determined grasping pose to a clearance pose.

[Claim 12]

The method any one of claims 9 to 11, wherein a tactile sensor mounted on the manipulator, and the method further comprises acquiring the object retrieval trajectory by:

identifying a center position of a plurality of contact points in a contact area between the tactile sensor and the target object;

determining a position vector for changing the center position of the plurality of contact points to a predetermined stable position;

outputting a robot control command based on the position vector to place the manipulator at the predetermined stable position; and

acquiring the object retrieval trajectory by executing an object retrieval task from the predetermined stable position.

[Claim 13]

The method of any one of claims 9 to 12, further comprising acquiring the object placement trajectory by:

downsampling the object retrieval trajectory to match a predetermined control policy frequency;
reversing the downsampled object retrieval trajectory into the object placement trajectory; and
generating the control policy model by training a policy learning model based on the object placement trajectory and the image of the scene including the target object.

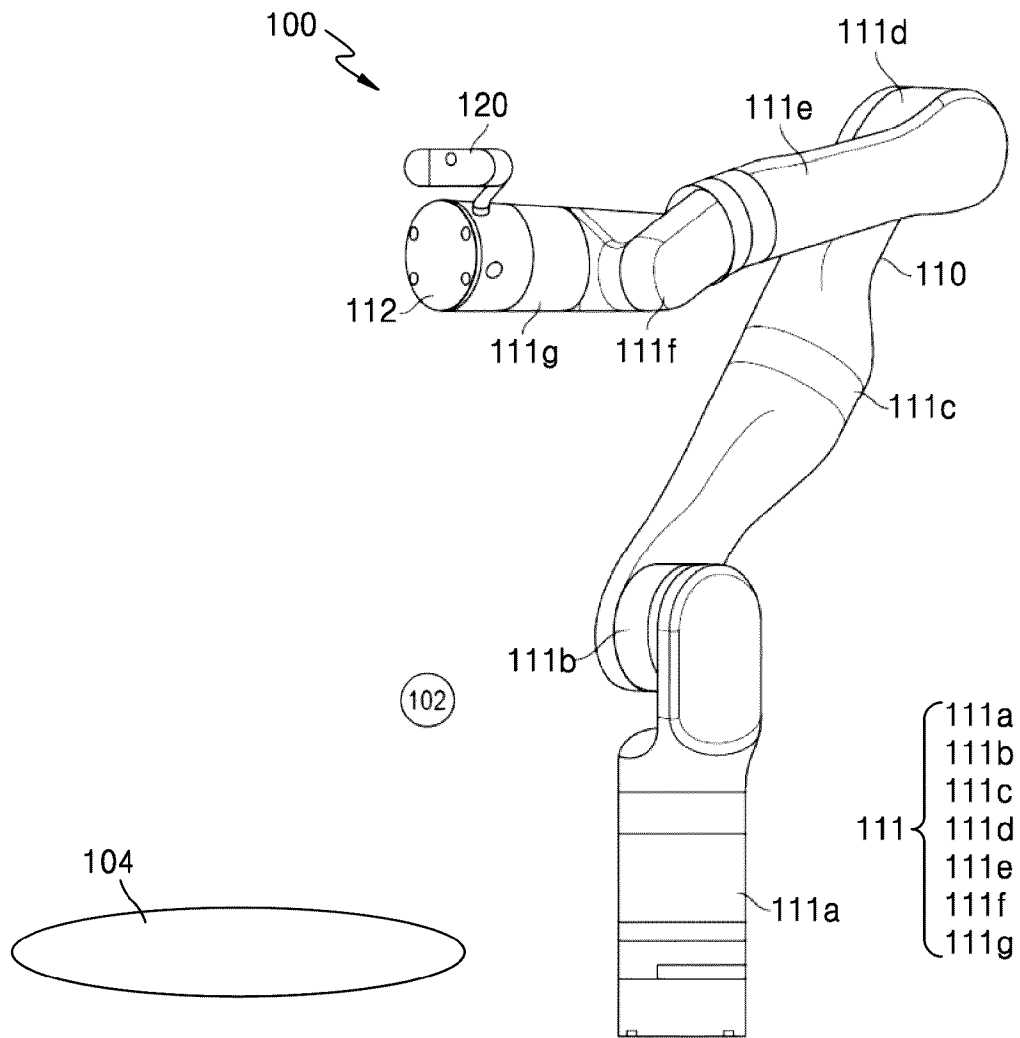
[Claim 14]

The method any one of claims 9 to 13, further comprising acquiring the object retrieval trajectory by:
commanding the robot to position the manipulator at a grasping position;
changing a rotational stiffness level and a translation stiffness level of the manipulator from original values to predetermined minimum values;
commanding the robot to close a gripper of the manipulator, and
acquiring the object retrieval trajectory while the rotational stiffness level and the translation stiffness level are set to the predetermined minimum values.

