ERGONOMICALLY INTELLIGENT TELEOPERATION AND PHYSICAL HUMAN-ROBOT INTERACTION

by

Mojtaba Yazdani

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mechanical Engineering

The University of Utah

December 2022

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Mojtaba Yazdani**

has been approved by the following supervisory committee members:

Andrew S. Merryweather ,	Chair	**14 Nov 2022** Date Approved
Tucker Hermans ,	Chair	**14 Nov 2022** Date Approved
Jake J. Abbott ,	Member	Date Approved
John M. Hollerbach ,	Member	Date Approved
Jason Wiese ,	Member	**15 Nov 2022** Date Approved

by **Bruce K. Gale** , Chair/Dean of

the Department/College/School of **Mechanical Engineering**

and by **David B. Kieda** , Dean of The Graduate School.

ABSTRACT

Workplace injury is a concerning issue affecting millions of workers worldwide annually. Work-related musculoskeletal disorders (WMSDs) are the second most significant cause of disabilities worldwide, and awkward postures are known to be the main contributor WMSDs. Robotics, tele-portation, and human-robot collaboration are well-suited alternatives for high-risk tasks. However, WMSDs are still common among human teleoperators and operators physically interacting with robots in industries. To address these issues, we propose ergonomically intelligent systems for teleoperation and physical human-robot interactions that cover postural estimation, assessment, and optimization.

First, we explore estimating the 3D posture of the human using the interacting robot as the sensor. Later we combine it with the OpenPose markerless posture estimation method to prove an occlusion-robust algorithm for posture estimation. We model the posture estimation problem as a partially observable dynamical system and use a particle filter to infer the 3d posture from the sensory observations. We perform a human subject study to evaluate our posture estimation approaches.

Next, we use deep neural networks to learn differentiable and continuous ergonomic models based on RULA and REBA, popular and well-studied risk assessment tools in the ergonomics community. We evaluate them by assessing the postures in different types of tasks from two datasets of human motions from subject studies. Finally, we introduce a framework for postural optimization, which includes solutions for different types of teleoperation and physical human-robot interaction tasks. Here, we use gradient-free and gradient-based approaches to solve the optimizations, which benefit from our differentiable ergonomic models. We develop a simulated environment for teleoperation and physical human-robot interaction in which a human performs interactive tasks with robots while accepting and applying the suggested postural corrections from the postural optimization algorithm.

To Roya

For her advice, her patience, and her faith,

because she always understood.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Since I was introduced to robotics, I have always been passionate about human-robot interaction. I have always aimed to make robots more thoughtful and more intelligent around humans. I am proud and grateful that I had the opportunity to work on this area of research during my Ph.D. studies at the University of Utah Robotics Center. I was lucky that I could use the resources from both the robotics and ergonomics and safety programs to define and conduct my research.

I am writing to express my sincere gratitude to my supervisors, Andrew Merryweather and Tucker Hermans, for the invaluable opportunity of learning from them. I appreciate Andrew's support, kindness, friendship, and sincere commitment to helping his students during this part of my research journey. Thank you for allowing me to explore and broaden my expertise in this exciting multidisciplinary area of research. Thank you for pushing me to engage myself with different related research communities.

Tucker has always been my resource of motivation, and I cannot express how motivated I got after each meeting with him. I admire his commitment to providing learning and thinking opportunities for his students and his endless efforts to clarify complex concepts for the students. Our weekly reading groups were among my best times during my Ph.D. research. In addition to technical matters, he taught me how to do research, ask the right questions, correctly think about, define and formulate research problems, and think out of the box to find a solution. These valuable skills will help me in my future research journey. What I am grateful for the most is his patience with me and his belief and trust in me, even when my confidence and belief in myself were diluted. I truly enjoyed working under his guidance.

I extend my gratitude to the other members of my committee, John Hollerbach, Jake Abbott, and Jason Wiese, for their support and guidance during my research. Their insightful feedback on my work and valuable suggestions from the beginning of my work has helped me complete and succeed in this dissertation.

My profound appreciation is for my family, particularly my parents and brother in my home country, for their love, support, and understanding during my studies. Although we were far from each other, I continuously received their love and felt their support.

I am also thankful to all my friends at the University of Utah. I would like to especially thank LL4MA lab members and alums Balakumar Sundaralingam, Adam Conkey, Qingkai Lu, Kautilya Chenna, Griffin Tabor, Michael Bently, Mohanraj Devendran Shanthi, Martin Matak, Mark Van der Merwe, Matthew Wilson, Rebecca Miles, Yixuan Huang, Iain Lee and Bao Thach for their constant support, friendship, and making such a motivating, respectful, and collaborative research environment in the lab. I can write pages about how much I learned from each of you. I already miss you a lot and wish to see you again soon. Also, I want to express my appreciation to other students and alums of the Utah Robotics Center, David Usevitch, Jacob Harris, Xiang He, Navid Fallahinia, and Ali Samareh Filsoofi, for all the friendships, collaborations, and fun we had during the classes and course projects.

Moreover, I want to express my appreciation toward the robotics and human-robot interaction communities, especially Georgia Chalvatzaki, Serena Ivaldi, Tom Williams, Kerstin Haring, Karinne Ramirez-Amaro, Bilge Mutlu, Daisuke Sakamoto, and Kate Tsui for being such a supportive family for me and helping me grow in the community.

My deepest gratitude is to my beloved wife, Roya Sabbagh Novin. Sharing my personal and research life with you was the most valuable experience in my life. I enjoyed every second of it in both challenging and happy times. I appreciate all the sincere love and constant support from you. Thank you for constantly pushing me toward success. Your belief in me has kept my spirits and motivation high during this process. Thank you for being my code debugger and editor. Nevertheless, most of all, thank you for being my best friend. I owe you everything.

Moreover, to my furry research buddy, Coco: I cannot even imagine how I could finish this journey without you. You brought pure joy and love, and happiness to our lives. Although I could not resist your puppy eyes and had to pause my work several times a day to play with you, those moments of joy and comfort were a great resource of relaxation and freshness for me. Thank you for being such a good boy!

Finally, this research was partly supported by DARPA under grant N66001-19-2-4035 and the National Institute for Occupational Safety and Health under award number T420H008414-10. I want to acknowledge them for their support.

CHAPTER 1

INTRODUCTION

Workplace injury is a concerning issue that affects millions of workers worldwide annually and costs businesses billions in revenue. Work-related musculoskeletal disorders (WMSDs) are the second most significant cause of disabilities worldwide [1], and awkward postures are the main contributor WMSDs [2]. The high rate of injuries in industry [3], [4] has further motivated studies on these areas.

Robotics, teleoperation, and human-robot collaboration are well-suited alternatives for high-risk tasks (e.g., construction and handling hazardous materials) with different levels of autonomy. Humans can teleoperate high-risk and hazardous tasks and design remote workstations ergonomically for the human operator [5], or collaborate with robots to move heavy objects.

However, WMSDs are still common among human teleoperators, even when they perform teleoperation without force feedback [6], [7], and physical collaboration with robots in the industry is not intelligent enough yet to consider human comfort in robot behaviors. The industry 4.0 and 5.0 initiatives [8], [9] and the recent growth in the development of collaborative environments where intelligent agents (e.g., collaborative robots and VR systems) physically collaborate with humans to perform different types of tasks, highlights the growing importance of human comfort and ergonomics in physical collaboration with the new interactive technologies. Moreover, the fact that those intelligent agents share a physical interaction point with the operators allows the development of novel technologies to consider human comfort in robot behaviors.

A common and practical solution to improve workplace ergonomics is to perform postural estimation and risk assessments and analyze human comfort during the task. Ergonomists usually perform this process for the worst posture during the task by taking measurements of human posture onsite or from recorded images or videos. They usually use risk assessment tools [10] to assess the task and provide interventions such as user training [11], workstation modification [12] or task rotation [13]. However, in most cases, it is not possible to have an ergonomist present during the task to perform a comprehensive risk assessment, which results in missing events or only a partial understanding of awkward postures. In addition, these sampling strategies fail to detect the most hazardous combinations of force and posture, increasing the inaccuracy of an assessment. Finally, although postural correction by ergonomists generally helps reduce the risk of injuries, in many cases, the corrections are sub-optimal or tailored to specific task requirements, which might not generalize across task variability.

Early improvements include using motion capture (MoCap) systems to estimate the posture still requires significant time and effort for set up [14]. Moreover, putting markers on human operators can be inconvenient. Alternative markerless techniques are more adaptable, minimally intrusive, and less expensive. However, they need calibration to deal with errors and uncertainties from the sensors [15], [16]. Lighting, background color, and even the user's clothing can perturb vision-based markerless methods [17]. In teleoperation and physical human-robot interaction, the human operator works in such proximity to the interactive robot that it may cause occlusion leading to a reduction in the accuracy of the estimated posture as mentioned in [18]. The above issues and challenges define the first research question of this dissertation:

Q(1): *"How can we have a low-cost, non-intrusive, easy to set up, with a minimal number of sensors, and occlusion-robust posture estimation in teleoperation and physical human-robot interaction?"*

Postural optimization in teleoperation and physical human-robot interaction has recently received attention in research. However, the literature is minimal and sparse in this area. Developing a human comfort model lies at the heart of effective postural optimization. RULA (rapid upper limb assessment) [19] and REBA (rapid entire body assessment) [20] are the popular risk assessment tools in ergonomics, both of which mainly depend on the human posture. However, the discrete scores and plateaus in these risk assessment tools [18] make it challenging to use them in postural optimization. Quadratic approximations of those models have been the standard approach in the literature [18], [21], [22], which are easy to use in gradient-based postural optimization. However, these approximations deviate far from the scientifically validated assessments, causing doubt that they can reliably provide the same ergonomic benefit. Here we observe the potential for a differentiable ergonomics model that is as informative as popular risk assessment tools in the ergonomics community. The above concerns, limitations, and research gaps bring up the second and the third research questions of this dissertation:

Q(2): *"How can we formalize a framework for postural optimization tailored for different types of teleoperation and physical human-robot interaction?"*

Q(3): *"How can we develop computational ergonomics and human-comfort models that are as informative and well-studied as the popular risk assessment tools in ergonomics community, while being continuous, differentiable, and easy to use in gradient-based postural optimization?"*

In this dissertation, we tackle most of the above challenges by introducing the general framework of *Ergonomically Intelligent Teleoperation and Physical Human-Robot Interaction Systems*, which includes three main parts, each answering one of the above research questions:

- **Postural Estimation:** We initially propose estimating the 3D posture of the human in teleoperation and physical human-robot interaction solely from the trajectory of the interacting robot, and we extend it to an occlusion-robust and multi-sensory 3D posture estimation using the trajectory of the interacting robot and the 2D postures from one of the most popular markerless posture estimation methods.

- **Postural Assessment:** We develop DULA and DEBA, two highly accurate neural network models for ergonomics, which are differentiable, continuous, and easy to use in gradient-based postural optimization.

- **Postural Optimization:** We propose a general framework for postural optimization that we tailor for different types of teleoperation and physical human-robot interaction and provide an optimization tool and simulator benefiting from our differentiable ergonomics models.

Figure 1.1 visualizes the ergonomically intelligent teleoperation systems.

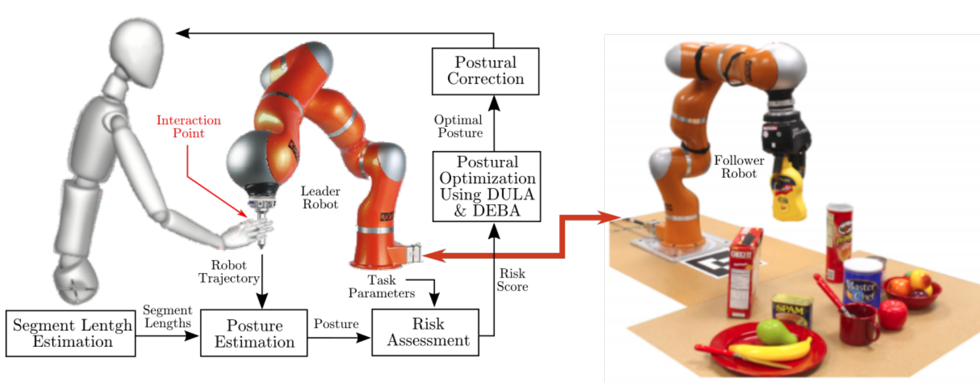


Figure 1.1: Ergonomically intelligent teleoperation systems.

Although in this dissertation, our implementation focus is on teleoperation as an exciting example of physical human-robot interaction, we can extend our solutions into other human-robot interaction applications with minor modifications.

1.1 Key contributions

This dissertation makes the following contributions:

1. We introduce the framework for ergonomically intelligent teleoperation and physical human-robot interaction system as an AI-based solution for reducing the risk of musculoskeletal disorders and improving workplace safety and ergonomics for human operators. The framework presents novel solutions for each part of the *estimation, assessment, and improvement* procedure in ergonomics.
2. We introduce the idea of using the trajectory of the interacting robot as the only sensor for estimating the human operator's segment lengths and joint angles. We use a system identification approach over some motion routines from the human operator to estimate the segment lengths. We formulate estimating the joint angles as a probabilistic partially observable dynamical system which we solve using a particle filter.
3. We propose a low-cost multi-sensory 3D posture estimation that is robust to occlusion. Our approach uses the observations from the interacting robot and a single RGB camera as the sensors. Here, we use the same probabilistic inference framework from the above.
4. We introduce DULA and DEBA, two neural-network models for ergonomic assessment, which are continuous, differentiable, and easy to use in gradient-based postural optimizations. We learned the models using a dataset from human postures labeled by the popular RULA and REBA ergonomic risk assessment tools. We use learned and generic models to assess human postures during different tasks.
5. We formulate a framework for postural optimization in teleoperation and physical human-robot interaction, which includes tailored approaches for different types of interactive tasks.
6. We develop a teleoperation and physical human-robot interaction simulated environment in ROS, which we use to evaluate our postural optimization approach.
7. We conduct two human-subject studies to evaluate our postural estimation and ergonomic assessment approaches.

We discuss all of the above contributions in Chapters 2–5. We follow up with a discussion and conclusion of this dissertation in Chapters 6 and 7.

1.2 References

[1] T. Vos, R. M. Barber, B. Bell, A. Bertozzi-Villa, S. Biryukov, I. Bolliger, F. Charlson, A. Davis, L. Degenhardt, D. Dicker *et al.*, “Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013,” *The Lancet*, vol. 386, no. 9995, pp. 743–800, 2015.

[2] B. R. da Costa and E. R. Vieira, “Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies,” *Amer. J. Ind. Medicine*, vol. 53, no. 3, pp. 285–323, 2010.

