ORIGINAL ARTICLE



Towards a balancing safety against performance approach in human–robot co-manipulation for door-closing emergencies

Chuande Liu¹ · Chuang Yu² · Bingtuan Gao¹ · Syed Awais Ali Shah¹ · Adriana Tapus²

Received: 15 March 2021 / Accepted: 1 June 2021 $\ensuremath{\textcircled{O}}$ The Author(s) 2021

Abstract

Telemanipulation in power stations commonly require robots first to open doors and then gain access to a new workspace. However, the opened doors can easily close by disturbances, interrupt the operations, and potentially lead to collision damages. Although existing telemanipulation is a highly efficient master–slave work pattern due to human-in-the-loop control, it is not trivial for a user to specify the optimal measures to guarantee safety. This paper investigates the safety-critical motion planning and control problem to balance robotic safety against manipulation performance during work emergencies. Based on a dynamic workspace released by door-closing, the interactions between the workspace and robot are analyzed using a partially observable Markov decision process, thereby making the balance mechanism executed as belief tree planning. To act the planning, apart from telemanipulation actions, we clarify other three safety-guaranteed actions: on guard, defense and escape for self-protection by estimating collision risk levels to trigger them. Besides, our experiments show that the proposed method is capable of determining multiple solutions for balancing robotic safety and work efficiency during telemanipulation tasks.

Keywords Human–robot interaction \cdot Robot–environment interaction \cdot Dynamic workspace \cdot Self-protective behaviors \cdot Motion planning

Introduction

People always pursue high-efficiency performance in emergencies while ensuring personal safety [1]. For example, in fire rescue work, putting out fires, smoke-diving and the handling of patients and heavy tools are typical tasks, in which good balance ability can be critical for safety-guaranteed

⊠ Bingtuan Gao gaobingtuan@seu.edu.com

> Chuande Liu liuchuande@seu.edu.cn

Chuang Yu chuang.yu@ensta-paris.fr

Syed Awais Ali Shah awais@seu.edu.cn

Adriana Tapus adriana.tapus@ensta-paris.fr

¹ School of Electrical Engineering, Southeast University, Nanjing 210096, Jiangsu, China

² Autonomous Systems and Robotics Lab, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, France and task productivity [2]. Uncertain fire conditions and the excessive use of protective equipment further increase the challenges placed on the balance control system. Although existing telemanipulation is a highly efficient master–slave work pattern because of human-in-the-loop control, it is not trivial for the human operators to specify the optimal measures to guarantee robotic safety [3–5].

The ability to balance safety against performance in humans is coordinated and fast even when a work-related accident is not anticipated. Unfortunately, the robots are always not well trained such knowledge, to take appropriate measures to deal with the potential emergency during telemanipulation tasks. Beyond the basic capabilities of moving and acting autonomously, it is also essential to assure the robots' survival to protect themselves from harmful states or collisions when physically interacting with their workspace [6–8]. It is really true, especially in the search-and-rescue process [9–11], robotic dexterous performing task is easy to encounter uncertainties coming from environments in real human–robot collaborative manipulation. Moreover, these conditions to perform a task may need robots to open doors first and then gain access to a new workspace for their end-



effectors operations [12–15]. For example, utilize robots to manipulate the electronic equipments in power stations [16-20]. Typical power station operations are involved a large number of refrigerator-like electric cabinets equipped with electronic monitoring, which need to be checked at close range or switching operated by hands after opening the cabinet door. Similar with many robot-environment interactive processes, leverage robots to further operations in the cabinet's internal workspace released by door-opening; this procedure is susceptible to environmental uncertainty such as the wind force. The uncertain force could significantly drive the opened door with close trends and then lead to collision damages, threatening the robots' mechanical safety and work performance. In such cases, the balance ability is essential and vital to interrupt the current manipulation and display the corresponding self-protection at appropriate times.

In this work, we propose a novel balancing robotic safety against manipulation performance approach by planning safety-critical motions and control during work emergencies in the door-closing scenario. Specifically, a dynamic disturbance model of the restricted workspace released by door-opening is established. And then, the workspace and robot interactions are analyzed using a partially observable Markov decision process (POMDP), thereby making the balance mechanism executed as belief tree planning. To perform the planning, besides the telemanipulation actions, we clarify other three types of safety-guaranteed actions: on guard, escape, and defense for self-protection by estimating collision risk levels to trigger them. Finally, we propose three motion controllers based on risk-time optimization to act the planned self-protective actions.