[Claim 15]

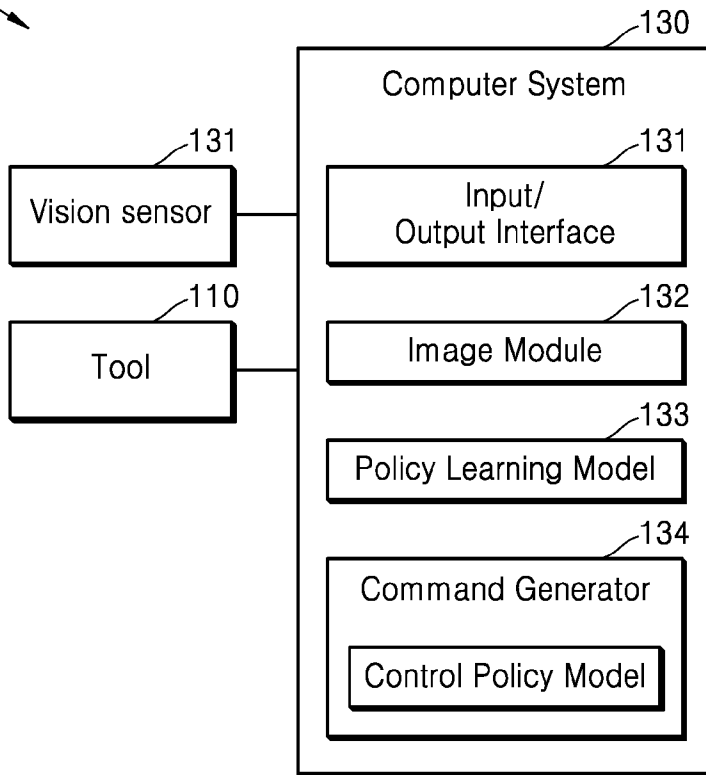
A computer-readable storage medium storing instructions, wherein the instructions, when executed by at least one processor, cause the at least one processor to perform the method of any one of claims 9 to 14.

[Fig. 1A]



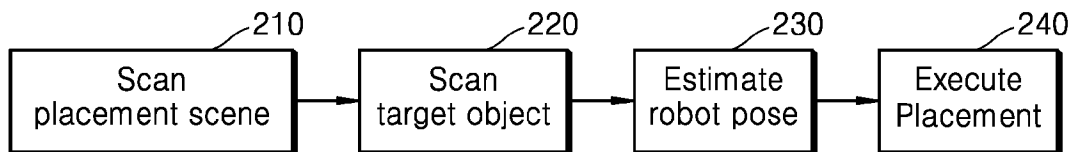
[Fig. 1B]