[3] S. P. Schneider, “Musculoskeletal injuries in construction: a review of the literature,” *Appl. Occupational Environmental Hygiene*, vol. 16, no. 11, pp. 1056–1064, 2001.

[4] W. Yu, T. Ignatius, Z. Li, X. Wang, T. Sun, H. Lin, S. Wan, H. Qiu, and S. Xie, “Work-related injuries and musculoskeletal disorders among factory workers in a major city of china,” *Accident Anal. Prevention*, vol. 48, pp. 457–463, 2012.

[5] P. G. Dempsey, L. M. Kocher, M. F. Nasarwanji, J. P. Pollard, and A. E. Whitson, “Emerging ergonomics issues and opportunities in mining,” *Int. J. Environmental Res. Public Health*, vol. 15, no. 11, p. 2449, 2018.

[6] L. Peternel, C. Fang, M. Laghi, A. Bicchi, N. Tsagarakis, and A. Ajoudani, “Human arm posture optimisation in bilateral teleoperation through interface reconfiguration,” in *IEEE RAS/EMBS Int. Conf. for Biomed. Robot. and Biomechatronics*, 2020.

[7] S. Yu, J. Lee, B. Park, K. Kim, and I. Cho, “Ergonomic analysis of a telemanipulation technique for a pyroprocess demonstration facility,” *Nucl. Eng. Technol.*, vol. 46, no. 4, pp. 489–500, 2014.

[8] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zühlke, “Human-machine-interaction in the industry 4.0 era,” in *2014 12th IEEE Int. Conf. on Ind. Inform. (INDIN)*. IEEE, 2014, pp. 289–294.

[9] P. K. R. Maddikunta, Q.-V. Pham, B. Prabadevi, N. Deepa, K. Dev, T. R. Gadekallu, R. Ruby, and M. Liyanage, “Industry 5.0: A survey on enabling technologies and potential applications,” *J. Ind. Inf. Integration*, p. 100257, 2021.

[10] M. Ramaganesh, R. Jayasuriyan, T. Rajpradeesh, S. Bathrinath, and R. Manikandan, “Ergonomics hazard analysis techniques-a technical review,” *Materials Today: Proc.*, vol. 46, pp. 7789–7797, 2021.

[11] A. Bazazan, I. Dianat, N. Feizollahi, Z. Mombeini, A. M. Shirazi, and H. I. Castellucci, “Effect of a posture correction–based intervention on musculoskeletal symptoms and fatigue among control room operators,” *Appl. Ergonom.*, vol. 76, pp. 12–19, 2019.

[12] A. A. Shikdar and M. A. Al-Hadhrami, “Smart workstation design: an ergonomics and methods engineering approach,” *Intl. J. Ind. Syst. Eng.*, vol. 2, no. 4, pp. 363–374, 2007.

[13] H. Motabar and A. D. Nimbarte, "The effect of task rotation on activation and fatigue response of rotator cuff muscles during overhead work," *Appl. Ergonom.*, vol. 97, p. 103461, 2021.

[14] M. Alvarez, D. Torricelli, A. J. del Ama, D. Pinto, J. Gonzalez-Vargas, J. C. Moreno, A. Gil-Agudo, and J. L. Pons, "Simultaneous estimation of human and exoskeleton motion: a simplified protocol," in *IEEE Int. Conf. on Rehabil. Robot.*, 2017, pp. 1431–1436.

[15] X. Xiao and S. Zarar, "A wearable system for articulated human pose tracking under uncertainty of sensor placement," in *IEEE Int. Conf. on Biomed. Robot. and Biomechatronics*, 2018, pp. 1144–1150.

[16] X. Chen and J. Davis, "Camera placement considering occlusion for robust motion capture," Computer Graphics Laboratory, Stanford University, Tech. Rep, Tech. Rep., 2000.

[17] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, and X. Twombly, "Vision-based hand pose estimation: A review," *Comput. Vision Image Understanding*, vol. 108, no. 1-2, pp. 52–73, 2007.

[18] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, "Postural optimization for an ergonomic human-robot interaction," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Syst. (IROS)*, 2017, pp. 2778–2785.

[19] L. McAtamney and E. N. Corlett, "RULA: a survey method for the investigation of work-related upper limb disorders," *Appl. Ergonom.*, vol. 24, no. 2, pp. 91–99, 1993.

[20] S. Hignett and L. McAtamney, "Rapid entire body assessment (REBA)," *Appl. Ergonom.*, vol. 31, no. 2, pp. 201–205, 2000.

[21] R. Rahal, G. Matarese, M. Gabiccini, A. Artoni, D. Prattichizzo, P. R. Giordano, and C. Pacchierotti, "Caring about the human operator: haptic shared control for enhanced user comfort in robotic telemanipulation," *IEEE Tran. Haptics*, vol. 13, no. 1, pp. 197–203, 2020.

[22] B. Busch, M. Toussaint, and M. Lopes, "Planning ergonomic sequences of actions in human-robot interaction," in *IEEE Int. Conf. on Robot. and Automat. (ICRA)*, 2018, pp. 1916–1923.

CHAPTER 2

IS THE LEADER ROBOT AN ADEQUATE SENSOR FOR POSTURE ESTIMATION AND ERGONOMIC ASSESSMENT OF A HUMAN TELEOPERATOR?

In this chapter and next, we explore posture estimation for ergonomically intelligent teleoperation and physical human-robot interaction (pHRI). In the first chapter, we investigate the hypothesis: "Is the leader robot an adequate sensor for posture estimation and ergonomic assessment of a human teleoperator?". This part of the work was presented at IEEE International Conference on Automation Science and Engineering (CASE) 2021, Lyon, France.

I primarily completed this work under the guidance of Prof. Tucker Hermans and Prof. Andrew Merryweather. Roya Sabbagh Novin assisted in system design and human subject studies.

2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)
August 23-27, 2021. Lyon, France

Is the Leader Robot an Adequate Sensor for Posture Estimation and Ergonomic Assessment of A Human Teleoperator?

Amir Yazdani*, Roya Sabbagh Novin*, Andrew Merryweather*, and Tucker Hermans†

2021 IEEE 17th International Conference on Automation Science and Engineering (CASE) | 978-1-6654-1873-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/CASE49439.2021.9551471

***Abstract*— Ergonomic assessment of human posture plays a vital role in understanding work-related safety and health. Current posture estimation approaches face occlusion challenges in teleoperation and physical human-robot interaction. We investigate if the leader robot is an adequate sensor for posture estimation in teleoperation and we introduce a new probabilistic approach that relies solely on the trajectory of the leader robot for generating observations. We model the human using a redundant, partially-observable dynamical system and we infer the posture using a standard particle filter. We compare our approach with postures from a commercial motion capture system and also two least-squares optimization approaches for human inverse kinematics. The results reveal that the proposed approach successfully estimates human postures and ergonomic risk scores comparable to those estimates from gold-standard motion capture.**

I. INTRODUCTION

Work-related musculoskeletal disorders (WMSDs) are the 2^{nd} largest cause of disabilities worldwide [1] and awkward postures are known to contribute to WMSDs. Teleoperation is a well suited alternative for high-risk tasks (e.g. construction and handling hazardous materials), since the the remote workstation for the human can be designed ergonomically [2]. However, WMSDs are still common among human operators, even when they perform teleoperation without force feedback [3], [4]. To improve ergonomics and lower the risk of WMSDs in teleoperation, we introduced *Ergonomically Intelligent Teleoperation Systems* in [5] and, in this paper, we focus on the problem of posture estimation to assess the ergonomics and risk of WMSDs in teleoperation tasks. Specifically, we investigate if the trajectory information of the leader robot provides adequate sensory information for probabilistic posture estimation for ergonomics assessment.

Posture estimation refers to the process of estimating the kinematic or skeletal configuration of the human body including segment lengths and joint angles. In addition to human biomechanics research, posture estimation is an important part of perception for smart agents (i.e. collaborative robots [6], companion mobile robots [7] and self driving cars [8]) interacting with the humans.

Collecting accurate and continuous posture data to determine a risk score over a work shift or entire task cycle can be tedious. Early improvements include using motion capture (MoCap) systems to estimate the posture and task parameters (e.g. frequency and duration of the task) still

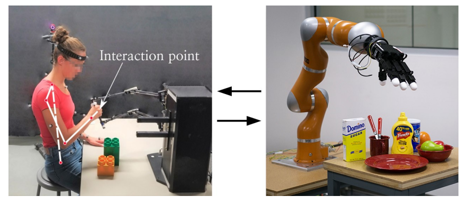


Fig. 1: Teleoperation setup for human subject experiments including a Quanser HD^2 haptic interface. Reflective markers on the participant's body are only used for comparison with a MoCap system.

require significant time and effort for set up [9]. Moreover, putting markers on human operators can be inconvenient. Alternative markerless techniques are more adaptable, minimally intrusive, and less expensive. However, they need calibration to deal with errors and uncertainties from the sensors [10], [11]. Vision-based markerless methods can also be perturbed by the lighting, background color and even the user's clothing [12]. In teleoperation, a human operator remotely controls the follower robot using the leader robot (see Fig. 1). Using the leader robot in such close proximity to the human increases occlusion and reduces the accurately in estimate posture as mentioned in [13].

Instead, we propose an alternative non-invasive and probabilistic approach that estimates the posture without using any additional sensors beyond the leader robot necessary for the teleoperation. Our approach could be used either standalone for monitoring the human teleoperator posture or in combination with other methods in a multi-modal sensory system to provide robust estimation during occlusions. In this paper, we formalize posture estimation as a probabilistic inference problem, in which we measure the leader robot's trajectory (pose and velocity) as the observation and infer the unobserved human posture (joint angles and angular velocities). We use the *circle point analysis* (CPA) [14] for segment length estimation of the human body. We also impose physical limits on joint angles and check the validity of the posture based on the posture-dependant ranges of human motion provided by [15]. We incorporate multiple observations over time enabling us to perform inference using a standard particle filter. Then, we use the estimated posture over the course of the task to assess the user's risk of WMSDs using RULA [16], a standard measure in the ergonomics and safety community. Moreover, we conduct a human subject study and compare the posture estimation results from the particle filter with the results from well-known deterministic solvers for least-squares optimization

*Department of Mechanical Eng. and Robotics Center, University of Utah, Salt Lake City, UT, USA, amir.yazdani@utah.edu
†School of Computing and Robotics Center, University of Utah, Salt Lake City, UT, USA, and Nvidia Corporation, Seattle, WA, USA

and the postures from the gold-standard MoCap system. Finally, we show that our posture estimation approach has a good accuracy in risk assessment, comparing to the postures from a MoCap system.

Dataset, code, and supplementary materials are available at https://sites.google.com/view/posture-estimation-in-teleop

II. Related Work

In physical human-robot interaction and teleoperation, researchers mainly have used external sensors (i.e. a vision system or IMU) to estimate a user's posture, especially for hand gestures [17]–[19]. The idea of solely using the leader robot's trajectory for posture estimation of human teleoperators has been introduced concurrently with this research by Rahal et al. in [20], where they solved the IK of the 7-DOF human arm. Unlike our probabilistic approach, which can encode a distribution of arm postures, they rely on heuristics to resolve the redundant IK. Their heuristic for redundancy resolution does not always hold across different tasks (e.g. some tasks might require the human teleoperator to use a working mode different from the working mode of the neutral posture) where the approach in [20] will fail.

The application of particle filters in human posture estimation is extensively discussed in the literature [21]–[23]. The sampling base of particle filters makes them well suited to human posture estimation due to their ability to handle the nonlinearities of human motion [24]. Moreover, the output estimation is a probabilistic distribution that can preserve different working modes of human posture. However, the high dimension of human motion requires a high number of particles to achieve accurate estimation.

Defining a model for human joint limits is challenging. Studies show that the range of motion (ROM) for a joint varies depending on the positions of other joints (inter-joint dependency) or other degrees-of-freedom in the same joint (intra-joint dependency) [15], [25] and vary by gender and person. Akhter et al. [26] used a dataset of recorded MoCap of human motion to develop a discontinous mathematical model for posture-dependant ROM and check the validity of a full-body posture. Jiang et al. [15] used the above model to label the validity of a set of randomly-generated postures and learned a differentiable neural network based on the generated data and used it as a constraint in the inverse kinematics optimization. Their arm model only includes shoulder and elbow and not the wrist. We use this learned network for checking the validity of the arm posture.

The literature highlights that evaluating ergonomics to improve working postures reduces the number of WMSDs [27], [28]. Among all the risk assessment tools, RULA [16] and REBA [29] rely mostly on the human posture (i.e. joint angles) and target the human upper body and whole body, respectively. This makes RULA more suitable for analyzing upper extremity tasks that are common during teleoperation.

III. Problem Statement

We seek to solve the problem of estimating the human joint-space trajectory in teleoperation using only the observed task-space poses and velocities of the leader robot. We

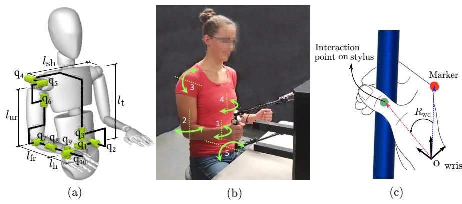


Fig. 2: (a) Kinematics model of human upper body, (b) Five motion routines for CPA segment length estimation, (c) Hand pose correction for MoCap in grasping the stylus using a fixed rigid-body transformation R_{wc}.

model the physical interaction between the human and the leader robot as an interaction point where the human kinematic chain makes contact with the robot's stylus (Fig. 1).

The state variables of the human include posture (joint angles) $\mathbf{q}$ and angular velocities $\dot{\mathbf{q}}$. They map into the state variables of the stylus through the kinematics of the human model parameterized by the segment length $\psi = [l_t, l_{sh}, l_{ur}, l_{fr}, l_h]$. We estimate ψ independently, prior to posture estimation. At each time step, the robot provides an observation as a task-space pose $\mathbf{z}$ and velocity $\dot{\mathbf{z}}$ of the stylus at the interaction point,

$$[\mathbf{z}_t; \dot{\mathbf{z}}_t] = h(\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi)) \tag{1}$$

where h is the observation function and ϕ is the forward kinematics of the human model. This defines only a partial observation of the human posture, because of redundancy in the human kinematics and a noisy measurement at the interaction point, which may change slightly during a task.

