3-D Fitts' Law for Performance Prediction of Human–Machine Teaming

Yuhui Wan and Chengxu Zhou

***Abstract*—Human machine teaming (HMT) applications, bolstered by advancements in control technology, are proliferating across industries. Despite this, accurately predicting HMT performance and time efficiency, particularly in open and potentially hazardous environments, remains challenging due to the complexity of modeling mission difficulty and human behavior. This study tackles the challenge by extending cursor-based Fitts' law into a 3-D context with six degrees of freedom. This represents the first attempt to model task difficulty in HMT by integrating both target translational and orientational components, effectively expanding Fitts' law beyond human–computer interaction to robots and machinery. This novel approach is tailored for HMT applications and tested through an excavator simulation and a quadruped manipulator robot experiment. The study showcases our model's capability to predict the performance of complex, multistep HMT tasks accurately. This research opens a novel pathway in HMT performance prediction, offering promising implications for real-world applications.**

***Index Terms*—Fitts' law, human–machine teaming (HMT), human–robot interaction, mobile manipulation.**

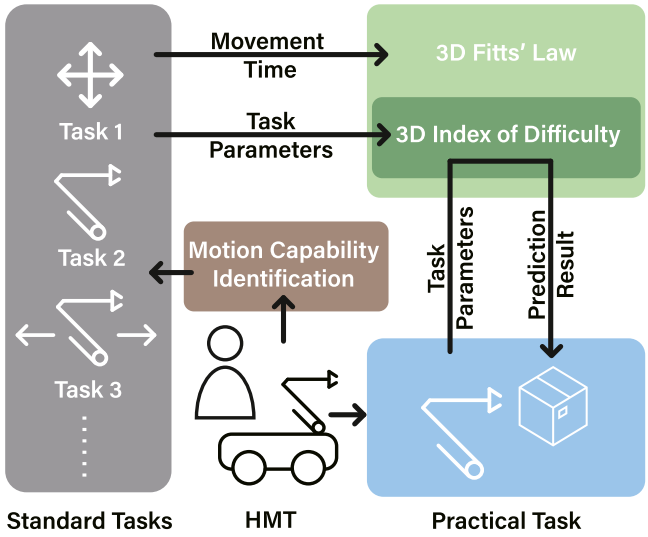


Fig. 1. Proposed method designs standard tasks based on the splitting motion capability of targeted HMT and feeds standard tasks' data into 3-D Fitts' law to understand the performance of HMT, then it predicts feature task performance based on new task parameters and previous HMT data.

I. Introduction

HUMAN–MACHINE teaming (HMT) represents a pivotal advancement in technology, facilitating collaborative operations between humans and machines across a variety of complex and dynamic environments. This synergy is particularly prominent in sectors, such as the manufacturing industry [1], state-of-the-art aerial transportation [2], aerospace [3], and emergency response scenarios [4], where the integration of sophisticated machines and skilled human operators contributes to systems performance. Performance prediction is critical as it directly influences decision-making and the overall effectiveness of human–machine interactions [5]. The ability to predict HMT performance under various mission parameters can significantly improve operational planning, minimize downtime, and enhance safety.

Despite advancements in technology, accurately forecasting the time required for HMT to perform specific tasks in real-world scenarios remains a persistent challenge. The prediction of HMT performance remains complex [1], [6], and the inherent unpredictability of missions in open environments [7], [8]. This study addresses these challenges by enhancing task difficulty and human skill modeling, thereby enabling more accurate and efficient outcomes.

Given the complexity of human–machine systems, considering every aspect, including machine parameters, human factors, and system control errors, can be arduous. Our approach leverages Fitts' law [9] as a framework to treat the HMT as an integrated system, where such factors are incorporated collectively within our predictive model as hidden variables. Furthermore, we extend Fitts' law into a 3-D context in six degrees of freedom (DoF) with three DoFs in translation and three DoFs in orientation for the spatial relationship between the machine

Received 9 September 2024; revised 7 February 2025; accepted 7 April 2025. This work was in part supported by the Royal Society under Grant RG\R2\232409 and in part by the Advanced Research and Innovation Agency under Grant SMRB-SE01-P06. Paper no. TII-24-4677. *(Corresponding author: Chengxu Zhou.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Business, Environment and Social Sciences Research Ethics Committee of the University of Leeds under Application No. MEEC 21-009.

Yuhui Wan is with the School of Mechanical Engineering, University of Leeds, LS2 9JT Leeds, U.K. (e-mail: wan@w9ybz.com).

Chengxu Zhou is with the Department of Computer Science, University College London, WC1E 6BT London, U.K. (e-mail: chengxu.zhou@ucl.ac.uk).

This article has supplementary downloadable material available at https://doi.org/10.1109/TII.2025.3563553, provided by the authors.

Digital Object Identifier 10.1109/TII.2025.3563553

agent and mission target, which enhances precisely modeling task difficulty and predicting future HMT performance, alongside with motion capability identification to generate standard tasks (see Fig. 1 for a depiction). By including the orientation of the target, our approach extends previous works from human–computer interaction [10], [11] to real-world HMT applications. This novel approach is the first of its kind, filling the critical gap in existing models by considering the multidimensional nature of HMT tasks in the real world.

Predictive models for HMT performance encompass a wide spectrum of approaches [12]. Numerous studies have concentrated on human-centric factors, such as modeling cognitive performance [13], while others have utilized operator models to semantically map working memory [14]. The notion of 'neglect tolerance', based on the operator's available time, has been employed to estimate the maximum number of robots an operator can manage [15]. One study utilizes a higher order Markov chain approach to model human assembly sequences alongside robots to predict future activities based on their historical sequence [16]. Furthermore, physiological measures, including brain activity and eye movement, have been leveraged to anticipate performance and situational awareness in teleoperated drone exploration [17]. While these models provide invaluable insights, they often overlook the specific characteristics and requirements of machines and tasks and lack a reliable benchmark for real-world HMT applications.

Fitts' law [9], a predominant model in human–computer interaction, holds the potential for predicting system performance by integrating human and computer interface elements. However, its applicability is currently constrained by its 1-D design that uses distance and target size to predict the MT of the cursor on the screen, thereby limiting its relevance outside virtual environments. Efforts to extend Fitts' law, introducing variables, such as cursor size and angular distance [18], [19], [20], [21], [22], have predominantly remained within the confines of virtual or simplified 2-D contexts. This poses a challenge in accurately modeling the complex, real-world tasks that are characteristic of HMT systems. While some studies have attempted to apply these adapted versions of Fitts' law to real-world HMT [23], [24], machines and robots are often viewed as conduits for human input rather than direct participants in task execution. Consequently, these approaches predominantly focus on human performance, neglecting the physical capabilities and parameters of the machinery involved.