The main contributions of this paper are summarized as follows:

- 1. To our knowledge, this paper yields the first evaluative framework to balance robotic safety against its operation performance during dynamic interactions in a door-closing workspace, within the collision risk consideration coming from environmental uncertainty;
- Apart from the manipulation actions, this paper clarifies other three safety-guaranteed actions: on guard, elbow defense the door, and escape out respectively to the collision risk with low, middle or high levels to act the balance policy, which is verified real true based on the experiments with our build-up robot platform;
- 3. Additionally, this paper is to provide guidance for the safe manipulation and deal with emergencies of a class of rescue robot operations and the upgrade of motion planning.

The rest of this paper is organized as follows. Related works are described in "Related work", "Workspace construction and problem formulation" explains the workspace



construction and problem formulation. A novel balancing safety against performance method is proposed in "Proposed method". "Experiments and results" validates the efficiency of the proposed method by experiments. Finally, conclusions are drawn in "Conclusions".

Related works

Related works about emergency measures, balance mechanism and workspace construction are introduced briefly in this section.

The studies on robotic emergency measures for selfprotection are both control and planning interesting. From the perspective of the reflex-based control, self-protective behaviors are categorized as the state-action association of behaviors, which traditionally depends on the subsumption architecture [21]. In this paradigm, the robot can quickly react to the stimulus since the sensory input from the dynamic environment directly triggers the coupled action from a wide variety of measures. Given this, many studies have focused on time-delay compensation [22,23] or reflex-based self-protective patterns and successfully applied to some humanoid robots, i.e., grasp reflex [24] and mainly slip [25]. For them, facing accidental collision risks, generating and maintaining stable controllers are their preventive measures. From the perspective of motion planning in some constrained environments [26], the self-protective response is to avoid one or more dynamic obstacles with uncertain motion patterns [27]. For them, it is necessary to plan smooth and collision-free orbits or trajectories to perform the desired task [28]. In the sense that, the self-protection to guarantee safety in the situation is based on around the dynamic obstacles to avoid collision paths.

Robotic balance mechanism is the prerequisite optimization policy-decision process for taking emergency measures, knowing when to pursue high-efficiency performance, or prefer a security guarantee. Traditionally, owing to the exclusive pursuit of the best performance value, this mechanism is not so flexible even redundant that it could be ignored in both subjective and objective aspects. Similarly, the exclusive pursuit in another extreme case is absolute security.

The aforementioned control and planning technologies are applied to a typical workspace released by door-opening [29,30], which has also received abundant attention during the last decades. When door-opening actions work in practice [31], the opened door matches external disturbances, such as uncontrolled rotational inertia to get closing trends, are unavoidably encountered. 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. In some cases, the researchers would suitably treat the unlocked door driven by external disturbances depending on the further task's difficulty. For easy tasks such as opening the door only to traverse it [32,33], it is no need to care too much about the unlocked door's state information due to a quick pass through after door-opening. However, complex tasks such as opening the door to get handwork inside [34,35], are generally more time-consuming and need more operating precision. We can not ignore the uncertain disturbances [36] coming from the unlocked door leading to a potential risk of collision damages. Compared with the above-mentioned simple task, cabinet handwork inside limits the end-effector's workspace and keeps the robot in the unlocked door's adverse influence range for a long time. To solve this problem, professional roboticists initially took a dual-arm mobile manipulator scheme [37,38]. More precisely, using one arm to defense the unlocked door's closing trends disturbances while planning another arm to work inside. They applied this theoretical pattern to an expensive PR2 (Personal Robot 2) to fetch a beer from a refrigerator [39]. 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 robot in human–robot collaborative manipulation to respond to emergencies. The unlocked door has closing disturbances during handwork after door-opening.

Workspace construction and problem formulation

Consider a time-varying workspace W(t) released from its door-opening action, is constrained by the door's frame $\mathcal{D}_{\text{frame}}$ and its leaf $\mathcal{D}_{\text{leaf}}$. In the top view to see W(t), Fig. 1 shows the dynamic interactive progress, which seems like a shrinking Chinese folding fan when $\mathcal{D}_{\text{leaf}}$ 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, $\mathcal{D}_{\text{leaf}}$ would stop close at a certain position p_n . After these, the state equation for $\mathcal{W}(t)$ can be written as:

$$\omega(t) = f\left[\theta(t), F_{W}(t) - F_{r}(t), t\right], \qquad (1)$$

where $\theta(t)$ denotes the angle between $\mathcal{D}_{\text{frame}}$ and D_{leaf} , $\omega(t)$ denotes the angular velocity and $f(\cdot)$ denotes a time-variation function.