We seek to estimate $\boldsymbol{\tau} = [\mathbf{q}_t; \dot{\mathbf{q}}_t]_{t=1:T}$ given the stylus trajectory $\mathcal{Z} = [\mathbf{z}_t; \dot{\mathbf{z}}_t]_{t=1:T}$ that predicts a stylus pose closest to the observed stylus pose and obeys the human motion model f (Eqs. 3 and 4):

$$\begin{aligned} \tau^* = \arg\min_{\tau} \sum_{t=1}^{T} & ||\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi) - [\mathbf{z}_t; \dot{\mathbf{z}}_t]||^2_{\Sigma_1} + \\ & || [\mathbf{q}_t, \dot{\mathbf{q}}_t] - f(\mathbf{q}_{t-1}, \dot{\mathbf{q}}_{t-1})||^2_{\Sigma_2} \\ s.t. \qquad & \mathbf{q}_{\min} \leq \mathbf{q} \leq \mathbf{q}_{\max} \end{aligned} \tag{2}$$

where $\mathbf{q}_{\min}$ and $\mathbf{q}_{\max}$ are the joint limits. The high degree-of-freedom in human kinematics makes this problem a *redundant* problem with an infinite number of solutions. We seek the solution closest to the true posture of the human teleoperator.

IV. Approach

In this section, we provide an approximate solution for Eq. 2 by modeling the problem as partially observable posture estimation in teleoperation and solving it using a particle filter. We provide the kinematics model of the human upper body with only one moving arm. Next, we discuss adopting a particle filter for inference in our problem. Finally, we detail using CPA for segment length estimation.

A. Human kinematics model

We use a 10-DOF kinematics model (Fig. 2(a)) to analyze the upper body motion of a human sitting on a chair and operating the leader robot. We assume that the chair is stationary with a known position w.r.t. the robot, and the human sits on the center of the chair. The parameters of this model (ψ) include the length of each segment in the upper body model. We compare 3 techniques for segment length estimation: (1) *Full measurement*: manually measuring the segment lengths from anatomical landmarks on participant bodies [30], [31], (2) *Height measurement*: measuring the height of the participants and selecting the segment lengths fitting to 50-percentile populations from the ANSUR II anthropometric model ([32]), (3) *CPA analysis* where the lengths are calculated from some calibration motion routines described in Sec. IV-C.

We define the human's *state variables* as $\mathbf{q} = [q_i]_{i=1:10}$, $\dot{\mathbf{q}} = [\dot{q}_i]_{i=1:10}$ where q_i represents the angle of joint i (shown in Fig. 2(a)). We assume that the user's hand stays attached to the leader robot's stylus, as such we can transfer the pose of the hand from the human's frame to the robot's frame.

We encode human motion limits in two ways. First, we use fixed limits on the joint angles based on the biomechanics literature [33], [34]. If an angle estimate exceeds its limits, we project the estimate to the closest limit. Second, we use the learned and posture-dependant joint limit model from [15] to ensure the validity of the posture.

The estimated posture of the torso has a high effect on the estimated posture of the arm. Using the full range of motion for the torso would cause challenges in our posture estimation due to the four degrees of redundancy in the kinematics model. To overcome this issue, other researchers assumed that the torso posture is fully known and they only consider the arm [20]. Instead, we include the torso in the posture estimation problem by assuming that the torso stays close to the vertical position with a low variance. We incorporate this as a perturbance of the torso posture in the problem. This assumption is reasonable and was confirmed in our workstation where the human teleoperator sits behind a table interacting with a haptic interface.

From the kinematics of human motion, we find joint angles and velocities based on the previous step as follows:

$$\dot{\mathbf{q}}_k = \dot{\mathbf{q}}_{k-1} + \ddot{\mathbf{q}}_{k-1}dt \tag{3}$$

$$\mathbf{q}_k = \mathbf{q}_{k-1} + \dot{\mathbf{q}}_k dt \tag{4}$$

We model joint accelerations generated from a Gaussian distribution $\ddot{\mathbf{q}}_{k-1} \sim \mathcal{N}(0, \tilde{\mathbf{\Sigma}}_v)$. Since dt is fixed, setting $\mathbf{\Sigma}_v = \tilde{\mathbf{\Sigma}}_v \cdot dt$ transforms Eq. (3) to:

$$p(\dot{\mathbf{q}}_k \mid \dot{\mathbf{q}}_{k-1}) \sim \mathcal{N}(\dot{\mathbf{q}}_{k-1}, \mathbf{\Sigma}_v) \tag{5}$$

We model the observation likelihood function by a Gaussian distribution over the hand's pose and velocity as the end-effector of the human kinematic chain:

$$p([\mathbf{z}_k, \dot{\mathbf{z}}_k] \mid [\mathbf{q}_k, \dot{\mathbf{q}}_k]) = \mathcal{N}(\phi(\mathbf{q}_k, \dot{\mathbf{q}}_k, \psi), \mathbf{\Sigma}_K) \tag{6}$$

in which $\mathbf{\Sigma}_K$ is the kinematic covariance matrix.

B. Particle Filter for Posture Estimation

We approximate the solution for the partially observable problem of posture estimation by using a particle filter [35] with some modifications. As the estimation of the 10-DOF human model from the trajectory of the leader robot has high ambiguity due to the redundancy, we add the joint angular velocities to our state variables and use the velocity of the leader robot's stylus in our observations. As a prior, we encode that the human *starts* the task in a static, neutral posture as shown in Fig. 2(b). We initialize M particles using a truncated normal distribution with the mean at the neutral posture $\mathbf{q}_{neutral}$ and set the initial angular velocities to zero:

$$\mathbf{q}_0^{[m]} \sim \mathcal{N}(\mathbf{q}_{neutral}, \Sigma_0), \quad \dot{\mathbf{q}}_0^{[m]} = 0 \qquad m = 1, ..., M \tag{7}$$

where $\Sigma_0 = 0.2 \times (\mathbf{q}_{\max} - \mathbf{q}_{\min})$ for each joint.

Each particle is propagated in time based on the kinematics of human motion using Eqs. 3 and 4. Then, the particles are weighted based on the observation likelihood function in Eq. (6) defined as the innovation error between the estimated pose of the stylus and the observed pose from the leader robot. We use the multivariate Gaussian distribution to define the likelihood weighting function:

$$\begin{aligned} w_k^{[m]} =& v_p \cdot \det(2\pi\mathbf{\Sigma}_K)^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}([\mathbf{z}_k, \dot{\mathbf{z}}_k] - \\ & \phi(\mathbf{q}_k, \dot{\mathbf{q}}_k, \psi))^T \mathbf{\Sigma}_K^{-1}([\mathbf{z}_k, \dot{\mathbf{z}}_k] - \phi(\mathbf{q}_k, \dot{\mathbf{q}}_k, \psi))\} \end{aligned} \tag{8}$$

where $v_p \in \mathbb{R}, 0 \leq v_p \leq 1$, encodes the validity of the posture as output of the learned neural network from [15].

C. Circle Point Analysis for Segment Length Estimation

We estimate the segment lengths ψ through a calibration procedure variant of circle point analysis (CPA) [14]. Starting from the neutral posture, the user performs five predefined motion patterns that only include motion in one of their joints. When following such a pattern, we assume that the human hand will move on a circle. Estimating the circle parameters defines the location of the active joint and its distance from the end-effector. From this we derive the lengths of each arm segment.

Fig. 2(b) presents the five motion patterns used in data generation for CPA: (1) wrist flexion/extension to estimate hand length; (2) upper arm external/internal rotation to estimate forearm length; (3) upper arm abduction/adduction to estimate upper arm length; (4) rotation from the hip to estimate shoulder length; and (5) lateral bending from the hip to estimate torso length. We note that in estimating the last two segments we use the previously estimated arm and hand segment lengths.

V. Implementation & Experimental Protocol

We conducted a human subject experiment in which participants interact with a 6-DOF Quanser HD^2 haptic interface as the leader robot (see Fig. 1), and we recorded their upper body motion using a 12-camera Optitrack [36] MoCap system for comparison. We recruited 8 participants (4 female, 4 male) with ages ranging from 25 to 33 years and heights in

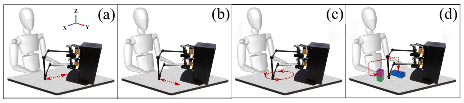


Fig. 3: The tasks for human subject study: repetitive motions following a straight line in the X direction (a), a straight line in the Y direction (b), a circular path (c), and repetitive motions between random sides of two blocks positioned at different heights, with an *unprescribed* motion and high range of hand rotation (d).

the range of 171 ± 21cm. Participants were graduate students from various programs and did not have any experience with teleoperation robots, and each received a 15-min training with the leader robot. Each participant performed 4 tasks visualized in Fig 3. We provided a printed visual guide on the table for the first three tasks, however, the participants were not required to follow the path accurately. The participants were not told what posture to initialize the task from and how high they should be above the table to do the task. The robot collected data from the participant's motion without exerting any force.

The gold-standard MoCap system estimates the upper-body posture for a 10-DOF torso, however, our human model only includes 3-DOF for the torso. To address this discrepancy, the MoCap posture is retargeted to our human model using the inverse kinematics. Moreover, the segment lengths are variable during a motion in MoCap data, while our model uses fixed lengths. This change is more visible in the forearm and upper arm, where we observed almost 2.3cm and 1.8cm of change, respectively, for a participant doing the circular task. This is mainly because the marker placement on the body will never be perfect, and motion is subject to some skin artifacts leading to this type of error [37]. Additionally, as shown in Fig. 2(c), the MoCap pose for the hand uses the segment from the wrist axis to the marker at the metacarpophalangeal joint of the index finger, while our human model uses the segment from the wrist joint to the interaction point. To correct for this, we use a fixed rigid-body transformation R_{wc} for the MoCap wrist joint calculated for each participant.

To compare with other well-known approaches, we solved the least-squares problem in Eq. (2) using two other deterministic methods: (1) boosted and bounded online least-squares IK optimization (*Online-IK*) in which we simply solve the inverse kinematics optimization independently at each time step by initializing it with the solution from the previous time step, and (2) boosted and bounded offline least-squares trajectory IK optimization (*Offline-TrajIK*) in which we solve the inverse kinematics problem for the whole trajectory by initializing it with the solution trajectory from the Online-IK. We used dogleg algorithm with rectangular trust regions from SciPy [38] as the optimization solver.

As neither marker-based nor markerless posture estimation techniques provide ground truth posture, we additionally provide qualitative analysis by overlaying the estimated posture on synchronized video frames. Fig. 9 shows the posture inferred by our approach aligns well with the MoCap

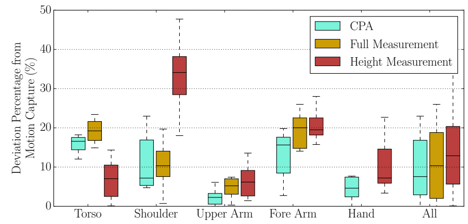


Fig. 4: Deviation of segments lengths from the MoCap lengths.

estimates. We see some error due to our fixed segment lengths assumption for MoCap motion. While other possible approaches of analysis exist (e.g. hand-labeling points [39]), these approaches are error prone and time-consuming.

In our implementation, we use a fixed number of particles (M=500), and used $\mathbf{\Sigma}_v = 0.01 \cdot \texttt{diag}(0.01, 0.01, 0.01, 0.05, 0.05, 0.05)$ and $\mathbf{\Sigma}_K = 0.01 \cdot \texttt{diag}(0.001, 0.001, 0.001, 0.05, 0.05, 0.05, 1, 1, 1, 10, 10, 10)$. To assess the risk, we developed a code based on RULA which uses both the estimated postures from our approach and the MoCap estimates as the input and outputs the RULA risk score. For that, we used the following assumptions for all tasks: the human is sitting on a chair, minimal intermittent force/load (less than 2.0Kg), muscle use occurrence less than 4x per minute, untwisted and vertical position for neck and torso, and supported legs and feet.

VI. Results & Discussion

This section provides results from our human subject experiments. We discuss the performance of the proposed posture estimation approach comparing with MoCap, as well as the estimated risk assessment results.

A. Segment Lengths Estimation

We compared the deviation[1] of estimated segment lengths from MoCap lengths using the various methods discussed in Section IV-A, among all participants in Fig. 4. The deviation for *full measurement* lengths of the hand is zero since the MoCap marker set did not provide a representative length for the hand. We used the *full measurement* value instead. The last three columns of the figure (All) shows the deviation for all of the segments. Statistical analyses reveal that *CPA* lengths deviate least from the MoCap lengths significantly[2]. The main reason that *CPA* deviates less than *full measurement* is that the anatomical landmarks used for manual measurement of the lengths are not necessarily on the joint axes calculated by the MoCap from the markers.

B. Initialization and Validation of Particles

Fig. 5 shows the evolution of all the particles in a trial of task (3) with circular motion, as well as the mean and standard deviation of their distributions. On the left, we

[1]We use the term “deviation” instead of “error” since the MoCap posture is also an estimate and not ground truth.

[2]We use $\alpha = 0.05$ for statistical analysis.

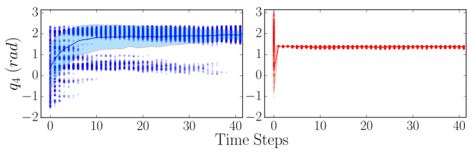


Fig. 5: Behaviour of particles through time for the shoulder abduction joint of subject 1 during the circular task. (left) Particles initialized uniformly over the ROM with higher diagonal values of Σ_K. (right) Particles initialized from a normal distribution with the mean at the neutral posture and lower diagonal values of Σ_K.

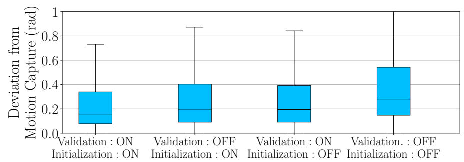


Fig. 6: The effect of initialization and posture validation on posture estimation in trials of the circular and the horizontal tasks.

initialized particles from a uniform distribution over a joint's ROM and used high values of kinematic covariance (0.1 for position, 0.5 for orientation). Here, particles from multiple modes are kept for about 40 steps, then they converge into a single mode. On the right, we initialize particles from a normal distribution around the neutral posture and the kinematic covariance matrix Σ_0 as described in the Section IV-A. The particles converge quickly to the correct mode. This shows the effect of initializing the particles around the neutral posture. Although we did not suggest the any initial posture to the participants, we observed that they all started from a posture close to their neutral posture independent of the type of the task they perform, which supports our idea of initialization of particles around the neutral posture.

Moreover, Fig. 6 shows the comparison between the effect of initialization around the neutral posture and validation of the posture based on the posture-dependant ROM on the deviation of the estimated postures from the MoCap postures. This is done for 4 trials of the task (1) and 4 trials of the task (3). We see that the combination of these two methods reduces the deviation, however, the effect of each one is almost equal. This implies that in some applications, it is safe to just use the initialization around the neutral posture instead of the computationally-expensive validation of particles.

