Automation in Construction 168 (2024) 105741

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Intelligent ergonomic optimization in bimanual worker-robot interaction: A Reinforcement Learning approach

Mani Amani [a], Reza Akhavian [a,b,*]

[a] Computational Science Research Center, San Diego State University, 5500 Campanile Drive, San Diego, 92182, CA, USA
[b] Department of Civil, Construction, and Environmental Engineering, San Diego State University, 5500 Campanile Drive, San Diego, 92182, CA, USA

ARTICLE INFO

Keywords:
Reinforcement Learning
Bimanual robot object handover
REBA
Robotics
Virtual Reality

ABSTRACT

Robots have the potential to enhance safety on construction job sites by assuming hazardous tasks. While existing safety research on physical human-robot interaction (pHRI) primarily addresses collision risks, ensuring inherently safe collaborative workflows is equally important. For example, ergonomic optimization in co-manipulation is an important safety consideration in pHRI. While frameworks such as Rapid Entire Body Assessment (REBA) have been an industry standard for these interventions, their lack of a rigorous mathematical structure poses challenges for using them with optimization algorithms. Previous works have tackled this gap by developing approximations or statistical approaches that are error-prone or data-dependent. This paper presents a framework using Reinforcement Learning for precise ergonomic optimization that generalizes to different types of tasks. To ensure practicality and safe experimentations, the training leverages Inverse Kinematics in virtual reality to simulate human movement mechanics. Results of a comparison between the developed framework and ergonomically naive approaches are presented.

1. Introduction

Production and assembly work in industries such as construction still relies heavily on human workers, causing a decrease in productivity and safety, in part due to physical stress and ergonomic tension [1]. Robotics development and research in construction has been a heavily evolving field in recent years, focusing on improving productivity, quality, and safety. However, physical human–robot interaction (pHRI) in construction, particularly with robots having a manipulator robotic arm has received very little attention in such contexts [2]. This is while other industries such as manufacturing have made significant strides in safe pHRI recently [3]. Methods such as the Rapid Entire Body Assessment (REBA) [4], Rapid Upper Limb Assessment (RULA) [5], and Ovako Working Posture Analysis (OWAS) [6] are industry standards for measuring and assessing the ergonomic safety of tasks to prevent Work-related Musculoskeletal Disorders (WMSDs) in fieldwork contexts. Much pHRI research also uses these industrial frameworks as a means to account for human safety in collaborative tasks [7].

Ergonomic optimization is a critically important safety research topic in pHRI. With the widespread adoption of sensors and computers, many researchers have attempted to create a safer work process and pHRI using algorithmic approaches that benefit from this enhanced connectivity and computational power. Studies have shown that collaborative robots have the ability to significantly enhance workplace safety and efficiency by supporting physically demanding tasks and reducing biomechanical stress on workers [8]. Furthermore, these ergonomic pHRI collaborations not only help with operator safety but also improve production efficiency [9]. Moreover, ergonomic optimizations are increasingly being applied to the development of exoskeletons to enhance human comfort and efficiency. This approach aims to reduce residual forces and improve usability [10]. These developments underscore the importance and potential of ergonomic optimization in HRI for improving worker safety and productivity.

REBA and RULA work based on a scoring system using predefined body angles and postures to determine the ergonomic risk to the participant given the activity. The range for these scores typically spans discrete positive integers. This scoring regimen creates an easy, fast, and transparent method of manually gauging the ergonomic safety of a task. However, this discreteness makes the scores non-differentiable, making mathematical optimization a difficult task for automated approaches in pHRI tasks. Another challenge in optimization for ergonomics in pHRI is the dependency of the approach on specific physical tasks. Traditional analytical and approximate solutions must account for the varying dynamics of the object and task, requiring frequent reevaluation and adjustments as tasks change. To overcome this, model-free approaches such as Reinforcement Learning (RL) are attractive options. However, even though this approach offers

* Corresponding author at: Computational Science Research Center, San Diego State University, 5500 Campanile Drive, San Diego, 92182, CA, USA.
E-mail address: rakhavian@sdsu.edu (R. Akhavian).

https://doi.org/10.1016/j.autcon.2024.105741
Received 22 March 2024; Received in revised form 29 August 2024; Accepted 30 August 2024
Available online 17 September 2024

task-independent optimization through statistical methods, its need for unique calibration for each task and participant poses practical challenges for real-world application.

1.1. Previous approaches

To overcome this obstacle, previous works have developed a wide array of techniques. Some researchers tackle this issue by collecting data from a pHRI task and then augmenting the dataset using methods such as regression, which closely smoothens the distribution while still maintaining the raw score reading [11]. This differentiable method can then be used to optimize ergonomic safety in tasks such as object co-manipulation in HRI by defining and optimizing the dynamics of the task. Other studies aim to create predictive techniques such as quadratic approximations [12]. Non-gradient-based optimization algorithms such as genetic algorithms have also proved to help improve ergonomic scores in HRI object handover tasks [13]. However, most current research has been focused on small and light objects through one-handed handover tasks. Many of these works also focus on isolated aspects of object handover, such as release time or correction during co-manipulation. These approaches mainly require human data collection and the use of previous task-specific data to model and optimize, while not being focused on practical applicability for the fieldworkers in a cluttered environment such as those seen in construction sites. The existing challenges for RL training stem from two main issues:

1. Designing accurate reward functions for REBA scores is challenging due to their discrete and bounded nature, making it difficult to account for small differences.
2. Real-world training with RL is impractical because it requires a robot and live readings of human limb and body angles for state iterations and reward function calculations.

Additionally, rapidly advancing in safety training and evaluation and robotics research, Virtual Reality (VR) tools offer innovative ways to study and optimize pHRI without the constraints of the physical world [14]. Current Inverse Kinematics (IK) developments have created high-fidelity pose and shape estimation of a parametric humanoid model [15]. Unity is a versatile game engine that supports robotics simulation and integration through the ROS# plugin [16]. Additionally, packages like Unity ML-Agents [17] provide a robust framework for conducting robotics simulations within Unity, which could be extended to pHRI testing and workforce training. There have been several previous works that use Unity as a tool for pHRI ergonomic optimization [18] and worker ergonomic evaluation in VR settings has already been a widely researched topic in the field of construction [19].

1.2. The developed solutions and contributions

RL operates by assessing the current state to determine an action, with the action's quality judged by a reward function. The learner iteratively selects actions based on a predefined action-selection policy, guided by the reward system. To address the challenges identified above, we propose the implementation of an RL algorithm designed to pinpoint the optimal location for bimanual object handover tasks. A high-fidelity 3D VR environment has been developed to simulate the dynamics of pHRI in human–robot object handover scenarios with great detail. With an RL approach, we create a task-agnostic collaborative robot training routine that as long as human kinematics are representative of real-life actions, the model can find optimal locations for personalized object handover. Another advantage of this methodology is that the need for differentiable scores has been eliminated; the RL algorithm will get the live REBA scores from the worker's digital twin inside the simulation and optimize for the ideal configuration. Once the learner has converged to the optimal position, the framework is ready to apply this in real-time with a simple transformation of dimensions and coordinates. Towards accomplishing the goals of this work, the main difficulties were to accurately simulate human kinematics, speeding conversion times and augmenting REBA in order to suit the needs of pHRI optimization. As such, the main contributions of this paper are as follows:

1. Designing a novel task-independent method of finding the best ergonomic location for object placement in bi-manual pHRI object handover in an industrial setting utilizing a fast-converging RL routine.
2. Presenting a VR framework to train a model for any given task tailored to the physical attributes of the worker collaborator before the model's deployment.
3. Introducing alternative approaches and circumventions towards ergonomic optimization derived from currently established frameworks for further validity.

2. Methodology

2.1. Q-learning

Q-Learning is a model-free RL algorithm for any finite Markov Decision Process (MDP) [20]. The independence of Q-Learning from mathematical models of the environment or the specific problem at hand makes it particularly suitable for HRI object handover ergonomic optimization. This eliminates the need for modeling dynamics or approximations that can have the potential for inaccuracies. The foundation of the model is based on the Bellman Equation [21], as expressed by Eq. (1). Where s is the state, a is the action, $R(s,a)$ is the reward, γ is the discount factor, $P(s' \mid s,a)$ represents the transition probabilities and $V(s')$ is the value of the next state.

$$V(s) = \max_a \left(R(s,a) + \gamma \sum_{s'} P(s'|s,a)V(s') \right) \tag{1}$$

The Q-Learning equation is as follows:

$$Q_{\text{new}}(s,a) = (1-\alpha)Q_{\text{Current}}(s,a) + \alpha \left(r + \gamma \max_{a'} Q(s',a') \right) \tag{2}$$

where α is the learning rate and $Q(s,a)$ represents the Q-value of the state. The Q-value illustrates the quality of the action that is taken given the current state. The action is defined by a previously defined action policy such as epsilon-greedy [22] or softmax [23] action policies.

The softmax action policy chooses an action using Boltzmann distribution and explores all possible actions in every visited state. The alternative candidate actions to be chosen are ranked according to their value estimation. It chooses action a on the tth iteration with the probability given by:

$$\xi(t) = \frac{e^{\frac{Q_i(t)}{\tau}}}{\sum_k e^{\frac{Q_k(t)}{\tau}}} \tag{3}$$

where τ is the temperature [24]. As $\tau \to 0$, the policy chooses exploitation over exploration, since the actions with lower values have a higher opportunity to be selected, and a higher τ tends to pick actions with a higher value. We chose to keep a high τ since the reward mechanism is based on a discrete score and it is harder to differentiate action qualities to a significant extent. We decrease the temperature as the score decreases past the score threshold of 5 in a step-wise fashion by 0.1 units, and increase τ by 0.1 in the case that the score worsens.