100

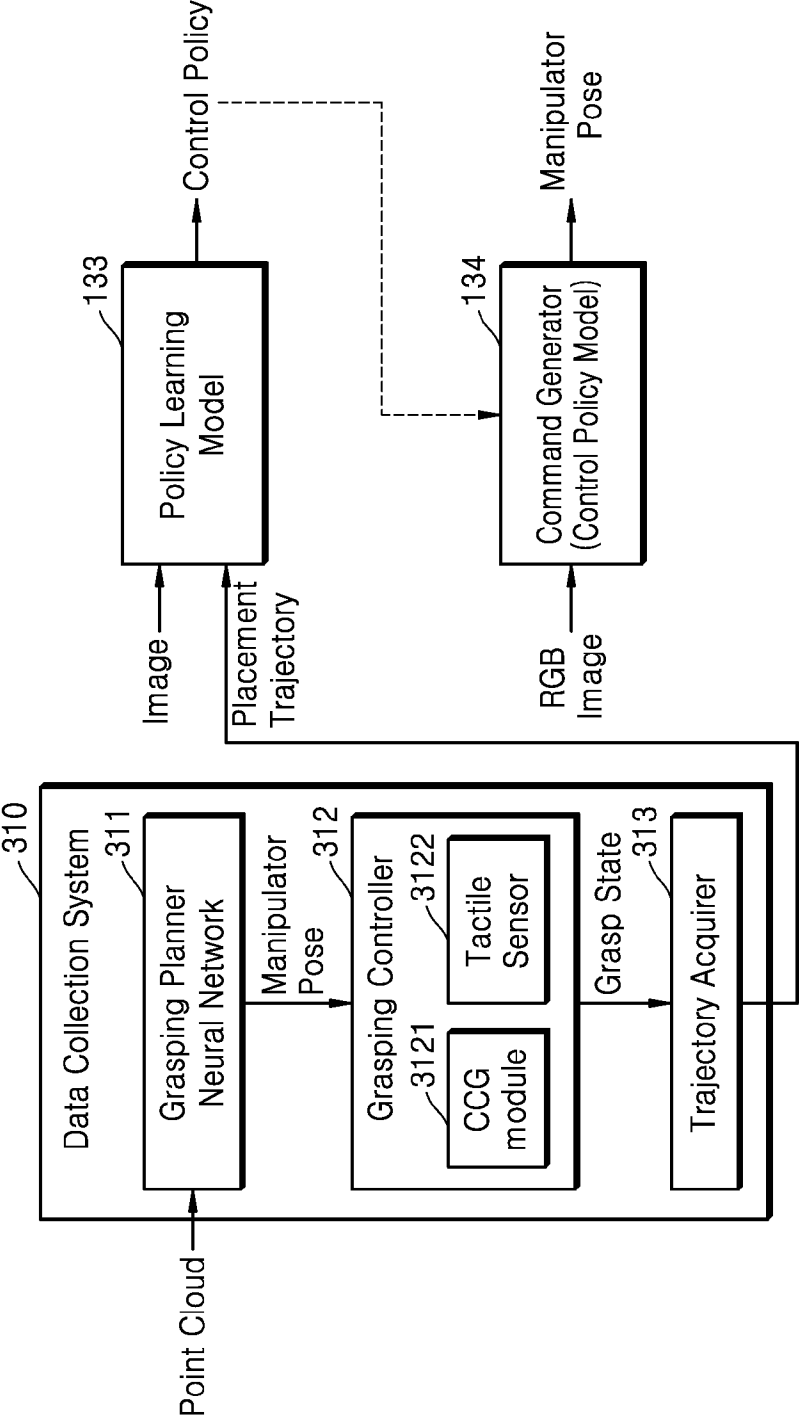


[Fig. 2]

