Self-protective motion planning for mobile manipulators in a dynamic door-closing workspace

Chuande Liu and Bingtuan Gao
Southeast University, Nanjing, China, and
Chuang Yu and Adriana Tapus
ENSTA ParisTech, Palaiseau, France

Abstract

Purpose – Many work conditions require manipulators to open cabinet doors and then gain access to the desired workspace. However, after opening, the unlocked doors can easily close, interrupt a task and potentially break the operating end-effectors. This paper aims to address a manipulator's behavior planning problem for responding to a dynamic workspace released by door opening.

Design/methodology/approach – A dynamic model of the restricted workspace released by an unlocked door is established. As a whole system to treat, the interactions between the workspace and robot are analyzed by using a partially observable Markov decision process. A self-protective policy decision executed as a belief tree is proposed. To respond to the policy, this study has designed three types of actions: stay on guard in the workspace, using an elbow joint to defense the door and linear escape out of the workspace for self-protection by observing collision risk levels to trigger them. Finally, this study proposes self-protective motion controllers based on risk time optimization to act to the planned actions.

Findings – The elbow defense could balance robotic safety and work efficiency by interrupting the end-effector's work and using the elbow joint to prevent the door-closing in an active collision way. Compared with the stay and escape action, the advantage of the elbow defense is having a predictable performance to quick callback the interrupted work after the risk was relieved.

Originality/value – This work provides guidance for the safe operation of a class of robot operations and the upgrade of motion planning.

Keywords Robot operation, Dynamic workspace, Physical interaction, Mobile manipulator, Self-protective behaviors, Motion planning, Markov decision

Paper type Research paper

1. Introduction

Self-protection is the most essential motor behavior to assure mobile manipulator robots' survival while performing a desired task in real human–robot coexistence environments (Shimizu *et al.*, 2012). Beyond the basic capabilities of moving and acting autonomously, the robots should protect themselves from harmful states or collisions when physically interacting with their workspace (Aoude *et al.*, 2013). The self-protective response in humans is fast and coordinated even when a collision is not anticipated. Figure 1 shows when people open a refrigerator door, facing the refrigerator, to fetch some foods inside. It is a natural behavior using the hands to absorb the impact of the opened refrigerator door and keep the door steady and far away from main body. Especially in such cases, both occupied by the current time-consuming internal manipulation, human upper limbs are adjusted to defense proactively the refrigerator door close, which could interrupt the work, by fast motor reflexes (Bauer *et al.*, 2010).

People have high hopes for robots to replace manual labor to the open refrigerator to fetch foods or similar task scenarios (Nemec *et al.*, 2017). These task conditions may need robots to open doors first (Su and Chen, 2019) and then gain access to the workspace for their end-effectors (Klingbeil *et al.*, 2010). For example, use robots to inspect the electronic equipment in power substations (Wang *et al.*, 2020). Typical substation inspections are involved a large number of refrigerator-like electric cabinets equipped with electronic monitoring, which need to be checked at close range or operated by hands after opening the cabinet door. Unfortunately, the mobile manipulator robots integrated intelligent cabinet inspection technologies are always not well trained such knowledge on standby, to take appropriate measures to deal with potential dynamic interference coming from the opened cabinet doors.

In this paper, we think that self-protection is not only crucial but necessary in refrigerator-like door opening activities to facilitate their further task success. We then attempt to plan and create self-protective actions for mobile manipulators in the cabinet's inner workspace to respond to the unlocked door closing based on human-like synergistic manner.

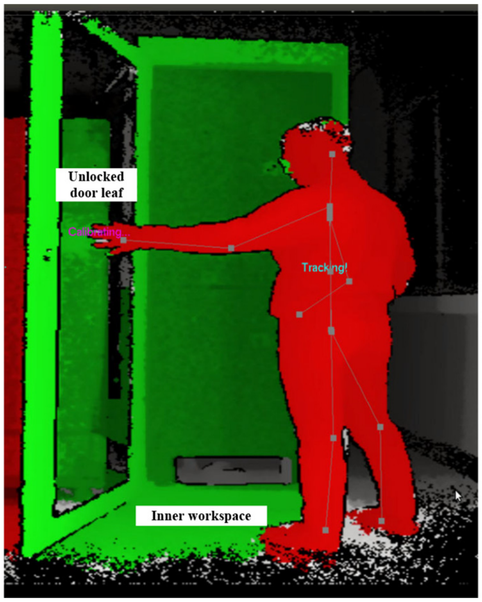
The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/0143-991X.htm**

Industrial Robot: the international journal of robotics research and application

[DOI 10.1108/IR-02-2021-0025]

Received 1 February 2021
Revised 28 February 2021
27 March 2021
Accepted 28 March 2021

Figure 1 Open refrigerator to fetch foods scenario



Our main contributions of this paper are as follows:

- To the best of our knowledge, this paper yields the first self-protective motion planning for dynamic interactions in a door-closing workspace, within the collision risk consideration coming from environmental uncertainty.
- This paper is to balance robotic safety and its work performance during the door-closing emergency by planning three self-protective actions: stay on guard in the workspace, using a joint elbow to defense the door, and linear escape out of the workspace, respectively, to the low, middle, and high collision risk levels, which is verified real true based on the experiments with our build-up robot platform.
- This paper is to provide guidance for the safe operation of a class of robot operations and the upgrade of motion planning.

The rest of this paper is organized as follows. Related works are described in Section 2; Section 3 explains the problem formulation. A novel self-protective motion planning is proposed in Section 4. Section 5 validates the efficiency of the proposed method by experiments. Finally, conclusions are drawn in Section 6.