Since our goal is to find the most ergonomically ideal final position for the object, we choose to apply actions directly to the object's position. The IK of the humanoid model then responds to this change in position. The ergonomic score of the position is calculated and used as a variable in the reward function. This technique circumvents the need to apply actions to the body itself. Instead, we rely on sufficiently developed and customizable IK solutions to simulate the required motions. The fidelity of these limb and body adjustments will depend on the specific IK configuration and object geometry that is used.

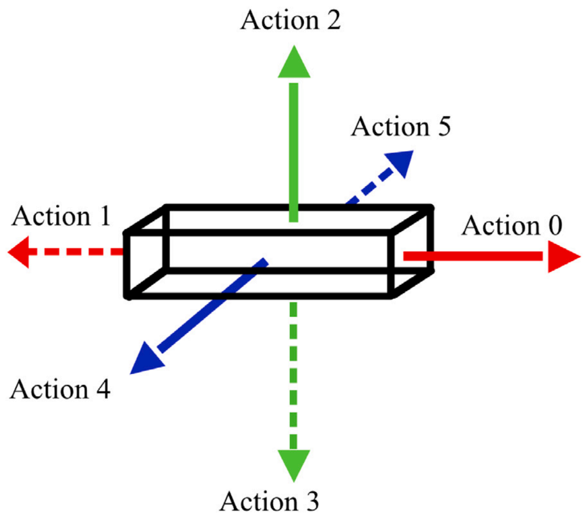

Fig. 1. Depiction of the Set of Actions Being Performed on The Object.

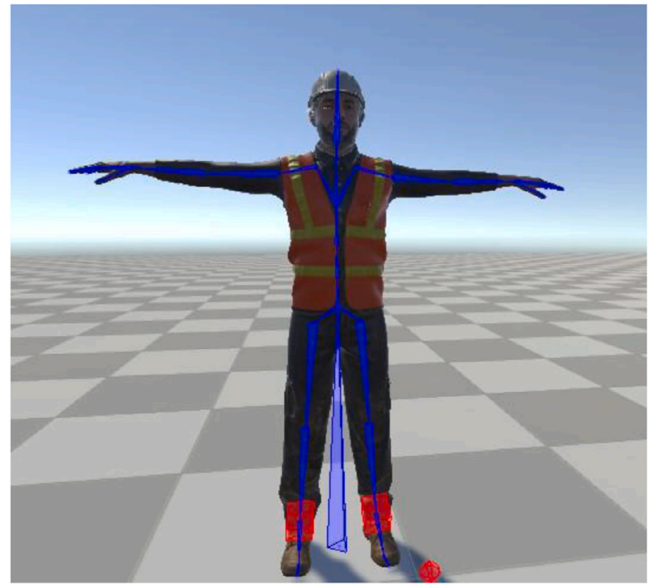

Fig. 2. Rigging of the 3D worker humanoid model.

In the case of this study, the learner's action set has 6 DOF with an action step size value of $\delta = 0.003$ (the equivalent of 0.3 cm), moving in a 3D space governed by Eq. (4) as shown in Fig. 1. The step size can also determine the speed and the quality of the convergence, depending on the nature of the distribution. The larger the step size the greater the agent moves towards the minimum. However, the global minimum can be inaccessible if the step size is not small enough to reach the desired position. Conversely, if the step size is small, the speed of convergence can be slow. However, there is a higher probability of reaching a global minimum depending on the nature of the distribution. In this particular scenario given body lengths and object dynamics, the step size of 0.3 cm provides us a great balance between speed and precision.

$$f_{\text{action}}(x, y, z, \delta, \text{action}) = \begin{cases} (x+\delta, y, z) & \text{if action } = 0 \\ (x-\delta, y, z) & \text{if action } = 1 \\ (x, y+\delta, z) & \text{if action } = 2 \\ (x, y-\delta, z) & \text{if action } = 3 \\ (x, y, z+\delta) & \text{if action } = 4 \\ (x, y, z-\delta) & \text{if action } = 5 \end{cases} \tag{4}$$

As the object moves in the direction that the chosen action is assigned, the humanoid model uses its hands in an attempt to receive the object for handover. At each state, the REBA score of the humanoid model is calculated and it is recorded. The current state's REBA reading is then used to calculate the reward of that state given that action and the RL model continues its process. In order to simulate the kinematics of the humanoid model in this scheme, we have developed a custom IK module.

2.2. VR and inverse kinematics

Calculating REBA scores and ergonomic assessment inside a VR environment has been the subject of previous studies [25]. Rigging is a procedure used in skeletal animation for representing a 3D character model using a series of interconnected digital bones [26]. This technique enables us to reproduce human kinematics at a high level. This framework will take the body measurements such as limb proportions, height, and width to generate a humanoid 3D model. While there is software that generates 3D models automatically, the model can be created manually using modeling software such as 3D Max and Blender. Once the model is created, we can rig the model to ensure accurate kinematics and mechanics to represent the worker in a virtual environment.

Previous works have also used Unity as a robotics engine as a means to develop robotic control and perception [27]. Our framework also uses the same philosophy in using Unity as a tool for both robotic control and visual representation of pHRI collaboration for increased transparency and multi-disciplinary accessibility. The most recent updates and developments have made this technique viable for seamlessly integrating robotics and game engines [28]. For our purposes, we utilize the Navigation heuristics, IK solvers, and Animation Rigging features provided by Unity (see Fig. 2).

IK is one of the main techniques for granular control over robotic effector manipulation [29]. Even though animating human biomechanics through IK is a difficult task [30], there have been promising results in determining pose and ergonomic evaluation in pHRI using 3D simulators and IK [31]. For a generalized postural ergonomic analysis, an accurate humanoid skeletal IK framework will suffice, and other dynamics such as muscular mechanics can be ignored.

Unity offers basic IK functionalities within the software; however, these functions are not readily available for simulation purposes. These basic functionalities are the building blocks to allow the developers to create the IK framework that they need. We created our own full-body IK specifically designed for object handover simulations using a series of scripts and functions. We used the "Two-Bone IK" and the "Multi-Aim" constraints offered in Unity. The Two-Bone IK function simulates the kinematics of two bones and the target. This makes it a great tool to simulate limb behavior such as arms and legs. The Multi-Aim tool is used to simulate the upper body motion in object handovers. Three Multi-Aim tools were applied to the head, chest, and spine, each with varying speeds and thresholds in order to get a smooth and realistic motion. A Raycast gravity script was used to simulate the effects of gravity on the model. In this technique, a Raycast is projected from the feet of the model onto the terrain to solve the IK problem between the geometry of the surface and the feet. This method grants smooth movement and alignment of the model with moving and stationary environments. The entire system is controlled by a rig controller script in order to be able to generalize it to different tasks, body types, and functions. One advantage of this setup is that it can be extended and expanded to any other handover tasks and objects. The setup is specifically designed to simulate different handover scenarios tuning the hyper-parameters to the new conditions. The arrangement of the IK framework on each body part can be seen in Fig. 3.

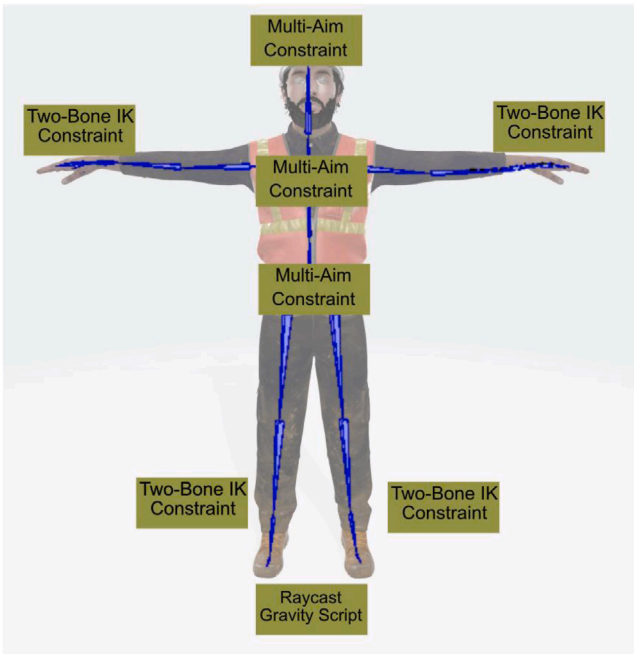

Fig. 3. Configuration of IK functions on the model.

2.3. REBA

Rapid Ergonomic Body Assessment or REBA is a standard industry-wide metric for overall posture assessment for industrial tasks. REBA is a reliable quantifiable method to evaluate work-related MSDs in the construction industry [32]. The framework defines a from 1–15 for the scores. However, this distribution does not apply to every task. The nature of the task and how a human engages with it will affect the score distribution [33]. Because REBA is a step-wise linear function, it is non-differentiable. Previous works have overcome this obstacle by introducing methods such as regression neural networks [11], as a sum of weighted polynomials [34] or a continuous linear function [35]. Approximation by nature will introduce uncertainty and errors to the process. These discrepancies, as opposed to using raw scores, can cause issues for the scientific validity in applied settings [11].