C. Posture Estimation

Fig. 7 represents the deviation between the posture from our approach and the posture from MoCap for all the participants, and trials across the tasks using different segment length estimation methods. Overall, the approach generally agrees with a median deviation less than 0.09rad (less than 5deg) and upper quartile less than 0.25rad (less than 15deg) considering the observation solely from the stylus trajectory and having no extra sensors. Statistical analysis for comparing the effect of all three segment length estimation

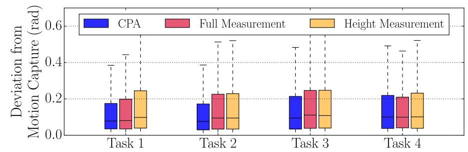


Fig. 7: Deviation of the posture estimated by the proposed approach from MoCap system for all the participants in different tasks.

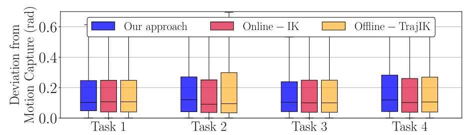


Fig. 8: Deviation of the posture estimated by our approach vs Online-IK and Offline-TrajIK methods among all the tasks.

methods on the posture estimation accuracy reveals that the estimated postures using *CPA* method has a significantly lower deviation from the MoCap postures, where there is no significant difference between the *full measurement* and *height measurement* methods. Hence, we used segment lengths from *CPA* method for the rest of the experiments.

Fig. 8 compares our posture estimation approach with two other least-squares IK solutions for redundant robots, for all participants among 4 tasks. From the statistical analysis of the results, we can conclude that our probabilistic *particle filter* approach performance is not significantly different than the other two methods.

Since there is no ground truth for human posture to compare with, we provide a qualitative evaluation using video frames of a participant during task (3) with overlaid reconstructed skeletons from MoCap (green) and our approach (red) in Fig. 9. The estimated postures aligns well with the MoCap postures during the entire task. Videos of overlaid skeletons for all four tasks are available in the supplementary video. Although the leader robot's trajectory is smooth, the videos show non-smooth estimations from our approach, which is due to the characteristics of the particle filter and plotting the skeleton for the most-probable posture.

D. Risk Assessment

To represent the benefits of our probabilistic approach, we plot the expected value of RULA score and its standard deviation over time during task (1) using the posture estimates from our particle filter in Fig. 10. Unlike the deterministic estimators, our probabilistic approach provides an estimated distribution over posture and hence, a distribution over estimated RULA scores which can be used by probabilistic human-aware planning methods in p-HRI and teleoperation.

Fig. 11 shows the maximum RULA score during a task (the one most often used in ergonomics) for the postures from our approach and MoCap postures. Our approach was successful in identifying all instances where the RULA score was higher than 2 (i.e. future investigation or change may be or is needed) in all 32 trials. The experiment resulted in the

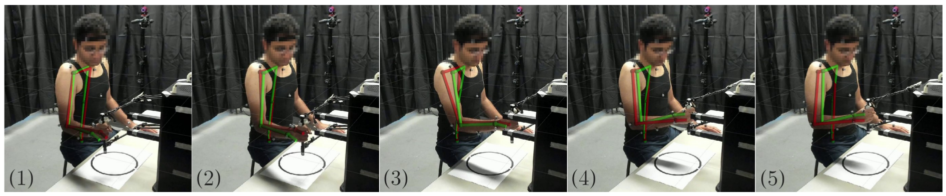


Fig. 9: Video-overlaid skeletons show the posture (most probable particle) estimated from our proposed approach (red) and MoCap (green).

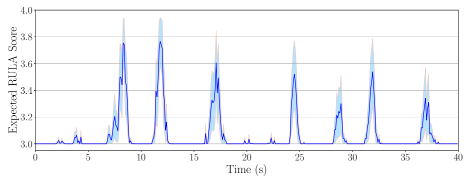


Fig. 10: Expected RULA score (solid line) and standard deviation (shaded region) over time for participant 1 performing task (1).

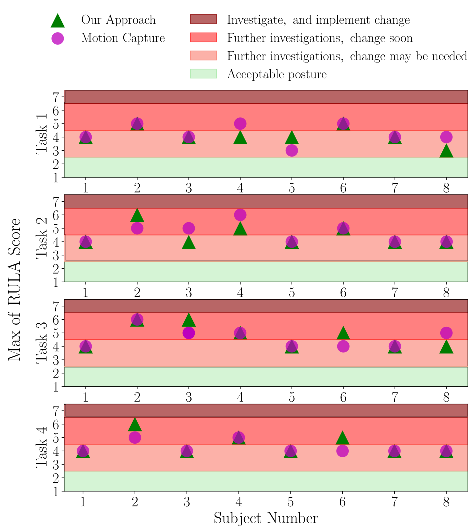


Fig. 11: Comparison of the maximum value of RULA scores of a task (using estimated posture from our approach and MoCap for all the participants and trials.

same interpretation of the RULA score in 27 trials (84.37%) and the same RULA score in 21 trials (65.63%). Our approach estimates the same RULA score (not the maximum value for task) across all trajectories and participants with median accuracy of more than 86.4% for tasks (1) and (2), and more than 74.7% for tasks (3) and (4).

VII. Discussion and Future Direction

Overall, the results showed that the proposed posture estimation solely from the leader robot has the potential to be used for continuous monitoring of ergonomics in teleoperation. It is accurate enough to rise alerts when further investigation or change in the task is required.

In this paper, we assumed a seated human teleoperator on a chair with known relative position to the leader robot, but we can extend our approach to a standing human by adding the foot pose as a state variable. Here, we focused on teleoperation applications, which has a high prevalence of WMSDs, However, it is possible to extend it to other physical human-robot interaction tasks such as co-manipulation and learning from demonstration by a human expert, in order to assess the risk of WMSDS and also to enhance the learning performance by including the human's posture and comfort during demonstration. Our approach also can be used for full arm representation in VR systems based on the pose of controllers and headset. It would be straightforward to combine our proposed approach with other sensing modalities (e.g., vision or MoCap) when available in a specific application context by integrating these additional observations into the particle filter weighting scheme.

Future work will focus on increasing the complexity of the human model (e.g. adding muscle activation). Additionally, we are examining the combination of our approach with OpenPose [40] using incremental smoothing method [41] in order to improve estimation results, and decrease runtime.

VIII. Conclusion

In this paper, we investigated if a leader robot is an adequate sensor to continuously monitor posture to assess the ergonomics of a human-teleoperated task. We described a probabilistic approach which is based solely on the data recorded from a leader robot's end-effector, that is already necessary to perform the teleoperation task. We compared our approach with well-known *Online-IK* and *Offline-TrajIK* methods for human inverse kinematics. We used CPA to estimate segment lengths and utilized RULA for risk assessment.

Our results show that the leader robot can be an adequate sensor for human posture estimation and ergonomic assessment in applications such as continuous monitoring of the human teleoperator, while our assumptions are valid. Moreover, the probabilistic distribution of postures and their corresponding risk scores can be used in human-aware planning for shared autonomy and pHRI.

IX. Aknowledment

This work is supported by DARPA under grant N66001-19-2-4035.

References

[1] T. Vos, R. M. Barber, B. Bell, A. Bertozzi-Villa, S. Biryukov, I. Bolliger, F. Charlson, A. Davis, L. Degenhardt, D. Dicker, *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013," *The Lancet*, vol. 386, no. 9995, pp. 743–800, 2015.

[2] P. G. Dempsey, L. M. Kocher, M. F. Nasarwanji, J. P. Pollard, and A. E. Whitson, "Emerging ergonomics issues and opportunities in mining," *International journal of environmental research and public health*, vol. 15, no. 11, p. 2449, 2018.

[3] L. Peternel, C. Fang, M. Laghi, A. Bicchi, N. Tsagarakis, and A. Ajoudani, "Human arm posture optimisation in bilateral teleoperation through interface reconfiguration," in *IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob 2020)*, 2020.

[4] S. Yu, J. Lee, B. Park, K. Kim, and I. Cho, "Ergonomic analysis of a telemanipulation technique for a pyroprocess demonstration facility," *Nuclear Engineering and Technology*, vol. 46, no. 4, pp. 489–500, 2014.

[5] A. Yazdani and R. Sabbagh Novin, "Posture Estimation and Optimization in Ergonomically Intelligent Teleoperation Systems," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 2021, pp. 604–606.

[6] A. M. Zanchettin, N. M. Ceriani, P. Rocco, H. Ding, and B. Matthias, "Safety in human-robot collaborative manufacturing environments: Metrics and control," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 2, pp. 882–893, 2016.

[7] D. Fridovich-Keil, A. Bajcsy, J. F. Fisac, S. L. Herbert, S. Wang, A. D. Dragan, and C. J. Tomlin, "Confidence-aware motion prediction for real-time collision avoidance1," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 250–265, 2020.

[8] K. Mangalam, E. Adeli, K.-H. Lee, A. Gaidon, and J. C. Niebles, "Disentangling human dynamics for pedestrian locomotion forecasting with noisy supervision," in *The IEEE Winter Conference on Applications of Computer Vision*, 2020, pp. 2784–2793.

[9] M. Alvarez, D. Torricelli, A. J. del Ama, D. Pinto, J. Gonzalez-Vargas, J. C. Moreno, A. Gil-Agudo, and J. L. Pons, "Simultaneous estimation of human and exoskeleton motion: a simplified protocol," in *IEEE Int. Conf. on Rehabilitation Robotics*, 2017, pp. 1431–1436.

[10] X. Xiao and S. Zarar, "A wearable system for articulated human pose tracking under uncertainty of sensor placement," in *IEEE Int. Conf. on Biomedical Robotics and Biomechatronics*, 2018, pp. 1144–1150.

[11] X. Chen and J. Davis, "Camera placement considering occlusion for robust motion capture," Computer Graphics Laboratory, Stanford University, Tech. Rep, Tech. Rep., 2000.

[12] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, and X. Twombly, "Vision-based hand pose estimation: A review," *Computer Vision and Image Understanding*, vol. 108, no. 1-2, pp. 52–73, 2007.

[13] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, "Postural optimization for an ergonomic human-robot interaction," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 2778–2785.

[14] B. W. Mooring, Z. S. Roth, and M. R. Driels, *Fundamentals of manipulator calibration*. Wiley New York, 1991.

[15] Y. Jiang and C. K. Liu, "Data-driven approach to simulating realistic human joint constraints," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1098–1103.

[16] L. McAtamney and E. N. Corlett, "RULA: a survey method for the investigation of work-related upper limb disorders," *Applied ergonomics*, vol. 24, no. 2, pp. 91–99, 1993.

[17] P. Vartholomeos, N. Katevas, A. Papadakis, and L. Sarakis, "Design of motion-tracking device for intuitive and safe human-robot physical interaction," in *IEEE Int. Conf. on Telecommunications and Multimedia*, 2016, pp. 1–6.

[18] J. Buzzi, E. De Momi, and I. Nisky, "An uncontrolled manifold analysis of arm joint variability in virtual planar position and orientation tele-manipulation," *IEEE Transactions on Biomedical Engineering*, 2018.

[19] G. Martinez, I. A. Kakadiaris, D. Magruder, and I. Magruder, "Teleoperating ROBONAUT: A case study," in *British Machine Vision Conference*, 2002, pp. 1–10.

[20] R. Rahal, G. Matarese, M. Gabiccini, A. Artoni, D. Prattichizzo, P. R. Giordano, and C. Pacchierotti, "Caring about the human operator: haptic shared control for enhanced user comfort in robotic telemanipulation," *IEEE Transactions on Haptics*, vol. 13, no. 1, pp. 197–203, 2020.

[21] R. J. Mozhdehi, Y. Reznichenko, A. Siddique, and H. Medeiros, "Deep convolutional particle filter with adaptive correlation maps for visual tracking," in *IEEE Int. Conf. on Image Processing*, 2018, pp. 798–802.

[22] T. Zhang, S. Liu, C. Xu, B. Liu, and M. Yang, "Correlation particle filter for visual tracking," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2676–2687, 2018.

[23] R. Van Der Merwe, A. Doucet, N. De Freitas, and E. A. Wan, "The Unscented Particle Filter," in *Advances in neural information processing systems*, 2001, pp. 584–590.

[24] R. Poppe, "Vision-based human motion analysis: An overview," *Computer vision and image understanding*, vol. 108, no. 1-2, pp. 4–18, 2007.

[25] X. Wang, M. Maurin, F. Mazet, N. D. C. Maia, K. Voinot, J. P. Verriest, and M. Fayet, "Three-dimensional modelling of the motion range of axial rotation of the upper arm," *Journal of biomechanics*, vol. 31, no. 10, pp. 899–908, 1998.

[26] I. Akhter and M. J. Black, "Pose-conditioned joint angle limits for 3D human pose reconstruction," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1446–1455.

[27] S. A. Ismail, S. B. M. Tamrin, M. R. Baharudin, M. A. M. Noor, M. H. Juni, J. Jalaludin, and Z. Hashim, "Evaluation of two ergonomics intervention programs in reducing ergonomic risk factors of musculoskeletal disorder among school children," *Res J Med Sci*, vol. 4, no. 1, pp. 1–10, 2010.

[28] Z. Khodabakhshi, S. A. Saadatmand, M. Anbarian, and R. Heydari Moghadam, "An ergonomic assessment of musculoskeletal disorders risk among the computer users by rula technique and effects of an eight-week corrective exercises program on reduction of musculoskeletal pain," *Iranian Journal of Ergonomics*, vol. 2, no. 3, pp. 44–56, 2014.

[29] S. Hignett and L. McAtamney, "Rapid entire body assessment (REBA)," *Applied ergonomics*, vol. 31, no. 2, pp. 201–205, 2000.

[30] S. Pheasant and C. M. Haslegrave, *Bodyspace: Anthropometry, ergonomics and the design of work*. CRC press, 2018.

[31] W. T. Dempster, "Space requirements of the seated operator, geometrical, kinematic, and mechanical aspects of the body with special reference to the limbs," Michigan State Univ East Lansing, Tech. Rep., 1955.

[32] R. C. Fromuth and M. B. Parkinson, "Predicting 5th and 95th percentile anthropometric segment lengths from population stature," in *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering*, 2009, pp. 581–588.

[33] P. K. Levangie and C. C. Norkin, "Joint structure and function; a comprehensive analysis. 3rd," *Philadelphia: FA. Davis Company*, 2000.

[34] "NASA man-system integration standard," Available at https://msis.jsc.nasa.gov/sections/section03.htm (09/16/2020).

[35] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.

[36] "NaturalPoint Inc., Corvalis, OR," https://optitrack.com/ (09/16/2020).