In this paper, note that we do not concern about W(t) having the enlarged dynamic space situation. Thus, standing in the fan-shaped area, the robot is always facing the potential collision risk. We assume that the robotic chassis is necessarily treated as a collision-free part due to equipped with some indispensable precision sensors. After these, the goal is to plan a policy π in W(t) to get the maximum value function $V(\pi)$ and then control to execute π 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:

$$\max\left\{ \left| V\left(\pi\right) \right| \pi: \boldsymbol{q}_{0} \to \boldsymbol{q}_{d} \in \mathbb{R}^{D} \right\} \text{ s.t. } \mathcal{W}\left(t\right),$$
(2)

where q is the robotic degree of freedom (DOF) and D is the number of the DOF.

Proposed method

In this section, Fig. 2 shows the proposed balancing safety against performance method, which is mainly completed by three aspects, i.e., balance mechanism, interaction estimators and responding measures, that will be presented in detail.

Balance mechanism

The balance mechanism is a collaborative control based on the risk estimators, choosing the manual or automated policy decisions to deal with the workspace.

The upper part of Fig. 2 shows the human–robot interaction for master–slave manipulation tasks after door-opening. The control system to generate action sequences involves an autonomous controller's network interaction with the human operators. Under the received manual policy and action commands, assuming no significant delays or communication issues occur between the master and the slave, the robot platform could perform dexterous manipulation in efficiencycritical applications such as turn on a power switch for the human in the control loop.

The lower part of Fig. 2 shows the robot–environment interaction for door-closing emergencies. In the policy decisions block, a partially observable Markov decision process (POMDP) [26] architecture simulates the interaction relationship between agents decisions and their environment, which models our robot acting in the partially observable stochastic compressed workspace. It is defined formally as a 7-tuple (S, A, Z, T, O, R, b_0), where:

S: indicates a state set of $\mathcal{D}_{\text{leaf}}$ at the current time;

 \mathcal{A} : indicates an action set that the robot will perform at the next moment;

 \mathcal{Z} : indicates an observation set of $\mathcal{D}_{\text{leaf}}$ at the current time; T: the function T(s, a, s') = p(s'|s, a) indicates the probabilistic state transition from $s \in \mathcal{S}$ to $s' \in \mathcal{S}$, when the robot in state $s \in \mathcal{S}$ takes an action $a \in \mathcal{A}$. It can model our imperfect states set of $\mathcal{D}_{\text{leaf}}$ changes and robot control;

O: the function O(s, a, z) = p(z|s, a) indicates a set of conditional observation probabilities currently observed, which can capture sensors noise;

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





Fig. 1 The dynamic model of the restricted workspace



Fig. 2 Pipeline in terms of the balancing safety against performance approach

As analyzed previously, the POMDP planning aims to choose a policy π that maximizes its value based on \mathcal{A} and \mathcal{S} , but \mathcal{S} is not known exactly due to imperfect observation. Instead, the robot maintains a belief, which is a probability distribution over \mathcal{S} . The robot starts with an initial belief b_0 . At time *t*, it infers a new belief, according to Bayes' rule [40], by incorporating information from the action a_t taken and the observation z_t received:

$$b_t(s') = \tau (b_{t-1}, a_t, z_t) = \eta O(s', a_t, z_t) \sum_{s \in S} T(s, a_t, s') b_{t-1}(s),$$
(3)

where η is a normalizing constant.



Figure 3 shows that a POMDP policy prescribes the action at a belief. With the policy π and an initial belief b_0 , the expected value function V_{π} can be written as:

$$V_{\pi}(b_0) = E\left(\sum_{t=0}^{\infty} \gamma^t R(s_t, a_{t+1}) | b_0, \pi\right),$$
(4)

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 [41], as shown in Fig. 3. Each node of a belief tree corresponds to a belief *b*. At each node, the tree branches on all actions in \mathcal{A} and all observations in \mathcal{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. To find an optimal plan for a POMDP, using Bellman's equation relationship [42], we traverse the belief tree from the bottom up and compute an optimal action recursively at each node:

$$V^{*}(b) \stackrel{\Delta}{=} \max_{\pi} V_{\pi}(b) = \max_{a \in A} \left\{ \sum_{s \in S} b(s) R(s, a) + \gamma \sum_{z \in Z} p(z \mid b, a) V^{*}(b') \right\}, \quad (5)$$

where we notice that every value function V_{π} that satisfies Eq. (5) is both necessary and sufficient for the induced policy to be optimal.