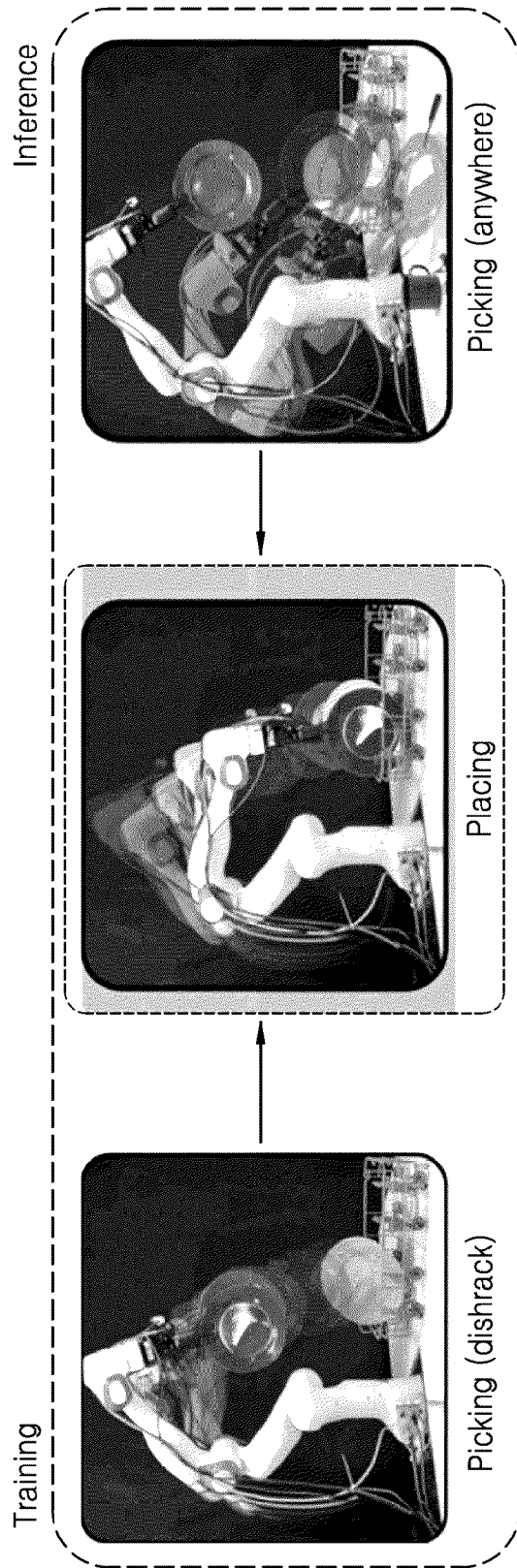
200



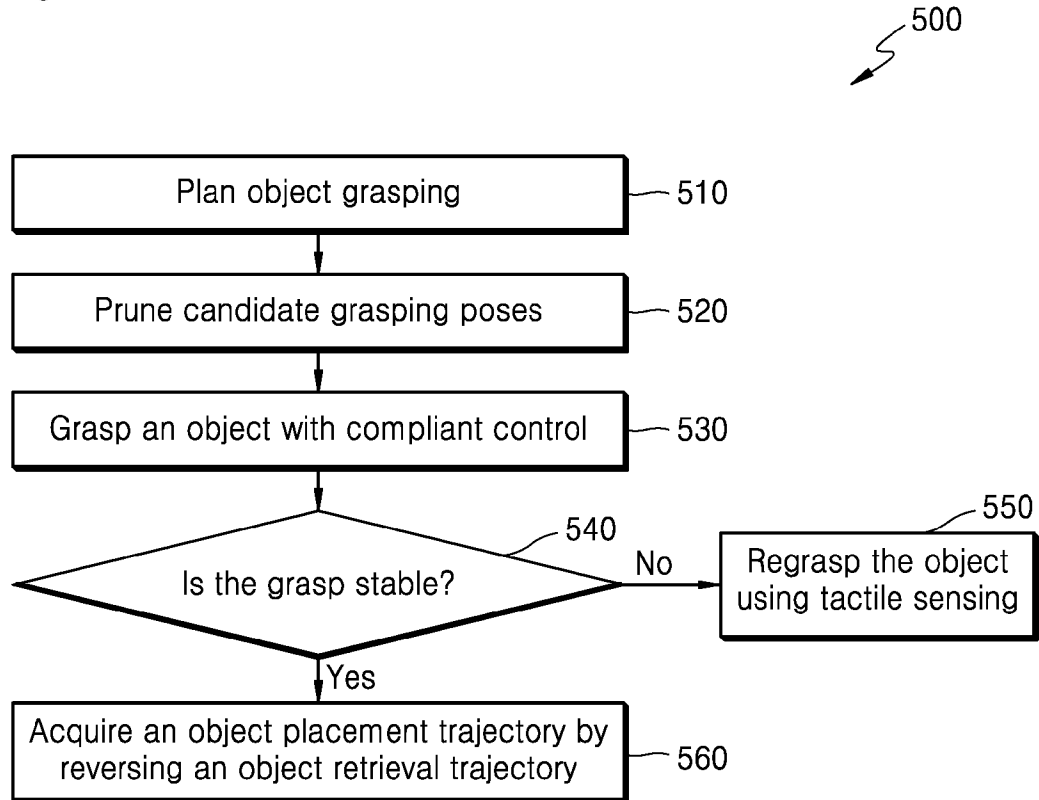
[Fig. 3]



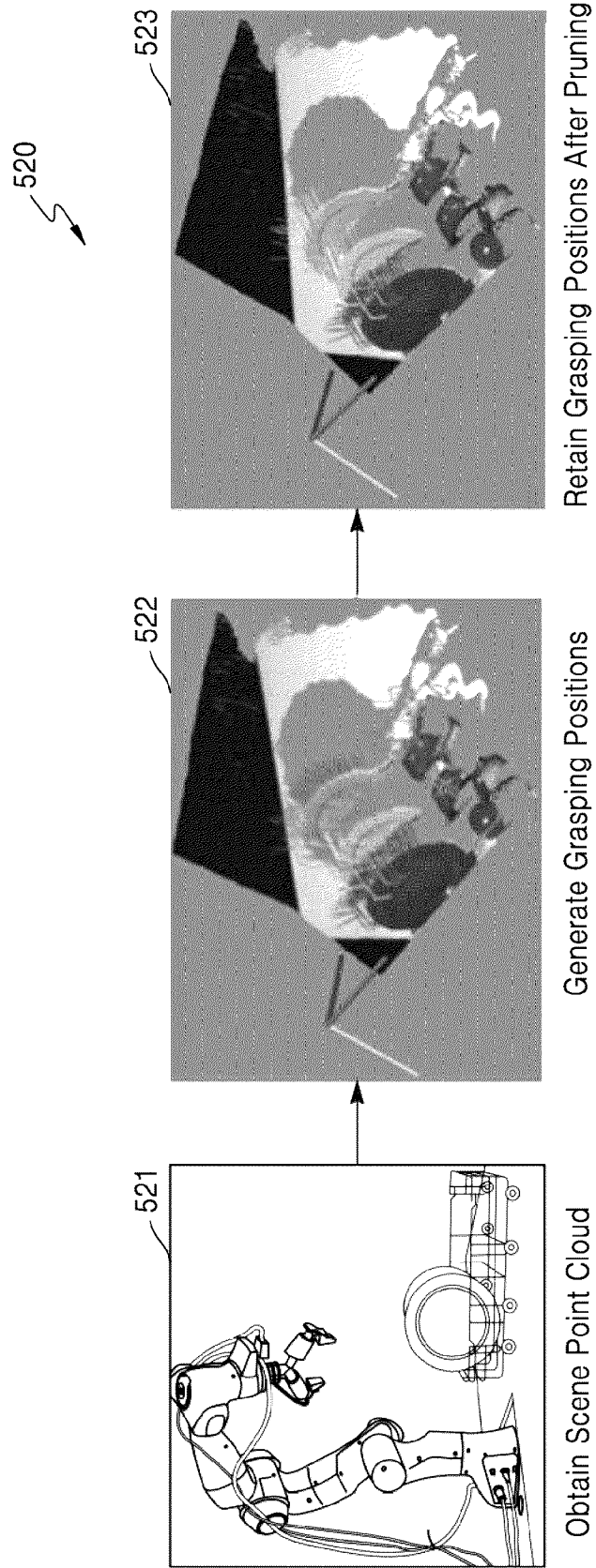
[Fig. 4]