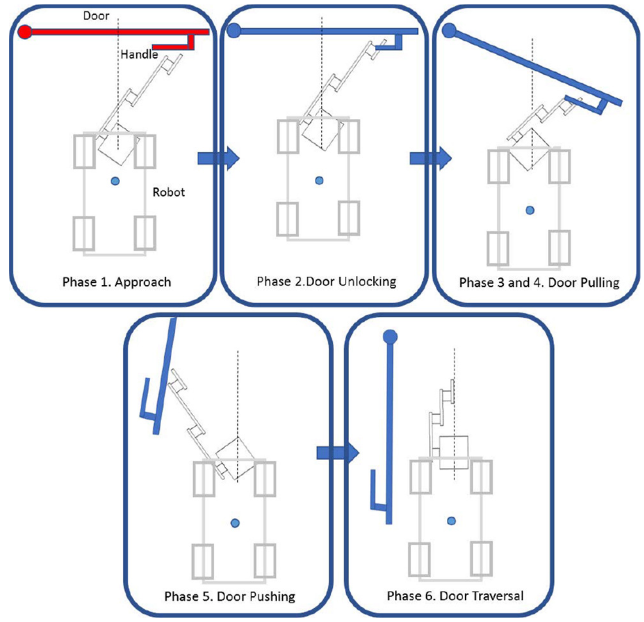
2. Related work

Like the task aforementioned to inspect the refrigerator-like electric cabinets, opening the door is the prerequisite and complicated part of the whole mission. During the last decades, the control design or machine learning (Yahya *et al.*, 2017) for the door-opening problem has received abundant attention, and there are many typical steps to open the door, as shown in Figure 2, including but not limited to the followings:

Step 1: locate to approach the door handle (Adiwahono *et al.*, 2013).

Step 2: rotate the handle to door unlocking (Kalakrishnan *et al.*, 2011).

Figure 2 Door-opening behavior to construct internal constrained workspace in top view (Prieto *et al.*, 2019)



Step 3: the robot pulls to release the door leaf from the frame at a small angle, e.g. 10°–15° (Chung *et al.*, 2009).

Step 4: and then move the robotic arm to the other side of the door (Milighetti *et al.*, 2012).

Step 5: push the door to enlarge the internal space between leaf and frame (Abdo *et al.*, 2013).

After these mentioned steps, the excellent performances of the door-opening technologies save valuable time for subsequent operations and create enough large workspace.

When door-opening technologies work in practice, note that *Steps 4* and *5* not only make the door open wider quickly and efficiently with a pull force but also match external disturbances causing the door to have uncontrolled rotational inertia. Unless dealt with in a proper way, they would deteriorate the performances of the following operations and even give rise to inconsistent task results, which leads to mission failure. It is very true, especially in the outdoor area (Chan *et al.*, 2019); the sudden wind can also cause the unlocked door leaf untimely close during the robot occupied by the current task. For involving task such as opening the door to get handwork inside (Rühr *et al.*, 2012), which generally limits the robot task space and keeps the robot in the unlocked door leaf's adverse influence range for a long time, we cannot ignore the uncertain disturbances (Kim, 2019) coming from the unlocked door leaf leading to a potential risk of collision damages. To solve this problem, professional roboticists initially took a dual-arm mobile manipulator scheme (Valner *et al.*, 2018). More precisely, using one arm to defense the door-closing disturbances while planning another arm to handwork inside. They applied this theoretical pattern to an expensive PR2 (Personal Robot 2) to fetch a beer from a refrigerator (Beer me, robot, 2010). Based on this pattern, the scheme mentioned above even could be used in multi-arm robot systems; unfortunately, it is not friendly for robots with only one arm.

In this work, we focus on a single-arm mobile manipulator to respond to door closing.

3. Problem formulation

Consider a time-varying cabinet workspace $W(t)$ released from its prerequisite door-opening action, is constrained by the door frame D_{frame}, as well as the unlocked door leaf D_{leaf}, e.g.:

$$\mathcal{D}_{leaf}[\theta(t), \omega(t), t] \times \mathcal{D}_{frame} \rightarrow \mathcal{W}(t) \tag{1}$$

where $\theta(t)$ and $\omega(t)$ denote the angle and the angular velocity respectively. From the top view to see $W(t)$, Figure 3 shows the dynamic interactive progress, which seems like a shrinking Chinese folding fan when D_{leaf} is driven by the force such as a sudden wind $F_w(t)$. Simultaneously, due to the resisting force $F_r(t)$ coming from rotation friction and air resistance, D_{leaf} would stop close at a certain position p_n. After these, the state equation for $W(t)$ is written as:

$$\omega(t) = f[\theta(t), F_w(t) - F_r(t), t] \tag{2}$$

where $f(\cdot)$ denotes a time-variation function.

Assumptions: $W(t)$ is a dynamic door-closing workspace; the mobile manipulator's chassis equipped with some indispensable precision sensors is a collision-free part; ignore the end-effector's specific operations and implementation. After these, the goal of self-protective motion planning is to plan a policy π and then control to act the motions between the start configuration $q_0 \in \mathbb{R}^D$ and the goal configuration $q_d \in \mathbb{R}^D$, which can be written as:

$$\max\left\{V(\pi)|\pi : q_0 \rightarrow q_d \in \mathbb{R}^D\right\} \quad s.t. \quad \mathcal{W}(t) \tag{3}$$

where $V(\pi)$ is the expected total reward with the policy, q denotes the degree of freedom (DOF) in a robotic manipulator and D is the number of the DOF.