REBA scores are calculated based on joint and body angle ranges which are associated with a score depending on the ergonomic risk. The current state of the art mainly uses image processing as a means of calculating scores for HRI and ergonomic intervention purposes. Other methods such as using inertial measurements [36] and motion capture technologies [37] have also gained traction. In this framework, the joint angles are directly calculated through the digital bones embedded into the 3D model using the established calculation regimen. This feature gives us exact measurements of joint angles based on the REBA framework as an extension. The proposed framework at this point does not utilize live REBA reading since the relatively optimal position for the specific task is acquired in advance. Once the optimal ergonomic position relative to the worker is acquired, the robot is equipped with a robust policy to ergonomically handover the object at the desired location provided the global position is known. However, during a real-world implementation, it would be prudent to calculate real-time ergonomic metrics and fail-safes to ensure worker safety.

3. Optimization

We chose to optimize for an auxiliary score based on the actual REBA calculations which we will call *postural scores* over the final REBA scores in the virtual simulation. The calculation of REBA as designed by its inventors consists of collecting limb and joint angles which then are processed through a predefined criterion to calculate each score. These scores are then tallied up and filtered utilizing two separate calculation tables into two different groups:

Group A: Analysis of the neck, trunk, and legs
Group B: Analysis of the arms and wrists

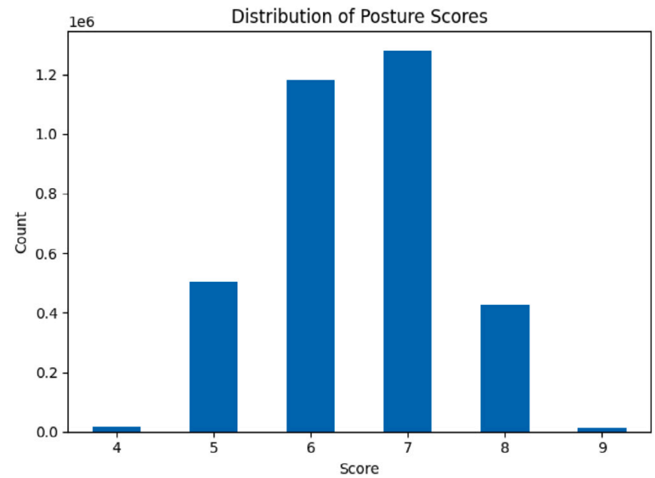

Fig. 4. Distribution of Postural Score (x-axis) given a bi-manual HRI Object Handover in all in-reach positions (y-axis) in the assigned boundary.

We use these two scores which already are part of the REBA calculation as *postural scores*. Using these two postural scores, we can calculate the final REBA score using Table C. However, these tables introduce an additional simplification and discreteness to the distribution. As it can be seen in Fig. 5, the final REBA score can be unchanged even when one of the two scores changes. For example, when the score for group A is 8 and the score for group B is 6, any increase in the value of score B between 6 and 10 would result in the same final REBA score. However, the ergonomic tax has increased even though the actual limb scores for that body group have worsened. This is expected since REBA was not built for precise and accurate ergonomic analysis, especially in VR. However, REBA was created to give an easy, fast, and transparent tool for industry practitioners to measure potential ergonomic harm in industrial settings [4,38]. The transformation tables also add discreteness to the distribution. Since our protocol utilizes a reward mechanism based on the improvement of an ergonomic score, the less sensitive score will prove as a harder goal to optimize for. With the postural scores, smaller changes in the positions would result in better or worse scores, helping guide the RL scheme to converge to the optimal position faster. We hypothesize that increased granularity would increase our chances of achieving the true optimal position more reliably. We also argue that the postural score is representative of the true REBA score. The postural score is based on the same score definitions given by REBA for its calculations, meaning it does not perform any different calculation that would be in contradiction with the REBA framework. Furthermore, the relationship between these scores and the REBA score is monotonic, meaning REBA will not decrease as the postural score increases. We further demonstrate that our scheme does improve ergonomic positions by calculating true REBA scores in our experiments.

The Unity-based simulation incorporates a 3D humanoid model of a construction worker, which is sourced from Mixamo—a platform offered by Adobe that facilitates the creation and rigging of animations and 3D models. This model's body proportions and sizes are adjustable and can be easily rigged with the help of Blender to accurately replicate individual human movements. Once the 3D model is fully rigged, it is imported into Unity. Within Unity, we have designed an IK framework

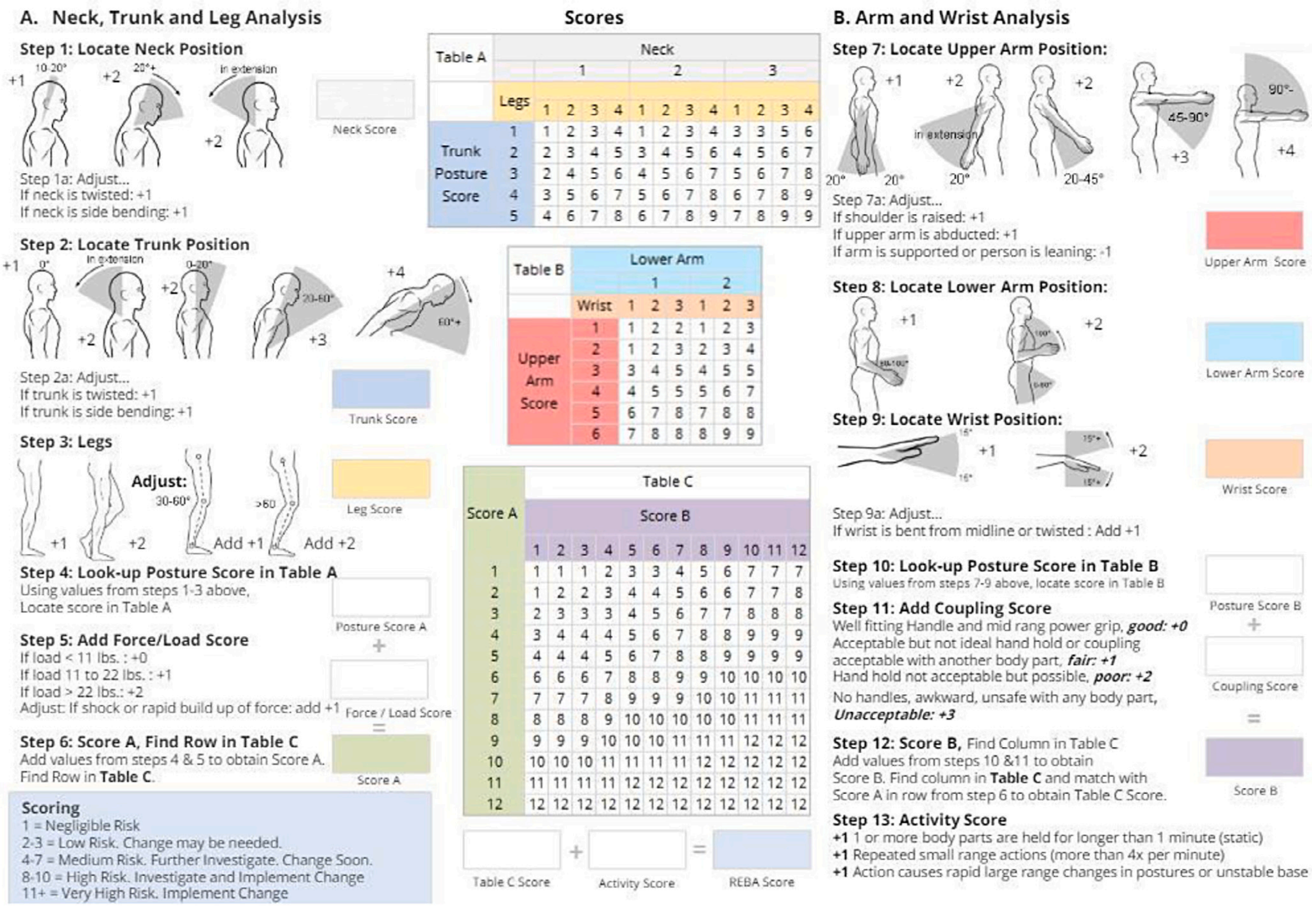

Fig. 5. REBA assessment worksheet [4].

for the digital twin, specifically designed for bi-manual object handovers. This framework leverages Unity's native functions and is further refined with custom C# code written by our team to enhance the IK system's precision and adaptability. The lower body and leg movement is controlled by proprietary code and Unity's IKFootSolver and two-bone IK constraints. All components and clashes are then handled by a rig controller script. For all other body shapes and sizes, the IK target and IK hint positions and rig weights will need to be aligned and adjusted according to the dimensions of the model. A postural score calculator script was created using the joint angles of the model following the guidelines defined by the framework. Since the only destination a robot can control and project towards is the position of its end effectors, each end effector would either need to be averaged to the theoretical center of mass of the object or solve and optimize for each effector separately. We chose to optimize for the center of the end effectors using Eq. (6).

