Computational Model of Robot Trust in Human Co-worker for Physical Human-Robot Collaboration

Qiao Wang[1], Dikai Liu[1], Marc G. Carmichael[1], Stefano Aldini[1], and Chin-Teng Lin[2]

***Abstract*—Trust is key to achieving successful Human-Robot Interaction (HRI). Besides trust of the human co-worker in the robot, trust of the robot in its human co-worker should also be considered. A computational model of a robot's trust in its human co-worker for physical human-robot collaboration (pHRC) is proposed. The trust model is a function of the human co-worker's performance which can be characterized by factors including safety, robot singularity, smoothness, physical performance and cognitive performance. Experiments with a collaborative robot are conducted to verify the developed trust model.**

***Index Terms*—Physical Human-Robot Interaction, human factors and human-in-the-loop, trust.**

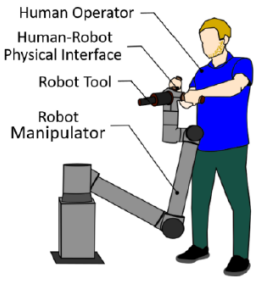

Fig. 1. A pHRC scenario with a human co-worker physically controlling a robot manipulator [4]

I. INTRODUCTION

WITH the advancement of collaborative robots (cobots) [1], these types of robots can now actively interact with human co-workers in the same workspace. The collaboration between humans and robots combines the strengths of both parties, the perception and decision-making of human co-workers and the power (e.g., force and torque) and endurance of robots.

There are various ways in which humans could actively interact with cobots. In this paper, we focus on physical Human-Robot Collaboration (pHRC). pHRC defines a human co-worker physically contacts or exchanges force continuously with a robot to accomplish a shared goal in the same workspace. There are various applications for pHRC, such as rehabilitation and material handling [2], [3]. A pHRC scenario is shown in Figure 1.

Trust is key to achieving successful Human-Robot Interaction (HRI) [5]. The human may disuse (under-reliance), or misuse (over-reliance) the robot without an appropriate level of trust [5]. Research on computational trust models in HRI has been conducted by several research groups. Lee and Moray [6] developed an auto-regressive moving average trust model based on operator's subjective rating scales of trust, the automatic controller's performance and fault. Xu [7] proposed a human-to-robot online probabilistic trust inference model which employs a dynamic bayesian network to estimate the trust based on the subjective human feedback. Saeidi proposed a trust-based control of a semi-autonomous mobile robot [8] and Sadrfaridpour proposed a trust-based control of a manipulator [9].

In human-human interaction (HHI), trust is bidirectional. In order to emulate HHI for HRI, models of robot trust in humans were also studied by several researchers. Rahman [10] proposed a robot-to-human trust model for handover tasks and Tran [11] proposed a robot-to-human confidence model based on the Fluid Stochastic Petri Net model in pHRC.

For the existing human-to-robot trust models [6]-[9], subjective questionnaires about trust are used to identify models' parameters. However, this cannot be realized in robot-to-human trust model development. Performance measurements in these robot-to-human trust models [10]-[11] are restricted to a particular task. The proposed approach in this research directly uses measurements from sensors during pHRC, which is objective and without human bias. In addition, the robot-to-human trust model takes into account many pHRC factors, including safety, robot singularity, smoothness, physical performance and cognitive performance, which evaluates human co-worker performance more comprehensively. Therefore, the main contribution of this paper is the development of a relatively accurate, objective and comprehensive computational model of robot trust in a human co-worker for physical human-robot collaboration. The organization of the paper is as follows. The computational model of robot trust in a human co-worker is described in Section II. The human co-worker performance modelling is presented in Section III. An experimental testbed and design of experiments are presented in Section IV. The results and discussion are shown in Section

Manuscript received: July, 10, 2021; Revised November 26, 2021; Accepted December 23, 2021.

This paper was recommended for publication by Editor Angelika Peer upon evaluation of the Associate Editor and Reviewers' comments. This work is supported in part by the Australian Research Council (ARC) Discovery Project Grant [DP210101093].

[1]Qiao Wang, Dikai Liu, Marc G. Carmichael, and Stefano Aldini are with the Robotics Institute, Faculty of Engineering and Information Technology, University of Technology Sydney, 81 Broadway, Ultimo NSW 2010, Australia {Qiao.Wang-1}@student.uts.edu.au;

[2] Chin-Teng Lin is with the Australian Artificial Intelligence Institute, School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney, 81 Broadway, Ultimo NSW 2010, Australia

Digital Object Identifier (DOI): see top of this page.

V and the conclusion and future work are in Section VI.

II. Computational Trust Model

The trust of humans in robots is dynamic and highly depends on the robot performance [12]. Therefore, We assumed that the trust could be estimated based on performance [12]. When developing a model of robot trust in humans, a similar dynamic model is used. We introduce a real-time computational robot-to-human trust model:

$$T[n] = \frac{\sum_{k=0}^{N} \beta^k p[n-k]}{\sum_{k=0}^{N} \beta^k} \tag{1}$$

$T \in [0,1]$ is the robot's trust in human co-worker. $T = 0$ or $T = 1$ represents no trust or complete trust respectively in the human co-worker. $p \in [0,1]$ is a normalised measure of human performance, which may be comprised of multiple performance measures. A discount factor $\beta \in [0,1]$ is introduced that reduces sensitivity to historic performance as more recent performance has greater impact on trust value [6]. β^k is the weighting of p at the time step $n-k$. When k is larger, β^k is smaller. N is the length of the moving time window and n is the current time step. β and N determine the sensitivity of T to past performance. The past performance has less effect on T if β and N are small.