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 a reason is the following child node of the belief tree branches on all possible actions in A.

Interaction estimators

In what follows, the observations and risk estimators block shown in Fig. 2 switch the control priority to trigger the mentioned manual or automated modes in detail.

Figure 4 shows the observation progress for the dynamic $\mathcal{D}_{\text{leaf}}$. Let *O* denote the robot's sensing position, P_i , P_{i+1} and *Q* denote three marked feature points on $\mathcal{D}_{\text{frame}}$ 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. Likewise, $|P_i Q|$ is parallel to d and $|P_i Q| = |P_{i+1}Q| = d$ 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:

$$\Delta \hat{\theta} = \angle P_i Q P_{i+1} = \angle P_i Q O - \angle P_{i+1} Q O, \tag{6}$$

where

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

For $\mathcal{D}_{\text{leaf}}$, the moment of inertia around the door axis is:

$$I = \frac{1}{3}md^2,\tag{7}$$

where *m* denotes the D_{leaf} mass. Based on Eqs. (6) and (7), the observed angular kinetic energy $\hat{E}_{D_{\text{leaf}}}$ around the door axis can be written as:

$$\hat{E}_{\mathcal{D}_{\text{leaf}}} = \frac{1}{2} I \hat{\omega}^2, \tag{8}$$

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 risk estimators block to switch and trigger the above balance mechanism. Combing with Eq. (8), we treat the risk levels coming from \mathcal{D}_{leaf} as inputs, train and divide them into four pre-defined parts (e.g., no risk, low risk, middle risk and high risk), which can be written as:

$$\mathcal{Z} = \begin{cases} z_0 = \text{no risk} & \widehat{E}_{\mathcal{D}_{\text{leaf}}} = 0\\ z_1 = \text{low risk} & 0 < \widehat{E}_{\mathcal{D}_{\text{leaf}}} \le E_{\min}\\ z_2 = \text{middle risk} & E_{\min} < \widehat{E}_{\mathcal{D}_{\text{leaf}}} < E_{\max}\\ z_3 = \text{high risk} & E_{\max} \le \widehat{E}_{\mathcal{D}_{\text{leaf}}} \end{cases}$$
(9)

where E_{\min} and E_{\max} denote the minimum energy and maximum energy to trigger the child node in belief tree (see Fig. 3). In addition, due to the resisting force $F_{r}(t)$ coming from rotation friction and air resistance, $\hat{E}_{D_{\text{leaf}}}$ could gradually decrease to zero in the door-closing progress.

Responding measures

Last, four types of responding measures shown in Fig. 5, i.e., telemanipulation actions and other three types of emergency actions for self-protection, are presented.

For the robot platform, to deal with emergencies in the limited workspace, the simultaneous multi-action between





Fig. 3 POMDP planning performs a lookahead search on the belief tree





chassis and arm part leads to complicated movements, even mission failure. To simplify the problem, we assume that action implementation related to the chassis and arm part is mutually exclusive. Based on this, there are four typical classes of actions $a \in A$ in the dynamic workspace:

$$\mathcal{A} = \begin{cases} a_0 = \{\text{telemanipulation}\}\\ a_1 = \{\text{ on guard}\}\\ a_2 = \{\text{defense}\}\\ a_3 = \{\text{escape}\} \end{cases}, \tag{10}$$

where

telemanipulation denotes a task-related action subset, which is well-behaved human-in-the-loop operations to deal with work, as shown in Fig. 5a;

on guard denotes to stop current actions and estimate the collision risk, ready to take the next action according to circumstances, as shown in Fig. 5b;

defense denotes to defense actively the risk of the collision damages using the dexterous arm part. Figure 5d shows the defense part might be the end-effector. Consider that the end**Fig. 5** Typical responding actions in interactive workspace



effector has a fragile structure to break and usually expensive, which is not suitable for actual applications. In contrast, using the elbow joint to defense plays a dominant role as active self-protection shown in Fig. 5e;

escape denotes to escape out of the workspace before collision damages, as shown in Fig. 5c.