As a result, there is a substantial gap in current predictive models, which struggle to offer a comprehensive framework that can encapsulate the diverse and intricate nature of HMT systems and the tasks they encounter. This includes challenges related to understanding the motion capabilities of machines and the complexity of tasks in real-world settings, underscoring the need for more holistic and multidimensional approaches in HMT performance prediction.

In response to these challenges, this article presents a pioneering framework for predicting HMT performance. The key contributions of our study can be summarized as follows:

1) We propose a systematic method to predict real-world HMT performance. Our approach comprises three key components as follows.
   a) The creation of an expansive mathematical model of HMT tasks by extending Shannon's index of difficulty (ID) formulation into 3-D space with six DoFs, considering both translation and orientation between the target and the agent.
   b) The development of principles for designing standardized benchmark tasks, which ascertain the motion capabilities of the targeted machine agent, laying the foundation for a comprehensive model for all subsequent tasks.
   c) The development of a formulated performance prediction model based on Fitts' law, employing standard tasks to understand the characteristics of the HMT and predict future system performance.
   d) Implementation of a method to identify the minimum set of standard tasks required for accurate prediction based on the machine agent's motion capabilities.
2) Following our theoretical contributions, we conducted two empirical validations to demonstrate the model's efficacy and applicability in diverse HMT systems.
   a) A web-based simulation involving 16 human participants using a remote control excavator, providing insights into virtual HMT interactions.
   b) A real-world experiment with a quadruped manipulator robot, involving 7 human users and 2 distinct human–machine interfaces, to assess the model in a tangible, real-life scenario.

The rest of this article is organized as follows. We begin with a review of Fitts' law and its adaptations in Section II. Our proposed 3-D Fitts' law is outlined in Section III. The validation method in Section IV, encompassing the simulation in Section IV-C and experimental in Section IV-D. The results and analysis are presented in Section V, followed by a discussion in Section VI. Finally, Section VII concludes this article.

II. Related Work

P. M. Fitts proposed a widely recognized method, known as Fitts' law, for predicting human–machine interface performance [9]. This model predicts the MT of a user moving a mouse cursor on the screen to a target location via an ID of the movement

$$\mathrm{MT} = a + b \cdot \mathrm{ID} \tag{1}$$

where constants a and b depict a linear polynomial line representing the relationship between MT and ID called the prediction line. The foundation of the model's accuracy is predicated on the accurate modeling of ID, originally encompassing only the linear distance (d) to the target and the width (w) of the target

$$\mathrm{ID} = \log_2\left(\frac{2d}{w}\right). \tag{2}$$

Over the years, numerous modifications and enhancements have been made to the original Fitts' law (2), with one of the most prevalent being the Shannon formulation [20]. This formulation mitigates the possibility of a negative ID to better align with empirical data by incrementing the ratio by one

$$\mathrm{ID} = \log_2\left(\frac{d}{w} + 1\right). \tag{3}$$

However, both the Shannon formulation and the original Fitts' law are limited to 1-D and only consider linear movements. Subsequent research explored the expansion of this model into a 2-D space. Most of these adaptations are predicated on the Shannon formulation (3). One model accounted for both the linear distance to the target position and the cursor's rotation distance to the target direction within the difficulty index by integrating the respective rotational distance (θ) and rotational tolerance (δ) [21]

$$\mathrm{ID} = \log_2\left(\frac{d}{w} + 1\right) + \log_2\left(\frac{\theta}{\delta} + 1\right). \tag{4}$$

Since these models predominantly focus on human control of a cursor, usually represented as a single point devoid of size or mass, there are complications when attempting to represent the machine system. An innovative experiment utilized a human finger as the cursor [10], factoring in the size of the human finger (f) for ID calculation

$$\mathrm{ID} = \log_2\left(\frac{2d}{w + f}\right). \tag{5}$$

This methodology has important implications for modeling the difficulty of HMT tasks since a single point cannot accurately represent most HMT systems. Previous research explored combining these modifications to evaluate the performance of a human–robot system [25].

Nevertheless, these methodologies remain confined to a 2-D space and face difficulties in accurately representing real-world HMT operations. Limited research explores the extension of Fitts' law into the 3-D realm [26] and its application to HMT [23]. Recent studies have combined translation distance and rotational distance [27], and also incorporated changes in target depths into the original Fitts' law [28]. Recent studies have expanded the application of Fitts' law to virtual reality (VR) environments, specifically focusing on cursor control [11], [29]. While these adaptations effectively evaluate human performance within human–computer interaction instead of machinery, their scope is generally limited to the cursor, which only has translation in 3-D space without target orientation, which is crucial in the context of machine operation. Consequently, such adaptations might not fully capture the complexities and nuanced performance demands of HMT in practical, real-world settings. Benefiting from the diversity of approaches evolving from Fitts' law, our proposed model optimally integrates different methods to develop a 3-D six DoFs prediction model to represent operational HMT.

There is a conspicuous lack of research focusing on developing Fitts' law into a 3-D model for predicting HMT performance. Such predictive models would benefit both the development process and real-world operation of HMTs, especially for mobile manipulation missions in uncertain environments.

III. Method

A. Process Overview

This section outlines a step-by-step overview of the algorithmic process involved in applying the proposed 3-D Fitts' law model for predicting HMT performance. The process, as shown in Fig. 1, can be articulated through the following stages.

1) *Motion Capability Identification:*
   a) Identify the motion capabilities of the system.
   b) Determine the minimum set of standard tasks required, as detailed in Section III-B.
2) *Standard Task Modeling:*
   a) Define parameters for each standard task, including distances, angles, and target sizes.
   b) Calculate the ID for each standard task using the extended 3-D formulation of Fitts' law, as detailed in Section III-C.
3) *Data Collection:*
   a) Conduct standard tasks through simulations or real-world experiments and record the performance.
4) *Practical Task Modeling:*
   a) Decompose the practical task into standard tasks.
   b) Calculate the ID of the practical task based on corresponding standard tasks.
5) *Performance Prediction:*
   a) Emphasise to derive prediction lines from the data collected for standard tasks.
   b) Apply the prediction model to estimate performance times for the practical task based on its ID calculated.

B. Motion Capability Identification

Standard tasks play an integral role in both understanding the system and constructing an effective predictive model. The key to building this model lies in the identification of the motion capabilities of the targeted machine agent. This identification process determines the minimum set of standard tasks (p) necessary to ensure the model's comprehensiveness, allowing it to accurately represent and predict a wide range of subsequent tasks. Essentially, the goal is to configure every potential task as a combination of these standard tasks.