III. Human Co-worker Performance Modelling

In order to evaluate the human co-worker performance p in the context of pHRC in Equation 1, factors that affect the human co-worker performance in pHRC need to be identified and quantified. In this work, the human co-worker performance is characterized by Safety Performance $p_S \in [0,1]$, Singularity Performance $p_{SP} \in [0,1]$, Smoothness Performance $p_{SM} \in [0,1]$, Physical Performance $p_{PW} \in [0,1]$ and Cognitive Performance $p_{CP} \in [0,1]$. These factors are important and common measurements of safe and intuitive pHRC. Details on justification of selecting these pHRC factors have been provided in each sub-section in Section III. Tran [11] proposed a method to incorporate different performance factors:

$$p = \prod_{A=1}^{N_c} p_A^c (C + (1-C) \times \sum_{A=1}^{N_{nc}} \gamma_A p_A^{nc}) \tag{2}$$

p_A^c and p_A^{nc} are the critical and non-critical performance factors, respectively. Critical factors p_A^c are strongly relevant with the continuation of a task, safety of the human co-worker and the robot. Non-critical factors p_A^{nc} are used for assessing the performance of the human co-worker that is not essential to the task and safety of the human co-worker and the robot. N_c and N_{nc} are the number of critical and non-critical performance factors. γ_A is the weighting coefficient and represents the relative importance of each p_A^{nc}, $\sum_{A=1}^{N_{nc}} \gamma_A = 1$. C represents the maximum contribution of the p_A^{nc}.

In this paper, the Safety Performance p_S and the Singularity Performance p_{SP} are safety-related which are regarded as p_A^c. The other three performance factors are not safety-related which are regarded as p_A^{nc}. Hence, based on Equation 2, the human co-worker performance is modelled as:

$$p[n] = p_S p_{SP}(C + (1-C)(\gamma_{SM} p_{SM} + \gamma_{PW} p_{PW} + \gamma_{CP} p_{CP})) \tag{3}$$

The weighting coefficients $\gamma_{SM} + \gamma_{PW} + \gamma_{CP} = 1$ are positive constants that could be adjusted based on the relative importance of the corresponding p_A^{nc} according to specific task requirements.

A. Safety Performance

In order to achieve safe pHRC, collisions between the robot and surrounding objects need to be avoided to prevent potential damage to both the robot and surrounding objects. As a result, a safety performance p_S is defined based on the possibility of collision between the robot and surrounding objects whilst under control by the human co-worker. The possibility of collision increases from low ($p_S = 1$) to high ($p_S = 0$). An interpolation function is employed to acquire a smooth curve of p_S [13] which is shown as:

$$p_S[n] = f(a[n], a^-, a^+, 1, 0) \tag{4}$$

f is a fifth-order polynomial with null first and second derivatives at a^- and a^+ and f is bounded in the range between 1 and 0.

$$a[n] = \frac{v^2[n] - v_0^2[n]}{2\Delta s[n]} \tag{5}$$

$a[n]$ is the magnitude of the constant deceleration required to stop the robot when it reaches the position to collide with an object which is based on the kinematic Equation 5. a^- is the threshold deceleration at which p_S starts to reduce as shown in Figure 2(a) and a^+ is the maximum deceleration allowed. $v[n]$ is the velocity of the robot when reaching the position to collide with an object which is assumed to be zero ($v[n] = 0m/s$). $\Delta s[n]$ is the distance between the robot and the object. $v_0[n]$ is the current robot velocity toward the object. Employing a takes consideration of both the robot velocity toward an object v_0 and the distance between the robot and an object Δs. When Δs is small or v_0 is large, a is large which represents the possibility of collision is high.

Figure 2(a) shows the change of Safety Performance with deceleration a. p_S starts to reduce once $a > a^-$ because the possibility of collision increases. p_S reaches the minimum value ($p_S = 0$) once $a \geq a^+$. p_S needs to be evaluated for each surrounding object i by using algorithm 1. N_{obj} is the total number of surrounding objects.

B. Singularity Performance

In the pHRC scenario showing in Figure 1, when the manipulator is close to a singular configuration, the joint velocities are large, causing the robot behaviour to be unpredictable and dangerous. Hence, Singularity Performance $p_{SP} \in [0,1]$ is proposed to quantify the possibility of entering a singular configuration. When $p_{SP} = 1$ (or $p_{SP} = 0$), the possibility of entering a singular configuration is extremely low (or high).

In this work, p_{SP} is defined as a function of the smallest singular value σ_{min}, which can be calculated through

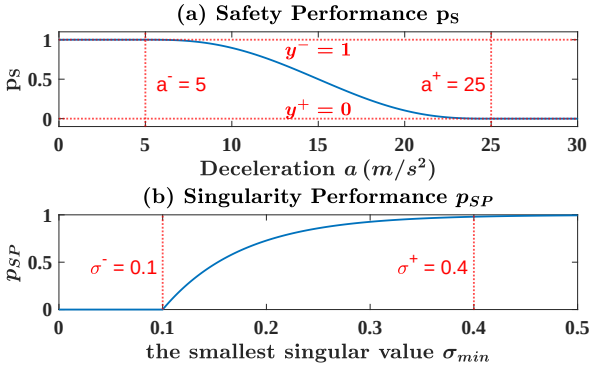

Fig. 2. (a) An example of the Safety Performance p_S versus deceleration a (Equation 4) when $x^- = a^- = 5$, $x^+ = a^+ = 25$, $y^- = 1$ and $y^+ = 0$. (b) An example of the Singularity Performance p_{SP} versus the smallest singular value σ_{min} (Equation 6) when $\sigma^+ = 0.4$, $\sigma^- = 0.1$ and $\varphi = 0.02$.