The rewards for taking a_i after z_i are pre-trained as the following Table 1. Let good = +1, ok = 0, and bad = -1. We treat $\pi(a_i, z_i|_{i=0,1,2,3})$ as balance policy between

safety and efficiency performance in the dynamic workspace. Among them, $\pi(a_0, z_0)$ and $\pi(a_3, z_3)$ are traditional research area to improve performance or stress reaction, which are the subset of our proposed balance method.

Note that the higher risk level, the less time can be used to act $\pi(a_i, z_i|_{i=1,2,3})$, which requires control based on risk time optimization. Let t_{z_i} denote the collision time in the risk level z_i without considering safety-critical measures. Obviously, we can get $t_{z_3} < t_{z_2} < t_{z_1}$, and the action controller



Table 1Actions at different risklevels and their rewards function

$\overline{R(a_i, z_j _{i,j=0,1,2,3})}$	z_0 no risk	z_1 low risk	z_2 middle risk	z3 high risk
a_0 telemanipulation	Good	Ok	Bad	Bad
<i>a</i> ₁ on guard	Bad	Good	Bad	Bad
<i>a</i> ₂ defense	Bad	Bad	Good	Bad
a ₃ escape	Bad	Bad	Ok	Good



Fig. 6 Schematic of the balance policy and control method

 $f_{a_i}(\cdot)$ can be written as:

$$\min_{0 < t_{a_i} < t_{z_i}} \left\{ f_{a_i}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t_{a_i}) : \boldsymbol{q_{a_0}} \to \boldsymbol{q_{a_i}} \in \mathbb{R}^D \right\} \ i = 1, 2, 3,$$
(11)

where t_{a_i} denotes the time to perform the telemanipulation configuration $q_{a_0} \in a_0$ to the desired configuration $q_{a_i} \in a_i$. It is switching control progress to self-protection during the telemanipulations. Based on these, the schematic of the proposed balance policy and control method is shown in Fig. 6.

Experiments and results

In this section, we will present our experimental conditions first and then set up four types of experiments to verify the proposed balance method's efficiency.



Experimental set-up

Figure 7 shows the outlook of the human-in-the-loop robot platform and the dynamic workspace, which is constructed by a standardized power cabinet. The robot platform is mainly composed of a chassis, a 6-DOFs arm, an end-effector and a Kinect, which faces the opened door and runs at 30 frames per second on the chassis. In the dynamic workspace, the specific telemanipulation task is to turn on a switch for electricity supply. Table 2 shows more detailed information and other components.

Figure 8 shows the robot platform's geometric relationship during switch work. In the top view, the chassis is partly standing in the fan-shaped area, whose escape is opposite to the end-effector's working orientation. Without considering safety-guaranteed measures, the collision would be at some position on the chassis after t_{z_i} , which could be acquired by using the fan with no, low, middle or full power for door-closing. After that, the 6-DOFs arm indicates the rela**Fig. 7** Human-in-the-loop robot platform and door-closing workspace in a power cabinet box



Table 2Robot platform andpower cabinet box equipmentlist

No.	Туре	Name	Description
1	Mobile chassis	None	Size: 0.698 m × 0.598 m × 0.284 m
2	Subcontroller	None	ROS/Moveit!, ubuntu 16.04, ROS kinectic
3	Power source	AC/DC	220VAC/48VDC
4	Depth camera	Kinect	Version 2, attached on chassis
5	Computation	On-board PC	i7-4720HQ, 8 GB RAM, 245 GB SSD
6	Robotic arm	AUBO i5	6-DOFs, max length: 1.0225 m
7	Network receiver	None	Wireless network receiver
8	Electric fan	AUX	A switch to low/middle/full power (130 W)
9	Fan-shaped area	None	Robot is in collision risk standing in this area
10	Power cabinet door	None	m: 15.2 kg, size: 2.05 m \times 0.72 m \times 0.05 m
11	End-effector	None	Two-finger gripper
12	RGB camera	UVC camera	Eye in hand with camera calibration
13	Switch	Power switch	The current handwork in the cabinet box
14	Joystick	None	Telemanipulation controller

tionship between each joint's coordinate system by utilizing 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. We acquire the arm's initial configuration $\boldsymbol{q}_{a_0^2}$ as

$$\boldsymbol{q}_{a_0^2} = [1.124, -0.947, -1.237, -0.1426, 0.4637, 0.6819]$$

rad.

which comes from the telemanipulation current configuration to reach the switch.