[37] C. Metcalf, C. Phillips, A. Forrester, J. Glodowski, K. Simpson, C. Everitt, A. Darekar, L. King, D. Warwick, and A. Dickinson, "Quantifying Soft Tissue Artefacts and Imaging Variability in Motion Capture of the Fingers," *Annals of Biomedical Engineering*, pp. 1–11, 2020.

[38] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, *et al.*, "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nature methods*, vol. 17, no. 3, pp. 261–272, 2020.

[39] V. Athitsos and S. Sclaroff, "Estimating 3D hand pose from a cluttered image," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, 2003, pp. II–432.

[40] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: realtime multi-person 2d pose estimation using part affinity fields," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 1, pp. 172–186, 2019.

[41] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216–235, 2012.

CHAPTER 3

OCCLUSION-ROBUST MULTISENSORY POSTURE ESTIMATION IN PHYSICAL HUMAN-ROBOT INTERACTION

3D human posture estimation is critical in human-related sciences such as biomechanics, ergonomics, sports medicine, entertainment, game development, and human-robot interaction. It requires a sufficiently accurate and robust motion tracker that can be easily used to evaluate human movements [1]. Traditionally, marker-based motion capture systems (MoCap) were the primary approach in biomechanics studies due to the high frequency of the sensors and their accuracy [2]. However, due to their personal, environmental, and technical constraints [3–8], new markerless methods are being developed that benefit from deep-learning techniques to extract human posture from images and videos [9–11]. Among them, OpenPose [12] is a popular markerless method for human motion analysis due to its ease of use and integration.

Markerless techniques are more adaptable, minimally intrusive, and less expensive than marker-based systems. However, they require time-consuming calibration to deal with errors and uncertainties from the sensors [13], [14]. Lighting, background color, and even the user's clothing can also perturb vision-based markerless methods [15], which will cause inaccurate segment lengths and joint angles. More importantly, occlusion is a significant issue, which leads to losing data and the inability to track human motion and collect accurate and continuous posture data [16], [17]. In posture estimation, occlusion might happen by an object in the environment masking a part of the camera view or a human body part masking other segments of the body (self-occlusion) [18]. Although using multiple cameras with optimal placements [19] can reduce occlusion, it requires a more sophisticated and expensive system, which needs to be tailored for specific working environments and applications.

One important application of 3D human posture estimation is human motion assessment in workplaces to reduce the risk of work-related musculoskeletal disorders (WMSDs) [20], [21].

WMSDs are the second-largest source of disabilities worldwide [22], and awkward postures are the main contributor to the work-related injuries [23–25]. Teleoperation, as an example of physical human-robot interaction (pHRI), is a well-suited alternative for high-risk tasks (e.g., construction and handling hazardous materials) since we can design the remote workstation for the human ergonomically [26]. However, WMSDs are still common among human operators, even when they perform teleoperation without force feedback [27], [28].

In teleoperation, a human operator remotely controls the follower robot using the leader robot (see Figure 3.1). Using the leader robot in such proximity to the human increases occlusion and reduces the accuracy of posture estimation via markerless methods as mentioned in [17]. However, it also provides the ability to use the leader robot stylus as an additional sensor, as we initially proposed in [29]. In this chapter, we focus on the teleoperation application of pHRI. However, we can extend our proposed approach and implementations with minor modifications to other applications with direct human and robot interactions.

To improve ergonomics and lower the risk of WMSDs in teleoperation, we introduced a *ergonomically intelligent teleoperation system* in [30], [31], in which the teleoperation system intelligently estimates the human posture during interaction with the leader robot (more details in [29]), assesses the posture using ergonomics risk assessment tools, and if the risk of WMSDs is high, provides an ergonomically optimal postural correction for the human to complete the task (more details in [32]). We also introduced different types of autonomous postural correction in teleoperation using the leader robot. Figure 3.1 illustrates the details of the framework.

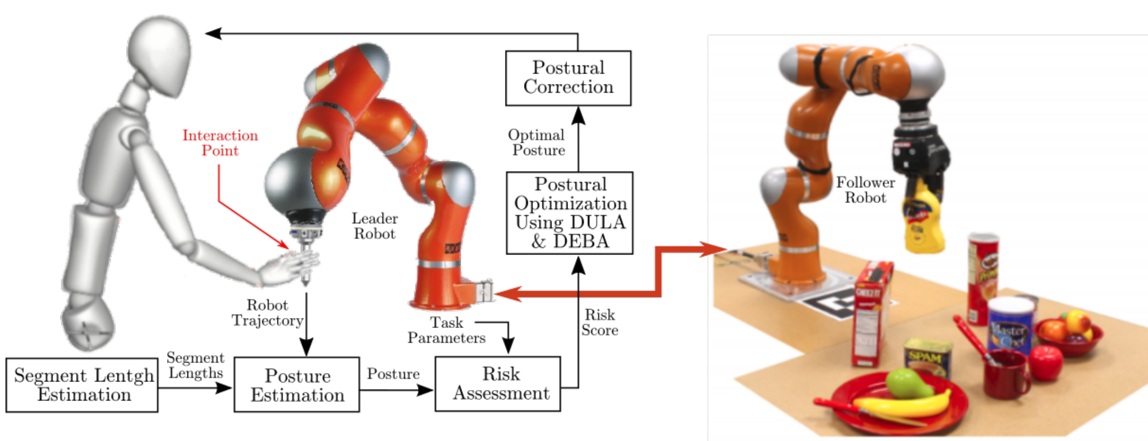


Figure 3.1: The framework for ergonomically intelligent teleoperation system. The same framework is true for other pHRI applications just by removing the follower robot.

This chapter focuses on improving posture estimation performance to assess the ergonomics and risk of WMSDs in teleoperation tasks. Previously [29], we showed that the leader robot is an adequate sensor for posture estimation in teleoperation, and we can use high-resolution sensory information from the robot encoders for continuous, low-cost, and occlusion-free monitoring of the 3D posture and risk of WMSDs in teleoperation with adequate accuracy. However, to deal with the under-determined and ambiguous solutions, we had to limit the torso range of motion based on preliminary information from the type of tasks and initialize the estimation approach around the neutral posture. Here, we improve our previous approach by combining it with a markerless posture estimation approach to provide occlusion-robust posture estimation with improved redundancy resolution.

We use the popular OpenPose [12] method with the input of single-view video frames from an RGB camera as our additional sensory information. OpenPose is a markerless posture estimation approach that estimates the 2D postures of the human in the camera coordinate frame from a single camera. They use a non-parametric representation, referred to as part affinity fields, to learn to associate body parts with individuals in the image. Figure 3.2 shows an example of estimated 2D postures from OpenPose. Although OpenPose estimates the postures in an online form, it suffers from sensor uncertainty, model accuracy, and occlusion [33]. Moreover, recovering the 3D postures from 2D OpenPose postures without additional sensory information would not be accurate because of the redundancy problem.

To benefit from both sensory information from the robot and OpenPose, we propose to combine

Figure 3.2: An example of estimated postures from OpenPose [12].

them as a multi-sensory posture estimation system to provide an accurate occlusion-robust estimation in teleoperation that can also handle the issue with redundancy. We formalize posture estimation as a probabilistic inference problem, in which we measure the leader robot's trajectory (pose and velocity) and the 2D position of anatomical landmarks of the body from OpenPose in the camera coordinate frame as the observation and infer the unobserved 3D human posture (joint angles and angular velocities). The combination of a simple RGB camera and the leader robot (which is already available from the teleoperation with no extra cost) would still be low-cost compared to MoCap systems, depth sensors, or multiple camera systems.

We compare the combined method with posture estimation solely from either the robot's trajectory or the 2D postures from OpenPose. Similar to our previous work, we use *circle point analysis (CPA)* for segment length estimation of the human body, as we showed that it has higher accuracy compared to other approaches [29]. We also impose physical limits on joint angles and check the validity of the posture based on the posture-dependant ranges of human motion provided by [34]. We incorporate multiple observations over time, enabling us to perform inference using a standard particle filter. Then, we use the estimated posture throughout the task to assess the user's risk of WMSDs using RULA [35], a standard ergonomics and safety measure in the community. Moreover, we conduct a human-subject study and compare the posture estimation results from the approaches mentioned above and the postures from the gold-standard Optitrack[1] MoCap system. Finally, we show that our posture estimation approach has good accuracy in risk assessment compared to the postures from a MoCap system.

We summarize the contributions of this chapter as the following:

- Modeling the 3D posture estimation in teleoperation using sensory information from the leader robot's trajectory and 2D postures from OpenPose as a partially observable dynamic system.
- Inferring the 3D posture from the observations using a particle filter.
- Conducting a human-subject study to evaluate and compare the performance of each approach in estimating the 3D posture and ergonomics assessment in teleoperation.

We structure the remainder of this chapter as follows. We first review the related literature in the field and focus on the recent state-of-the-art work. We then formalize the 3D posture estimation

[1] https://optitrack.com/

using three types of observations: solely from the robot, solely from OpenPose, and multisensory observation from the robot and OpenPose. Later, we introduce our approach to solve the partially observable posture estimation problem and provide details on our human-subject study. To conclude, we present and discuss the results of our experiment.

3.1 Related work

This section provides an overview of the literature on posture estimation in human-robot interaction and its key challenges and components.

3.1.1 Posture estimation in human-robot interaction

Human posture estimation has been an essential tool in various areas of human-robot interaction such as socially assistive robots [36], [37], pHRI [38], [39], teleoperation [40–42], engagement analysis [43], human-aware decision making and planning [44], and situation awareness [45]. Posture estimation in pHRI has been used mainly as a tool to derive other metrics for HRI evaluation. For instance, researchers have discussed ergonomics assessment and safety-based factors such as joint overloading [46], [47], the separation distance between the robot and human [48], and muscular fatigue [49]. In teleoperation, estimated posture has been used to directly control a remote robot [50–52]. The idea of estimating the 3D posture of a human teleoperator only from the trajectory of the leader robot was introduced concurrently in our previous work [29] and by Rahal et al. [53]. Rahal et al. [53] solved the inverse kinematics (IK) of the 7-DOF human arm holding the stylus of a leader robot in teleoperation, assuming the known position of the human's shoulder. Later, Vianello et al. [39] used the same idea to predict human posture in pHRI, given the robot trajectory executed in a collaborative scenario. They formalized the problem as predicting the human's joint velocities given the current posture and robot end-effector velocity over demonstrated human movements.

3.1.2 Redundancy resolution of the human skeletal model

The inverse kinematics problem of human kinematics has infinite solutions due to the redundancy of the kinematic chain. General methods for inverse kinematics of redundant systems use the pseudo-inverse of the Jacobian matrix to minimize the end-effector error using, e.g., the Newton-Raphson method [54]. Several studies have focused on the improvement of the solution to the problem. For instance, Kim et al. [55] formulated the inverse kinematics problem by defining the swivel angle: the rotation angle of the plane defined by the upper and lower arm around a virtual axis that connects the shoulder and wrist joints. They analyzed reaching tasks recorded with a motion capture system indicating that the swivel angle is selected such that when the elbow joint is flexed, the palm points to the head. They used this criterion to resolve the redundancy of the human arm. Rahal et al. [53] used the same formulation of swivel angle and used a heuristic that minimizes the deviation from neutral posture to resolve the redundancy. The above approaches for redundancy resolution do not always hold across different tasks and, therefore, cannot be generalized to all applications. Another researcher tried to include the intrinsic principles of human arm motion, including both skeletal kinematics and muscle strength properties [56]. Although their approach results in bio-mimetic arm postures, the relationship between elements of joint torques in their skeletal model needs more investigation, and the muscle effort in their model is realistic only for static postures.

Recently, learning-based approaches have been used to resolve the redundancy of the human arm and generate human-like motions [39], [57–59]. For instance, authors in [39] learned the distribution of the null space of the Jacobian and the weights of the weighted pseudo-inverse from demonstrated human movements. In [60], authors designed a one-class support-vector-machine model to classify human-like postures. Then, they used the redundancy characteristic of a 7-DOF robotic arm with a linear regression model to enhance the search for human-like postures.

In all of the above literature, the redundancy of the 7-DOF human arm model was the target. However, in our posture estimation in teleoperation, we consider the redundancy of the full 10-DOF upper-extremity of the human skeleton, including the torso and arm. Although going deep into the redundancy resolution of human arm models is not the scope of our work, in our previous work [29], we used the prior information about the task, including the torso being almost vertical and initiating the neutral posture to resolve the redundancy. Here in this chapter, we explore the effect of using low-cost multiple sensory information to disambiguate the redundancy of the human skeletal model.

3.1.3 Human joint range of motion

Another challenge in human posture estimation is defining the joint ranges of motion. The literature shows that the range of motion for a joint varies depending on the positions of other joints (inter-joint dependency) or other degrees-of-freedom in the same joint (intra-joint dependency) and varies by gender and person [34], [61]. Akhter et al. [62] developed a discontinuous mathematical model for the posture-dependant range of motion based on recorded MoCap of human motions and checked the validity of a full-body posture. Later, Jiang et al. [34] used the above model to label the validity of a set of randomly generated postures, learned a differentiable neural network based on the generated data and used it as a constraint in the inverse-kinematics optimization. However, their arm model is limited to the shoulder and elbow and does not include the human wrist. Our work uses their learned model on top of the standard human range of motion from biomechanics [63], [64].

3.1.4 Postural assessment in human-robot interaction

Ergonomics risk assessment of tasks and workplaces is an important part of reducing the number of WMSDs [65], [66]. Ergonomists have provided simple and easy-to-calculate risk assessment tools, which are common in practice. NASA TLX [67], RULA (rapid upper limb assessment) [35], REBA (rapid entire body assessment) [68], strain index [69], and ACGIH TLV [70] are among the common risk assessment tools in ergonomics. These models are supported by extensive human-subject studies that validate their effectiveness in reducing ergonomic risk factors [66]. pHRI researchers have also proposed computational models for human comfort and ergonomics, including peripersonal space [71], muscle fatigue [72], joint overloading [73], and muscular comfort [71]. RULA and REBA are among the most popular risk assessment tools in the ergonomics community, which both rely mostly on the human posture (i.e., joint angles) and provide quantitative scores. The inputs to both models are the joint angles of the human, and they output discrete, integer scores in which a higher number indicates a higher risk of WMSDs, i.e., lower human comfort [74].

Moreover, there are interpretations for different ranges of RULA scores that provide the final assessment results. Those features make RULA the most suitable assessment tool to assess the common upper extremity tasks in teleoperation and pHRI. We also proposed DULA and DEBA, the differentiable and continuous versions of RULA and REBA risk assessment tools in our previous work [32] to be used in gradient-based postural optimization in pHRI as a part of our ergonomically intelligent system framework. We use the generic RULA for assessing posture in this chapter.