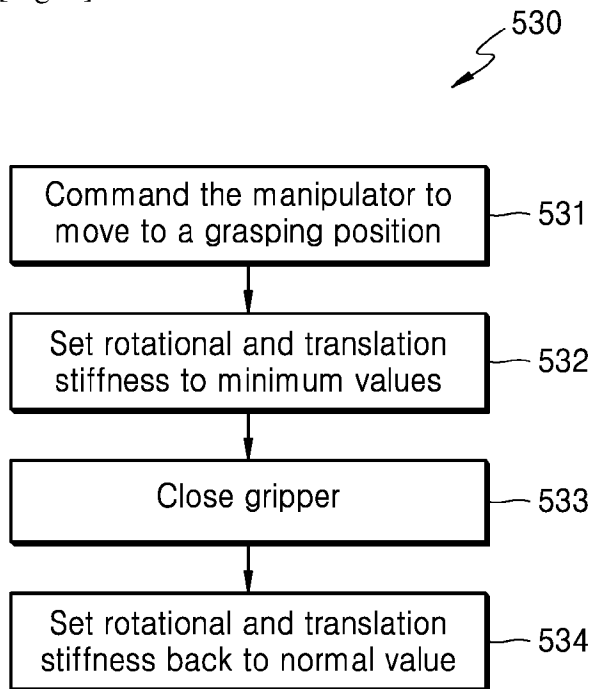
[Fig. 5]



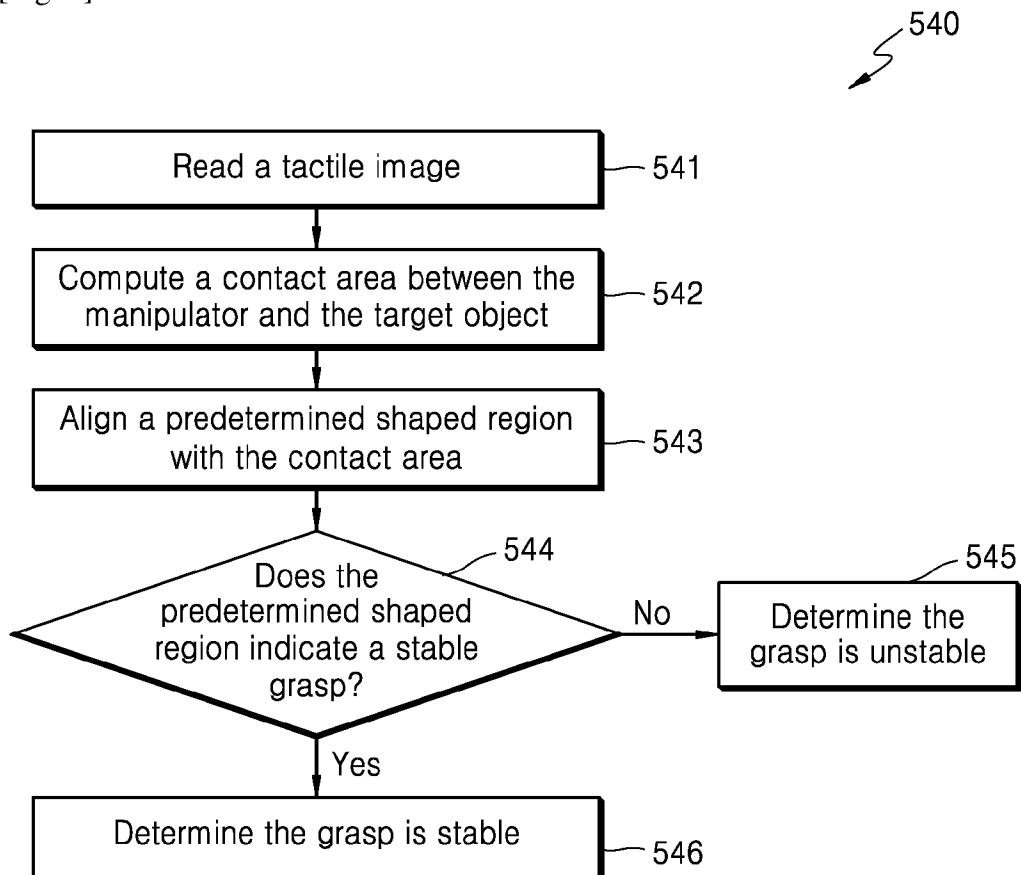
[Fig. 6]