Algorithm 1 Safety Performance Calculation

for $i \leftarrow 1$ **to** N_{obj} **do**
  $\vec{d_i} = \vec{s_i} - \vec{s}_{robot}$ // Position vector from the position of robot $\vec{s}_{robot}$ to the position of object i $\vec{s_i}$
  $\vec{v_i} = 0$ // The velocity toward object i
  // If the angle between vector of robot velocity $\vec{v}_{robot}$ and d_i is less than or equal to 90 degree, calculate the velocity toward the object i
  if $\vec{v}_{robot} \cdot \vec{d_i} \geq 0$ **then**
    $\vec{v_i} = proj_{\vec{d_i}} \vec{v}_{robot}$
  end if
  $a_i = \frac{0^2 - \|\vec{v_i}\|^2}{2 \times \|\vec{d_i}\|}$ // Deceleration for object i based on Equation 5.
  $p_S^i = f(|a_i|, a^-, a^+, 1, 0)$ // Safety Performance for object i based on Equation 4.
end for
$p_S = \min(p_S^i), \forall i \in 1, 2, 3...N_{obj}$ // Safety Performance p_S is defined as the smallest safety performance among all the surrounding objects

singular values decomposition (SVD) of the Jacobian matrix (J). Inspired from [14], an exponentially-shaped function was introduced to scale the singular value.

$$p_{SP}[n] = \begin{cases} 1 - \varphi^{\frac{\sigma_{min}[n] - \sigma^-}{\sigma^+ - \sigma^-}} & \sigma_{min}[n] > \sigma^- \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

σ^+ is the smallest singular value threshold at which p_{SP} starts to reduce as shown in Figure 2(b). σ^- is the minimum value allowed for the smallest singular value ($\sigma^+ > \sigma^- > 0$). The difference between σ^+ and σ^- should not be too small to ensure a smooth transition of p_{SP} with respect to σ_{min}. φ determines the smoothness of transition between σ^- and σ^+ ($1 >> \varphi > 0$). φ should be tuned to ensure both smooth transition and $p_{SP}(\sigma^+) \approx 1$ as shown in Figure 2(b).

From Figure 2(b), when $\sigma_{min} > \sigma^+$, $p_{SP} \approx 1$ which indicates that the possibility of entering a singular configuration is low. When $\sigma_{min} < \sigma^-$, $p_{SP} = 0$, indicates that the possibility of entering a singular configuration is high. In addition, whether the manipulator is approaching a singular configuration is also another important factor to be considered when calculating p_{SP}. p_{SP} should be higher when the manipulator is heading away from a singular configuration because the possibility of entering a singular configuration is lower compared with heading toward a singular configuration. Based on [14], if the desired robot end-effector velocity in Cartesian space x_d causes the decrease of smallest singular value σ_{min}, it indicates that the manipulator is approaching a singular configuration and vice versa.

The Jacobian matrix $J[n+1]$ at the next time step is calculated based on the robot's current pose and desired velocity. $\sigma_{min}[n+1]$ is calculated based on $J[n+1]$ through SVD. Therefore, $\sigma^+ = \sigma^+_{AW}$ and $\sigma^- = \sigma^-_{AW}$ when the manipulator is leaving a singular configuration ($\sigma_{min}[n+1] \geq \sigma_{min}[n]$) and $\sigma^+ = \sigma^+_{AP}$ and $\sigma^- = \sigma^-_{AP}$ when approaching a singular configuration ($\sigma_{min}[n+1] < \sigma_{min}[n]$):

$$\sigma^+, \sigma^- = \begin{cases} \sigma^+_{AW}, \sigma^-_{AW} & \text{if } \sigma_{min}[n+1] \geq \sigma_{min}[n] \\ \sigma^+_{AP}, \sigma^-_{AP} & \text{if } \sigma_{min}[n+1] < \sigma_{min}[n] \end{cases} \quad (7)$$

The parameters in Equation 7 should be tuned that $\sigma^+_{AP} > \sigma^+_{AW} > 0$ and $\sigma^-_{AP} > \sigma^-_{AW} > 0$ to ensure $p_{SP}(\sigma_{min}|\sigma^+_{AW}, \sigma^-_{AW}) > p_{SP}(\sigma_{min}|\sigma^+_{AP}, \sigma^-_{AP})$ to reflect that the possibility of entering a singular configuration is lower when heading away a singular configuration compared to heading toward a singular configuration.

C. Smoothness Performance

An experienced human co-worker normally operates the robot with smoother movements when compared with novice users. One way to measure smoothness is to use jerk, the first time derivative of acceleration [15]. As a result, smoothness sm can be calculated as:

$$sm[n] = \|\dddot{\vec{v}}[n]\| \quad (8)$$

$\vec{v}[n]$ is the vector of the velocity of the end-effector of the robot in Cartesian space. And $\|.\|$ is the Euclidean norm.

Smoothness Performance $p_{SM} \in [0, 1]$ is used to quantify the degree of smoothness when the human co-worker moves the robot during performing a pHRC task. $p_{SM} = 0$ is defined as an extremely unsmooth movement. $p_{SM} = 1$ corresponds to when the smoothness of human co-worker's movement is within an acceptable range. The interpolation function in Equation 4 is used to measure a p_{SM}.

$$p_{SM}[n] = f(sm[n], sm^-, sm^+, 1, 0) \quad (9)$$

sm^- is the threshold smoothness at which p_{SM} starts to reduce from 1 and sm^+ is the maximum smoothness allowed.

D. Physical Performance

The human co-worker's performance is also affected by the physical workload the human co-worker has to take during a period of time. Heavy workload causes muscle fatigue and thus affects performance. Sadrfaridpour [9] proposed a human physical performance model for collaborative manufacturing which can be used for pHRC because the force applied by the

human co-worker could be measured through a force-torque sensor in real-time. The human physical performance, p_{PW}, is calculated by Equation 10. (Note that the human co-worker physical performance is extremely complex to model, the physical performance model in this paper is just a simplified model).

$$p_{PW}[n] = \frac{F_{max,iso}[n] - F_{th}}{MVC - F_{th}} \tag{10}$$

F_{th} is the equilibrium point at which the fatigue and recovery balance out. Maximum Voluntary Contraction (MVC) is the maximum isometric force with zero level of fatigue ($F_{max,iso} = MVC$). $F_{max,iso}[n]$ is the maximum isometric force which will reduce when a human co-worker's muscle applies a force for some time due to the level of fatigue increases. Hence, the physical performance p_{PW} degrades correspondingly when the $F_{max,iso}$ decreases as shown in Equation 10. $F_{max,iso}$ is shown as:

$$\begin{aligned} F_{max,iso}[n+1] = F_{max,iso}[n] \\ -C_f F_{max,iso}[n]\frac{\|F[n]\|}{MVC} + C_r(MVC - F_{max,iso}[n]) \end{aligned} \tag{11}$$