The calculation of this minimum set of standard tasks, p, is directly linked to the total number of distinct motion capabilities (m) of the machine agent. This total encompasses both the individual motion capabilities and their possible combinations (C). Mathematically, this relationship can be expressed as

$$p = \sum_{r=1}^{m} C_m^r = \sum_{r=1}^{m} \frac{m!}{r!(m-r)!} = 2^m - 1. \tag{6}$$

To illustrate, consider a wheeled-legged humanoid robot with two distinct types of locomotion capabilities (legs and wheels) and two manipulation capabilities (each arm). For this robot, the total number of motion capabilities (m) is 4. Therefore,

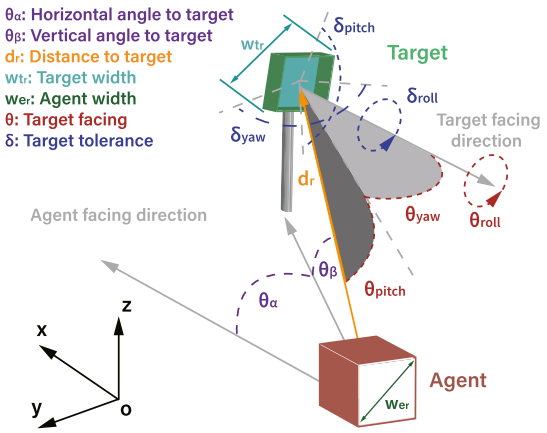


Fig. 2. Giving the position of the machine agent and the target, the definition of each parameter used in calculating the 3-D ID in a polar coordinate system.

according to our model, the required number of standard tasks (p) for accurate prediction would be 15.

C. Extending Fitts' Law to 3-D

Our proposed model extends Fitts' law to predict task performance by quantifying the motion time required to complete tasks with a HMT system, taking into account the task's IDs. The key advancement in our model lies in the comprehensive consideration of task difficulty. Unlike most previous works that treat the agent solely as a point and target as a volume in space and only account for translational motion, our model introduces both translation and orientation in 3-D, considering all six DoFs. This makes our approach the first to extend Fitts' law to true 3-D space, accounting for three translational and three orientational DoFs, as illustrated in Fig. 2. It is critical to consider the target's orientation for tasks in the real world, including pressing buttons and assembling parts. Consequently, the ID in our 3-D Fitts' law is represented by both a translation ID ($\mathrm{ID}_{\mathrm{trans}}$) and an orientation ID ($\mathrm{ID}_{\mathrm{ori}}$), which can be expressed as

$$\mathrm{ID} = \mathrm{ID}_{\mathrm{trans}} + \mathrm{ID}_{\mathrm{ori}}. \tag{7}$$

1) Modeling Distances: Inspired by (4) and (5), there are two fundamental types of distance: linear distance and rotational distance. In this work, we define the ID of linear distance as

$$\mathrm{ID}_{\mathrm{linear}} = \log_2\left(\frac{d}{w_t \pm w_e} + 1\right) \tag{8}$$

where d is the linear distance, and w_t and w_e represent the sizes of the target and agent, respectively. The relationship (addition or subtraction) between the target and agent sizes depends on the specific task requirements.

Next, we define the ID for rotational distance as

$$\mathrm{ID}_{\mathrm{rot}} = \log_2\left(\frac{\theta}{\delta} + 1\right) \tag{9}$$

where θ represents the rotational distance and δ represents the rotational tolerance or limit.

By combining these two forms of distance, we can describe both the translation and orientation of the target relative to the machine agent in 3-D space.

2) Translation Index of Difficulty: For the translation ID ($\mathrm{ID}_{\mathrm{trans}}$), we establish a coordinate system with the agent's starting location at the origin. In this coordinate system, the agent faces the $+x$ direction, with $+y$ pointing to the left and $+z$ pointing upwards. Either a Cartesian or polar coordinate system can be used to describe the position of the target relative to the agent. For example, a Cartesian robot would be better suited to a Cartesian coordinate system, whereas an articulated robot may prefer a polar coordinate system.

In a Cartesian coordinate system, the translation is described by linear distances along the x, y, and z axes. Thus, the $\mathrm{ID}_{\mathrm{trans}}$ can be expressed as the sum of three $\mathrm{ID}_{\mathrm{linear}}$ components

$$\mathrm{ID}_{\mathrm{trans}} = \sum_{i\in\{x,y,z\}} \mathrm{ID}_{\mathrm{linear}_i} = \sum_{i\in\{x,y,z\}} \log_2\left(\frac{d_i}{w_{t_i} \pm w_{e_i}} + 1\right). \tag{10}$$

In a polar coordinate system, the position is described by two rotational distances (α and β) and one linear distance (r). The azimuthal angle (θ_α) is defined as the angle between the positive x-axis and the projection of the point onto the xy-plane, while the polar angle (θ_β) is the angle between the positive xy-plane and the line from the agent to the target, as depicted in Fig. 2. These angles have corresponding maximum movement ranges (δ_α, $\delta_\beta \in [0°, 180°]$), depending on the agent's configuration. For an agent capable of rotation 360°, a maximum range of 180° will be used on either side, depending on which side is closer to the target.

Thus, the $\mathrm{ID}_{\mathrm{trans}}$ in a polar coordinate system can be represented as the sum of two $\mathrm{ID}_{\mathrm{rot}}$ components and one $\mathrm{ID}_{\mathrm{linear}}$ component

$$\begin{aligned}\mathrm{ID}_{\mathrm{trans}} &= \sum_{i\in\{\alpha,\beta\}} \mathrm{ID}_{\mathrm{rot}_i} + \mathrm{ID}_{\mathrm{linear}_r} \\ &= \sum_{i\in\{\alpha,\beta\}} \log_2\left(\frac{\theta_i}{\delta_i}+1\right) + \log_2\left(\frac{d_r}{w_{t_r} \pm w_{e_r}}+1\right)\end{aligned} \tag{11}$$

where the target and agent sizes (w_t and w_e) are represented by the diagonal widths on the contacting surface. For example, in locomotion tasks, this is the ground, and the size is the diagonal length of the vertical projection area. In manipulation tasks, where the contacting surface is the front of the end-effector, the size is measured on this contacting surface.

3) Orientation Index of Difficulty: The orientation ID ($\mathrm{ID}_{\mathrm{ori}}$) accounts for the target's approach tolerance (δ) and the difference from the optimal approach angle (θ), where the optimal angle corresponds to the target's facing direction. We divide δ and θ into yaw, pitch, and roll components, as shown in Fig. 2. Thus, the $\mathrm{ID}_{\mathrm{ori}}$ comprises three $\mathrm{ID}_{\mathrm{rot}}$ components, which can be calculated as

$$\mathrm{ID}_{\mathrm{ori}} = \sum_{i\in\{\mathrm{yaw,pitch,roll}\}} \mathrm{ID}_{\mathrm{rot}_i} = \sum_{i\in\{\mathrm{yaw,pitch,roll}\}} \log_2\left(\frac{\theta_i}{\delta_i} + 1\right). \tag{12}$$

4) Generalization of the Model: HMTs may exhibit different motion capabilities across various DoFs, and even within the same motion, the difficulty can vary depending on the direction or type of movement. To accommodate these variations and ensure that our model remains generalisable to a wide range of applications, we introduce weighting factors (k) for each component of the ID.