Figure 9 show the Kinect camera's view and the eye view on hand. Based on the two views, a well-trained human operator could drive our robot platform to activate the policy $\pi(a_0, z_0)$. In this case, a_0 denote action $(a_0^i \in a_0|_{i=1,\dots,7})$ step as: move-in workspace; reach, clamp, rotary and loosen the switch; take arm back, and move-out workspace (see Fig. 7). In Kinect camera view, the two-dimensional barcodes are detected and measured by point cloud, which are marked positions *P* and *Q* to get the distances $|P_i O|$, $|P_{i+1} O|$, and |OQ| (see Fig. 4). In the following experiments, we only use the Kinect camera as the risk estimator.









Results and analysis

Based on the mentioned experimental conditions, $\pi(a_i, z_i | i=0,1,2,3)$ were implemented on the robot platform against the door-closing, as shown in Fig. 10.

Figure 10a shows a balance policy sequences to humanrobot collaborative experiments with $\pi(a_0^2|\mathbf{q}_{a_0^2}, z_0)$, $\pi(\text{on-guard} \in \mathbf{a}_1, z_1)$ and $\pi(\text{elbow-defense} \in \mathbf{a}_2, z_2)$. The responding results are shown in Fig. 11. In the left column of Fig. 11, |OQ| is a constant because the chassis is stationary in the workspace; |OP| and $\angle P_i QO$ are gradually decreasing with wind force $F_w(t)$ after time t_2 , and interrupt the change when the defense collision happens. We use the local maximums in $\hat{E}_{\mathcal{D}_{\text{leaf}}}$ to judge z_i changes. The judgment is true when z_i is changed for the first time from z_0 to another higher level. Let $E_{\min} = 0.2J$ and $E_{\max} = 0.4J$,



we get the time t_2 ($E_{\min} < \hat{E}_{D_{\text{leaf}}} < E_{\max}$) to trigger π (elbow-defense $\in a_2, z_2$). During the time ($t_2 - t_1$), robot platform hold on the current configuration $q_{a_0^2}$ to do next action a_0^3 , with estimating the $\hat{E}_{D_{\text{leaf}}}$ to ensure no more than $E_{\min} = 0.2J$. In other words, the robot performed vigilant self-protective awareness compared with the artificial stop or pause in telemanipulation.

Risk getting higher after t_2 , the robot platform would get the damages more than 1.5 J, which is avoided by elbowdefense. Responding to z_i , the switch operations that require precise and small-scale motion are assigned to the endeffector, while the large-scale action for self-protection is fast and carried out by the chassis or arm. We let the end-effector's orientation remain to face the switch and keep end-effector horizontal movement (see Fig. 9) to pre-trained defense con-





figuration, hoping to continue current work quickly after the π (elbow-defense, z_2). Based on this, the control based on risk-time optimization is treated as a linear move control in the end-effector's workspace with full speed. In the right column of Fig. 11, all the arm joints are related to the configuration's execution and have significant changes. Their angular velocities, ω_{q1} and ω_{q5} , to get the full speed in a short time with their physical constraints. The acquired finial defense configuration q_{a_2} is

 $\boldsymbol{q}_{a_2} = [-0.96, -0.9715, -1.212, -0.1234, 2.5467, 0.914]$ rad

Figure 10b shows other balance policy sequences to human–robot collaborative experiments with $\pi(a_0^2 | \boldsymbol{q}_{a_0^2}, z_0)$,

 π (on-guard $\in a_1, z_1$) and π (line-escape $\in a_3, z_3$). The responding results are shown in Fig. 12. In Fig. 12, bring into correspondence with three policies in Fig. 11, t_1 and t_2 are also the states (z_0, z_1, z_3) change time and the responding action's start time. |OP| and $\angle P_i QO$ both get the horizontal curves after t_2 . What the difference is, the horizontal curves in the elbow-defense case, the defense collision stopped the door close. But in the line-escape case, the reason is that observation is in the camera's blind vision when escape out of the workspace. The local maximum at t_2 indicates the robot is at high risk, which triggers the chassis to escape out straight with chassis' max speed. Additionally, Compared with the line-escape action, the on-guard and elbow defense's advan-





Fig. 12 The collaborative experiments of the telemanipulation, on-guard and line-escape

tage is having a predictable performance to quick callback the interrupted work after the risk was relieved, without the time-consuming cost of re-planning or re-doing move-in the workspace.