To find the optimal ergonomic score, we utilize a Tabular Q-learning methodology given by Eq. (2) with a softmax action selection given by Eq. (3). A solution space for the RL scheme was chosen given by the body frame of the 3D model. The size and shape of this distribution set will affect the distribution of the postural scores and by extension the convergence quality and speed. We chose a boundary given by arm extension in the z-axis, from knee height to the top of the head for the y-axis and shoulder width for the x-axis. The reward mechanism utilizes an inverse quadratic relationship of the ergonomic score given by Eq. (5). By opting for an action step size of 0.003 units, this solution space gives us about 3.4 million solutions for the Q-learner to explore. Per Fig. 4, we can see that most of the distribution belongs to Posture scores 6–7, and only 0.004% of the solution space belongs to the score 4, which coupled with the discrete nature of the distribution, renders optimization and RL objectives extremely difficult. Recording every possible score for every position took approximately 3 h to calculate. REBA, which the calculated postural score is a subset of, and most other rapid ergonomic scores do not account for human subjective preferences. This creates a possibility of obtaining the same minimum optimal score at multiple different positions, causing the algorithm to converge and stop exploring different locations at sub-optimal locations at the global scale. We decided to incorporate a reward value (6) for the symmetry of the box about the body contact frame given that humans prefer symmetry [39]. This score is then incorporated back into the initial reward function given by Eq. (7). Given the discrete distribution of ergonomic scores, the Q-Learning algorithm is prone to convergence on local minima. This often occurs because, in instances where a new position does not yield a change in the ergonomic score, the algorithm's reward becomes dependent solely on symmetry changes within the unchanged score region. As a result, Q-Learning may settle on suboptimal solutions. To counteract this tendency and promote exploration within the model, we opted to keep the softmax action selection temperature relatively high.

$$r = \frac{1}{E^2} \tag{5}$$

$$S = X_{\mathrm{LE}} + (X_{\mathrm{RE}} - X_{\mathrm{LE}}) \tag{6}$$

$$r = \frac{1}{E^2} + S \tag{7}$$

After the model identifies the lowest postural score, the relative position or positions are saved into a CSV file ready to be used. In a real-world translation, a simple dimension conversion can be introduced to be able to deploy this framework to pHRI settings. Each unit used in the

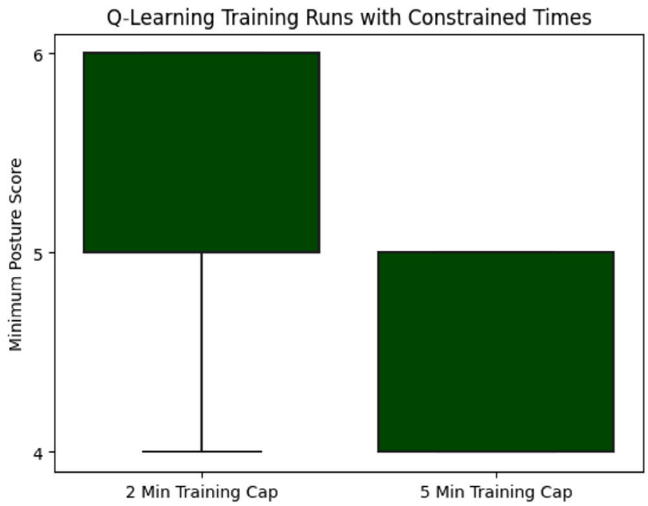

Fig. 6. Minimum postural score from 10 separate training runs with the Q-learning algorithm.

Unity game engine translates to 1 meter in the real world. This standardized approach allows us to be able to convert the optimal locations that are in the "Unity unit" to real-world dimensions depending on the deployment and sensor readings from the deployed robot.

4. Training results

To better test the convergence of the model, we ran the algorithm with a time cap. The reasoning behind this is twofold: to evaluate the time needed for convergence and to prevent the algorithm from stopping at a local minimum. To promote exploration and avoid getting stuck in a local minimum, we used a high τ value for the Boltzmann action selection. We tested the algorithm with both a 2-min and a 5-min cap across 10 iterations. These time caps were chosen for testing purposes. They can be adjusted at the user's discretion depending on scenario and operational factors. While exploring the entire handover solution space takes about 3 h, the model was able to converge in a matter of minutes. Both time caps achieved the lowest global postural score for the bi-manual object handover location. Results can be seen in Fig. 6. After acquiring the data, we chose the minimum postural score and the relative position associated with the score. From here we created a script that moves the box to the optimal location. To contrast this with a baseline, another script was created to move the object to the closest distance in the boundary so that the model can reasonably receive the box. This hand-off location has been previously used in the literature, albeit in the context of motion planning and handover approaches [40].

We chose 5 locations with different X, Y, and Z coordinates outside the model's reach. We then collected the postural score associated with each location. The results show that the optimized approach was able to achieve the minimum postural score at every initial starting coordinate with no variation. This is to be expected since the destination of the framework is independent of the starting coordinates. The shortest distance approach achieved worse REBA at every starting point. The maximum REBA score was 7 and the minimum REBA score was 5. All results are depicted in Table 1 and Figs. 6 and 7. Fig. 8 shows a sample of the results in the virtual environment. The Unity engine was running on a computer with AMD Ryzen 7 5800 8 Core Processor, 80 GB of installed RAM, and an NVIDIA GeForce RTX 12 GB 3060 GPU.

Table 1
Comparison of optimized and ergonomically naive methods.

Relative position	Optimized location	Shortest distance
(x = 0.028, y = 1.122, z = 1.354)	4	5
(x = 0.263, y = 1.122, z = 1.354)	4	6
(x = −0.423, y = 1.122, z = 1.354)	4	5
(x = 0.028, y = 0.472, z = 0.6)	4	7
(x = 0.028, y = 2.292, z = 1.354)	4	5

5. Experiments

To further investigate the efficacy of this optimization regimen, we conducted within-subject intra-lab experiments (N = 7). A VR scenario was created to test the improvements attributed to the framework. In the scenario, the participants are placed in a room with a box positioned in front of them. We simulated a stationary object handover scenario. A snapshot of the VR environment is depicted in Fig. 9. While similar studies used a sample size as small as 4 participants [11,41], we also conducted a power analysis using Cohen's d to measure if the sample size was enough. The power analysis is the calculation used to estimate the small sample size needed for an experiment given significance level, power, and effect size. Since this is a within-subject experiment, we utilize Cohen's d for our effect size. Cohen's d acts as an effect size to conduct the power analysis.

This scenario required the human to be stationary and the robot to transport the object towards the human for the handover. For this purpose, the participants were free to move and rotate their bodies as long as they did not physically move around (with the exception of not raising their legs off the ground). Two different object starting points were tested.

1. The box was directly aligned to the center of the participant's frame
2. The box was offset by 1 in-game unit to the left on the x-axis of the participant

After signaling their readiness, the box in the VR simulation is transported towards the participant using the two different (optimized and unoptimized) end locations. After reaching its destination, the box changes color to signal to the participant that is ready for the handoff. The participants are instructed to reach out to simulate grabbing and receiving the box from existing handles on the sides of the object. The participants were asked to hold the position while we took photos of their real-world postures to calculate the corresponding REBA score. This procedure is done for both of the two different end locations and start locations. An example of this process is shown in Fig. 10.

One shortcoming of assessing ergonomic optimization in VR is the tunnel-visioned worldview that is a byproduct of the headset [42]. The field of view of individuals using VR is significantly less as compared to a natural field of view. The reduced field of view can force the users to rotate and change neck position considerably more as opposed to a real-world scenario, which can negatively affect REBA scores due to the higher injury risk associated with neck angles. For consistency and simplicity, we assume the lowest risk neck score between both scenarios to remove the potential confounding factor from our analysis. Fig. 11 shows an example of a recorded data point of the experiment.

As it can be seen in Fig. 12, in both cases, the optimized position ends up in the same coordinates since that has been predetermined by the framework.

At x = 0, after optimization, the mean result for the REBA score is 1.57 ± 0.90. In contrast, the unoptimized group recorded a significantly higher mean of 4.00 ± 1.20 for the REBA score. The statistical analysis confirms the effectiveness of the optimization with a p-value of 0.002 and Cohen's d of 3.20, indicating a statistically significant difference at 95% confidence level.

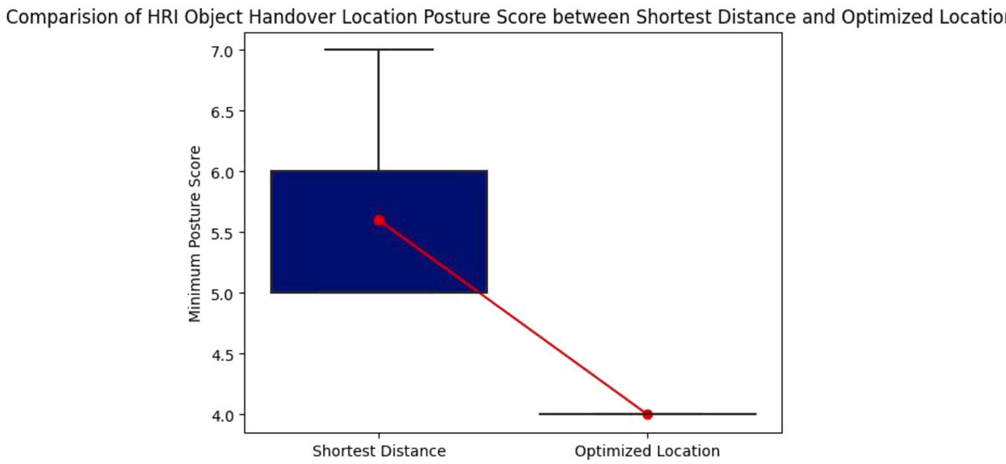

Fig. 7. Comparison between the shortest distance in the boundary vs. optimal location heuristic.

(a) Test Case 2, Shortest Distance (b) All Test Cases, Optimized Location (c) Test Case 4, Shortest Distance

Fig. 8. Posture scores based on heuristic and starting location.

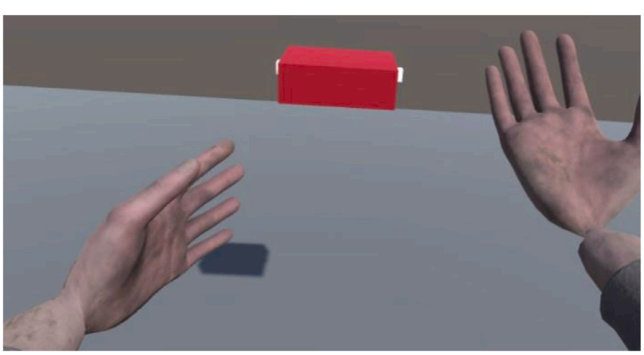

Fig. 9. VR environment included a box positioned in front of the human for a simulated handover scenario.