4. Proposed method

Figure 4 shows a partially observable Markov decision process (POMDP) architecture, which simulates the interaction relationship between agents decisions and their environment (Luo *et al.*, 2019), which models our mobile manipulator acting in $W(t)$. It is defined formally as a 7-tuple (S,A,Z,T,O,R,b_0), where:

S: indicates a states set of D_{leaf} at the current time;

A: indicates an actions set that the mobile manipulator will perform at the next moment;

Z: indicates an observations set of D_{leaf} at the current time;

Figure 3 Dynamic model of the restricted workspace

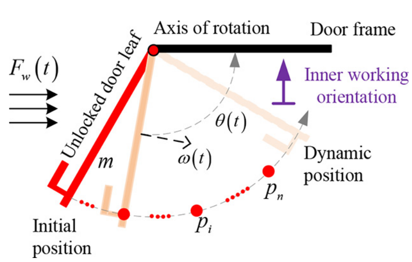
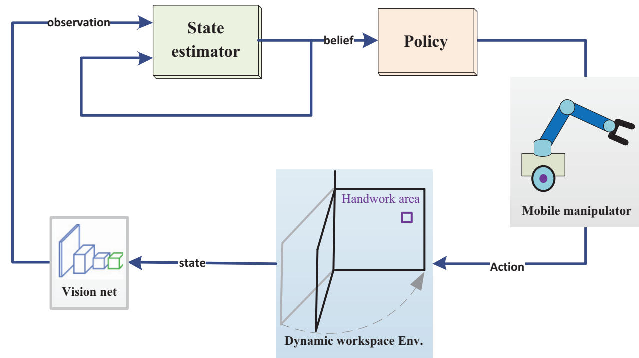


Figure 4 Pipeline in terms of the POMDP estimation and policy decision architecture



T: the function $T(s,a,s') = p(s'|s,a)$ indicates the probabilistic state transition from $s \in S$ to $s' \in S$, when the mobile manipulator in state $s \in S$ takes an action $a \in A$;

O: the function $O(s,a,z) = p(z|s,a)$ indicates a set of conditional observation probabilities currently observed;

R: the function $R(s, a)$ defines a real-valued reward for the mobile manipulator when it takes action $a \in A$ in state $s \in S$.

4.1 Decision-making to self-protective actions

As analyzed previously, the POMDP planning aims to choose a policy π that maximizes its value based on A and S, but S is not known exactly because of imperfect observation. Instead, the mobile manipulator maintains a belief, which is a probability distribution over S. The mobile manipulator starts with an initial belief b_0. At time t, it infers a new belief, according to Bayes' rule (Wang *et al.*, 2019), by incorporating information from the action a_t taken and the observation z_t received:

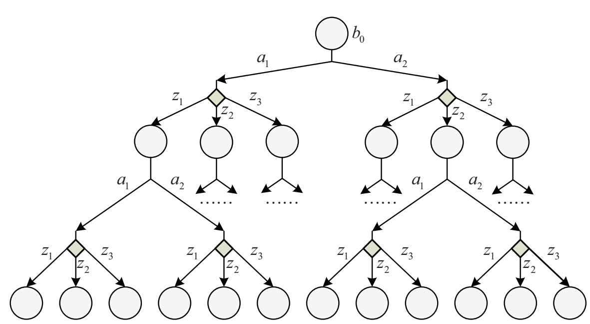
$$\begin{aligned} b_t(s') &= \tau(b_{t-1}, a_t, z_t) \\ &= \eta O(s', a_t, z_t) \sum_{s \in \mathcal{S}} T(s, a_t, s') b_{t-1}(s) \end{aligned} \tag{4}$$

where η is a normalizing constant.

Figure 5 shows that a POMDP policy prescribes the action at a belief. With the policy π and an initial belief b_0, the expected total discounted reward V can be written as:

$$V_\pi(b_0) = \left(\sum_{t=0}^{\infty} \gamma^t R(s_t, a_{t+1}) | b_0, \pi\right) \tag{5}$$

Figure 5 POMDP planning performs lookahead search on a belief tree



where s_t is the state at time t, $a_{t+1} = \pi(b_t)$ is the action that the policy π chooses at time t, and $\gamma \in [0,1]$ is a discount factor. The expectation V is taken over the sequence of uncertain state transitions and observations over time.

A key idea in POMDP planning is the belief tree (Kaelbling and Lozano-Pérez, 2013), as shown in Figure 5. Each node of a belief tree corresponds to a belief b. At each node, the tree branches on all actions in A and all observations in Z. If a node with belief b has a child node with belief b', then $b' = \pi(b,a,z)$. Conceptually, we may think of POMDP planning as a tree search in the belief space, the space of all possible beliefs that the mobile manipulator may encounter (Osa, 2020). To find an optimal plan for a POMDP, we traverse the belief tree from the bottom up and compute an optimal action recursively at each node using Bellman's equation (Thrun, 2002):

$$V^*(b) = \max_{a \in \mathcal{A}} \left\{ \sum_{s \in \mathcal{S}} b(s) R(s,a) + \gamma \sum_{z \in \mathcal{Z}} p(z|b,a) V^*(\tau(b,a,z)) \right\} \tag{6}$$

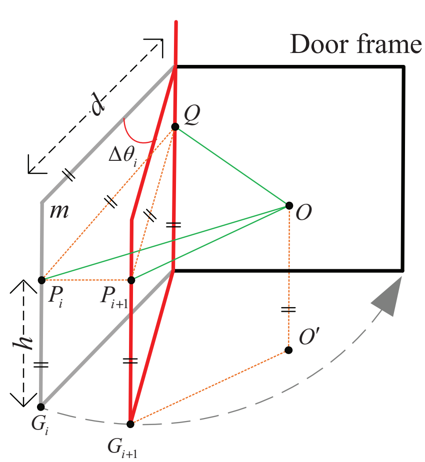
Based on the above discussions, in the sense that, our POMDP planning is a special case of belief space planning. In other words, the belief space planning is more general and does not require the planning model to satisfy the mathematical structure of POMDPs. For example, the reward function R may depend on the belief b and not just on S and A.