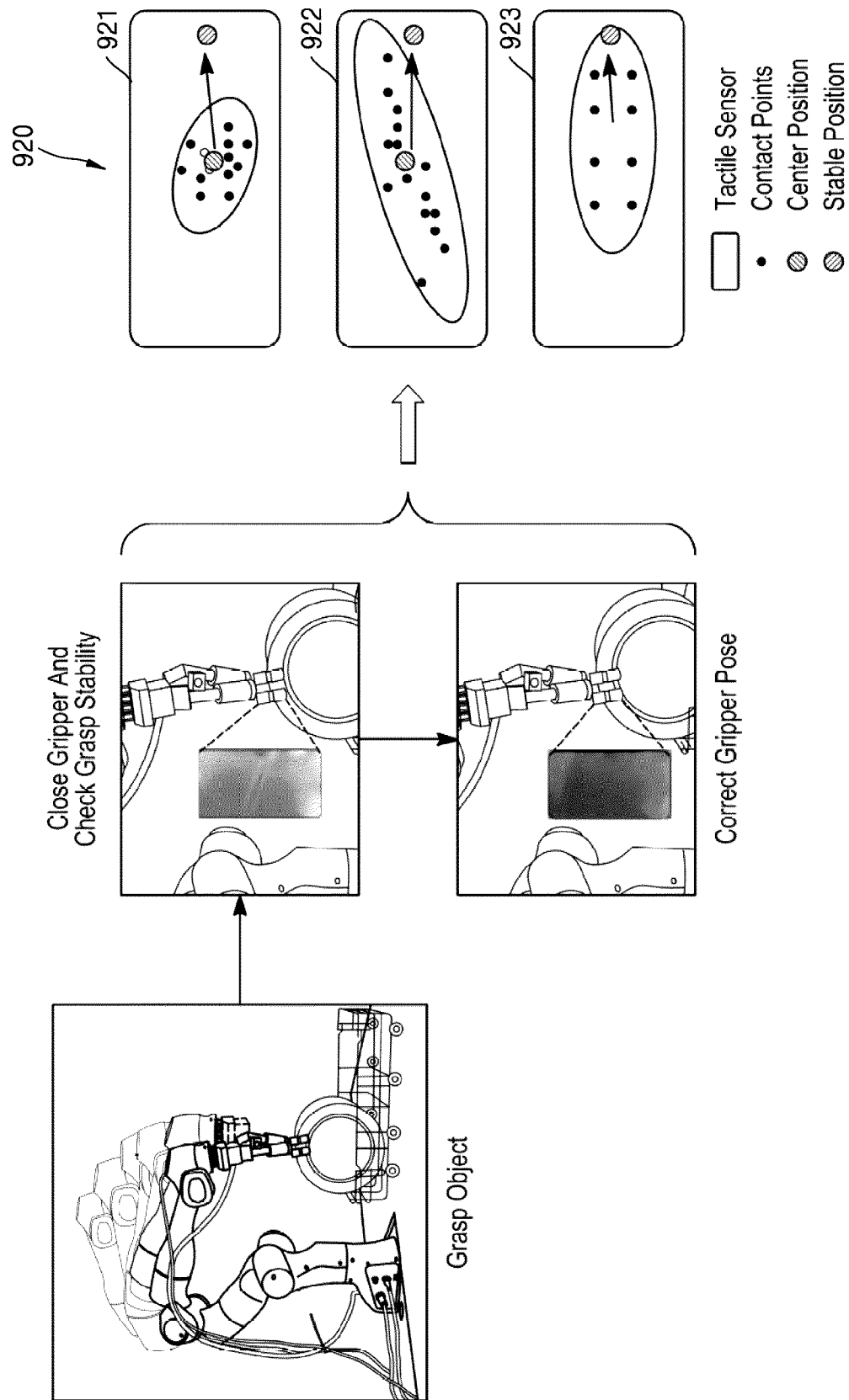
[Fig. 7]



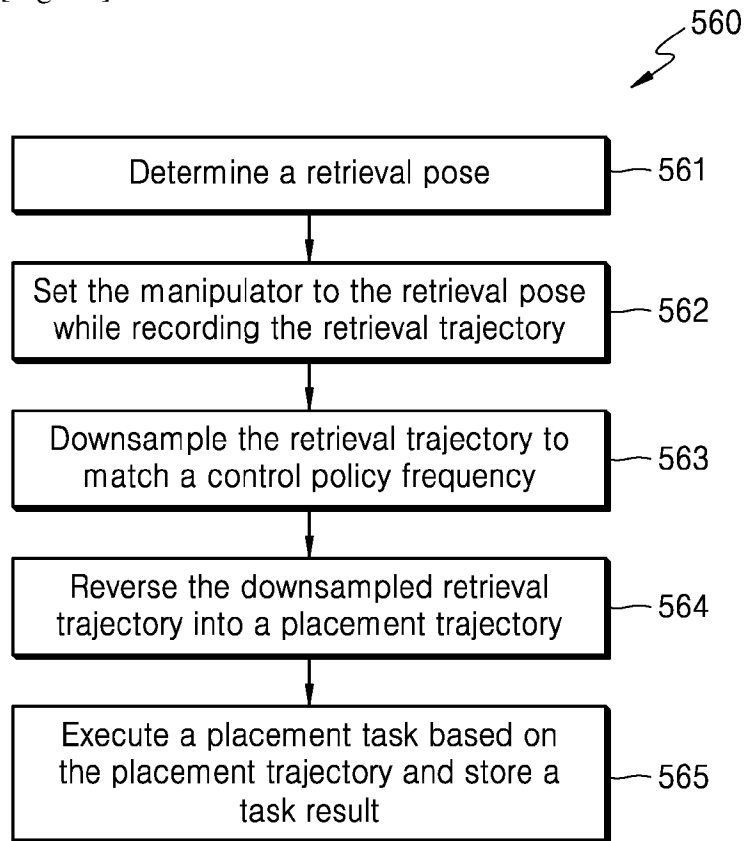
[Fig. 8]