These weights reflect the intrinsic motion capabilities of the HMT across different DoFs, allowing us to adapt the model to account for any differences in difficulty encountered during translation or orientation. For instance, an HMT may perform differently in linear versus rotational movements, in which case different weights can be applied to the translational and orientational indices of difficulty. In this case, (7) is modified as follows:

$$\mathrm{ID} = \boldsymbol{k}_{\mathrm{trans}} \cdot \mathrm{ID}_{\mathrm{trans}} + \boldsymbol{k}_{\mathrm{ori}} \cdot \mathrm{ID}_{\mathrm{ori}} \tag{13}$$

where $\boldsymbol{k}_{\mathrm{trans}}$ and $\boldsymbol{k}_{\mathrm{ori}}$ are weight vectors representing the relative difficulty of translation and orientation, respectively, based on the HMT's specific motion capabilities.

By introducing these weighting factors, we ensure that the model can be tailored to different HMTs, accounting for their inherent strengths and limitations in various DoF. This makes the model more flexible and generalizable across a wide range of tasks and systems without compromising its predictive accuracy. For simplicity, in the article, we assume an evenly capable HMT, and therefore all DoFs share the same weights i.e., $k = 1$ for all DoFs.

5) Practical Task Modeling: In real-world applications, practical tasks are often composed of multiple subtasks. Regardless of the number of subtasks involved, as long as they are within the HMT's identified motion capability (see Section III-B), each practical task can be decomposed into individual subtasks, and each subtask's difficulty can be quantified using the weighted ID as described in (13).

Thus, for any feasible practical task, the overall ID for the proposed 3-D Fitts' law can be expressed as

$$\mathrm{ID}_{\mathrm{prac}} = \sum_{i=1}^{n} k_i \cdot \mathrm{ID}_i \tag{14}$$

where n is the number of subtasks in the practical task and k_i represents the weight applied to each subtask based on its specific characteristics and the HMT's motion capabilities.

This decomposition allows us to apply the proposed 3-D Fitts' law to a wide range of complex tasks, ensuring that the difficulty of each motion component is accounted for in relation to the HMT's capabilities. The result is a comprehensive model that can predict task performance even in complex, real-world scenarios involving multiple DoF and varying task difficulty.

IV. Validation

This section presents a comprehensive validation process for the proposed prediction model, including simulation and real-world experiments. The validation aims to assess the model's versatility and effectiveness, both in a simulation environment with a construction machinery agent and in real-world scenarios with a legged robot agent.

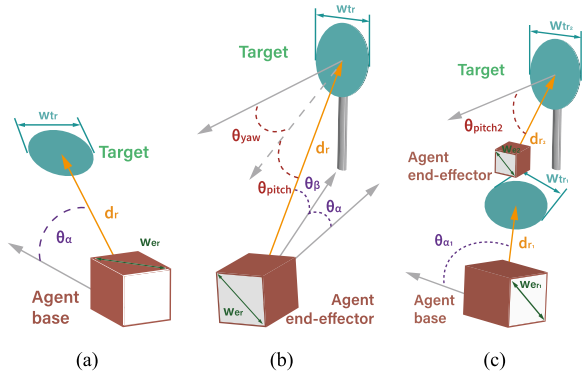


Fig. 3. Example of parameters for the selected platform in three required standard tasks: (a) locomotion task. (b) manipulation task. and (c) combined task. (a) T_{mob}. (b) T_{mani}. (c) T_{comb}.

The performance parameters of standard tasks were used to predict the HMT performance of an advanced multistep practical task, as depicted in Fig. 1. The predicted time result was compared with the actual measured value to evaluate the effectiveness of the proposed model.

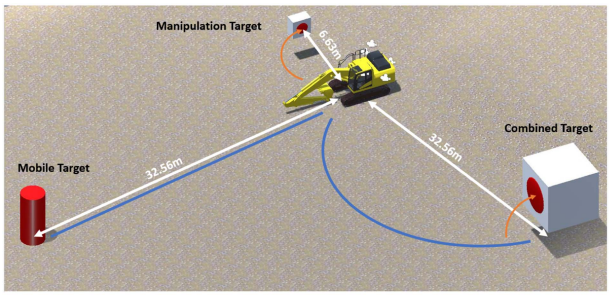
A. Design Principle

We select articulated machines and robots due to their broad applicability in industrial settings. Consequently, a polar coordinate system is employed to align with these applications through (11). In detail, we use a mobile base with an articulated arm as the agent platform. Such platforms benefit from merging locomotion with manipulation capability, allowing them to navigate diverse environments while manipulating objects. They are frequently used in various industries, such as construction, logistics, manufacturing, and service robotics. Specifically, we selected an excavator for the simulation and a legged mobile manipulator robot for the real-world experiment. Although these platforms share a common structure, their control strategies each have unique characteristics.

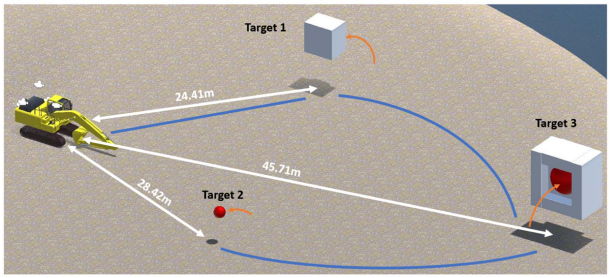
Such a platform has two motion capabilities ($m = 2$): one locomotion capability and one manipulation capability. According to (6), three standard tasks are needed ($p = 3$). The first standard task measures locomotion capability only, the second task assesses manipulation ability exclusively, and the third task integrates both capabilities. Any future tasks can be represented by combinations of these three standard tasks. The data from these tasks model the performance characteristics of the HMT in different motions, which were used to predict future practical tasks. Although we generalized our model in Section III-C4, to be conservative, the system is simplified to have the same weight ($k = 1$) for all tasks and motions.

B. Modeling Tasks

The three standard tasks in the selected mobile manipulation platform are locomotion task (T_{mob}), manipulation task (T_{mani}), and combined task (T_{comb}), as shown in Fig. 3. The practical



(a)



(b)

Fig. 4. Simulation environment setup. The white lines show the relative distance to the targets. The blue and orange lines show the potential locomotion and manipulation paths. (a) Standard task targets with markers indicating trajectories. (b) Practical task targets with markers representing trajectory of target order 1-3-2.

task (T_{prac}) is a complex explosive ordnance disposal (EOD) mission formed of a series of different subtasks.

1) Locomotion Task: This task necessitates only the locomotion capability. It contains only locomotion movement, as shown in Fig. 3(a). It has ID (ID_{mob}) calculated as (7).

2) Manipulation Task: This task utilizes only the manipulation capability and involves a single task, as shown in Fig. 3(b). Thus, its ID (ID_{mani}) can be calculated as (7) as well.