C_f and C_r are the fatigue and recovery coefficients which are individual-dependent. $\|F[n]\|$ is the magnitude of force applied by the human co-worker. F_{th} is defined as:

$$F_{th} = MVC\frac{C_r}{2C_f}(-1 + \sqrt{1 + \frac{4C_f}{C_r}}) \tag{12}$$

E. Cognitive Performance

Cognitive workload is a measure of mental work when executing a task. Therefore, it is believed to have an effect on human performance [9]. Sadrfaridpour [9] proposed using the Yerkes-Dodson (YD) law to describe a human co-worker's cognitive performance. The YD law [16] stated that when the level of arousal increases, the human cognitive performance increases correspondingly. However, it will only happen before the level of arousal increases up to the point that is known as the optimal level of arousal (OLA). After this point, the human cognitive performance reduces as the level of arousal increases. The OLA value is smaller for a more difficult task. Saeidi [8] proposed a dynamic model, which describes the relationship between human arousal, task difficulty and performance based on the YD law:

$$p_{CP}[n] = (p^+_{CP} - p^-_{CP})(\frac{r[n]}{\beta_{CP}})^{\beta_{CP}}(\frac{1 - r[n]}{1 - \beta_{CP}})^{1-\beta_{CP}} + p^-_{CP} \tag{13}$$

$r[n] \in [0, 1]$ is the utilization ratio of the human co-worker, which represents the amount of time that the human co-worker has been controlling the robot. p^+_{CP} and p^-_{CP} are the upper and lower limits of the Cognitive Performance p_{CP} which are individual-specific. The value of β_{CP} is determined by task difficulty. A larger β_{CP} represents a less difficult task. The most difficult task is when $\beta_{CP} = 0$. Figure 3 shows the human performance compared with different task difficulty β_{CP}. When the utilization ratio $r[n]$ gradually increases, p_{CP} increases up to the highest point (OLA point) then decreases. In addition, the OLA point shifts to the right when the task becomes easier ($\beta_{CP} \to 1$).

$$\begin{aligned} r[n+1] &= a_r r[n] + b_r M[n] \\ a_r &= 1 - \frac{1}{\tau} \in (0,1) \quad b_r = \frac{1}{\tau} \in (0,1) \end{aligned} \tag{14}$$

τ is the time constant that defines the sensitivity of the next time step utilization ratio $r[n+1]$ to the current utilization ratio $r[n]$. Larger τ causes less variation in utilization ratio r. $M[n]$ is the control mode. $M[n] = 1$ refers to manual control mode and $M[n] = 0$ represents autonomous control mode [8].

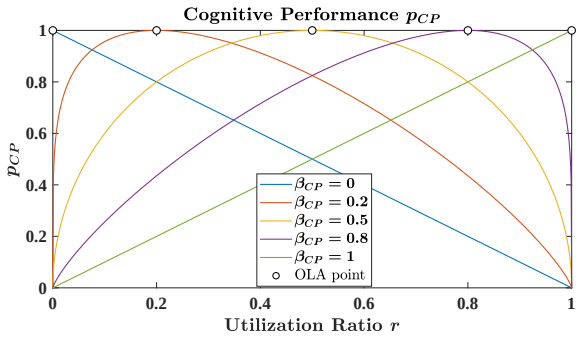

Fig. 3. Examples of the Cognitive Performance p_{CP} versus utilization ratio r with different levels of task difficulty β_{CP} (Equation 13).

IV. Experimental Evaluation

This section explains the design of three experiments to verify the proposed trust model by using a human-robot collaborative system.

A. Experimental Testbed

The experiment testbed is called ANBOT [17] (Figure 4a) which is a collaborative robotic system for human-robot collaborative operation. ANBOT consists of a UR10 arm from Universal Robots and a six-axis force-torque sensor mounted between the robot end-effector and the robotic arm to measure the interaction forces applied by the human co-worker. The human co-worker constantly hold the handlebar mounted on the robot end-effector.

There is a monitor to display the actual trajectory of the robot end-effector, as shown in Figure 4b. In this experiment, the movement of the robot is constrained into two dimensions to reduce the complexity of experiment. Hence, the human co-worker can only move the robot in vertical and horizontal directions, which are parallel to the monitor.

B. Design of Experiments

Three experiments are designed for verifying the proposed computational trust model. Figure 5a shows the trajectory that the robot end-effector needs to track in Experiments 1 and 2. Figure 5b is for Experiment 3. The big filled red circles represent objects (or obstacles) located on the trajectory. The white line is the desired trajectory that needs to be followed.

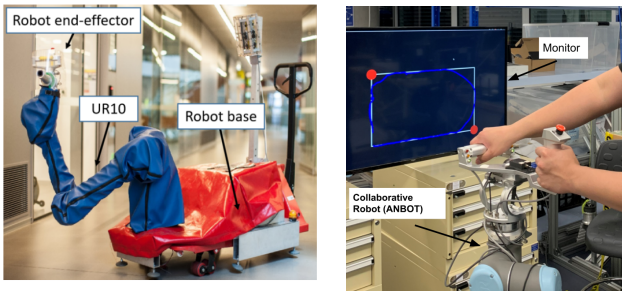

Fig. 4. (a) Experimental Testbed - ANBOT. (b) The human co-worker is operating the ANBOT to follow a desired trajectory.

Each of the experiments starts and ends at the bottom left corner of the rectangle with clockwise movement.