At $x = -1$, after optimization, the mean REBA score is 1.57 ± 0.90. In contrast, the unoptimized group recorded a significantly higher mean REBA score of 5.71 ± 1.48. The statistical analysis confirms the effectiveness of the optimization with a p-value of < 0.001 and Cohen's d of 3.87.

5.1. *Construction object case studies*

In this section, we explore ergonomic optimization using construction objects by replacing the generic object. As long as certain constraints such as grabbing point location and object search bounds are defined, the RL algorithm can optimize for the ideal ergonomic handover location.

5.1.1. *Shovel*

We used the shovel as a common construction tool that can be a point of interest in bimanual object handovers. This is a proof-of-concept example even though in reality shovels may be handed over with only one hand. The setup was changed in order to adjust for the new grip goals and geometry of the object. An example of the new IK setup can be seen in Fig. 13.

The entire handover range of REBA scores for this item was calculated which is depicted in Fig. 14.

As it can be seen in Fig. 14, the minimum constitutes about 0.1% of the entire distribution. The distribution loosely resembles the distribution of a generic object. This is due to the IK goals depending on the geometry of the shovel being relatively similar to the goals of the generic object. We then attempted to search for the ideal ergonomic position using the Q-Learner. In a 5-min training cap, the Q-learner achieved the absolute minimum 2 times. The first one was found in 155 s and the last one in 160 s. The configuration of the framework using the first found position is shown in Fig. 15

5.1.2. *Container box*

Another common construction equipment example that was used is a container box. Similar to the shovel, we adjusted the model and the IK to the new object as seen in Fig. 16.

We calculated the entire handover range of REBA scores for this item as well, as shown in Fig. 17.

The absolute minimum of this distribution is a score of 5, which accounts for about 9% of the distribution. This is different from the previous objects due to the object handles being wider, changing the handover due to the geometry of the object. This facet highlights the flexibility of the framework in handling different conditions.

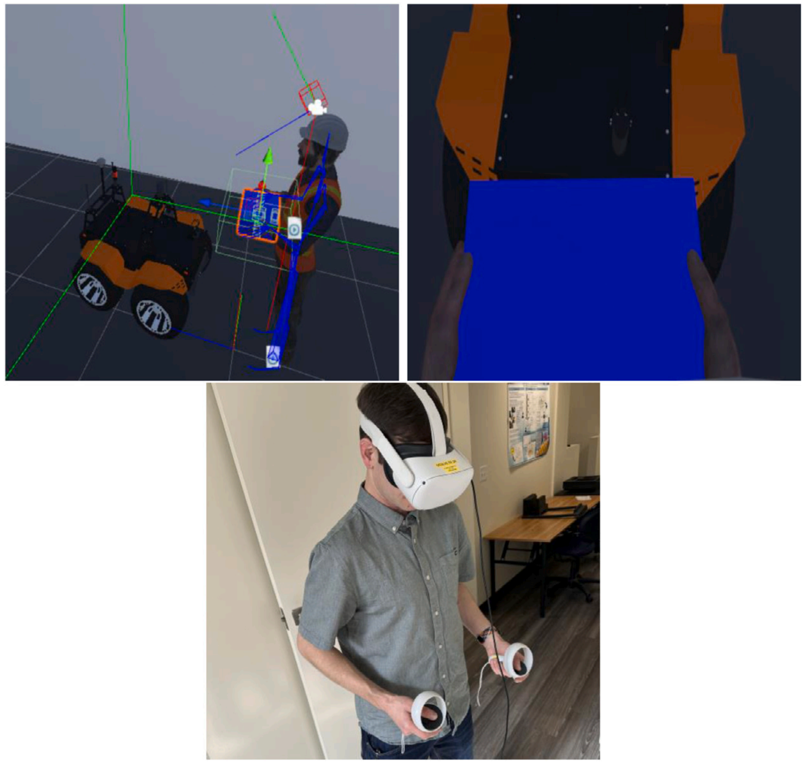

Fig. 10. Participant attempting to receive the box in the VR simulation and the real-world body posture manifestation.

Fig. 11. Manifested postures of the optimized framework (left) and unoptimized framework (right).

The absolute minimum of this distribution is a score of 5, which accounts for about 9% of the distribution. This is different from the previous objects due to the object handles being wider, changing the handover due to the geometry of the object. This facet highlights the flexibility of the framework in handling different conditions. After applying the Q-learner to the new setup with a 5-min cap, the absolute minimum was found 9 times with the first one in 4 s and the last one in 147 s. The configuration of the framework using the first found position is depicted in Fig. 18.

6. Discussion

This analysis promotes the framework's versatility. Any 3D object can be optimized for ideal placement using this method. This framework can be leveraged significantly in construction projects. The existence of rich building information models (BIM) can prove as an asset since the families used in the BIMs can be exported as 3D files such as FBX and OBJ files. The extension of this framework is not just for bimanual handover. Further ergonomic assessments and education can help developers reduce the risk of WMSDs on job sites by cataloging high-risk and low-risk items in use through optimization in the virtual environment.

One of the most prominent obstacles to practical ergonomic optimization for intelligent applications such as pHRI can be attributed to the existing metrics with mathematical complexities. REBA calculates the ergonomic risk mainly through ranges of the angle of the specific joint in a step-wise manner. Most ergonomic metrics have two (usually contradictory) qualities: generality and accuracy [43]. The former refers to the metric's capacity to be applicable across various tasks, contexts, and rates of operation, signifying its universality. The latter pertains to the metric's exactness in quantifying the genuine ergonomic hazards associated with a particular task as determined by the framework. Recent sensitivity analyses indicate that the REBA methodology

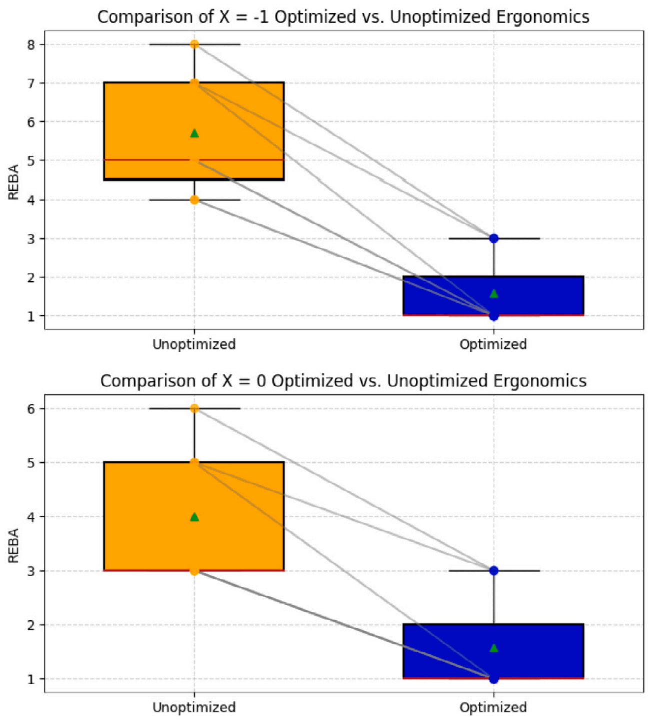

Fig. 12. Average of REBA scores between unoptimized and optimized schemes depending on starting positions.

Fig. 13. Alignment of the model with the shovel object.

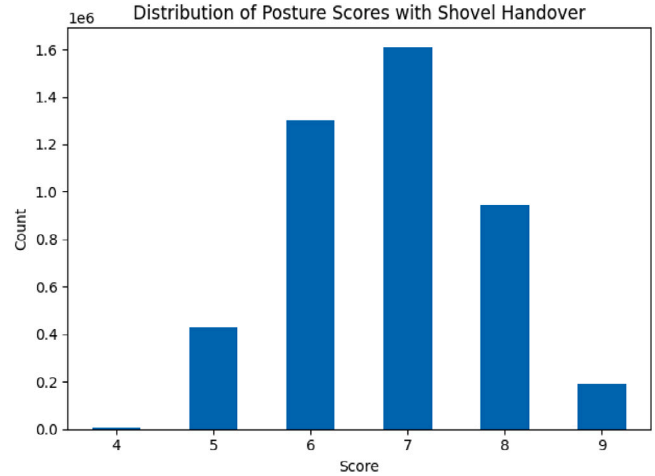

Fig. 14. Distribution of postural score (x-axis) given a bi-manual HRI object handover in all in-reach positions (y-axis) in the assigned boundary for the shovel object.

Fig. 15. Optimized location for the container box.