3) Combined Task: The last standard task requires both locomotion and manipulation capabilities, as shown in Fig. 3(c). As such, the total ID (ID_{comb}) is modelled as a sum of two subtasks ($n = 2$), as per (14)

$$ID_{comb} = ID_{comb_1} + ID_{comb_2}. \tag{15}$$

In the first step, the machine travels to the target location in subtask 1 (T_{comb_1}). In the second step, the machine employs the arm to touch the target in subtask 2 (T_{comb_2}), where the agent is expected to already stand directly in front of the target, with the direction angle and the target's best approach angle both set to zero ($\theta_\alpha = 0$ and $\theta_{yaw} = 0$).

4) Practical Task: As stated in Section III-C, all practical tasks that are feasible to the HMT can be decomposed into a combination of one or more standard tasks with (14). Therefore, a practical task (T_{prac}) has its ID (ID_{prac}) to be calculated by summation of these subtasks' difficulty.

C. Simulation

As depicted in Fig. 4, we have developed a web-based physical simulation of a teleoperated excavator. The excavator, one of the most commonly utilized HMTs in industrial applications due to its multiple DoFs, serves as an excellent testing ground for the proposed prediction method. The excavator has two motion capabilities: locomotion and manipulation. As per Section IV-A, the first standard task solely measures the locomotion capability, the second evaluates the manipulation capability, and the third combines both motions.

1) Simulation Design: Participants engaged in the study by playing the web-based excavator physical simulation. This identical simulation is hosted online via the WebGL application programming interface (API), allowing the collection of data from participants worldwide. Once the simulation has fully loaded, the WebGL API operates on local computer hardware and does not require or depend on an internet connection until the final result is submitted. This feature removes the potential disturbances caused by network instability, while still enabling participants to access the simulation remotely.

The simulation comprises three standard tasks (T^{sim}_{mob}, T^{sim}_{mani}, and T^{sim}_{comb}) and a practical task (T^{sim}_{prac}) with multiple subtasks, with parameters displayed in Table I. Since the excavator cannot move its bucket in the roll direction, there is no difficulty related to roll motion ($ID^{roll}_{ori} = 0$). The standard tasks require the excavator to touch corresponding targets, as shown in Fig. 4(a). In T^{sim}_{mob}, the distance is from the center of the excavator to the centre of the target, and the target can be approached from any direction ($\delta_{yaw} = 360$, and $\theta_{yaw} = 0$). Also, the ground is a flat surface, which does not have an orientation component in pitch direction ($\theta_{pitch} = 0$) for all locomotion tasks and subtasks. In T^{sim}_{mani}, the distance is from the end-effector of the excavator's arm to the centre of the target. To control variable, the interface is locked to the corresponding control only in T^{sim}_{mob} and T^{sim}_{mani}.

The T^{sim}_{prac} compromises three subtasks with three targets, each requiring a combination of locomotion and manipulation motions, as depicted in Fig. 4(b). To enhance generalizability, maintain user engagement, and avoid bias, participants complete the three subtasks in the randomized order of target appearance. This results in task difficulties ranging from 12.29 (Order: Target $2 \Rightarrow 3 \Rightarrow 1$) to 16.05 (Order: Target $3 \Rightarrow 2 \Rightarrow 1$). Parameters from the start point to all three targets are given in Table I.

The system automatically records the duration of each task. Upon completing all tasks, the participants submit their record data along with their background information to a cloud server. To encourage participation, the simulation is also equipped with a leaderboard. The recorded data of the selected 9 participants revealed that the randomly generated practical tasks had an average ID of 11.53.

2) Participants: A total of 16 worldwide participants' data were collected for the study. Among these, 9 participants (2 female, 5 male, and 2 unreported) completed the simulation tasks within a reasonable time frame without failure. These participants' ages ranged from 21 to 29 years old (mean age: 25.8 years; SD: 2.4 years), and they reported an average computer game experience of 7.6 (on a scale of 0 to 9).

D. Experiment

We also conducted experiments in the real world. A separate group of participants performed the designed tasks using a

TABLE I
PARAMETERS OF TASKS AND THEIR SUBTASKS WITH CALCULATED ID, WHERE d AND w IN METERS, θ AND δ IN DEGREES

		Distance	Target size	Agent size	Direction horizontal	Direction vertical	Tolerance yaw	Best angle yaw	Tolerance pitch	Best angle pitch	Task difficulty
		d	w_t	w_e	θ_α	θ_β	δ_{yaw}	θ_{yaw}	δ_{pitch}	θ_{pitch}	**ID**
Simulation	$T^{\text{sim}}_{\text{mob}}$	32.56	2	10.44	10.62	0	360	0	-	0	1.94
	$T^{\text{sim}}_{\text{mani}}$	6.63	1	1.41	71.57	45	180	18.43	180	45	3.44
	$T^{\text{sim}}_{\text{comb}_{\text{mob}}}$	25.18	3	10.44	80.84	0	180	83.16	-	0	4.90
	$T^{\text{sim}}_{\text{comb}_{\text{mani}}}$	5	3	1.41	0	63.43	180	0	180	63.43	
	$T^{\text{sim}}_{\text{prac}_{\text{taeget1.mob}}}$	24.41	2	10.44	55.01	0	180	124.99	-	0	avg:13.42 min:12.29 max:16.05
	$T^{\text{sim}}_{\text{prac}_{\text{taeget1.mani}}}$	2	2	1.41	0	63.43	180	0	180	63.43	
	$T^{\text{sim}}_{\text{prac}_{\text{taeget2.mob}}}$	28.42	1	10.44	39.29	0	360	0	-	0	
	$T^{\text{sim}}_{\text{prac}_{\text{taeget2.mani}}}$	0	1	1.41	0	0	360	0	360	0	
	$T^{\text{sim}}_{\text{prac}_{\text{taeget3.mob}}}$	45.71	3	10.44	10.08	0	120	79.92	-	0	
	$T^{\text{sim}}_{\text{prac}_{\text{taeget3.mani}}}$	2.86	3	1.41	0	70.73	120	0	120	70.73	
Experiment	$T^{\text{exp}}_{\text{mob}}$	2	0.1	0.69	45	0	360	0	-	0	2.14
	$T^{\text{exp}}_{\text{mani}}$	0.75	0.20	0.13	45	75	180	45	180	75	3.37
	$T^{\text{exp}}_{\text{comb}_{\text{mob}}}$	1.12	0.20	0	70	0	360	20	-	0	6.00
	$T^{\text{exp}}_{\text{comb}_{\text{mani}}}$	0.75	0.20	0.13	0	75	180	0	180	75	
	$T^{\text{exp}}_{\text{prac}_{\text{step1}}}$	1.45	0.14	0	5	0	90	45	-	0	13.63
	$T^{\text{exp}}_{\text{prac}_{\text{step2}}}$	0.75	0.14	0.06	0	10	90	0	90	10	
	$T^{\text{exp}}_{\text{prac}_{\text{step3}}}$	0.75	0.08	0.06	0	10	60	0	60	10	

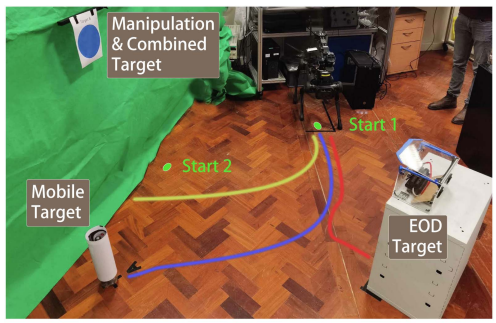
Also, $\delta_\beta = 90°$ for simulation, $\delta_\beta = 180°$ for experiment, $\delta_\alpha = 180°$ for both validation.

quadruped manipulator robot agent through two interfaces: a gamepad and a wearable motion capture suit (WMCS).