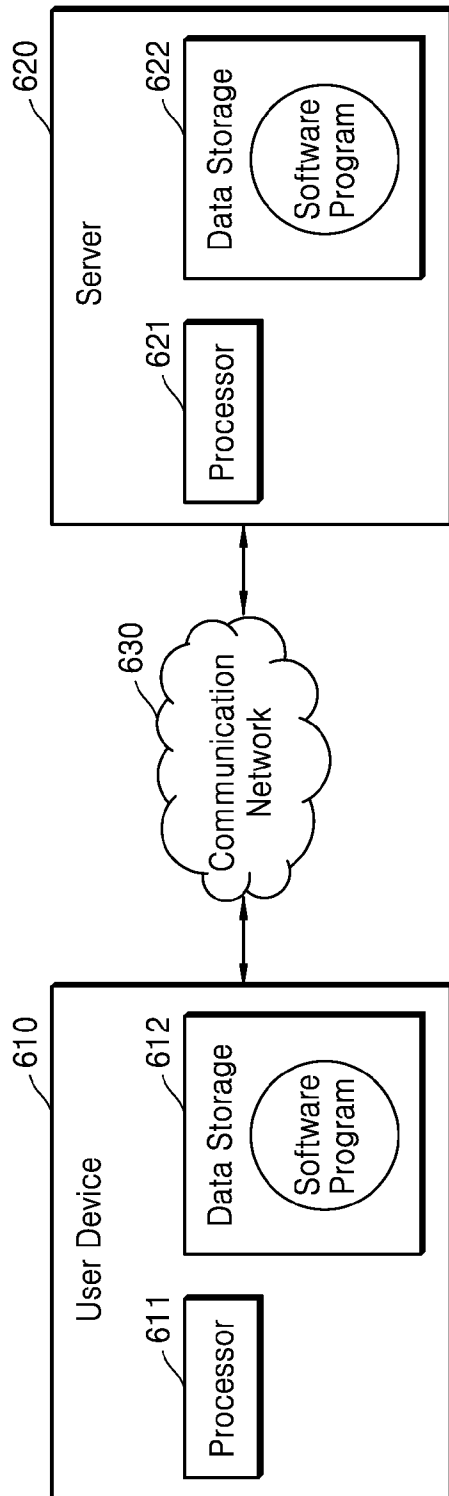
[Fig. 9]



[Fig. 10]