Additionally, at each node, all observations in Z are key points for the searching progress, for the reason is the following child node of the belief tree branches on all possible actions in A.

4.2 Observations design to the door-closing workspace

Figure 6 shows all the observations in Z. Let O denote the robot's sensing position, which is attached on the mobile manipulator. P_i, P_{i+1} and Q denote three marked feature points on the door frame and they are coplanar with O. $|OO'|$ is parallel to $|P_{i+1}G_{i+1}|$ and $|OO'| = |P_{i+1}G_{i+1}| = |P_iG_i| = h$ where h denotes the height between the marked point and the ground.

Figure 6 Observation progress for the dynamic states of the unlocked door leaf D_{leaf}



Likewise, $|P_iQ|$ is parallel to d and $|P_iO|, |P_{i+1}O| = d$ where d denotes the unlocked door leaf's width.

In such case, we can get $|P_iO|, |P_{i+1}O|$ and $|OQ|$ by measurement. According to the geometric relationship, the observed rotation angle $\Delta\hat{\theta}_i$ can be written as:

$$\Delta\hat{\theta} = \angle P_iQP_{i+1} = \angle P_iQO - \angle P_{i+1}QO \tag{7}$$

where:

$$\angle P_iQO = \arccos \frac{|P_iQ|^2 + |OQ|^2 - |OP_i|^2}{2|P_iQ| \cdot |OQ|}$$

$$\angle P_{i+1}QO = \arccos \frac{|P_{i+1}Q|^2 + |OQ|^2 - |OP_i|^2}{2|P_{i+1}Q| \cdot |OQ|}$$

For D_{leaf}, the moment of inertia around the door axis is:

$$I = \frac{1}{3} md^2 \tag{8}$$

where m denotes the D_{leaf} mass. Based on equations (7) and (8), the observed angular kinetic energy $\hat{E}_{\mathcal{D}_{leaf}}$ around the door axis can be written as:

$$\hat{E}_{\mathcal{D}_{leaf}} = \frac{1}{2} I \hat{\omega}^2 \tag{9}$$

where $\hat{\omega} = \Delta\hat{\theta} / \Delta t$ and Δt denotes the observation of time unit.

In this paper, $\hat{E}_{\mathcal{D}_{leaf}}$ indicates the so-called risk to cause collision damages. Combing with equation (9), we treat the risk levels as inputs and divide them into three parts by the Bang-Bang controller, which can be written as:

$$\mathcal{Z} = \begin{cases} z_1 = low\ risk & \hat{E}_{\mathcal{D}_{leaf}} \le E_{\min} \\ z_2 = middle\ risk & E_{\min} < \hat{E}_{\mathcal{D}_{leaf}} < E_{\max} \\ z_3 = high\ risk & E_{\max} \le \hat{E}_{\mathcal{D}_{leaf}} \end{cases} \tag{10}$$

where $E_{\min}$ and $E_{\max}$ denote the desired minimum and maximum energy to trigger the child node in the belief tree.

4.3 Actions design and control for self-protection

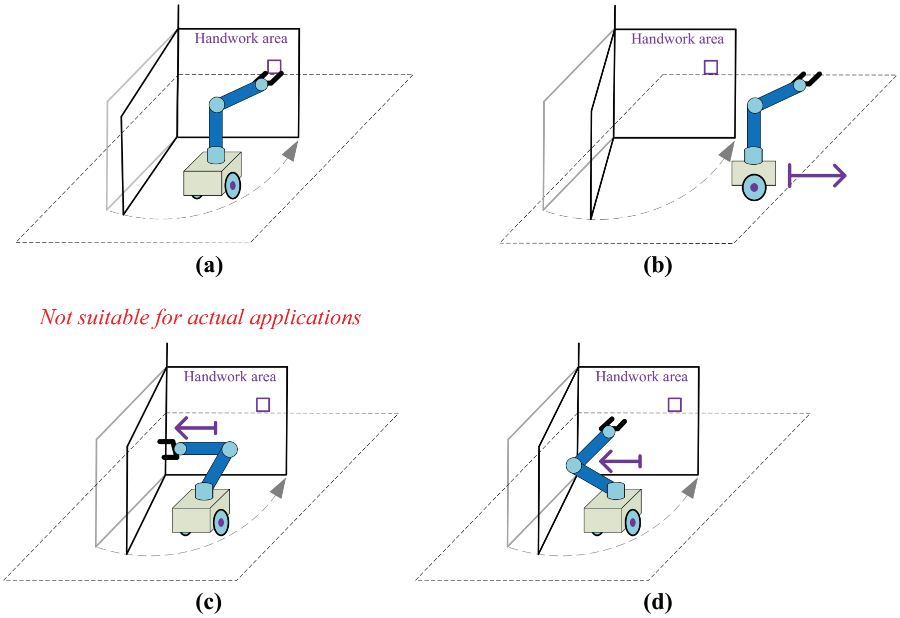
Assume that the chassis and arm part are mutually exclusive to implement the self-protective actions. Based on this, there are three typical classes of actions $a_i \in A$ in $W(t)$:

$$\mathcal{A} = \{a_1, a_2, a_3\} = \{\mathbf{stay}, \mathbf{escape}, \mathbf{defense}\} \tag{11}$$

where:**stay** denotes stay in the fan-shaped area and ignore the collision risk from door-closing, which is a conservative way to deal with danger, as shown in Figure 7(a);**escape** denotes escape out of the workspace before collision damages, as shown in Figure 7(b). It is time-consuming performance to return to re-start the work after escaping;**defense** denotes defense actively the collision risk using the dexterous elbow joint, as shown in Figure 7(*d*).