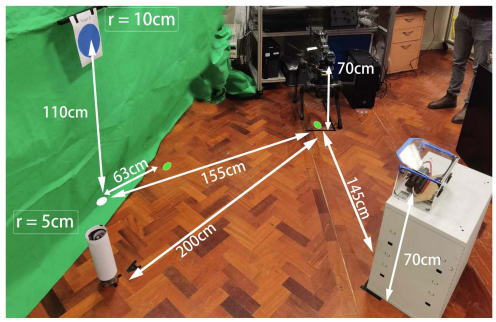
1) Experiment Design: Based on the principle from Section IV-B, three standard tasks ($T^{\text{exp}}_{\text{mob}}$, $T^{\text{exp}}_{\text{mani}}$, and $T^{\text{exp}}_{\text{comb}}$) and one practical task ($T^{\text{exp}}_{\text{prac}}$) are designed for the experiment, with parameters in Table I. Since the robot cannot move its end-effector in the roll direction, there is no difficulty related to roll motion ($ID^{\text{roll}}_{\text{ori}} = 0$). In $T^{\text{exp}}_{\text{mob}}$, $T^{\text{exp}}_{\text{comb}}$, and $T^{\text{exp}}_{\text{prac}}$, the robot start from "Start 1", and in $T^{\text{exp}}_{\text{mani}}$, the robot start from "Start 2". Then, the robot completes tasks by touching corresponding targets, as shown in Fig. 5. In detail, the practical task ($T^{\text{exp}}_{\text{prac}}$) simulates a bomb disposal task on target "EOD" with three subtasks. First, the robot walks towards the target "EOD". After reaching the target, it uses its end-effector to open a box and then pulls out a red wire on the "bomb" to disable it.

Similar to simulation, in $T^{\text{exp}}_{\text{mob}}$, the target can be approach from any direction ($\delta_{\text{yaw}} = 360$, and $\theta_{\text{yaw}} = 0$). The ground is flat with no orientation component in pitch direction ($\theta_{\text{pitch}} = 0$) for all locomotion motion tasks and subtasks. However, unlike the simulation, the robot operator in the experiment needs to avoid collision between the robot and some of the hard targets. Therefore, the robot is required to stop at a safe distance from the target during locomotion subtasks ($T^{\text{exp}}_{\text{comb}_{\text{mob}}}$ and $T^{\text{exp}}_{\text{prac}_{\text{step1}}}$) therefore, the robot is treated as a point ($w_e = 0$) and needs to stop at the designated safe locations.

2) Experiment Hardware: The quadruped manipulator robot used in this experiment consists of a Unitree AlienGo quadruped robot as mobile base, weight 21.5 kg, and a modified ViperX 300 robot arm, weight 2.5 kg, as the manipulator [30]. Participants



(a)



(b)

Fig. 5. Experiment setup: In (a), blue, yellow, and red trajectories represent the locomotion path of $T^{\text{exp}}_{\text{mob}}$, $T^{\text{exp}}_{\text{comb}}$, and $T^{\text{exp}}_{\text{prac}}$. In (b), the white lines show the relative positions. (a) Target and potential paths. (b) Paramater of the setup.

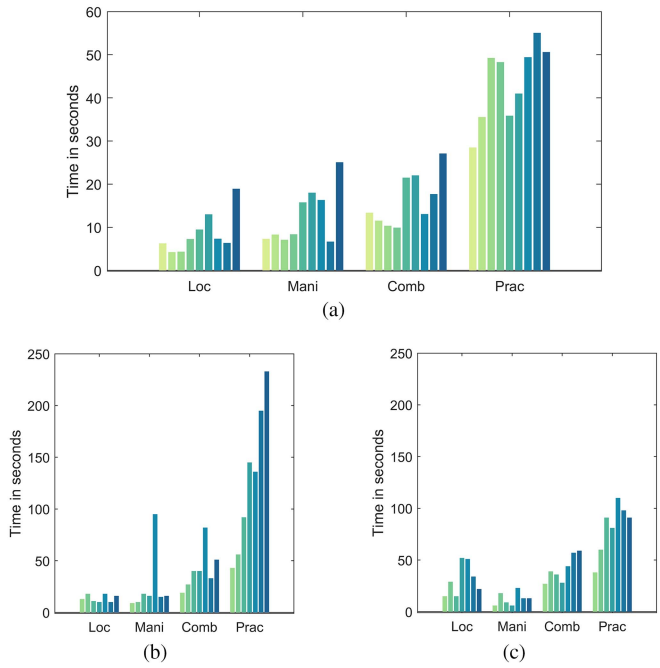


Fig. 6. Motion time the users took to complete the standard tasks and practical tasks in (a) simulation and in the experiment with the (b) gamepad and (c) WMCS, where each bar represents time for a participant. (a) Simulation. (b) Gamepad. (c) Wearable motion capture suit.