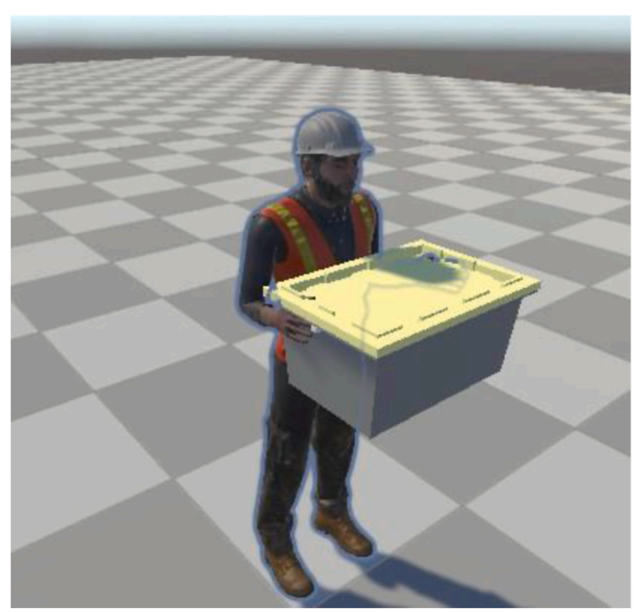

Fig. 16. Alignment of the model with the container box object.

exhibits limited suitability for material handling tasks [38], which constitute a significant aspect of pHRI object handovers. Instead, its efficacy is more pronounced in interventions targeting overall postural ergonomics. The speed and ease of use have been the strongest catalyst to REBA's popularity in the industry since its inception in the year 2000. However, the technological constraints of the past obscured the inadequacies of REBA for mathematical optimization tasks. Now, with advancements in image processing and sensor accuracy, we can gather and interpret limb and body angles for real-time ergonomic analysis — an application perhaps not originally intended by REBA's creators.

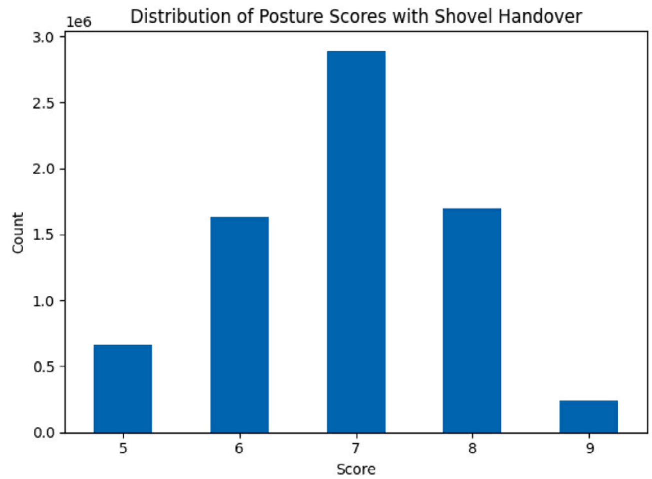

Fig. 17. Distribution of Postural Score (x-axis) given a bi-manual HRI Object Handover in all in-reach positions (y-axis) in the assigned boundary for the Container Box Object.

Fig. 18. Optimized location for the container box.

We encounter these shortcomings in our experiments. We use the tables given by Fig. 5 to calculate the REBA score of the participants. The REBA method has two separate scores that need to be calculated which then can be used through a series of tables which results in a final REBA score. In the worksheet, Score A is related to the "Neck, Trunk, and Leg Analysis", and Score B is related to the "Arm and Wrist Analysis". For the majority of the cases, at initial position $x = 0$, we calculate a score of 2 for Score A and a score of 5 for Score B. By using Table C in the worksheet and accounting for the activity score, we end up with a REBA score of 4. For initial position $x = -1$, the framework returns a score of 3 for Score A and a Score of 5 for Score B. This main difference between the REBA scores of the two initial positions is due to the twisted trunk induced by the asymmetry of the final object placement concerning the body by the scheme. However given Table C, we still result in a final REBA score of 4. However, there is a significant difference in comfort and physical strain between the two iterations. This difference is somewhat represented in the simple summation of the postural score but not inside the final valuation of the REBA score. This further demonstrates the shortcomings of REBA and the benefit of using more granule scoring schemes such as the postural scores presented in this paper which are derived from REBA itself.

Furthermore, REBA's consideration of weight differences is a linear addition to the score. In the REBA framework, first, the angles are considered, and a score is assigned. After that assignment, only then is the weight considered. This process has the ideal angles of the body to be independent of the weight. As the weight increases, the total ergonomic score increases depending on pre-defined thresholds. Factors such as center of mass, shape, and size of the object in relation to the weight are not considered. During the optimization, this does not greatly impact the final optimization position since the weight consideration adds scalar scores to the baseline amongst every position.

Moreover, the ideal REBA score is very close to sensitive parts of the human body such as the neck and trunk area. This can prove as a hazard given robotic end-effectors will be near those body parts. This will be true for any optimization metric that uses REBA or similar scores as their optimization metric. The current standard metric does not take this fact into account and mainly focuses on pure ergonomic optimization. These shortcomings can be taken into account in new pHRI-specific ergonomic metrics or certain handle locations if the object's geometry allows it.

One advantage of our framework is its flexibility in using different optimization metrics. While it currently employs REBA scores, it is not restricted to this metric alone. As new ergonomic pHRI metrics are developed, our framework can be easily adjusted to these developments. This interchangeability is a significant contribution, as it allows the framework to adapt and remain relevant with advancements in the field, using the metric as a tool for the RL reward function.

We can see as shown in the results section, that the postural score serves as an adequate optimization goal for this method. This was to be expected since the postural score is REBA in a more granule form. VR and computer simulations have the luxury of being able to calculate high-precision data with ease. The postural score approach also grants more flexibility in scoring each limb differently. By not limiting itself to the table presented in REBA, developers can create personalized scores and optimizations depending on the specific needs of the worker. Scores can be assigned with different numbers and values depending on the nature of the task and the health history of the individual. However, this approach demands further investigation to ensure its scientific validity and real-life efficacy. This further demonstrates the need for mathematical approaches to improve ergonomic optimization outside the scope of REBA (as demonstrated in this study) for not only pHRI contexts but also potential applications in VR. Additionally, more specialized and sensitive frameworks for pHRI applications in industrial settings could be devised.

Our method optimizes for the most ergonomic location of the end-effectors for bimanual or unimanual object handover regardless of the object as opposed to the nuances from the handover itself. Recent state-of-the-art research has approached the nuances of unimodal pHRI handover using whole-hand tactile sensors [44]. This RL framework can work in conjunction with the current handover works by serving as a planning mechanism for the end goal of the handover agents. Yu et al. utilized a learning-from-demonstration (LfD) method for implementing human grip state-reactive behavior norms in robots. Our framework can be incorporated with such frameworks to create ergonomically safe and seamless grip and handover scenarios. Furthermore, the LfD system indirectly encapsulates ergonomic safety to an extent. Through the demonstrations from workers, the optimal positions will indirectly have a level of comfort as a part of the features. While this alone might not be comprehensive of complete ergonomic safety, it presents future opportunities for the development of comprehensive ergonomic policies.

Current robotic research has enabled developers to have an accurate visual representation of localization and dimension perception through LiDAR technology [45], stereo vision [46], and SLAM-based algorithms [47] for a variety of different applications. Due to this measurement strategy, it is of paramount importance that the worker 3D humanoid model and its associated IK framework be as representative as possible of the worker's true body mechanics to decrease the potential divergence of the training and interaction optimality.

7. Conclusion

The implications of the findings are paramount for the pHRI practice in industries such as construction, manufacturing, and retail warehousing with substantial fieldwork. Robots are introduced to industrial applications to increase the safety of manual work, among others, through automating and taking over labor-intensive and repetitive tasks that are often not ergonomic. While their presence in the field should not pose additional safety hazards (e.g., struck-by accidents), their operations should also guarantee worker safety. As the only comprehensive framework to assess ergonomics of body posture, REBA has features that limit its functionality for real-time, feedback-enabled mechanisms needed for on-spot robot perception and reaction. As such, the developed methodology enhances the applicability of REBA for adaptive, safe, and fast implementation of REBA in bi-manual material handover situations between robots and field workers.

From a robotic implementation standpoint, this framework would require the exact knowledge of the global positions of the worker and the robot. Since the RL scheme returns an optimized position relative to the worker, it is necessary to know where the person is to transport the box to the relative position. Previous works have shown that worker localization via Indoor GPS can be achieved with high accuracy [48]. Such methods can be implemented in potential frameworks in order to enable the RL scheme in an applied setting.

The VR training method also opens up the groundwork for expedited framework testing outside of the job site. By testing pHRI frameworks and interactions in a virtual environment, we reduce the risk of potentially hazardous failure in a setting where humans are present. Many edge environmental edge cases can be analyzed and simulated which would pose a potential risk in real deployments.

8. Limitations and future work

While this paper presents an approach towards using ergonomics safety measures in pHRI, there are a few limitations that can be addressed in future work. The reward policy can be modified to better guide the Q-learner to reach the optimal position. Rewarding different features such as body ratios and joint angles can be able to assist the Q-learner's convergence. However, this would discourage the current model-free nature of the learner, since we are guiding it to the ideal position of the task by adding task-specific information into the training regimen. The boundary space can also be further limited to encourage faster convergence. Exploring the entire boundary space took a significantly longer duration compared to the RL algorithm. By further truncating the possible solution spaces, the odds of converging to a local minimum decrease, and the higher probability of the RL algorithm reaching the global optima faster, especially given the fact the softmax action selector's temperature has been set to a high value. However limiting the boundary would decrease the routine's generalizability due to the reason that given different tasks, configurations, and worker demographics, the REBA distribution can be highly variable. More sophisticated RL algorithms can be implemented for faster convergence. Additionally, increasing the action subset can prove beneficial. Currently, we use 6 different actions to navigate all general directions in a three-dimensional space. However, increasing the action function with more fine-grained actions can speed up convergence by accessing states that would need multiple other actions otherwise. We are considering the introduction of angular motions to our action subset, which could be particularly advantageous for the ergonomic optimization of handling complex-shaped objects. However, it is important to note that increasing the number of actions might also extend the training duration, as the model will need to explore and learn from a more complex action space.