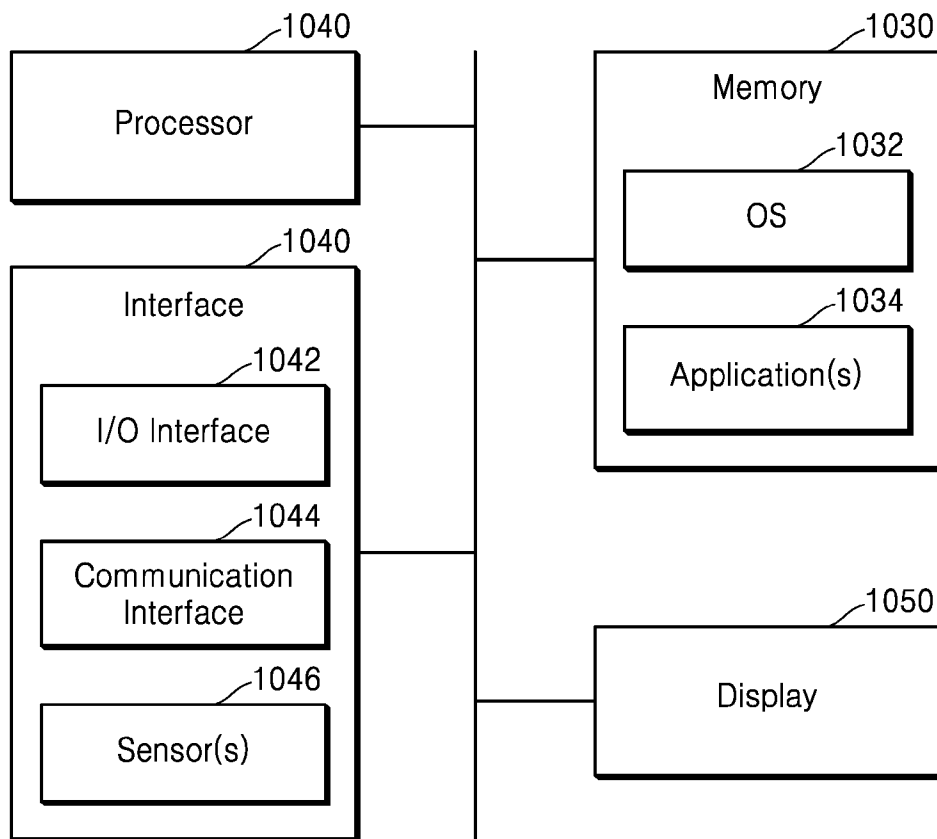


[Fig. 11]



[Fig. 12]

1000



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2024/095301

A. CLASSIFICATION OF SUBJECT MATTER B25J 13/08(2006.01)i; B25J 9/16(2006.01)i; B25J 19/02(2006.01)i 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) B25J 13/08(2006.01); B23P 21/00(2006.01); B25J 15/06(2006.01); B25J 9/16(2006.01); B25J 9/22(2006.01); G05B 19/418(2006.01); G06F 16/28(2019.01) Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal) & Keywords: robot, manipulator, image, control policy model, object placement pose, object retrieval trajectory, object placement trajectory, neural network		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	US 2022-0072706 A1 (BERKSHIRE GREY, INC.) 10 March 2022 (2022-03-10) paragraphs [0030]-[0053] and figures 1-7	1-3,9-10,15 4-8,11-14
Y	US 2022-0075355 A1 (HITACHI, LTD.) 10 March 2022 (2022-03-10) claim 1	1-3,9-10,15
A	US 2022-0355474 A1 (MUJIN, INC.) 10 November 2022 (2022-11-10) paragraphs [0021]-[0084], claims 1-3, and figure 3A	1-15
A	JP 2021-169149 A (FANUC LTD.) 28 October 2021 (2021-10-28) claim 1 and figure 9	1-15
A	JP 2021-003779 A (HITACHI LTD.) 14 January 2021 (2021-01-14) claim 1 and figures 1, 8	1-15
☐ 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 “D” document cited by the applicant in the international application “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 24 May 2024		Date of mailing of the international search report 24 May 2024
Name and mailing address of the ISA/KR Korean Intellectual Property Office 189 Cheongsa-ro, Seo-gu, Daejeon 35208, Republic of Korea Facsimile No. +82-42-481-8578		Authorized officer PARK, Tae Wook Telephone No. +82-42-481-3405

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2024/095301

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2022-0072706	A1	10 March 2022	CA	3014049	A1	17 August 2017
				CA	3014049	C	22 June 2021
				CA	3117170	A1	17 August 2017
				CA	3117170	C	29 August 2023
				CA	3178174	A1	17 August 2017
				CN	108778636	A	09 November 2018
				CN	108778636	B	19 November 2021
				CN	113894826	A	07 January 2022
				EP	3414060	A1	19 December 2018
				EP	3414060	A4	18 September 2019
				EP	3414060	B1	25 May 2022
				EP	3988257	A1	27 April 2022
				EP	3988257	B1	03 May 2023
				ES	2924496	T3	07 October 2022
				ES	2949949	T3	04 October 2023
				US	10350755	B2	16 July 2019
				US	11123866	B2	21 September 2021
				US	11213949	B2	04 January 2022
				US	11724394	B2	15 August 2023
				US	2017-0225330	A1	10 August 2017
				US	2019-0314988	A1	17 October 2019
				US	2019-0329410	A1	31 October 2019
				US	2023-0347515	A1	02 November 2023
				WO	2017-139330	A1	17 August 2017
US	2022-0075355	A1	10 March 2022	JP	2022-044127	A	17 March 2022
				JP	7422632	B2	26 January 2024
US	2022-0355474	A1	10 November 2022	CN	115697645	A	03 February 2023
				CN	116638526	A	25 August 2023
				EP	4334092	A1	13 March 2024
				JP	2022-172465	A	16 November 2022
				JP	2023-041731	A	24 March 2023
				JP	7272568	B2	12 May 2023
				WO	2022-235658	A1	10 November 2022
JP	2021-169149	A	28 October 2021	CN	113524168	A	22 October 2021
				DE	102021109195	A1	21 October 2021
				US	11813756	B2	14 November 2023
				US	2021-0323154	A1	21 October 2021
JP	2021-003779	A	14 January 2021	JP	7246267	B2	27 March 2023
				US	11707845	B2	25 July 2023
				US	2020-0406460	A1	31 December 2020