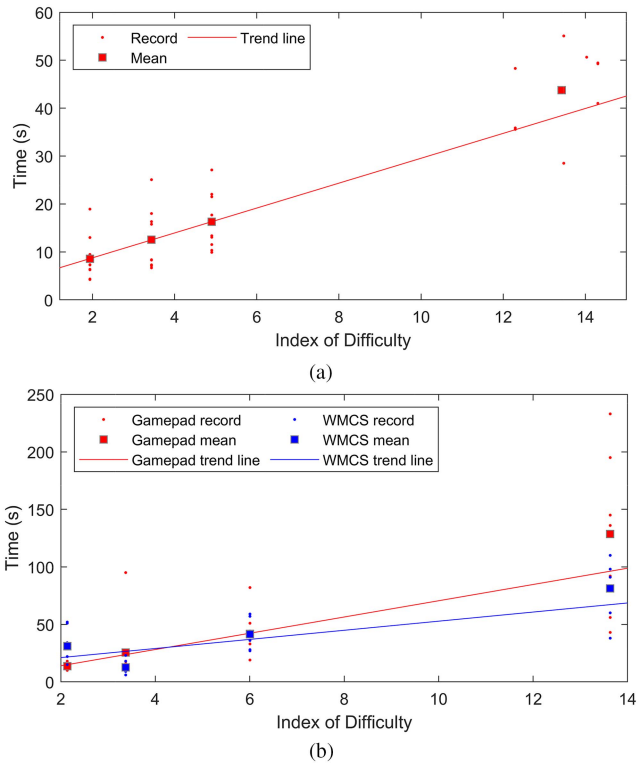


Fig. 7. Motion time of users took to complete each task versus different ID value from Table I, and the fitted prediction line to average motion time of standard tasks. (a) Simulation. (b) Experiment.

controlled the robot using a Logitech F710 wireless gamepad and a noitom perception neuron WMCS. To maintain consistency across control mechanisms, both interfaces utilized the same set of teleoperation strategies. The robot arm's manipulation is under position control, while the base quadruped robot's locomotion is under velocity control. For detailed information on teleoperation strategies, refer to [25].

3) Participants: To avoid bias originating from the characteristics and background of the individual operators, we incorporated multiple participants into the experiments. This approach better reflects the HMT performance within the targeted user group. A total of 7 participants (3 female, 4 male) volunteered for the experiment. Their ages ranged from 21 to 28 years (mean age: 26.3 years; standard deviation: 4.2), and they reported an average robot experience of 3.5 (on a scale of 0 to 9). Each participant received only basic instruction without practice to mitigate bias in the learning curves. In addition, the order of operation for the two interfaces was randomized.

V. Results

The 3-D Fitts' law predicts system performance based on the relationship between motion time and the complexity of the motion. We analyzed the modelled task difficulty and measured task time using MATLAB and Microsoft Excel. The ID for all tasks have been computed in Table I. The motion time for each trial was also gathered via automatic recording during the simulation and video recording for the experiment, as demonstrated in Fig. 6. For simulation, data from 9 users who completed tasks in a reasonable time were selected to test the proposed system, as illustrated in Fig. 6(a). In the experiment, 7 volunteers who completed all four tasks with both interfaces were selected, as shown in Fig. 6(b) and (c). To verify the accuracy of the proposed prediction model, we utilized the collected performance of standard tasks to forecast the performance of practical tasks.

TABLE II
Parameters in the Prediction Lines

Constant	Simulation	Experiment	
		GP	WMCS
a	3.57	−0.12	13.18
b	2.60	7.07	3.96
RMSE	0.06	2.32	16.61

A. Prediction

The average measured motion time, along with the ID for each task from Table I for both the simulation and experiment, have been plotted in Fig. 7. The relationship between task time and task difficulty can be expressed through a fitted prediction line, represented by (1), with constant and root-mean-square deviation (RMSE) found in Table II.

B. Verification

The prediction line is used to predict the time of practical tasks in both the simulation and experiment. In the simulation, the predicted time for the practical task is 38.42 s with the average ID of 13.42, and the measured average time to complete the

practical task is 43.7 s, which results in a prediction error of 12% . In the experiment, the predicted time for the practical task with the ID of 13.63 for the gamepad and WMCS controls are 96.19 s and 67.16 s, and the measured average times with the gamepad and WMCS are 128.6 s and 81.3 s, respectively. The prediction errors are 25% for the gamepad control and 17% for the WMCS control.

To assess how our model differs from earlier Fitts-based approaches, we also tested previous popular methods under the same task conditions of simulation and experiment. The results show that these prior methods exhibit significantly higher error rates, with the [29] method averaging 168%, the [21] method 82%, and the [27] model 72% . The primary reason for their failure in accurate modeling and prediction lies in their cursor-based design, as highlighted in Section II. This underscores the necessity of our method in robot and machinery applications.

VI. Discussion

The results of our study affirm the efficacy of the proposed method in predicting HMT performance with an error margin within 25% . This level of accuracy is consistent across both heavy machinery simulations and real-world robot experiments, underlining the robustness and wide applicability of our model.

An advantage of the proposed method is its ability to consider the HMT system as a whole, where all related factors are reflected in standard tasks. For instance, in our experiments, the gamepad exhibited higher variability in manipulation tasks, as shown in Fig. 6. Participants reported difficulties in achieving fine control with the gamepad, yet the WMCS consistently outperformed the gamepad even without prior experience. This suggests that while certain parameters, such as interface familiarity, are challenging to quantify manually, they are effectively captured within our model.

The RMSE values detailed in Table II indicate that the prediction line correlates more closely with simulation data than with experimental data. This discrepancy can primarily be attributed to increased disturbances and environmental uncertainties inherent in real-world settings. For instance, the stopping distances in robot locomotion tasks varied with changes in the legged robot's center of mass due to different arm positions. This variability also highlights the sensitivity of the prediction model to the physical machine configuration of the HMT system, suggesting that different base types (legged versus wheeled) might yield divergent prediction results.

Variations in participants' experience with the interfaces used also significantly influenced the results. Participants without prior experience with gamepads typically demonstrated lower performance. This suggests that in HMT design, different interfaces may be employed for different user groups.

An interesting observation was that the predicted motion time for both the simulation and experiment was consistently shorter than the actual time users took. This discrepancy is likely due to the learning curve effect, where participants became increasingly familiar with the machine system and environment through repeated practice of standard tasks, thereby enhancing their performance in practical tasks.

Feedback from participants also highlighted limitations in the camera view in simulation, which restricted their ability to accurately judge distances to objects, affecting performance. This underscores the potential oversight in the design of visual feedback mechanisms within HMT systems, suggesting that enhancements in camera technology and interface design could improve user experience and system accuracy.

VII. Conclusion

Evaluating and predicting the performance of HMTs has wide-ranging applications and potential, particularly for HMTs operating in open environments. Our study successfully demonstrates the practical applicability and accuracy of the proposed 3-D Fitts' law-based model for predicting HMT performance. The model's efficacy is empirically validated through simulation and experiments. These results validate the model's capability to accurately predict the performance of complex, multistep HMT tasks. The findings from our study open new pathways in HMT performance prediction and offer promising implications for real-world applications in diverse sectors, where human–machine collaboration is essential.

Our future research will extend the current linear framework to a nonlinear model, which is crucial in scenarios where human physical and cognitive fatigue influences performance during high-stress missions. Besides manually modeling human fatigue, artificial intelligence can be used to consider more complex human factors to provide a more accurate prediction in the real world. On the other hand, the quantified results of the proposed modeling can be leveraged by AI algorithms to tune parameters in human–machine interface design, as well as to refine task setups based on predictive feedback.

Moreover, the scope of our research will expand to encompass fine manipulation manoeuvres post-target contact. These include detailed tasks, such as opening a box, rotating a handle, or relocating a bottle. Incorporating these advanced manoeuvres into our model will enhance its predictive accuracy for complex tasks and offer insights into how human factors, such as fatigue, can impact task execution in HMT systems.