Experiment 1: This experiment is designed for verifying the Smoothness Performance model in Section III-C and the Singularity Performance model in Section III-B. In this experiment, the human subject is asked to move the robot end-effector to follow the path (shown in Figure 5a) as accurately and smoothly as possible. There is no limitation on time to complete the entire path. During the path following, the robot manipulator will approach and leave singular configurations. Therefore the Singular Performance model can be verified.

Experiment 2: This experiment is for verifying the Physical Performance model in Section III-D and Cognitive Performance Model in Section III-E. The Physical Performance p_{PW} and Cognitive Performance p_{CP} are mainly affected by the duration of operating the robot. p_{PW} is also affected by the magnitude of interaction force. Therefore, the human subject is asked to move along the path (Figure 5a), which is similar to Experiment 1, but as quickly as possible to ensure that the human subject exerts a large force while executing the task. Also, the human subject is asked to move around the path continuously for five loops to ensure that the duration of Experiment 2 is longer than other two experiments.

Experiment 3: This experiment is designed for verifying the Safety Performance model in Section III-A by introducing two objects/obstacles on the same path in Experiments 1 and 2. Besides avoiding the objects, the requirement on the human subject is the same as in Experiment 1. However, the human subject is required to move as fast as possible from the top right corner of the path toward the top left corner to demonstrate the effect of the velocity component in Equation 5 on Safety Performance p_S. This experiment is also conducted for verifying the combined models of human co-worker performance in Section III and the computational trust model in Section II because the variation of all the performance factors can be observed.

V. Results and Discussion

All the experiments are conducted by one human subject.

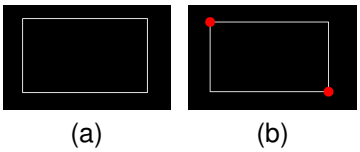

Fig. 5. (a) Trajectory tracking path for Experiments 1 and 2. (b) Trajectory tracking path for Experiment 3.

A. Experiment 1 – verifying the Smoothness Performance model and the Singularity Performance model

1) Singularity Performance: The values of the parameters in Singularity Performance model are $\varphi = 0.02$, $\sigma^+_{AW} = 0.25$, $\sigma^-_{AW} = 0.15$, $\sigma^+_{AP} = 0.35$, $\sigma^-_{AP} = 0.25$ [14]. The desired workspace and sensor noise need to be considered when determining the values of σ^+_{AW}, σ^-_{AW}, σ^+_{AP}, and σ^-_{AP}. When the values are larger, the amount of configurations regarded as unsafe movement is larger (Figure 2(b)), which results in decrease of the desired workspace. Moreover, if the sensor's noise is large, those values should be larger to increase the robustness.

Figure 6 shows the smallest singular value σ_{min} and Singularity Performance p_{SP} in Experiment 1. Before 8.5s, σ_{min} is larger than $\sigma^+_{AP} = 0.35$, hence, $p_{SP} \approx 1$ (Equation 6). At around 8.5s, p_{SP} starts to decrease because $\sigma_{min} < \sigma^+_{AP}$ until $p_{SP} = 0$ due to $\sigma_{min} < \sigma^-_{AP}$ at around 14.5s. After this time, the robot starts to head away from a singular configuration (yellow shaded area) which can be seen that σ_{min} starts to increase, then p_{SP} increases up to 1 instantaneously due to $\sigma_{min} > \sigma^+_{AW}$. It is noted that Equation 7 successfully predicts whether the manipulator is heading toward a singular configuration.

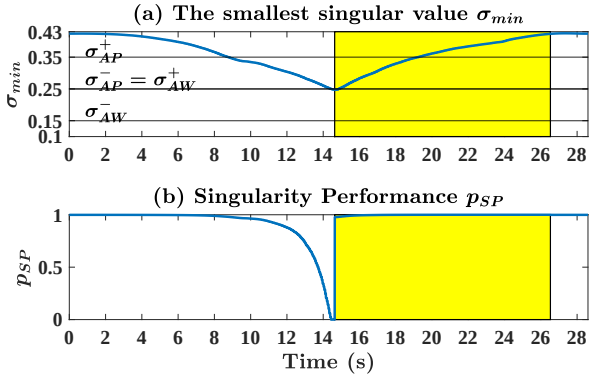

Fig. 6. (a) The smallest singular value σ_{min}. (b) The Singularity Performance p_{SP}. Yellow shaded area is the prediction on whether the robot is heading away a singular configuration based on Equation 7.

2) Smoothness Performance: The smoothness sm in Equation 8 is calculated through the second time derivative of second degree polynomial curve fitting of the velocity using the least-squares method with a time window of 1 second for noise reduction.

Figure 8(a) shows that sm is much higher at the beginning of the experiment and at the turning points (corners) of the desired path (Figure 7). The human co-worker needs to

decelerate before a turning point and accelerate after passing the turning point, which causes the unsmooth movement. sm is high at the beginning of the experiment because the human co-worker starts to move from rest. In order to show the details of Figure 8(a), sm is bounded between $[0, 1]m/s^3$.

Because a human co-worker tends to move with minimum-jerkiness-profile. sm^- should be set to as small as possible ($sm^- \to 0$). sm^+ depends on whether the requirement on the smoothness is high or not. In this experiment, the requirement for smoothness of the human co-worker's movement is high. Therefore, $sm^- = 0.1m/s^3$ and $sm^+ = 0.9m/s^3$ (Equation 9). In Figure 8(b), the Smoothness Performance p_{SM} decreases correspondingly when sm increases as long as $sm > sm^-$. It can also be seen that $p_{SM} = 0$ when $sm >= sm^+$ at the beginning of the experiment.

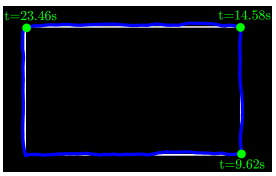

Fig. 7. The blue line corresponds to the actual robot end-effector Trajectory. The labelled time corresponds to the time when the human co-worker changes the directions on the corners of the trajectory.