3.2 Problem statement

Estimating the human posture includes two main parts: estimating the segment lengths and estimating the joint angles. In this section, we formulate these problems based on different types of observations and define their corresponding parameters. First, we define the kinematics of the human skeletal model. Then, we formulate the 3D posture estimation from different types of sensory information. Finally, we define the problem of estimation of segment lengths.

3.2.1 Human kinematics model

We model the kinematics of a human sitting on a chair and interacting with the leader robot with a 10-DOF kinematics chain, starting from the torso and ending with the right hand, as shown in Figure 3.3. We assume that the chair is stationary with a known relative position to the robot, and the human sits in the center of the chair.

We use the circle point analysis (CPA) from our previous work [29] to estimate the segment lengths. We define the human's *state variables* as $\mathbf{q} = [q^j]_{j=1:10}$, $\dot{\mathbf{q}} = [\dot{q}^j]_{j=1:10}$ where q^j represents the angle of joint j (shown in Figure 3.3). We assume that the user's hand stays attached to the leader robot's stylus; as such, we can use a fixed transformation matrix to transfer the pose of the hand from the human's frame to the robot's frame.

From the kinematics of human motion, we find joint angles and velocities based on the previous step from function f as follows:

$$\begin{bmatrix} \mathbf{q}_t \\ \dot{\mathbf{q}}_t \end{bmatrix} = f([\mathbf{q}_{t-1}; \dot{\mathbf{q}}_{t-1}; \ddot{\mathbf{q}}_{t-1}]) = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{q}_{t-1} \\ \dot{\mathbf{q}}_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \ddot{\mathbf{q}}_{t-1} \Delta t \end{bmatrix} \tag{3.1}$$

We model joint accelerations generated from a Gaussian distribution $\ddot{\mathbf{q}}_{t-1} \sim \mathcal{N}(0, \tilde{\boldsymbol{\Sigma}}_v)$. Since dt is fixed, setting $\boldsymbol{\Sigma}_v = \tilde{\boldsymbol{\Sigma}}_v \cdot dt$ results in:

$$p(\dot{\mathbf{q}}_t \mid \dot{\mathbf{q}}_{t-1}) \sim \mathcal{N}(\dot{\mathbf{q}}_{t-1}, \boldsymbol{\Sigma}_v) \tag{3.2}$$

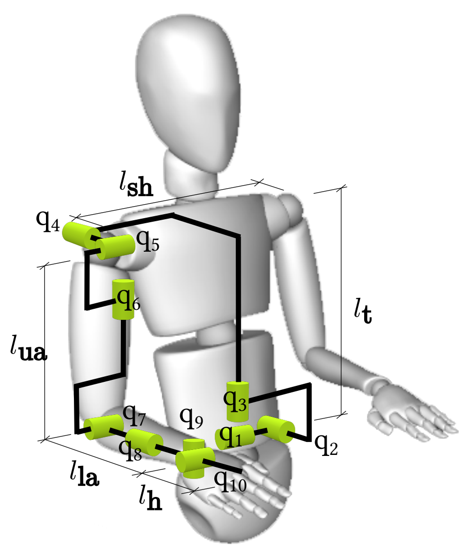


Figure 3.3: Kinematics model of the 10-DOF human upper body, adapted from [29].

3.2.2 3D posture estimation solely from OpenPose 2D postures

At each time step, OpenPose provides the 2D position of the keypoints corresponding to anatomical landmarks on the human body in the camera coordinate frame. To recover the 3D posture, we seek to find the 3D posture trajectory $\tau = [\mathbf{q}_t; \dot{\mathbf{q}}_t]_{t=1:T}$ such that its projection into the 2D camera frame matches the posture from OpenPose $(\mathbf{u}, \mathbf{v})$ and obeys the human motion model during the task. We can formulate the 3D posture estimation from 2D OpenPose postures as

$$\begin{aligned} \tau^* = \arg\min_{\tau} \sum_{t=1}^{T} (\sum_{m=1}^{M} & ||\mathcal{P} \cdot \phi(\mathbf{q}_t^{1:m}, \psi) - (u^m, v^m)||_{\Sigma_1}^2) + \\ & || [\mathbf{q}_t; \dot{\mathbf{q}}_t] - f([\mathbf{q}_{t-1}; \dot{\mathbf{q}}_{t-1}; \ddot{\mathbf{q}}_{t-1}])||_{\Sigma_2}^2, \\ s.t. \qquad \mathbf{q}_{\min}^j \leq \mathbf{q}^j \leq \mathbf{q}_{\max}^j \qquad & \text{for } j \text{ in } [1, 2, \ \ldots, \text{ number of joints}] \end{aligned} \tag{3.3}$$

where $\psi = [l_{\text{torso}}, l_{\text{shoulder}}, l_{\text{upper arm}}, l_{\text{lower arm}}, l_{\text{hand}}]$ are the segment lengths depicted in Figure 3.3, $\phi(\mathbf{q}_t, \psi)$ is the forward kinematics of human based up to the grasping point at the posture $\mathbf{q}$, f is the human motion model (see Eq. 3.1 for more details), Σ_1 is the Gaussian covariance for keypoint noise in the image space, Σ_2 is the covariance matrix for joint angles and their angular velocities, and $\mathbf{q}_{\min}^j$ and $\mathbf{q}_{\max}^j$ are the joint limits for joint j. We encode the human range of motion as truncated fixed limits on the joint angles based on the biomechanics literature [**?**], [64]. Moreover, we use the learned and posture-dependant joint-limit model from [34] to ensure the validity of the posture. Unlike our previous work, here we use the full range of motion from the biomechanics literature and do not narrow the torso range of motion.

If an occlusion occurs in the scene, OpenPose still tries to provide the estimated position of occluded keypoints based on the position of other keypoints unless it is impossible. We visualize one of those moments where the wrist and lower arm ‘are occluded in Section 3.5.1. These wrong estimated 2D keypoints will result in a high error in the estimated 3D posture.

3.2.3 3D posture estimation solely from robot’s trajectory

When the leader robot (or the interacting robot in general pHRI) is the only sensor for posture estimation, we seek to solve the problem of estimating the human joint-space trajectory in teleoperation using only the observed task-space poses and velocities of the leader robot. Another observation available by default at each time step is the estimated velocity at the previous time step. We model the physical interaction between the human and the leader robot as an interaction point where the human kinematic chain makes contact with the robot’s stylus. Figure 3.4 shows the physical interaction point between the human and the robot.

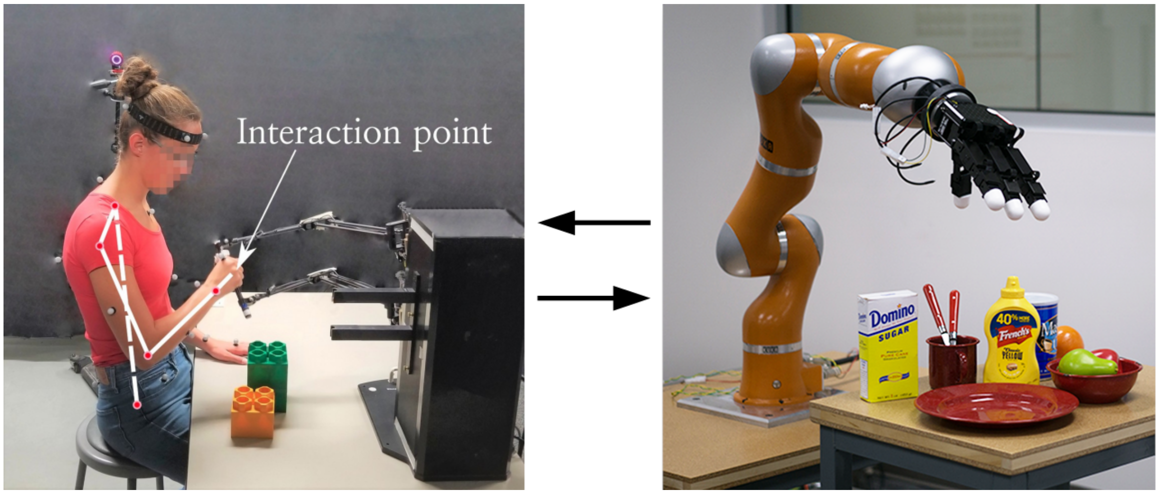


Figure 3.4: Teleoperation setup for human-subject experiments including a Quanser HD^2 haptic interface, adapted from [29].

At each time step, the robot provides an observation as a task-space pose $\mathbf{z}$ and velocity $\dot{\mathbf{z}}$ of the stylus at the interaction point,

$$[\mathbf{z}_t; \dot{\mathbf{z}}_t] = h(\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi)) \tag{3.4}$$

where h is the observation function, including the transfer of the human pose into the corresponding robot stylus pose in the robot's coordinate frame. The above formulation defines only a partial observation of the human posture because of redundancy in the human kinematics and a noisy measurement at the interaction point, which may change slightly during a task.

We seek to estimate $\tau = [\mathbf{q}_t; \dot{\mathbf{q}}_t]_{t=1:T}$ given the stylus trajectory $\mathcal{Z} = [\mathbf{z}_t; \dot{\mathbf{z}}_t]_{t=1:T}$ that predicts a stylus pose closest to the observed stylus pose and obeys the human motion model f (Eq. 3.1):

$$\begin{aligned}
\tau^* = \arg\min_{\tau} \sum_{t=1}^{T} & \|\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi) - [\mathbf{z}_t; \dot{\mathbf{z}}_t]\|^2_{\Sigma_3} + \\
& \| [\mathbf{q}_t; \dot{\mathbf{q}}_t] - f([\mathbf{q}_{t-1}; \dot{\mathbf{q}}_{t-1}; \ddot{\mathbf{q}}_{t-1}])\|^2_{\Sigma_2} + \\
& \|\dot{\mathbf{q}}_t - \dot{\mathbf{q}}_{t-1}\|^2_{\Sigma_4}, \\
s.t. \qquad & \mathbf{q}^j_{\min} \leq \mathbf{q}^j \leq \mathbf{q}^j_{\max} \qquad \text{for } j \text{ in } [1, 2, \ \ldots, \text{ number of joints}]
\end{aligned} \tag{3.5}$$

where Σ_3 is the covariance matrix for position and orientation elements, and Σ_4 is the covariance matrix for joint angular velocity elements. The high degree of freedom in human kinematics makes this problem a *redundant* problem with infinite solutions. We seek the solution closest to the true posture of the human teleoperator.

3.2.4 Multisensory 3D posture estimation from OpenPose and robot's trajectory

We model the multisensory posture estimation from a single RGB camera, and the robot as the combination of the above formulations:

$$
\begin{aligned}
\tau^* = \arg\min_{\tau} \sum_{t=1}^{T} & ||\phi([\mathbf{q}_t;\dot{\mathbf{q}}_t],\psi) - [\mathbf{z}_t;\dot{\mathbf{z}}_t]||^2_{\Sigma_1} + \\
& \sum_{m=1}^{M} (||\mathcal{P}\cdot\phi(\mathbf{q}_t^{1:m},\psi) - (u^m, v^m)||^2_{\Sigma_3}) + \\
& ||\,[\mathbf{q}_t;\dot{\mathbf{q}}_t] - f([\mathbf{q}_{t-1};\dot{\mathbf{q}}_{t-1};\ddot{\mathbf{q}}_{t-1}])||^2_{\Sigma_2} + \\
& ||\dot{\mathbf{q}}_t - \dot{\mathbf{q}}_{t-1}||^2_{\Sigma_4}, \\
s.t. \qquad \mathbf{q}^j_{\min} \le \mathbf{q}^j \le \mathbf{q}^j_{\max} \qquad & \text{for } j \text{ in } [1, 2, \ \ldots, \ \text{number of joints}]
\end{aligned}
\tag{3.6}
$$

All of the above problems are *redundant* problems, which generally have infinite solutions. Our goal is to find the solutions closest to the human's actual posture.

3.2.5 Segment length estimation

Segment length estimation using vision-based systems for 3D posture estimation is straightforward. Each segment length is calculated based on the 3D position of the tracked markers at the end of the segment. However, this process still suffers from inaccuracy due to imperfect marker placement and skin artifacts in MoCap systems [8], and model, sensor, and marker tracking uncertainties in markerless approaches [33].

We can define the estimation of segments lengths solely from the robot as

$$
\psi^* = \arg\min_{\psi} \sum_{t=1}^{T} |\phi(\mathbf{q}_t,\psi) - \mathbf{z}_t|^2, \tag{3.7}
$$

When the 3D posture from OpenPose is the only sensory information, we can formulate the segment length estimation as

$$
\psi^* = \arg\min_{\psi} \sum_{t=1}^{T}\sum_{m=1}^{M} |\mathcal{P}\cdot\phi(\mathbf{q}_t^{1:m},\psi) - (u^m, v^m)|^2, \tag{3.8}
$$

where M is the number of related keypoints from OpenPose, $\mathbf{q}_t^{1:m}$ is the set of joint angles contributing to the pose of the keypoint m, and $\mathcal{P}$ is the projection matrix of the camera that transfers 3D positions of keypoints in the world frame into 2D pixels (u, v) in the camera frame.

Solving the above optimizations is challenging and cumbersome when the posture $\mathbf{q}$ is unknown. Hence usually, they are solved for segment lengths and joint angles simultaneously. We assume that the segment lengths are known for the following parts of this section.

3.3 Approach

In this section, we introduce approximate solutions for Eqs. 3.3, 3.5, and 3.6 by solving the partially observable posture-estimation problem using a particle filter. Initially, we model the observation from the sensors. Later, we discuss adopting a particle filter for inference in posture estimation.

3.3.1 Observation likelihood

When we use the robot trajectory as the observation, we model the observation likelihood function by a Gaussian distribution over the hand's pose and velocity as the end-effector of the human kinematic chain:

$$p([\mathbf{z}_t, \dot{\mathbf{z}}_t] \mid [\mathbf{q}_t, \dot{\mathbf{q}}_t]) = \mathcal{N}(\phi(\mathbf{q}_t, \dot{\mathbf{q}}_t, \psi), \mathbf{\Sigma}_K) \tag{3.9}$$

in which $\mathbf{\Sigma}_K$ is the kinematic covariance matrix.

When using the 2D posture from OpenPose as the observation, we define the observation likelihood function by a Gaussian distribution over the 2D projected position of each keypoint:

$$p([u_t^m; v_t^m] \mid [\mathbf{q}_t; \dot{\mathbf{q}}_t]) = \mathcal{N}(\mathcal{P} \cdot \phi(\mathbf{q}_t^{1:m}, \psi), \mathbf{\Sigma}_P) \tag{3.10}$$

where $\mathbf{\Sigma}_P$ is the projection covariance matrix.