In contrast, Figure 7(c) shows a typical defense way looks more like human behavior in Figure 1. However, we do not want to use the way for actual applications, considering that the end-effector has a fragile structure to break and usually

Figure 7 Typical self-protective actions in dynamic interactive workspace



Notes: (a) Stay; (b) escape; (c) end-effector defense; (d) elbow defense

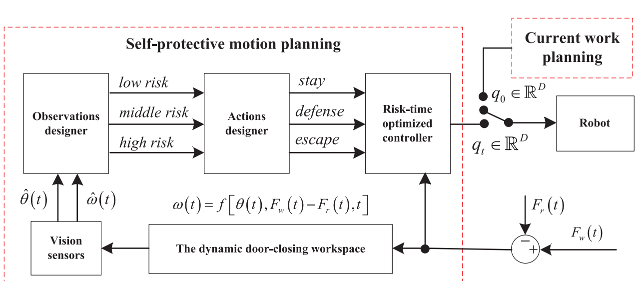
expensive. Note that **stay**, **escape** and **defense** are all possible actions to respond every risk level in Z. Based on equations (10) and (11), the rewards for taking $a_i \in A$ after $z_i \in Z$ are designed as Table 1.

We treat stay on guard in low risk, defense door-closing using the elbow in middle risk, and escape linear out in high risk as three self-protective actions in the door-closing workspace. Let t_{z_i} denote the collision time in the risk level z_i without considering the self-protective motion planning. Obviously, we can get $t_{z_3} < t_{z_2} < t_{z_1}$. Figure 8 shows the schematic of the proposed controller. With the switch changing to the proposed self-protective motion planning, q_0 comes from the last time of

Table 1 Actions at different risk levels and their rewards function

Rewards	Low risk	High risk	Middle risk
Stay	Good	Bad	Bad
Escape	Bad	Good	Ok
Defense	Ok	Ok	Good

Figure 8 Schematic of the planning for self-protective motions



current work planning. We design the corresponding q_d coming from three self-protective actions commands, their risk-time optimized controller $f_{a_i}(\cdot)$ can be written as:

$$\min_{0<t_{a_i}<t_{z_i}} \left\{ f_{a_i}(q, \dot{q}, t_{a_i}) : q_0 \to q_d \in \mathbb{R}^D \right\} \quad i = 1, 2, 3 \tag{12}$$

where t_{a_i} denotes the time to perform the a_i action.

5. Experiments and results

5.1 Experimental set-up

Figure 9 shows the outlook of the robot platform and the door-closing workspace; their initial conditions and geometric relationship are shown in Figure 10(a).

For the 6-DOFs arm, Figure 10(b) indicates the relationship between each joint's coordinate system by using red, green, and blue respectively denote the coordinate axis x_i, y_i, and z_i. The base coordinate system x_0, y_0, and z_0 is attached to the chassis. Base on this, as shown in Figure 10(c), the initial configuration q_0 of the 6-DOFs arm is:

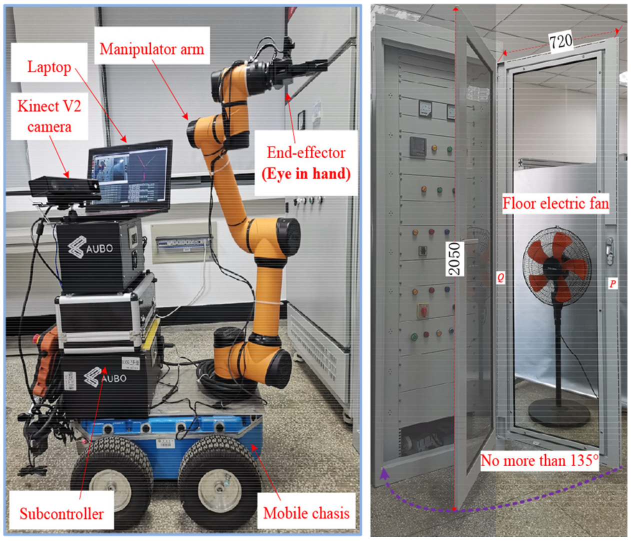
$$q_0 = [1.124, -0.947, -1.237, -0.1426, 0.4637, 0.6819] rad \tag{13}$$

A suitable defense configuration is the sufficient condition to defense success. The desired configuration q_d of the 6-DOFs arm is predefined as:

$$q_d = [-0.96, -0.9715, -1.212, -0.1234, 2.5467, 0.914] rad \tag{14}$$

Figure 11 shows the views of the Kinect camera and the eye in hand. The two-dimensional barcodes are detected and

Figure 9 Robot platform and internal dynamic workspace in a power cabinet box



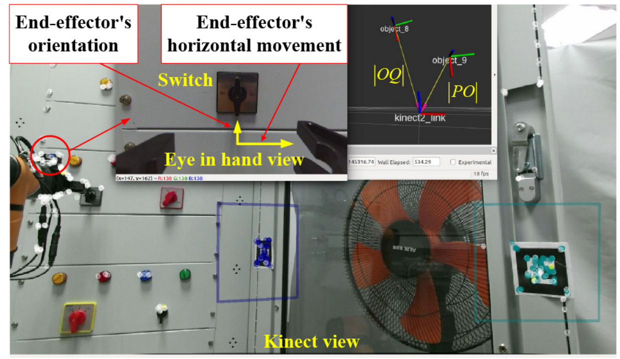
measured by point cloud, which are marked positions P and Q to get $|P_iO|$,$|P_{i+1}O|$, and $|OQ|$ (Figure 6). In the following experiments, we only use the Kinect camera as the observation.

5.2 Results and analysis

Based on the mentioned experimental set-up, three types of the self-protective actions were implemented on the robot platform against the sudden door-closing, as shown in Figure 12.