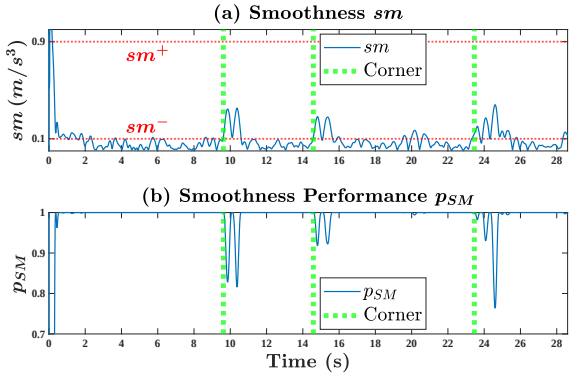

Fig. 8. (a) Smoothness sm.(b) Smoothness Performance p_{SM}. The green vertical lines correspond to the time when human co-worker changes the directions at the corners of the trajectory.

B. Experiment 2 – verifying the Physical Performance model and the Cognitive Performance model

1) Physical Performance: The values of parameters in the physical performance model are set $MVC = 200$, $F_{th} = 151.9$, $C_f = 10^{-4}$ and $C_r = 2.4 \times 10^{-4}$ which are similar to the literature [9]. However, the parameters are individual-dependent. The detailed method for measuring those parameters can be found in [9].

Figure 9 shows the magnitude of interaction force $\|F\|$, the maximum isometric force $F_{max,iso}$ and physical performance p_{PW} versus time in Experiments 1, 2 and 3. In Experiment 2, the human subject applied a much larger force and executed the task for the longest time (around 40.6s) because the human

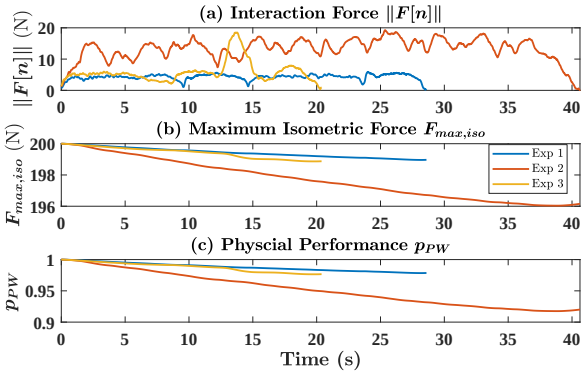

Fig. 9. (a) Magnitude of interaction force $\|F\|$ applied by human co-worker and measured by force-torque sensor. (b) The maximum isometric force $F_{max,iso}$. (c) Physical Performance p_{PW}.

subject is required to move along the trajectory as fast as possible for five loops. It can be seen that $F_{max,iso} = 196.16N$ and $p_{PW} = 0.92$ are the lowest at the end of experiments which represent the highest level of fatigue. The results demonstrate that the Physical Performance model could measure the human co-worker fatigue level effectively.

Figure 9(b) shows that the value of maximum isometric force $F_{max,iso}$ starts to reduce from maximum voluntary contraction ($MVC = 200$) at the beginning of the experiments. At the end of the experiments, $F_{max,iso} = 198.97N, 196.16N, 198.88N$ in Experiments 1, 2 and 3 and those values are close to 200 because the durations of all the experiments are short (28.6s, 40.6s, 20.4s), which indicates that the fatigue level of human co-worker increases slightly.

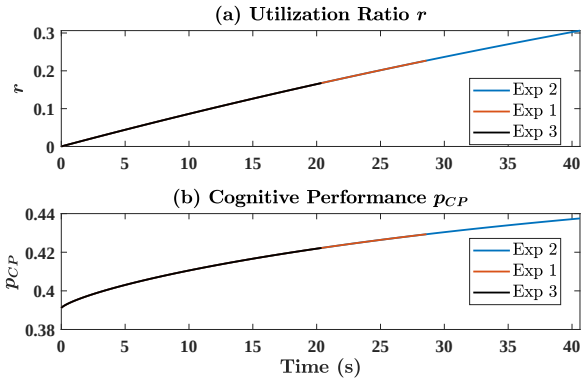

Fig. 10. (a) Utilization Ratio r (b) Cognitive Performance p_{CP} in Experiments 1, 2 and 3.

2) Cognitive Performance: The parameters in the Cognitive Performance model are subject and task-dependent. The details for selecting the parameters can be found in [8]. Based on [8], the parameter values used in this experiment are: $p_{CP}^- = 0.391$, $p_{CP}^+ = 0.4602$, $\beta_{CP} = 0.74$ and $a_r = 0.9991$. a_r depends on the system sampling rate. The sampling rate in [8] is 10Hz, and the sampling rate for this paper is 125Hz, therefore, $a_r = 0.999928$. Because the human subject constantly moves the robot during the experiment, the control mode is always 1 ($M = 1$ in Equation 14).