When combining the two observations as a multisensory posture estimation, then

$$p(\begin{bmatrix} \begin{bmatrix} u_t^1 \\ v_t^1 \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_t^M \\ v_t^M \end{bmatrix} \\ \begin{bmatrix} \mathbf{z}_t \\ \dot{\mathbf{z}}_t \end{bmatrix} \end{bmatrix} \mid \begin{bmatrix} \mathbf{q}_t \\ \dot{\mathbf{q}}_t \end{bmatrix}) = \mathcal{N}(\begin{bmatrix} \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:1}, \psi) \\ \vdots \\ \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:M}, \psi) \\ \phi(\mathbf{q}_t, \dot{\mathbf{q}}_t, \psi) \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma}_P \\ \vdots \\ \mathbf{\Sigma}_P \\ \mathbf{\Sigma}_K \end{bmatrix}) \tag{3.11}$$

3.3.2 Particle filter for posture estimation

We approximate the solution for the partially observable posture estimation problems in Section 3.2 by using a particle filter [75]. We initialize particles from a uniform distribution over the joint range of motion and set the initial angular velocities to zero. For the rest of the chapter, the superscript in the bracket over the parameters represents the corresponding particle number.

$$\mathbf{q}_0^{[n]} \sim \mathcal{U}(\mathbf{q}_{\min}, \mathbf{q}_{\max}), \quad \dot{\mathbf{q}}_0^{[n]} = 0 \qquad n = 1, ..., N, \tag{3.12}$$

where N is the number of particles.

Each particle is propagated in time based on the kinematics of human motion using Eq. 3.1. Then, the particles are weighted based on their corresponding observation likelihood functions from Eqs. 3.9, 3.10, or 3.11 defined as the *innovation error* between the estimated pose of the stylus and the observed pose from the leader robot, and/or the projected 2D estimate of keypoints and observed keypoints from OpenPose. We use the multivariate Gaussian distribution to define the likelihood weighting function for each case:

- Observation solely from the robot:

$$\begin{aligned} w_t^{[n]} =& v_p \cdot \det(2\pi\mathbf{\Sigma}_K)^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}([\mathbf{z}_t, \dot{\mathbf{z}}_t]- \\ & \phi(\mathbf{q}_t, \dot{\mathbf{q}}_t, \psi))^T \mathbf{\Sigma}_K^{-1}([\mathbf{z}_t, \dot{\mathbf{z}}_t] - \phi(\mathbf{q}_t, \dot{\mathbf{q}}_t, \psi))\} \end{aligned} \tag{3.13}$$

- Observation solely from OpenPose:

$$\begin{aligned} w_t^{[n]} =& v_p \cdot \det(2\pi\mathbf{\Sigma}_P)^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\begin{bmatrix} \begin{bmatrix} u_t^1 \\ v_t^1 \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_t^M \\ v_t^M \end{bmatrix} \end{bmatrix} - \\ & \begin{bmatrix} \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:1}, \psi) \\ \vdots \\ \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:M}, \psi) \end{bmatrix})^T \mathbf{\Sigma}_P^{-1}(\begin{bmatrix} \begin{bmatrix} u_t^1 \\ v_t^1 \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_t^M \\ v_t^M \end{bmatrix} \end{bmatrix} - \begin{bmatrix} \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:1}, \psi) \\ \vdots \\ \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:M}, \psi) \end{bmatrix})\} \end{aligned} \tag{3.14}$$

- Observation from the robot and OpenPose

$$\begin{aligned} w_t^{[n]} =& v_p \cdot \det(2\pi \begin{bmatrix} \mathbf{\Sigma}_P \\ \vdots \\ \mathbf{\Sigma}_P \\ \mathbf{\Sigma}_K \end{bmatrix})^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\begin{bmatrix} \begin{bmatrix} u_t^1 \\ v_t^1 \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_t^M \\ v_t^M \end{bmatrix} \\ \begin{bmatrix} \mathbf{z}_t \\ \dot{\mathbf{z}}_t \end{bmatrix} \end{bmatrix} - \\ & \begin{bmatrix} \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:1}, \psi) \\ \vdots \\ \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:M}, \psi) \\ \phi(\mathbf{q}_t, \dot{\mathbf{q}}_t, \psi) \end{bmatrix})^T \begin{bmatrix} \mathbf{\Sigma}_P \\ \vdots \\ \mathbf{\Sigma}_P \\ \mathbf{\Sigma}_K \end{bmatrix}^{-1} (\begin{bmatrix} \begin{bmatrix} u_t^1 \\ v_t^1 \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_t^M \\ v_t^M \end{bmatrix} \\ \begin{bmatrix} \mathbf{z}_t \\ \dot{\mathbf{z}}_t \end{bmatrix} \end{bmatrix} - \begin{bmatrix} \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:1}, \psi) \\ \vdots \\ \mathcal{P} \cdot \phi(\mathbf{q}_t^{1:M}, \psi) \\ \phi(\mathbf{q}_t, \dot{\mathbf{q}}_t, \psi) \end{bmatrix})\} \end{aligned} \tag{3.15}$$

where $v_p \in R$, $0 \leq v_p \leq 1$, encodes the validity of the posture as the output of the learned neural network from [34].

Each particle is weighted by the observation likelihood function when any observation is received. After normalizing the weights, we re-sample particles based on their weights; we generate more new particles around the particles with higher weights. This process repeats during the tasks, and we report the *most probable particle* as the estimated posture at each time.

3.4 Implementation and experimental protocol

Here, we provide details on our approach, experimental setup, and human-subject study to estimate the 3D posture of a human operator performing teleoperation.

3.4.1 Human-subject study

We conducted a human-subject experiment in which participants interacted with a 6-DOF Quanser HD^2 haptic interface[2] as the leader robot (see Figure 3.4). We recorded their upper-body motion using a 12-camera Optitrack MoCap system as the gold standard posture estimation method. We also used a single RGB camera to record videos of the participants and used OpenPose on top of that to generate the 2D posture projected into the camera frame.

We recruited eight participants (four females, four males) with ages ranging from 25 to 33 years and heights in the range of 150 to 192cm. Participants were graduate students from various programs at the University of Utah. They were informed about the study through a recruiting email sent to graduate students, and they volunteered to participate in our study. The study did not include any financial compensation for the participants. Participants participated under informed consent. Then, we measured participants' segment lengths by hand and put reflective markers on the anatomical landmarks on their bodies by the researchers trained by one of the PIs, an expert in biomechanics and MoCap systems.

None of the participants had any experience with the teleoperation of robots, and each one received a 15-min training with the teleoperation system. Then, to perform the segment length estimation by the CPA method, they followed five motion routines discussed in [29] while sitting on the chair and holding the robot's stylus in hand.

Finally, we asked each participant to perform four different tasks visualized in Figure 3.5. As the

[2]https://www.quanser.com/products/hd2-high-definition-haptic-device/

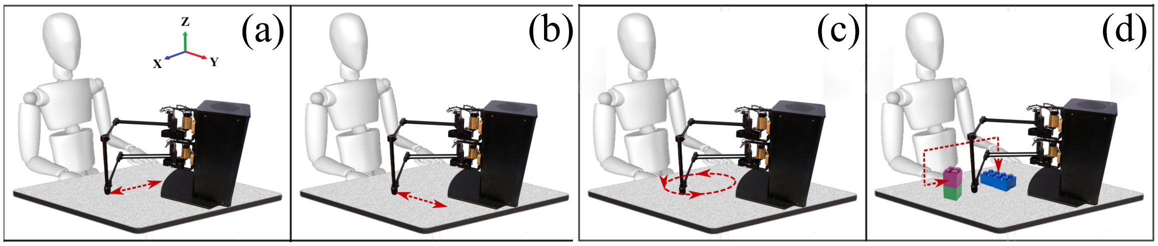


Figure 3.5: The tasks for human-subject study: repetitive motions following a straight line in the *X* direction (a), a straight line in the *Y* direction (b), a circular path (c), and repetitive motions between random sides of two blocks positioned at different heights, with an *unprescribed* motion and high range of hand rotation (d) [29].

study focuses on the interaction between the human and the leader robot, we just provided a printed visual guide on the table in front of the participant for the first three tasks to provide a guideline to follow in order to perform the required motion. However, the participants were not required to follow the path accurately. We did not tell the participants what posture to initialize the task from and how high they should be above the table to do the task. They chose those based on their comfort. The robot collected data from the participant's motions without exerting force on their hands.

As neither marker-based nor markerless posture estimation techniques provide ground truth posture, we provide qualitative analysis by overlaying the estimated postures and the MoCap postures on synchronized video frames shown in Section 3.5.1.

The human skeleton used in the MoCap system is much more complicated than the 10-DOF kinematics model we use, especially in the torso area. To address this discrepancy, we re-targeted the MoCap posture onto the simple 10-DOF kinematic model by calculating the pose of a virtual torso link from the hip to the neck of the MoCap skeleton. However, our model uses fixed segment lengths, and the MoCap skeleton uses variable segment lengths. MoCap calculates the segment lengths from the relative position of markers at each time, which are prone to change due to skin artifacts [8] and imperfect marker placement. Because of this, we see that the skeleton representing the re-targeted MoCap posture does not always follow the human arm in the video frames. So it should be considered when evaluating the qualitative results in Section 3.5.1.

3.4.2 OpenPose and camera calibration

Our developed ROS package runs OpenPose in the background and provides the keypoint positions in the camera frame based on the BODY-25 output format shown in Figure 3.6. For this research, we use keypoint #1, 2, 3, 4, and 8, representing the human torso and the right arm.

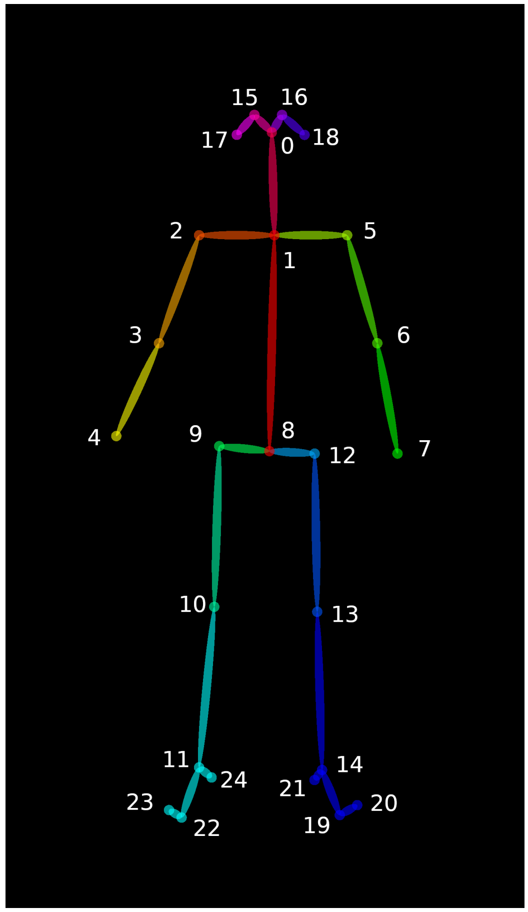
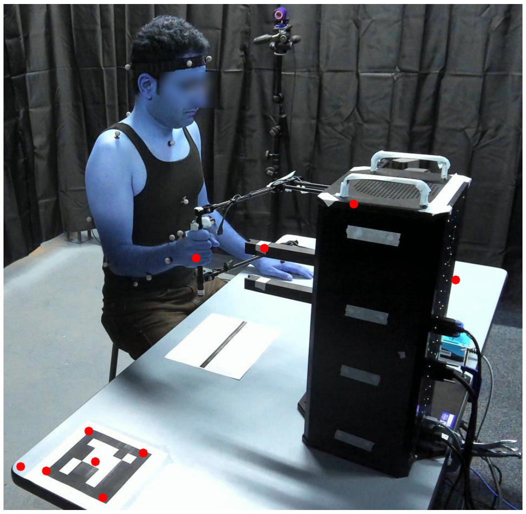
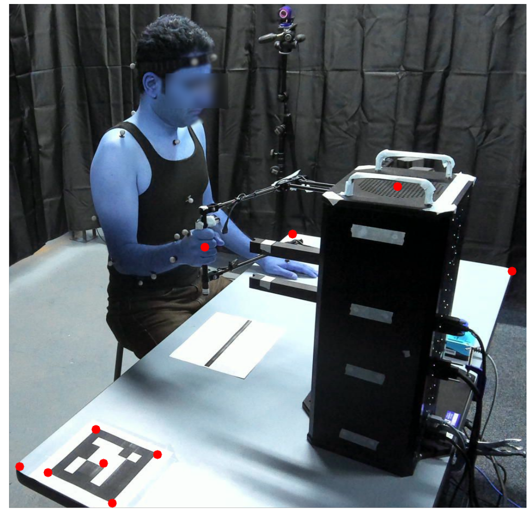


Figure 3.6: Set of keypoints for BODY-25 model in OpenPose [76].

Camera calibration is critical in calculating the particle weights based on the deviation between the projected anatomical landmarks with the OpenPose keypoints. We calibrate the camera intrinsics using a checkerboard. For extrinsic calibration, we used two approaches: (1) using an ArUco marker attached to the table with a known position and (2) using the known 3D position of 10 points in the scene to calibrate the camera extrinsic parameters based on the common pose estimation method from [77]. Figure 3.7 illustrates the projection of some points in the scene with known 3D positions (including table corners, robot top corners, the center of the stylus, and corners on the ArUco marker) into the camera frame using different extrinsic calibration approaches. You can see that the projected points using the known 3D points match better with the original points in the scene, so we use that calibrated extrinsic parameters for our work. Using this, we calculated the projection matrix $\mathcal{P} = P_I T_E$, where P_I and T_e are intrinsic and extrinsic parameters of the camera, respectively. We used the above projection matrix in Sections 3.2 and 3.3.



(a) Camera calibration by the ArUco marker.



(b) Camera calibration by points with the known 3D position.

Figure 3.7: Projected points from the scene with known 3D points into the camera image using different calibration methods.

To evaluate the robustness to occlusion for 3D posture estimation solely from OpenPose and our multisensory approach, We re-estimate the postures. At the same time, we occluded some parts of the scene covering using OpenCV in the observation pipeline to edit the video frames fed to the OpenPose. Figure 3.8 visualizes the artificial occlusion in the video frames and its effect on 2D posture estimated by OpenPose.