References

[1] Y. Ren and G.-P. Li, "An interactive and adaptive learning cyber physical human system for manufacturing with a case study in worker machine interactions," *IEEE Trans. Ind. Inform.*, vol. 18, no. 10, pp. 6723–6732, Oct. 2022.

[2] M. Xu, A. Hu, and H. Wang, "Visual-impedance-based human–robot cotransportation with a tethered aerial vehicle," *IEEE Trans. Ind. Inform.*, vol. 19, no. 10, pp. 10356–10365, Oct. 2023.

[3] C.-Y. Weng, Q. Yuan, F. Suárez-Ruiz, and I. -M. Chen, "A telemanipulation-based human–robot collaboration method to teach aerospace masking skills," *IEEE Trans. Ind. Inform.*, vol. 16, no. 5, pp. 3076–3084, May 2020.

[4] C. D. Bellicoso et al., "Advances in real-world applications for legged robots," *J. Field Robot.*, vol. 35, no. 8, pp. 1311–1326, 2018.

[5] P. Damacharla, A. Y. Javaid, J. J. Gallimore, and V. K. Devabhaktuni, "Common metrics to benchmark human-machine teams: A review," *IEEE Access*, vol. 6, pp. 38637–38655, 2018.

[6] Y. Wan and C. Zhou, "Predicting human-robot team performance based on cognitive fatigue," in *Proc. Int. Conf. Automat. Comput.*, 2023, pp. 575–580.

[7] C. R. Kube, "Task modelling in collective robotics," *Auton. Robots*, vol. 4, pp. 53–72, 1997.

[8] D. Y. Y. Sim and C. K. Loo, "Extensive assessment and evaluation methodologies on assistive social robots for modelling human–robot interaction–A review," *Inf. Sci.*, vol. 301, pp. 305–344, 2015.
[9] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, 1954, Art. no. 381.
[10] Y. Cha and R. Myung, "Extended Fitts' law for 3D pointing tasks using 3D target arrangements," *Int. J. Ind. Ergon.*, vol. 43, no. 4, pp. 350–355, 2013.
[11] L. D. Clark et al., "Extending fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology," *Int. J. Hum.- Comput. Stud.*, vol. 139, 2020, Art. no. 102413.
[12] B. Rohrer, "A developmental agent for learning features, environment models, and general robotics tasks," in *Proc. Front. Comput. Neurosci. Conf.*, 2011, doi: 10.3389/conf.fncom.2011.52.00001.
[13] A. Hasselberg and D. Söffker, "Petri-net-based modeling of human operator's planning for the evaluation of task performance using the example of air traffic control," *IEEE Trans Hum Mach Syst*, vol. 45, no. 6, pp. 676–685, Dec. 2015.
[14] M. Freed and R. Remington, "A conceptual framework for predicting error in complex human-machine environments," in *Proc. Annu. Meeting Cogn. Sci. Soc.*, 2022, pp. 356–361.
[15] J. Crandall, M. A. Goodrich, D. R. Olsen, and C. W. Nielsen, "Validating human-robot interaction schemes in multitasking environments," *IEEE Trans. Syst. Man Cybern. A Syst. Hum.*, vol. 35, no. 4, pp. 438–449, Jul. 2005.
[16] A. M. Zanchettin, A. Casalino, L. Piroddi, and P. Rocco, "Prediction of human activity patterns for human–robot collaborative assembly tasks," *IEEE Trans. Ind. Inform.*, vol. 15, no. 7, pp. 3934–3942, Jul. 2019.
[17] A. H. Memar and E. T. Esfahani, "Physiological measures for human performance analysis in human-robot teamwork: Case of tele-exploration," *IEEE Access*, vol. 6, pp. 3694–3705, 2018.
[18] H. Drewes, "Only one Fitts' law formula please!," in *Proc. SIGCHI Conf. Hum. Factor Comput. Syst*, 2010, pp. 2813–2822.
[19] A. T. Welford, *Fundamentals of Skill*. England, U.K: Methuen, 1968.
[20] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Hum.-Comput. Interact.*, vol. 7, pp. 91–139, 1992.
[21] M. F. Stoelen and D. L. Akin, "Assessment of Fitts' law for quantifying combined rotational and translational movements," *Hum. Factors*, vol. 52, no. 1, pp. 63–77, 2010.
[22] C. Radix et al., "Extension of fitts' law to modeling motion performance in man-machine interfaces," *IEEE Trans. Syst. Man Cybern. A Syst. Hum.*, vol. 29, no. 2, pp. 205–209, Mar. 1999.
[23] T. Petrič et al., "Cooperative human-robot control based on Fitts' law," in *Proc. IEEE Int. Conf. Humanoid Robots*, 2016, pp. 345–350.
[24] S. Sutjipto et al., "Fitts' law in the presence of interface inertia," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2020, pp. 4749–4752.
[25] Y. Wan et al., "Performance and usability evaluation scheme for mobile manipulator teleoperation," *IEEE Trans Hum Mach Syst*, vol. 53, no. 5, pp. 844–854, Oct. 2023.
[26] E. Triantafyllidis and Z. Li, "The challenges in modeling human performance in 3D space with Fitts' law," in *Proc. SIGCHI Conf. Hum. Factor Comput. Syst.*, 2021, pp. 1–9.
[27] A. Kulik, A. Kunert, and B. Froehlich, "On motor performance in virtual 3D object manipulation," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 5, pp. 2041–2050, May 2020.
[28] M. D. Barrera Machuca and W. Stuerzlinger, "The effect of stereo display deficiencies on virtual hand pointing," in *Proc. SIGCHI Conf. Hum. Factor Comput. Syst.*, 2019, pp. 1–14.
[29] C. Ware and R. Balakrishnan, "Reaching for objects in VR displays: Lag and frame rate," *ACM Trans Comput Hum Interact*, vol. 1, no. 4, pp. 331–356, 1994.
[30] C. Zhou et al., "Advancing teleoperation for legged manipulation with wearable motion capture," *Front. Robot. AI*, vol. 11, 2024, Art. no. 1430842.

Yuhui Wan received the M.S. degree in mechanical engineering from Purdue University, West Lafayette, Indiana. He is currently working toward the Ph.D. degree with the School of Mechanical Engineering, University of Leeds, Leeds, U.K.

His research interests include shared autonomy and embodied intelligence.

Chengxu Zhou received the Ph.D. degree in robotics from the Italian Institute of Technology, Genoa, Italy, in 2016.

He is an Associate Professor of Robotics and AI with the Department of Computer Science, University College London, London, U.K. His research interests include intelligent motion generation for legged robots, with particular emphasis on dynamic motion planning and whole-body control of articulated systems, leveraging optimization and machine learning techniques.