Conclusions

In this paper, a balancing safety against performance approach for door-closing emergencies in human-robot collaborative manipulation has been proposed. Specifically, We first established a dynamic disturbance model of the restricted workspace released by door-opening. And then, the workspace and robot interactions are analyzed using a partially observable Markov decision process (POMDP), thereby making the balance mechanism executed as belief tree planning. Responding to the policy, besides the telemanipulation actions, we clarify other three safety-guaranteed actions: on guard, escape, and defense for self-protection by estimating collision risk levels to trigger them. Finally, we propose a motion controller based on risk time optimization to act the planned self-protective actions. Our build-up robot platform and a power cabinet inner dynamic constrained workspace were setup to verify the validity and efficiency of the pro-



posed planning and control. This paper is to provide guidance for the safe manipulation and deal with emergencies of a class of robot operations and the upgrade of motion planning.

Funding Funding was provided by China Scholarship Council (Grant no. 201906090196).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecomm ons.org/licenses/by/4.0/.

References

 Geng N, Chen Z, Nguyen QA, Gong D (2021) Particle swarm optimization algorithm for the optimization of rescue task allocation with uncertain time constraints. Syst Complex Intell. https://doi. org/10.1007/s40747-020-00252-2

- Punakallio A, Lusa S, Luukkonen R (2005) Predictive values of balance tests for physical work ability in fire and rescue work. Int Congr Ser 1280:301–303
- 3. Rahal R, Matarese G, Gabiccini M et al (2020) Caring about the human operator: haptic shared control for enhanced user comfort in robotic telemanipulation. IEEE Trans Haptics 13(1):197–203
- Aboutalebian B, AliTalebi H, Etedali S, Suratgar A (2020) Adaptive control of teleoperation system based on nonlinear disturbance observer. Eur J Control 53:109–116
- Livatino S, Guastella D, Muscato G et al (2021) Intuitive robot teleoperation through multi-sensor informed mixed reality visual aids. IEEE Access 9:25795–25808
- Lee S, Chwa D (2020) Dynamic image-based visual servoing of monocular camera mounted omnidirectional mobile robots considering actuators and target motion via fuzzy integral sliding mode control. IEEE Trans Fuzzy Syst. https://doi.org/10.1109/TFUZZ. 2020.2985931
- Yahya A, Li A, Kalakrishnan M, Chebotar Y, Levine S (2017) Collective robot reinforcement learning with distributed asynchronous guided policy search. IROS 2017:79–86
- Shimizu T, Saegusa R, Ikemoto S, Ishiguro H, Metta G (2012) Self-protective whole body motion for humanoid robots based on synergy of global reaction and local reflex. Neural Netw 32:109– 118
- Geng N, Meng Q, Gong D, Chung PWH (2019) How good are distributed allocation algorithms for solving urban search and rescue problems? A comparative study with centralized algorithms. IEEE Trans Autom Sci Eng 16(1):478–485
- Balta H, Bedkowski J, Govindaraj S, Majek K, Musialik P, Serrano D (2016) Integrated data management for a fleet of search-andrescue robots. J Field Robot 34(3):539–582
- Lv H, Yang G, Zhou H et al (2020) Teleoperation of collaborative robot for remote dementia care in home environments. IEEE J Transl Eng Health Med 8:2168–2372
- Nemec B, Zlajpah L, Ude A (2017) Door opening by joining reinforcement learning and intelligent control. ICRA 2017:222–228
- Quintana B, Prieto S, Adan A, Bosche F (2018) Door detection in 3D coloured point clouds of indoor environments. Autom Constr 85:146–166
- Su H, Chen K (2019) Design and implementation of a mobile robot with autonomous door opening ability. IEEE Trans Fuzzy Syst 21(1):333–342
- Liu C, Gao B, Yu C, Tapus A (2021) Self-protective motion planning for mobile manipulators in a dynamic door-closing workspace. Ind Robot. https://doi.org/10.1108/IR-02-2021-0025
- Wang C, Yin L, Zhao Q, Wang W, Li C, Luo B (2020) An intelligent robot for indoor substation inspection. Ind Robot 47(5):705–712
- Zhang H, Su B, Meng H (2017) Development and implementation of a robotic inspection system for power substations. Ind Robot 44(3):333–342
- Zhao X, Liu Z, Liu Y et al (2017) Structure design and application of combination track intelligent inspection robot used in substation indoor. Procedia Comput Sci 107:190–195
- Lu S, Li Y, Zhang T (2009) Design and implement of control system for power substation equipment inspection robot. IROS 2009:93– 96
- Liu C, He J, Gao B (2019) Movement planning and control of an overhead power transmission line inspection bionic-robot. IEEE Cyber 2019:25–29
- 21. Brooks RA (1991) Intelligence without representation. Artif Intell 47(1-3):139–159
- Zakerimanesh A, Sharifi M, Hashemzadeh F, Tavakoli M (2021) Delay-robust nonlinear control of bounded-input telerobotic systems with synchronization enhancement. IEEE Robot Autom Lett 6(2):2493–2500