For future projects, we plan to incorporate Deep Q-learning methods [49] which utilize neural networks to represent the Q-function as opposed to the table of Q-values, and methods such as Soft Actor-Critic algorithms [50]. However, the biggest challenge to overcome is, once again, the discrete nature of the ergonomic frameworks.

Although the IK framework does provide an adequate representation of postural kinematics, a gap between true kinematics and the simulated ones is to be expected. We selected Unity for simulating kinematics over other scientific simulators because we aim to utilize this game engine in future HRI projects. Specifically, we intend to use Unity for robot logic and to integrate VR and AR for remote work and worker training applications. The objective of our approach is not to replicate human postural kinematics with high fidelity but rather to establish and refine a framework that is informed by ergonomic principles for use in HRI contexts. Options such as OpenSim would enable us to model and simulate the biomechanics and ergonomics of humans to a much more precise scale. However, scientific simulators and their integration with robotic controls and generalizability in a virtual environment or a construction job site can prove challenging.

For the experiment, the humanoid model in the game engine was not tailored to every single participant. We expect that this factor has reduced the maximum benefits that the regimen has to offer. However, this decision was deliberate since we were planning to investigate the generalizability of the scheme. We still were able to improve ergonomic scores while also achieving similar REBA scores among all participants. These results show a significant level of generalizability. However, we predict that these scores are not the true optimal positions for every participant.

For future work, we plan to present streamlined ways to tailor the humanoid models to the individuals. This customization enables the framework to learn personalized traits outside of operating hours. Once the robot has gained the relative position for a specific task with the specific worker, it can perform ergonomically optimized pHRI object handover with or without the Unity game engine. The learner optimizes using pixel units inside the game engine. These ratios are simply a change of dimensions relative to the body measurements of the worker imported into the 3D environment, resulting in a representation of body area and volume in a digitized setting. Subsequently, the real-world implementation would just consist of a dimension transformation for the robot to perform the movement with or without the game engine.

It would be ideal to personalize the humanoid model to every construction worker in an applied setting for maximum ergonomic optimization. The discrepancy in real-life manifestations of the human posture can be seen in Fig. 19. Even though there is a difference between the qualitative postures, this method still provides significant benefits as compared to ergonomically agnostic heuristics. Furthermore, such positions outside of the scope of the study can have either exact or marginally different REBA scores. However, it is not hard to infer that these postures can have different comfort levels and injury risks. This is another piece of evidence that points towards the need for sensitive and mathematical ergonomic and safety frameworks for pHRI tasks.

Finally, this object handover method is considered a non-adaptive approach [40]. A non-adaptive approach means that the robot simply moves the object to a pre-defined path without any adaptation to the potential changes of the receiver. For future work, we plan to incorporate adaptive corrections to the object handover for stronger generalizability in a dynamic environment which is paramount to scalable deployment of construction robots. Moreover, we would like to test this framework in a dynamic construction environment using physical robots. By defining the task and the human model of the construction worker. VR has been previously used to assist construction workers in training for HRC [51]. This work further opens up the possibilities of utilizing VR to tackle the pressing construction problems. One major obstacle to the deployment of construction robots can be attributed to the lack of trust in the feasibility and efficacy of the robots. Pairing training and ergonomic considerations inside and outside of the job site.

Fig. 19. Difference in manifested ergonomic posture on human subjects when using the identical RL optimized location.

CRediT authorship contribution statement

Mani Amani: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation. **Reza Akhavian:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Reza Akhavian reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The presented work has been supported by the U.S. National Science Foundation (NSF) CAREER Award through grant No. CMMI 2047138 and a scholarship funded by grant No. DUE 1930546, as well as SDSU Instructional Technology Services (ITS) Virtual Immersive Teaching and Learning (VITaL) Research Center. The authors gratefully acknowledge the support from the NSF and SDSU ITS. Any opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily represent those of the NSF or SDSU ITS. The authors would also like to thank Dr. Nguyen-Truc-Dao Nguyen, Dr. Miguel Dumett, and Mr. Jeffrey Xing for their advice and support as well as our colleagues at the Data-informed Construction (DiCE) Lab at San Diego State University for their voluntary assistance in the evaluation of the framework.

References

[1] J. Krüger, T. Lien, A. Verl, Cooperation of human and machines in assembly lines, CIRP Ann 58 (2) (2009) 628–646, http://dx.doi.org/10.1016/j.cirp.2009.09.009, URL https://www.sciencedirect.com/science/article/pii/S0007850609001760.

[2] P.B. Rodrigues, R. Singh, M. Oytun, P. Adami, P.J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, G.M. Lucas, A multidimensional taxonomy for human-robot interaction in construction, Autom. Constr. 150 (2023) 104845, http://dx.doi.org/10.1016/j.autcon.2023.104845, URL https://www.sciencedirect.com/science/article/pii/S092658052300105X.

[3] R. Jahanmahin, S. Masoud, J. Rickli, A. Djuric, Human-robot interactions in manufacturing: A survey of human behavior modeling, Robot. Comput.-Integr. Manuf. 78 (2022) 102404, http://dx.doi.org/10.1016/j.rcim.2022.102404, URL https://www.sciencedirect.com/science/article/pii/S0736584522000916.

[4] S. Hignett, L. McAtamney, Rapid entire body assessment (REBA), Appl. Ergon. 31 (2) (2000) 201–205, http://dx.doi.org/10.1016/s0003-6870(99)00039-3.

[5] L. McAtamney, N.E. Corlett, RULA: a survey method for the investigation of work-related upper limb disorders, Appl. Ergon. 24 (2) (1993) 91–99, http://dx.doi.org/10.1016/0003-6870(93)90080-s.

[6] O. Karhu, P. Kansi, I. Kuorinka, Correcting working postures in industry: A practical method for analysis, Appl. Ergon. 8 (4) (1977) 199–201, http://dx.doi.org/10.1016/0003-6870(77)90164-8, URL https://www.sciencedirect.com/science/article/pii/0003687077901648.

[7] M. Hamad, S. Nertinger, R.J. Kirschner, L. Figueredo, A. Naceri, S. Haddadin, A concise overview of safety aspects in human-robot interaction, 2023, http://dx.doi.org/10.48550/arXiv.2309.09936, arXiv:2309.09936.

[8] M. Lorenzini, M. Lagomarsino, L. Fortini, S. Gholami, A. Ajoudani, Ergonomic human-robot collaboration in industry: A review, Front. Robotics AI 9 (2023) 813907, http://dx.doi.org/10.3389/frobt.2022.813907.

[9] L. Gualtieri, E. Rauch, R. Vidoni, D.T. Matt, Safety, ergonomics and efficiency in human-robot collaborative assembly: design guidelines and requirements, Procedia CIRP 91 (2020) 367–372, http://dx.doi.org/10.1016/j.procir.2020.02.188.

[10] A. Schiele, F.C.T. van der Helm, Kinematic design to improve ergonomics in human machine interaction, IEEE Trans. Neural Syst. Rehabil. Eng. 14 (4) (2006) 456–469, http://dx.doi.org/10.1109/TNSRE.2006.881565.

[11] A. Yazdani, R.S. Novin, A. Merryweather, T. Hermans, DULA and DEBA: Differentiable ergonomic risk models for postural assessment and optimization in ergonomically intelligent pHRI, 2022, http://dx.doi.org/10.1109/IROS47612.2022.9981528.

[12] T. Stouraitis, M. Gienger, Predictive and robust robot assistance for sequential manipulation, IEEE Robot. Autom. Lett. (2023) 1–8, http://dx.doi.org/10.1109/LRA.2023.3320029.

[13] M. Omidi, G. Van de Perre, R. Kumar Hota, H.-L. Cao, J. Saldien, B. Vanderborght, I. El Makrini, Improving postural ergonomics during human–robot collaboration using particle swarm optimization: A study in virtual environment, Appl. Sci. 13 (9) (2023) 5385, http://dx.doi.org/10.3390/app13095385.

[14] P. Wang, P. Wu, J. Wang, H.-L. Chi, X. Wang, A critical review of the use of virtual reality in construction engineering education and training, Int. J. Environ. Res. Public Health 15 (6) (2018) 1204, http://dx.doi.org/10.3390/ijerph15061204.

[15] J. Li, C. Xu, Z. Chen, S. Bian, L. Yang, C. Lu, Hybrik: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2021, pp. 3382–3392, http://dx.doi.org/10.1109/CVPR46437.2021.00339.

[16] Unity Technologies, Simulating robots with ROS and unity, 2023, https://resources.unity.com/unitenow/onlinesessions/simulating-robots-with-ros-and-unity. (Accessed 14 October 2023).

[17] A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar, D. Lange, Unity: A general platform for intelligent agents, 2020, http://dx.doi.org/10.48550/arXiv.1809.02627, arXiv preprint arXiv:1809.02627, URL https://arxiv.org/pdf/1809.02627.pdf.

[18] I. El Makrini, G. Mathijssen, S. Verhaegen, T. Verstraten, B. Vanderborght, A virtual element-based postural optimization method for improved ergonomics during human-robot collaboration, IEEE Trans. Autom. Sci. Eng. 19 (3) (2022) 1772–1783, http://dx.doi.org/10.1109/TASE.2022.3147702.

[19] A.A. Akanmu, J. Olayiwola, O. Ogunseiju, D. McFeeters, Cyber-physical postural training system for construction workers, Autom. Constr. 117 (2020) 103272, http://dx.doi.org/10.1016/j.autcon.2020.103272, URL https://www.sciencedirect.com/science/article/pii/S0926580519313408.

[20] B. Luo, D. Liu, T. Huang, D. Wang, Model-free optimal tracking control via critic-only Q-learning, IEEE Trans. Neural Netw. Learn. Syst. 27 (10) (2016) 2134–2144, http://dx.doi.org/10.1109/TNNLS.2016.2585520.