Figure 12(a) shows the stay on guard progress for self-protection with the responding results shown in Figure 13(a). $|OP|$,$|OQ|$, and $\angle P_iQO$ can hardly change due to the existent $F_r(t)$ and weak $F_w(t)$ from the fan's low power, which means D_{leaf} was motionless to some degree. In other words, D_{leaf} would not cause damages with low $\hat{\omega}$ to the robot at the moment but could rotate further; the robot stayed in the workspace at the cost of collision risk, preferring continued work. Holding on to the current pose by observing the $\hat{E}_{\mathcal{D}_{leaf}}$ to

Figure 11 Observation to the marked points on the dynamic door leaf



ensure no more than $E_{\min} = 0.2\mathcal{J}$, the robot performed vigilant self-protective awareness during the switch work interruption.

Figure 12(b) shows the linear escape progress for a fast door-closing emergency, and the results, compared with the situation the robot stays on guard in the workspace, as shown in Figure 13(b). In this case, without self-protection measures to stay in the workspace, the robot platform would get the damages brought by $\hat{E}_{\mathcal{D}_{leaf}}$ more than $14\mathcal{J}$. We use the local maximums in $\hat{E}_{\mathcal{D}_{leaf}}$ to judge the environmental risk level changes. The judgment is true when the risk level is changed for the first time from low to middle or high. In the second line of Figure 13(b), the local maximum (larger than $E_{\max} = 0.4\mathcal{J}$) at 1.42 s indicates the robot is at high risk, which triggers the chassis to escape out straight. The escape movement with chassis' max speed and the real progress to succeed escape is within 2 s; that is to say, the total time $< t_{z_3}$ reveals successful self-protection, which is no more than 3.42 s. Note that trigger time to start escape would exist variance among multiple experiments for imperfect physical observation, the linear escape movement does not succeed every time in high risk. The success rate calculated by multiple linear escape experiments is 86%.

Figure 12(c) shows the elbow defense progress for door-closing, and the responding results are shown in Figure 14. With the medium $F_w(t)$ from the fan's middle power, the local maximum ($E_{\min} < \hat{E}_{\mathcal{D}_{leaf}} < E_{\max}$) at 1.23 s indicates the robot is

Figure 10 Initial conditions of the robot platform and the power cabinet box

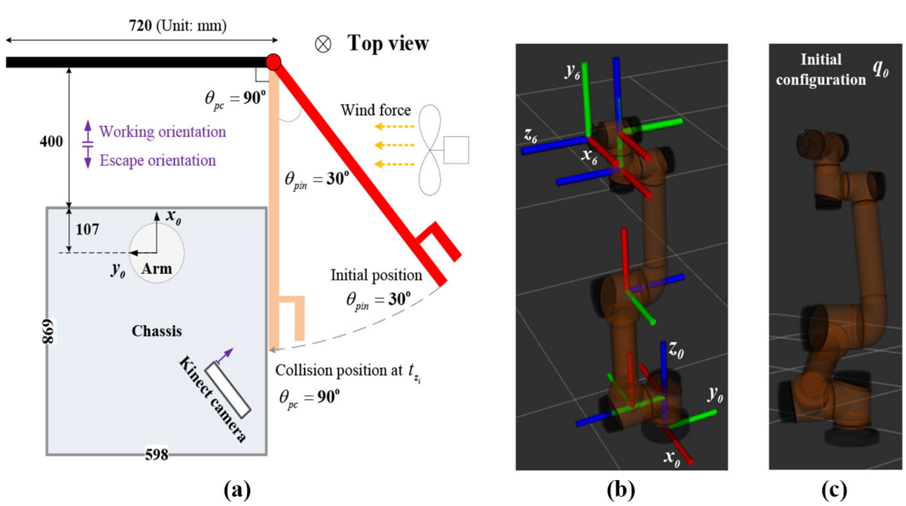


Figure 12 Execution of the self-protective actions against the sudden door-closing