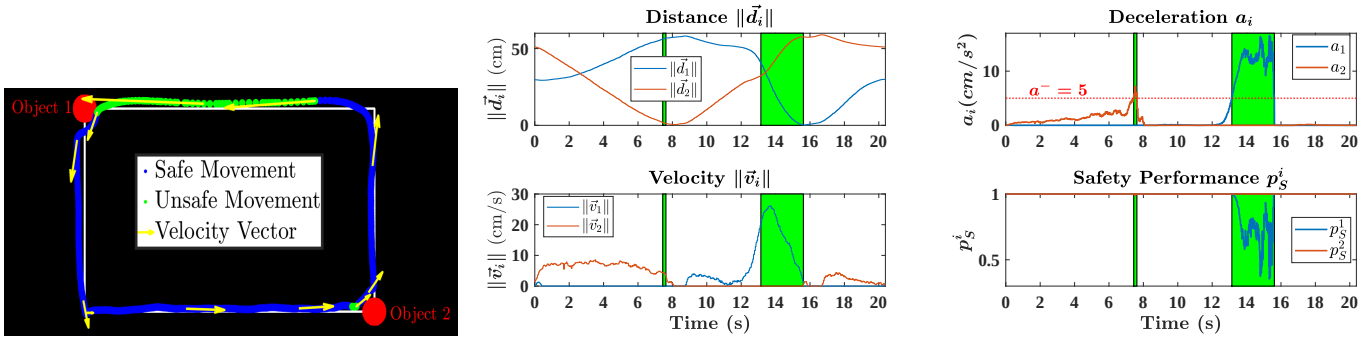

Fig. 11. (a) The actual robot end-effector trajectory along the designed trajectory. The blue trajectory refers to the safe movement ($\max(a_1, a_2) \leq a^-$) and the green trajectory refers to the unsafe movement ($\max(a_1, a_2) > a^-$) which corresponds to the shaded green regions in Figures 11b and 11c. The yellow vectors are the robot velocity vector. (b) Top: The distance between the robot and Object 1 $\|\vec{d}_1\|$ (Object 2 $\|\vec{d}_2\|$). Bottom: the velocity of the robot toward Object 1 $\|\vec{v}_1\|$ (Object 2 $\|\vec{v}_2\|$). (c) Top: The magnitude of the constant deceleration required to stop the robot when it reaches the position to collide with Object 1 a_1 (Object 2 a_2). Bottom: The safety performance for Object 1 p_S^1 (Object 2 p_S^2).

As shown in Figure 10(a), the utilization ratio r starts to increase from 0 when the human subject starts to operate the robot. r constantly increases during the experiment because r represents the amount of time the human co-worker has operated the robot. At the end of the experiments, $r \approx 0.3$ in Experiment 2 is the highest because the time the human subject have controlled the robot is the longest (around 40.6s). In Figure 10(b), Cognitive Performance p_{CP} starts to increase correspondingly from the lower threshold of cognitive performance $p_{CP}^- = 0.391$. The reason why p_{CP} constantly increases during the experiment is that r is smaller than the optimal level of arousal (OLA), which is when $\beta_{CP} = 0.74$. The reason why r and p_{CP} are the same between $0s$ and $20.4s$ for the three experiments due to all the cognitive performance parameters are the same ($a_r, p_{CP}^-, p_{CP}^+, \beta_{CP}$) and the human subject constantly operates the robot ($M = 1$ in Equation 14).

C. Experiment 3 – verifying the Safety Performance model, Human Co-worker Performance and the Computational Trust Model

1) Safety Performance: In Figure 11a, the green section indicates the period in which the movement of a human co-worker is regarded as unsafe. The blue section is regarded as a safe movement. The period of unsafe movement for approaching Object 2 (0.18s) is much shorter compared to Object 1 (2.47s) because the velocity toward Object 1 is much larger than that of Object 2 as shown in Figure 11b. It can also be seen in Figure 11a that the length of the robot velocity vector is much longer in green section when approaching Object 1.

The movement is identified as unsafe only when the robot is very close to Object 2 (1.15cm). Therefore, when the velocity is low, the distance will be the major factor determining the safety of movement.

Another essential factor for determining the safety movement is the direction of motion of the robot. At around 15.62s, the robot is heading away from Object 1 as shown the yellow arrow in Figure 11a. The movement is regarded as safe even though the distance between robot and Object 1 is small (0.22cm) as shown in Figure 11b.

The parameter values selected include $a^- = 5cm/s^2$ and $a^+ = 25cm/s^2$ (Equation 4). Similar considerations from singularity performance (Section V-A1) can be applied on selecting a^- and a^+. In Figure 11c, the safety performance p_S^i decreases correspondingly when $a_i > a^-$, indicating the possibility of collision starts to increase. It is hard to observe the decrease of p_S^2 because a_2 is a little over the a^-.

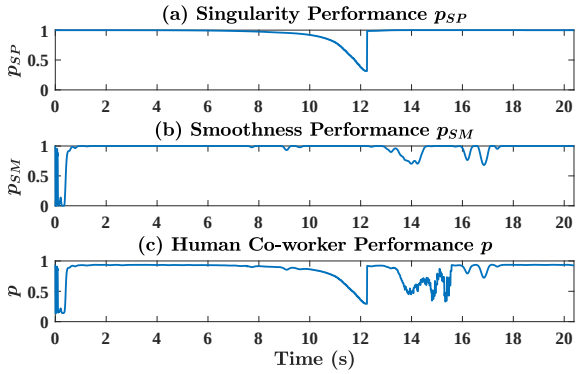

Fig. 12. (a) Singularity Performance p_{SP}. (b) Smoothness Performance p_{SM}. (b) The combined Human co-worker performance p

2) The combined Human Co-worker's Performance: γ_{SM}, γ_{PW} and γ_{CP} in Equation 3 are determined by the relative importance. For this experiment, the requirement on the smoothness of movement is high. In addition, the duration of experiment is short which result in the variation of p_{PW} and p_{CP} is subtle. Hence, we set $\gamma_{SM} = 0.8$, $\gamma_{PW} = 0.1$ and $\gamma_{CP} = 0.1$. p_A^{nc} is important because the smoothness requirement of human co-worker is high, therefore, $C = 0$.

Figure 12 shows that p is low at the beginning of the experiment because $p_{SM} = 0$ and $\gamma_{SM} = 0.8$. It can be observed that p decreases immediately when p_S (Figure 11c) or p_{SP} reduces because they are critical performance factors (Equation 3). $p < 1$ all the time due to p_{CP} bounds within

the interval $[0.391, 0.4602]$ along the experiment as shown in Figure 10. The effect of p_{PW} on p is trivial because the variation of p_{PW} is small (Figure 9) and $\gamma_{PW} = 0.1$.

3) The Computational Trust Model: The trust with different values of β and N is simulated as shown in Figure 13 based on the combined human co-worker performance (Figure 12). The time interval of p is bounded into [10 14]s to show the comparisons more clearly. Larger values of β and N will result in T becoming smoother. A good choice of values for β and N should balance the smoothness and responsiveness of T to p.