[21] E. Barron, H. Ishii, The bellman equation for minimizing the maximum cost, Nonlinear Anal. TMA 13 (9) (1989) 1067–1090, http://dx.doi.org/10.1016/0362-546X(89)90096-5, URL https://www.sciencedirect.com/science/article/pii/0362546X89900965.

[22] M. Wunder, M.L. Littman, M. Babes, Classes of multiagent q-learning dynamics with epsilon-greedy exploration, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), ISBN: 9781-60558-907-7, 2010, pp. 1167–1174, URL https://icml.cc/Conferences/2010/papers/191.pdf.

[23] Z. Song, R. Parr, L. Carin, Revisiting the softmax bellman operator: New benefits and new perspective, in: International Conference on Machine Learning, PMLR, 2019, pp. 5916–5925, http://dx.doi.org/10.48550/arXiv.1812.00456.

[24] S. Syafiie, F. Tadeo, E. Martinez, Softmax and ϵ-greedy policies applied to process control, IFAC Proc. Vol. 37 (2004) 729–734, http://dx.doi.org/10.1016/S1474-6670(17)31556-2.

[25] R.D. Barkokebas, X. Li, Use of virtual reality to assess the ergonomic risk of industrialized construction tasks, J. Constr. Eng. Manag. 147 (3) (2021) 04020183, http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0001997.

[26] A. Lugilde-López, F.A. Mikic-Fonte, M.C. Rodríguez, L.E.F. Santiago, M. Llamas-Nistal, Biarc spline interpolation with procedural grid and segments with armature to support colonosopy simulators in unity, in: 2022 International Symposium on Computers in Education, SIIE, 2022, pp. 1–6, http://dx.doi.org/10.1109/SIIE56031.2022.9982356.

[27] C. Bartneck, M. Soucy, K. Fleuret, E.B. Sandoval, The robot engine — Making the unity 3D game engine work for HRI, in: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2015, pp. 431–437, http://dx.doi.org/10.1109/ROMAN.2015.7333561.

[28] E. Coronado, S. Itadera, I.G. Ramirez-Alpizar, Integrating virtual, mixed, and augmented reality to human–robot interaction applications using game engines: A brief review of accessible software tools and frameworks, Appl. Sci. 13 (3) (2023) http://dx.doi.org/10.3390/app13031292, URL https://www.mdpi.com/2076-3417/13/3/1292.

[29] A. Goldenberg, B. Benhabib, R. Fenton, A complete generalized solution to the inverse kinematics of robots, IEEE J. Robot. Autom. 1 (1) (1985) 14–20, http://dx.doi.org/10.1109/JRA.1985.1086995.

[30] A. Aristidou, J. Lasenby, Y. Chrysanthou, A. Shamir, Inverse kinematics techniques in computer graphics: A survey, in: Computer Graphics Forum, Vol. 37, Wiley Online Library, 2018, pp. 35–58, http://dx.doi.org/10.1111/cgf.13310.

[31] L. Rapetti, C. Sartore, M. Elobaid, Y. Tirupachuri, F. Draicchio, T. Kawakami, T. Yoshiike, D. Pucci, A control approach for human-robot ergonomic payload lifting, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, 2023, pp. 7504–7510, http://dx.doi.org/10.1109/ICRA48891.2023.10161454.

[32] Y. Yu, X. Yang, H. Li, X. Luo, H. Guo, Q. Fang, Joint-level vision-based ergonomic assessment tool for construction workers, J. Constr. Eng. Manag. 145 (5) (2019) 04019025, http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0001647.

[33] D. Rabbani, S. Ahmed, Ergonomic analysis of material handling for a residential building at rourkela, J. Inst. Eng. (India): Ser. A 101 (2020) 1–11, http://dx.doi.org/10.1007/s40030-020-00477-x.

[34] B. Busch, M. Toussaint, M. Lopes, Planning ergonomic sequences of actions in human-robot interaction, in: 2018 IEEE International Conference on Robotics and Automation, ICRA, 2018, pp. 1916–1923, http://dx.doi.org/10.1109/ICRA.2018.8462927.

[35] Z. Xie, Improving Safety During Driving and Human-Robot Collaboration Through Deep Learning and Reinforcement Learning (Ph.D. thesis), North Carolina State University, ISBN: 9798379879617, 2023, p. 102, URL http://libproxy.sdsu.edu/login?url=https://www.proquest.com/dissertations-theses/improving-safety-during-driving-human-robot/docview/2838438495/se-2.

[36] S. Zelck, S. Verwulgen, L. Denteneer, H. Bossche, S. Scataglini, Combining a wearable IMU mocap system with REBA and RULA for ergonomic assessment of container lashing teams, 2022, pp. 462–465, http://dx.doi.org/10.1007/978-3-030-74614-8_59.

[37] M. Yunus, M. Jaafar, A.S.A. Mohamed, N. Azraai, T.D.M. Cchem Mrsc, Implementation of kinetic and kinematic variables in ergonomic risk assessment using motion capture simulation: A review, Int. J. Environ. Res. Public Health 18 (2021) 8342, http://dx.doi.org/10.3390/ijerph18168342.

[38] M. Joshi, V. Deshpande, Investigative study and sensitivity analysis of rapid entire body assessment (REBA), Int. J. Ind. Ergon. 79 (2020) 103004, http://dx.doi.org/10.1016/j.ergon.2020.103004, URL https://www.sciencedirect.com/science/article/pii/S0169814120301013.

[39] D. Evans, P. Orr, S. Lazar, D. Breton, J. Gerard, D. Ledbetter, K. Janosco, J. Dotts, H. Batchelder, Human preferences for symmetry: subjective experience, cognitive conflict and cortical brain activity, PLoS One 7 (6) (2012) e38966, http://dx.doi.org/10.1371/journal.pone.0038966, Epub 2012 Jun 13.

[40] M. Käppler, I. Mamaev, H. Alagi, T. Stein, B. Deml, Optimizing human-robot handovers: the impact of adaptive transport methods, Front. Robotics AI 10 (2023) http://dx.doi.org/10.3389/frobt.2023.1155143.

[41] L.v. der Spaa, M. Gienger, T. Bates, J. Kober, Predicting and optimizing ergonomics in physical human-robot cooperation tasks, in: 2020 IEEE International Conference on Robotics and Automation, ICRA, 2020, pp. 1799–1805, http://dx.doi.org/10.1109/ICRA40945.2020.9197296.

[42] Tom's Hardware, FOV (field of view) definition, 2023, URL https://www.tomshardware.com/reviews/fov-field-of-view-definition,5740.html. (Online Accessed 31 December 2023).

[43] C. Fransson-Hall, R. Gloria, A. Kilbom, J. Winkel, L. Karlqvist, C. Wiktorin, S.M..S. Group, A portable ergonomic observation method (PEO) for computerized on-line recording of postures and manual handling, Appl. Ergon. 26 (2) (1995) 93–100, http://dx.doi.org/10.1016/0003-6870(95)00003-u.

[44] H. Yu, V.R. Kamat, C.C. Menassa, W. McGee, Y. Guo, H. Lee, Mutual physical state-aware object handover in full-contact collaborative human-robot construction work, Autom. Constr. 150 (2023) 104829, http://dx.doi.org/10.1016/j.autcon.2023.104829, URL https://www.sciencedirect.com/science/article/pii/S0926580523000894.

[45] J.P. Queralta, F. Yuhong, L. Salomaa, L. Qingqing, T.N. Gia, Z. Zou, H. Tenhunen, T. Westerlund, FPGA-based architecture for a low-cost 3D lidar design and implementation from multiple rotating 2D lidars with ROS, in: 2019 IEEE SENSORS, 2019, pp. 1–4, http://dx.doi.org/10.1109/SENSORS43011.2019.8956928.

[46] A. Angulo, L. Pari, E. Espinoza, Y. Silva Vidal, E.D. Colquehuanca, Assisted operation of a robotic arm based on stereo vision for positioning near an explosive device, Robotics 11 (2022) 100, http://dx.doi.org/10.3390/robotics11050100.

[47] H. Taheri, Z.C. Xia, Slam; definition and evolution, Eng. Appl. Artif. Intell. 97 (2021) 104032, http://dx.doi.org/10.1016/j.engappai.2020.104032, URL https://www.sciencedirect.com/science/article/pii/S0952197620303092.

[48] H.M. Khoury, V.R. Kamat, Evaluation of position tracking technologies for user localization in indoor construction environments, Autom. Constr. 18 (4) (2009) 444–457, http://dx.doi.org/10.1016/j.autcon.2008.10.011, URL https://www.sciencedirect.com/science/article/pii/S0926580508001647.

[49] J. Fan, Z. Wang, Y. Xie, Z. Yang, A theoretical analysis of deep Q-learning, 2020, http://dx.doi.org/10.48550/arXiv.1901.00137, arXiv:1901.00137, URL https://arxiv.org/abs/1901.00137.

[50] T. Haarnoja, A. Zhou, K. Hartikainen, G. Tucker, S. Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel, S. Levine, Soft actor-critic algorithms and applications, 2019, http://dx.doi.org/10.48550/arXiv.1812.05905, arXiv:1812.05905, URL https://arxiv.org/abs/1812.05905.

[51] P. Adami, P.B. Rodrigues, P.J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, G. Lucas, Impact of VR-based training on human–robot interaction for remote operating construction robots, J. Comput. Civ. Eng. 36 (3) (2022) 04022006, http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0001016, URL https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29CP.1943-5487.0001016.