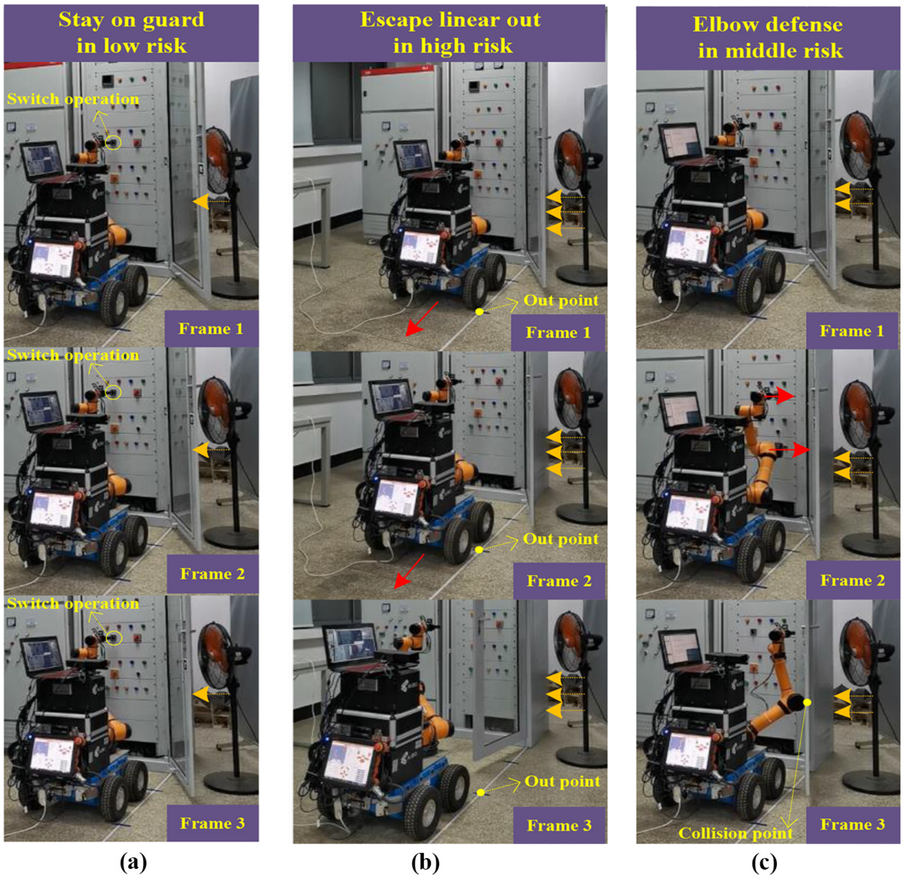
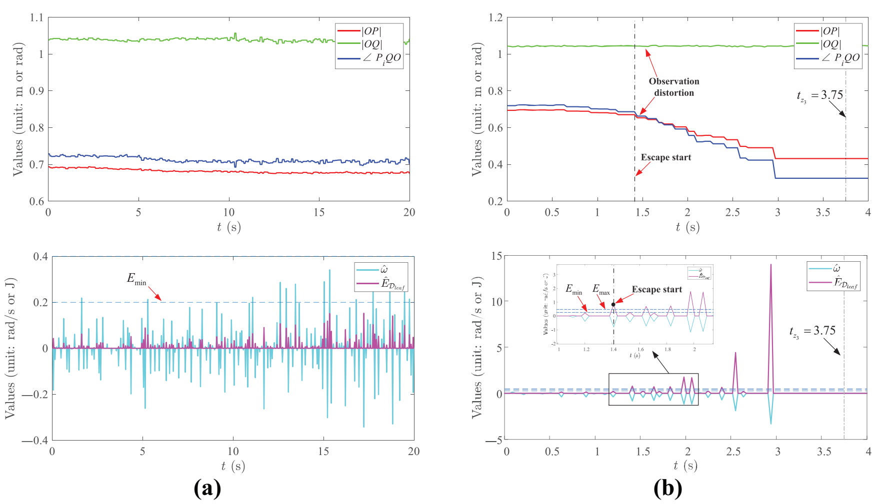
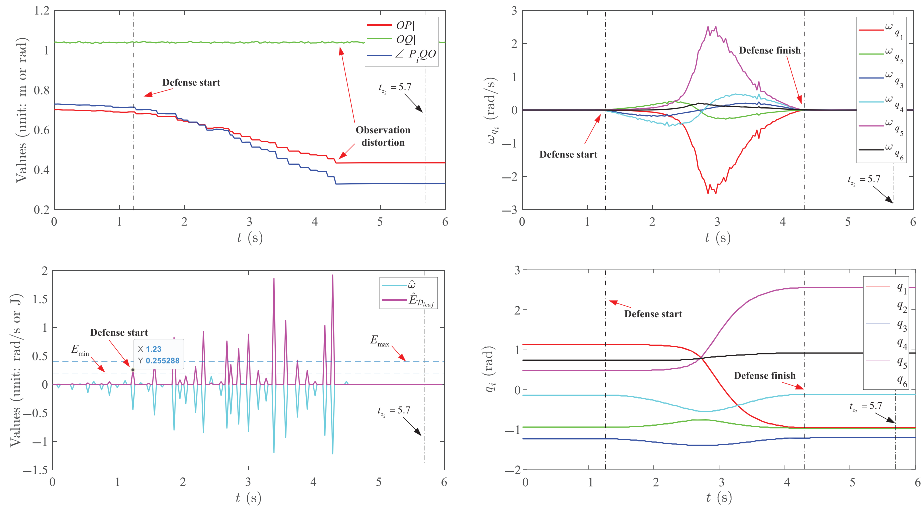


Figure 13 Interactions of the stay on guard and straight escape experiments



initially changing to face the middle risk from the low risk, which triggers the robotic arm to start defense by the elbow joint. In this case, the risk-time optimized control can be treated as a linear move with full speed controller. We let the end-effector's orientation remain face to the switch and keep end-effector horizontal movement during the elbow defense performing, hoping to continue current work quickly after the defense. Based on this, the arm joints, especially q_1 and q_5, are strongly related to the mentioned configuration's execution and have large changes; their angular velocities, ω_{q1} and ω_{q5}, to get the full speed in a short time with their physical constraints, are shown in the right column in Figure 14. The defense finish time

Figure 14 Interactions of the elbow defense experiment



is 4.3 s, less than t_{z2}, which means the robotic arm achieved to protect the chassis from a sudden collision. Note that $|OP|$ and $|OQ|$ get the observation distortion when the $\angle P_iQO = 0.33$ rad for hardly detecting them in the camera blind vision, but it does not affect to perform the elbow defense responding the middle (or high) risk for self-protection. Additionally, by interrupting the end-effector's work and utilizing the elbow joint to prevent the door-closing in an active collision way, the elbow defense can balance robotic safety and work efficiency. Compared with the stay on guard and linear escape action, the elbow defense's advantage is having a predictable performance to quick callback the interrupted work after the risk was relieved.

6. Conclusions

In this paper, a motion planning for the procedure of three self-protection behaviors in a dynamic constrained workspace by a mobile manipulator has been proposed. We first established a dynamic model of the restricted workspace released by an unlocked door. To treat as a whole system, the interactions between the dynamic workspace and robot were analyzed by using a POMDP. We proposed a self-protective policy decision executed as a belief tree planning. Responding to the policy decisions, we designed three types of actions: stay on guard in low risk, defense door-closing using the elbow in middle risk, and escape linear out in high risk for robot self-protection in the dynamic environment by observing collision risk levels to trigger them. Finally, we proposed a self-protective motion controller based on risk time optimization to act the planned actions. A mobile manipulator platform and a power cabinet inner dynamic constrained workspace were setup to verify the validity and efficiency of the proposed planning and control.