- Natori K, Tsuji T, Ohnishi K, Hace A, Jezernik K (2010) Timedelay compensation by communication disturbance observer for bilateral teleoperation under time-varying delay. IEEE Trans Ind Electron 57(3):1050–1062
- Bauer C, Milighetti G, Yan W, Mikut R (2010) Human-like reflexes for robotic manipulation using leaky integrate-and-fire neurons. IROS 2010:2572–2577
- Renner R, Behnke S (2006) Instability detection and fall avoidance for a humanoid using attitude sensors and reflexes. IROS 2006:2967–2973
- Osa T (2020) Multimodal trajectory optimization for motion planning. Int J Robot Res 39(8):983–1001
- Aoude GS, Luders BD, Joseph JM, Roy N, How JP (2013) Probabilistically safe motion planning to avoid dynamic obstacles with uncertain motion patterns. Auton Robot 35(1):51–76
- Luo Y, Bai H, Hsu D, Lee WS (2019) Importance sampling for online planning under uncertainty. Int J Robot Res 38(2–3):162– 181
- Kalakrishnan M, Righetti L, Pastor P, Schaal S (2011) Learning force control policies for compliant manipulation. IROS 2011:4639–4644
- Abdo N, Kretzschmar H, Spinello L, Stachniss C (2013) Learning manipulation actions from a few demonstrations. ICRA 2013:1268–1275
- 31. Chan W, Mizohana H, Chen X, Shiigi Y, Yamanoue Y, Nagatsuka M, Inaba M (2019) Multimodal sensing and active continuous closed-loop feedback for achieving reliable manipulation in the outdoor physical world. J Field Robot 36(1):17–33
- 32. Chung W, Rhee C, Shim Y, Lee H, Park S (2009) Door-opening control of a service robot using the multifingered robot hand. IEEE Trans Ind Electron 56(10):3975–3984
- Prieto SA, Adán A, Vázquez AS, Quintana B (2019) Passing through open/closed doors: a solution for 3d scanning robots. Sensors 19(21):4740–4753
- Rühr T, Sturm J, Pangercic D, Beetz M, Cremers D (2012) A generalized framework for opening doors and drawers in kitchen environments. ICRA 2012:3852–3858
- 35. Englert P (2018) Learning manipulation skills from a single demonstration. Int J Robot Res 37(1):137–154
- Kim J (2019) Trajectory generation of a two-wheeled mobile robot in an uncertain environment. IEEE Trans Ind Electron 67(7):5586– 5594
- Valner R, Vunder V, Zelenak A, Pryor M, Aabloo A, Kruusamäe K (2018) Intuitive 'human-on-the-loop' interface for tele-operating remote mobile manipulator robots. In: i-SAIRAS 2018 pp. 1–8. https://robotics.estec.esa.int/i-SAIRAS/isairas2018/Papers/Session%205b/1_valner_isairas2018_final_inline-50-32-Kruusam%C3%A4e-Karl.pdf
- Wang X, Chen L (2020) A vision-based coordinated motion scheme for dual-arm robots. J Intell Robot Syst 97:67–79
- Willow garage inc (2010) Beer me, robot. http://www. willowgarage.com/blog/2010/07/06. Accessed 13 July 2012
- Wang F, Liu Y, Zhang Y, Gao Y, Xiao L, Wu C (2019) Research on the shared control technology for robotic wheelchairs based on topological map. Ind Robot 47(6):825–835
- Kaelbling LP, Lozano-Pérez T (2013) Integrated task and motion planning in belief space. Int J Robot Res 32(9–10):1194–1227
- 42. Thrun S (2002) Probabilistic robotics. Commun ACM 45(3):52-57

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