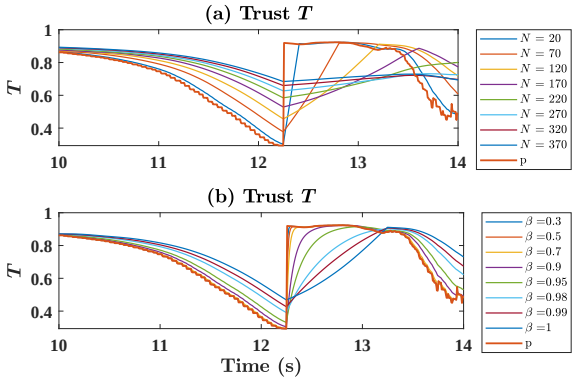

Fig. 13. (a) Trust plots for different values of time window size N when $\beta = 0.999$. (b) Trust plots for different values of β when $N = 125$.

VI. CONCLUSIONS AND FUTURE WORK

A computational model of robot trust in a human co-worker was proposed. This model takes into account many factors in physical human-robot collaboration, including robot safety, robot singularity, smoothness of robot motion, and human physical and cognitive performance. Three experiments were conducted for verification of all the factors of the model. The experimental results show that the modelling is appropriate.

Future work includes how the proposed trust model can be fit into a larger context of trust in human-robot interaction. When the human-to-robot trust is low, the human's dependence on the robot is low. If at the same time, the robot trust in human is low, the robot may not rely on the human. This may therefore result in a conflict of intents between the human and the robot and cause human trust in the robot to reduce further, leading to a reduction in performance of the combined human-robot system. Experimental validation of the proposed framework applied to real-world scenarios is required.

Real-world evaluation requires the proposed trust model to be integrated with robot control, such as trust-based impedance or admittance control, or trust-based role arbitration. For example, in trust-based admittance control, the damping in admittance control could be adapted based on the level of robot-to-human trust. When the trust decreases, the damping can be increased to impede or slow down the human co-worker's movement. For trust-based role arbitration, when the robot trust in the human co-worker decreases to a certain level, the robot may gradually take control of the collaboration.

The trust model could also be applied to other human-robot interaction applications such as a teleoperated medical robot. Smoothness performance could be used to evaluate human co-worker performance in control accuracy. Cognitive performance might be an essential factor due to the high demand for attention during a medical task.

ACKNOWLEDGMENT

This study is supported in part by the Australian Research Council (ARC) Discovery Project Grant [DP210101093]. The authors would like to thank Jonathan Woolfrey, Yujun Lai, Tiancheng Li and Sheila Sutjipto for their assistance.

REFERENCES

[1] J. E. Colgate, W. Wannasuphoprasit, and M. A. Peshkin, "Cobots: robots for collaboration with human operators," *American Society of Mechanical Engineers, Dynamic Systems and Control Division (Publication) DSC*, vol. 58, pp. 433–439, 1996.

[2] H. Yu, S. Huang, G. Chen, Y. Pan, and Z. Guo, "Human-robot interaction control of rehabilitation robots with series elastic actuators," *IEEE Transactions on Robotics*, vol. 31, pp. 1089–1100, 2015.

[3] E. Gambao, M. Hernando, and D. Surdilovic, "A new generation of collaborative robots for material handling," *2012 Proceedings of the 29th International Symposium of Automation and Robotics in Construction, ISARC 2012*, 2012.

[4] S. Aldini, A. Akella, A. K. Singh, Y. K. Wang, M. Carmichael, D. Liu, and C. T. Lin, "Effect of mechanical resistance on cognitive conflict in physical human-robot collaboration," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2019-May, pp. 6137–6143, 2019.

[5] M. Lewis, K. Sycara, and P. Walker, *The Role of Trust in Human-Robot Interaction*, pp. 135–159. Cham: Springer International Publishing, 2018.

[6] J. Lee and N. Moray, "Trust, control strategies and allocation of function in human-machine systems," *Ergonomics*, vol. 35, no. 10, pp. 1243–1270, 1992.

[7] A. Xu and G. Dudek, "Optimo: Online probabilistic trust inference model for asymmetric human-robot collaborations," *ACM/IEEE International Conference on Human-Robot Interaction*, vol. 2015-March, pp. 221–228, 2015.

[8] S. Hamed, *Trust-Based Control of (Semi)Autonomous Mobile Robotic Systems*. PhD thesis, Clemson University, 2016.

[9] B. Sadrfaridpour, *Trust-Based Control of Robotic Manipulators in Collaborative Assembly in Manufacturing*. PhD thesis, Clemson University, 2018.

[10] S. M. Mizanoor Rahman, Y. Wang, I. D. Walker, L. Mears, R. Pak, and S. Remy, "Trust-based compliant robot-human handovers of payloads in collaborative assembly in flexible manufacturing," in *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 355–360, 2016.

[11] A. Tran, "Robot confidence modeling and role change in physical human-robot collaboration," 2019.

[12] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.

[13] B. Navarro, A. Cherubini, A. Fonte, G. Poisson, and P. Fraisse, "A framework for intuitive collaboration with a mobile manipulator," pp. 6293–6298, 09 2017.

[14] M. G. Carmichael, D. Liu, and K. J. Waldron, "A framework for singularity-robust manipulator control during physical human-robot interaction," *International Journal of Robotics Research*, 2017.

[15] N. Hogan and D. Sternad, "Sensitivity of smoothness measures to movement duration, amplitude, and arrests," *Journal of motor behavior*, vol. 41, pp. 529–34, 11 2009.

[16] J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation in the kitten.," *Journal of Animal Behavior*, vol. 5, pp. 330–336.

[17] M. G. Carmichael, S. Aldini, R. Khonasty, A. Tran, C. Reeks, D. Liu, K. J. Waldron, and G. Dissanayake, "The anbot: An intelligent robotic co-worker for industrial abrasive blasting," *IEEE International Conference on Intelligent Robots and Systems*, pp. 8026–8033, 2019.