References

Abdo, N., Kretzschmar, H., Spinello, L. and Stachniss, C. (2013), "Learning manipulation actions from a few demonstrations", *2013 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1268-1275.

Adiwahono, A.H., Chua, Y., Tee, K.P. and Liu, B. (2013), "Automated door opening scheme for non-holonomic mobile manipulator", *13th International Conference on Control, Automation and Systems (ICCAS 2013)*, pp. 839-844.

Aoude, G.S., Luders, B.D., Joseph, J.M., Roy, N. and How, J. P. (2013), "Probabilistically safe motion planning to avoid dynamic obstacles with uncertain motion patterns", *Autonomous Robots*, Vol. 35 No. 1, pp. 51-76.

Bauer, C., Milighetti, G., Yan, W. and Mikut, R. (2010), "Human-like reflexes for robotic manipulation using leaky integrate-and-fire neurons", *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2572-2577.

Beer me, robot (2010), "PR2 robot fetches beer from the refrigerator", *Willow Garage Inc.*, available at: www.willowgarage.com/blog/2010/07/06/beer-me-robot

Chan, W.P., Mizohana, H., Chen, X., Shiigi, Y., Yamanoue, Y., Nagatsuka, M. and Inaba, M. (2019), "Multimodal sensing and active continuous closed-loop feedback for achieving reliable manipulation in the outdoor physical world", *Journal of Field Robotics*, Vol. 36 No. 1, pp. 17-33.

Chung, W., Rhee, C., Shim, Y., Lee, H. and Park, S. (2009), "Door-opening control of a service robot using the multifingered robot hand", *IEEE Transactions on Industrial Electronics*, Vol. 56 No. 10, pp. 3975-3984, doi: 10.1109/TIE.2009.2025296.

Kaelbling, L.P. and Lozano-Pérez, T. (2013), "Integrated task and motion planning in belief space", *The International Journal of Robotics Research*, Vol. 32 Nos 9/10, pp. 1194-1227.

Kalakrishnan, M., Righetti, L., Pastor, P. and Schaal, S. (2011), "Learning force control policies for compliant manipulation", *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4639-4644.

Kim, J. (2019), "Trajectory generation of a two-wheeled mobile robot in an uncertain environment", *IEEE Transactions on Industrial Electronics*, Vol. 67 No. 7, pp. 5586-5594.

Klingbeil, E., Saxena, A. and Ng, A.Y. (2010), *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2751-2757.

Luo, Y., Bai, H., Hsu, D. and Lee, W.S. (2019), "Importance sampling for online planning under uncertainty", *The International Journal of Robotics Research*, Vol. 38 Nos 2/3, pp. 162-181.

Milighetti, G., Hoffmann, E., Fetzner, A. and Kuntze, H.B. (2012), "Visually and force controlled opening and closing of doors by means of a mobile robot arm", *7th German Conference on Robotics (ROBOTIK 2012)*, pp. 1-6. 978-3-8007-3418-4.

Nemec, B., Žlajpah, L. and Ude, A. (2017), "Door opening by joining reinforcement learning and intelligent control", *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 222-228.

Osa, T. (2020), "Multimodal trajectory optimization for motion planning", *The International Journal of Robotics Research*, Vol. 39 No. 8, pp. 983-1001.

Prieto, S.A., Adán, A., Vázquez, A.S. and Quintana, B. (2019), "Passing through open/closed doors: a solution for 3d scanning robots", *Sensors*, Vol. 19 No. 21, p. 4740.

Rühr, T., Sturm, J., Pangercic, D., Beetz, M. and Cremers, D. (2012), "A generalized framework for opening doors and drawers in kitchen environments", *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3852-3858.

Shimizu, T., Saegusa, R., Ikemoto, S., Ishiguro, H. and Metta, G. (2012), "Self-protective whole body motion for humanoid robots based on synergy of global reaction and local reflex", *Neural Networks*, Vol. 32, pp. 109-118.

Su, H.-R. and Chen, K.-Y. (2019), "Design and implementation of a mobile robot with autonomous door opening ability", *International Journal of Fuzzy Systems*, Vol. 21 No. 1, pp. 333-342.

Thrun, S. (2002), "Probabilistic robotics", *Communications of the ACM*, Vol. 45 No. 3, pp. 52-57.

Valner, R., Vunder, V., Zelenak, A., Pryor, M., Aabloo, A. and Kruusamäe, K. (2018), "Intuitive 'human-on-the-loop' interface for tele-operating remote mobile manipulator robots", *International symposium on artificial intelligence, robotics, and automation in space (i-SAIRAS)*, pp. 1-8. available at: https://robotics.estec.esa.int/i-SAIRAS/isairas2018/Papers

Wang, C., Yin, L., Zhao, Q., Wang, W., Li, C. and Luo, B. (2020), "An intelligent robot for indoor substation inspection", *Industrial Robot: The International Journal of Robotics Research and Application*, Vol. 47 No. 5, pp. 705-712.

Wang, F., Liu, Y., Zhang, Y., Gao, Y., Xiao, L. and Wu, C. (2019), "Research on the shared control technology for robotic wheelchairs based on topological map", *Industrial Robot: The International Journal of Robotics Research and Application*, Vol. 47 No. 6, pp. 825-835.

Yahya, A., Li, A., Kalakrishnan, M., Chebotar, Y. and Levine, S. (2017), "Collective robot reinforcement learning with distributed asynchronous guided policy search", *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 79-86.

Corresponding author

Bingtuan Gao can be contacted at: gaobingtuan@seu.edu.cn