3.4.3 Particle filter details

In our implementation, we used a fixed number of particles (N=500) and used the following parameters:

$$\mathbf{\Sigma}_v = \mathtt{diag}(2, 2, 2, 5, 5, 5, 5, 5, 5, 5) \quad (\frac{deg}{sec})$$

$$\mathbf{\Sigma}_K = 0.01 \cdot \mathtt{diag}(0.01, 0.01, 0.01, 0.01, 0.01, 0.05, 0.1, 0.1, 0.1, 10, 10, 10) \quad ([m, rad, \frac{m}{sec}, \frac{rad}{sec}])$$

$$\mathbf{\Sigma}_P = 0.01 \cdot \mathtt{diag}(3, 3, 3, 3, 3) \quad (pixel)$$

We achieved the above numbers by tuning the particle filters on one trial of the subject (#1) performing task (a), and we used the same numbers for all other trials.

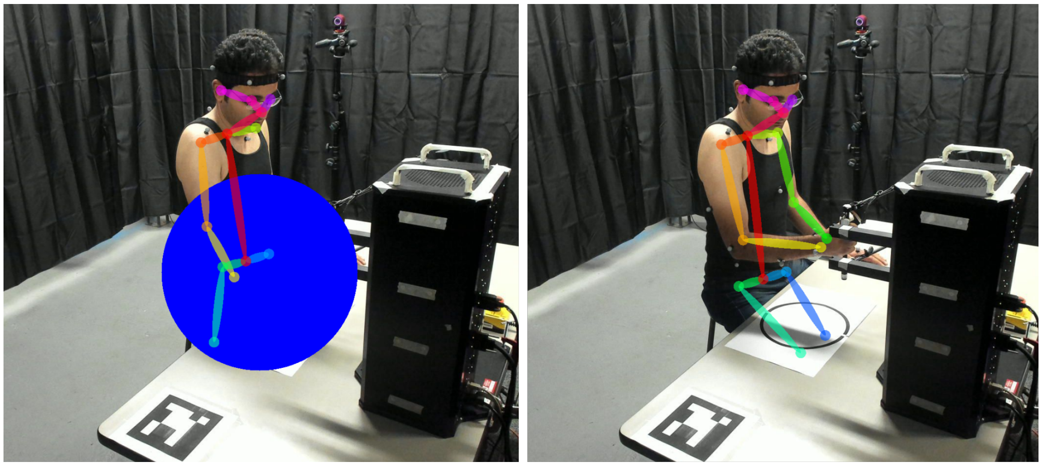


Figure 3.8: To evaluate the robustness to occlusion, we edited the video frames to make artificial occlusion in the video input for OpenPose. You can see the 2D OpenPose posture from the occluded video frame (left) and non-occluded video frame (right).

3.4.4 Risk assessment

We compute RULA risk assessment scores using the estimated postures from our approach and the MoCap estimates. In calculating the RULA score, we use the following assumptions for all tasks: the human is sitting on a chair, the subject receives minimal intermittent force/load (less than 2.0Kg), muscle use occurs less than 4x per minute, untwisted and vertical position for neck and torso, and supported legs and feet. We performed all the processes, including data collection, synchronization, posture estimation, and ergonomics risk assessment using ROS on a Linux machine with a Core-i9 CPU, 32 GB of RAM, and an NVIDIA 1070Ti GPU.

3.5 Results and discussion

In this section, we present the results from our human subject experiments. We discuss the performance of the proposed multisensory 3D posture estimation approach compared with the posture estimation solely from the robot's trajectory and posture estimation solely from OpenPose. We compare them with the estimated postures from the gold-standard MoCap system and evaluate ergonomics risk assessment using any of the above three approaches.

3.5.1 Posture estimation

Initially, we compare the posture estimation accuracy of three approaches by calculating the deviation[3] of their posture from the MoCap postures. Figure 3.9 presents the deviation for three approaches for all the participants and trials across our four tasks. As expected, our multisensory approach has a significantly lower deviation than the other two approaches, and it is true among all of the tasks. The median deviation for our multisensory approach is less than 5 deg with 75% quantile at 10 deg and upper whisker at 23 deg, which reveals the accuracy of our approach.

Moreover, we see that the deviation for estimated posture solely from the robot trajectory is lower in tasks (b, c, d) compared to task (a). Since tasks (b, c, d) include more motion in the arm, the robot trajectory provides broader observation ranges; hence, the particle filter can resolve the redundancy better and ends with less deviation. Similar phenomena occur for the estimated postures from OpenPose. During tasks (c, d), the motion of the human causes more significant movement for the OpenPose keypoints; hence the deviations in those tasks are less than in tasks (a, b). Higher whiskers for estimated postures from OpenPose also show that the estimated posture was wrong in some periods during the task, while the projected anatomical landmarks on the human skeleton were close to their corresponding keypoints from OpenPose. This issue is due to the redundancy in projecting 3D points into 2D points in the camera frame. Figure 3.10 shows two examples of those moments. Each example shows that the estimated posture (pink skeleton) matches well with the MoCap posture (green skeleton) in the video frame. At the same time, the view from another angle

[3] We use the term “deviation” instead of “error” since the MoCap posture is also an estimate and not ground truth.

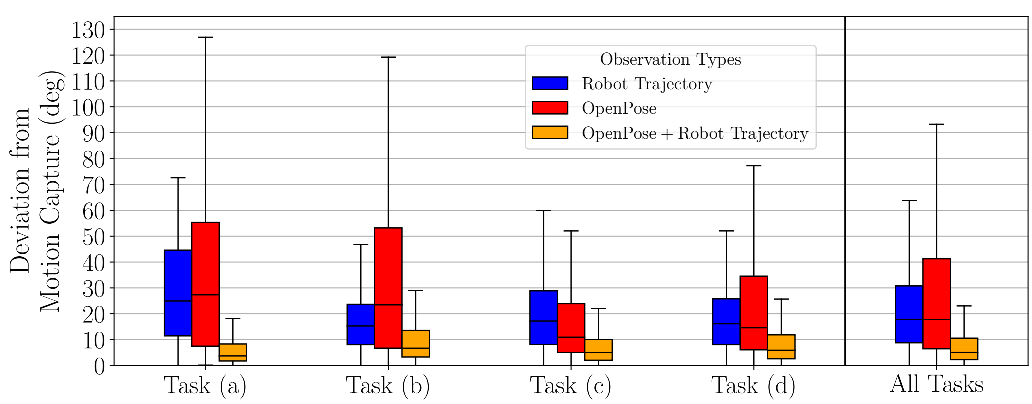


Figure 3.9: Deviation of the estimated postures by three different approaches from MoCap posture among tasks.

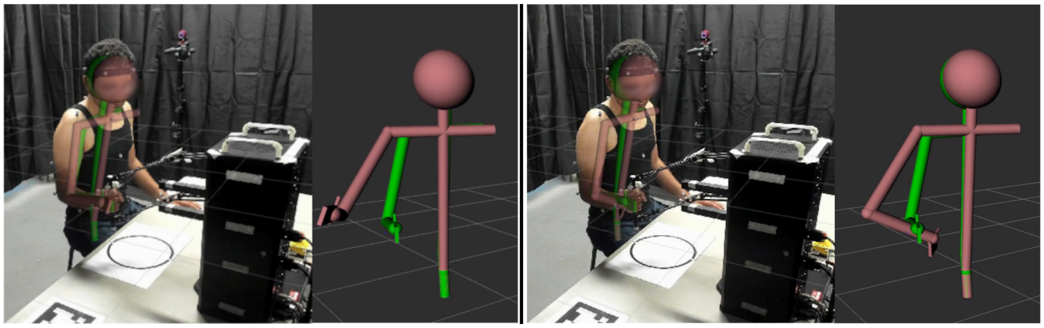


Figure 3.10: Two examples of the wrong posture estimated by OpenPose. Although the estimated posture (pink skeleton) matches well with the MoCap posture (green skeleton) in the camera view, a view from a different angle shows that those postures deviate highly from each other.

shows they are very different.

When we compare the results from the multisensory approach with the results from our previous work [29], in which we had a median of deviation less than 5 deg and upper quantile less than 15 deg while we used the limited range of motion for the torso and initializing particles around the neutral posture; we can conclude that by our multisensory approach is slightly more accurate, while it does not need any knowledge on the range of motion used in the task or any prior information about the posture which might reduce the unreliability of our approach.

Next, we compare the deviation of each posture estimation approach from MoCap posture among the human joints in Figure 3.11. We see that the deviation for our multisensory approach is very low among all the joints. At the same time, estimated postures from OpenPose have a high deviation in upper arm abduction (q_4), upper arm vertical rotation (q_5), and upper arm horizontal rotation (q_6). These high deviations also match what we discussed in Figure 3.10, and the same reasoning holds here too.

Finally, to show how the particles for two joints change through time, we plot the distribution of the joint angle of particles for a subject performing task (a) in Figure 3.12. Focusing on the initial five steps, we see that the standard deviation of the particles' joint angles from the multisensory approach shrinks faster than the other two approaches, which means that the particles converged to *a solution* faster than the other two approaches. Also, note that the converged *solution* from the multisensory approach is closer to the estimated joint angles from the MoCap (green line).

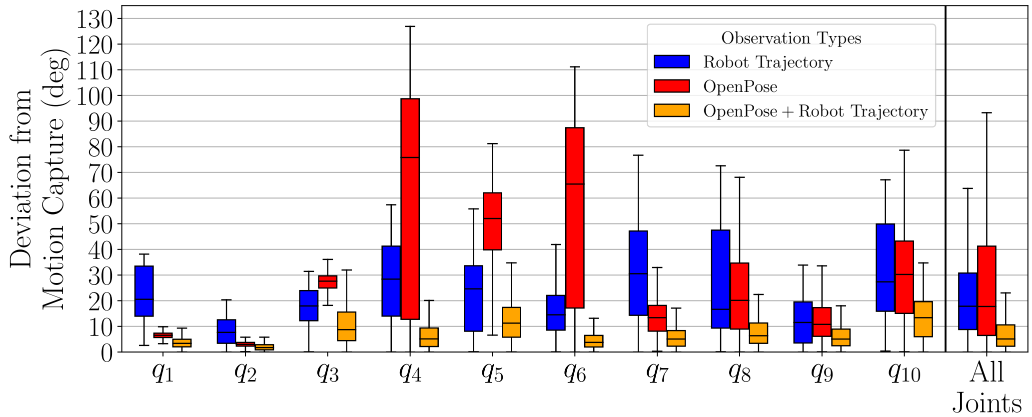


Figure 3.11: Deviation of the posture estimated by the proposed approach from MoCap postures for all the participants and different human joints.

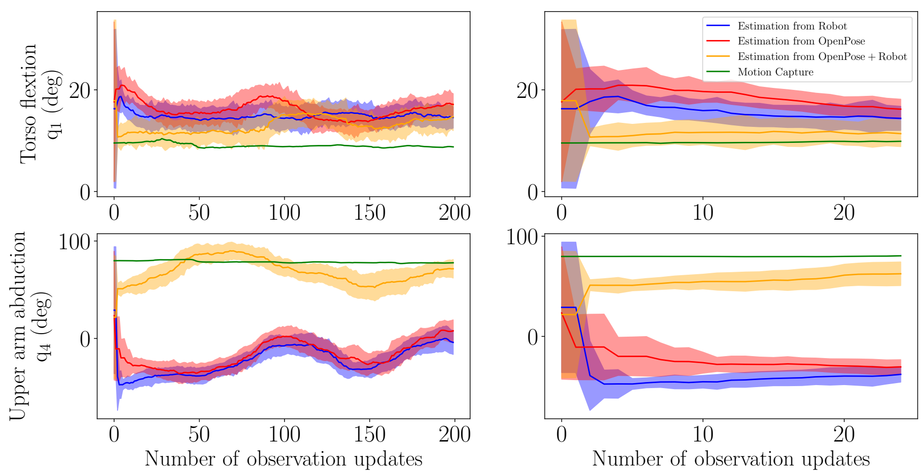


Figure 3.12: The distribution of particles over time. The figure's orange, red and blue lines are the average joint angles, and the shaded area around them shows one standard deviation around the average value. The left column of the figure focuses on the full range of observation steps during the task, while the right column shows the first 24 steps of observation.

3.5.2 Robustness to occlusion

We synthetically occluded video frames of a scenario to evaluate the robustness of the OpenPose-based posture estimation approaches. Figure 3.13 visualizes the deviation of estimated postures from motion capture for each joint for the scenario with occlusion and without occlusion. The plot shows that the multisensory posture estimation approach from the robot and the OpenPose has a slightly higher deviation than the same approach in non-occluded scenarios. This result reveals the robustness of our approach to occlusion compared to the approach solely using OpenPose.

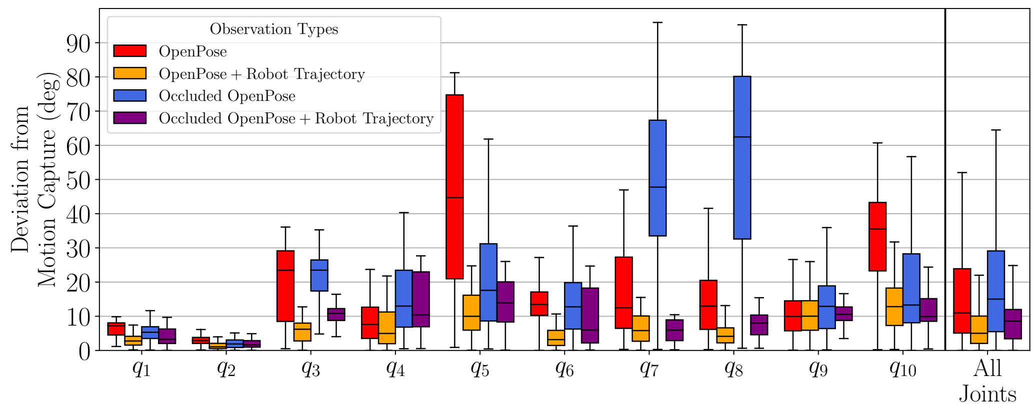


Figure 3.13: Deviation of the postures estimated by the proposed approaches from MoCap postures in occluded and non-occluded scenarios. The plot is for one subject performing task (c).

3.5.3 Risk assessment

Figure 3.14 shows the maximum RULA score during the tasks (the one most often used in ergonomics) for the postures estimated from the three approaches and MoCap postures. We also summarized them in Table 3.1. All the approaches successfully identified all instances where the RULA score was higher than 2, i.e., places where "further investigation or change may be or is needed". Results prove that the multisensory approach outperforms the two alternative approaches, and its accuracy of it is better than the approach from our previous work in which we had the observation from the robot with a limited range of motion for the torso and initializing the particles around the neutral posture. Once more, it proves that using the extra camera and 2D OpenPose predictions as the additional observation removes the requirement for having prior information about the task in posture estimation.

3.6 Conclusion and future directions

In this chapter, we proposed a multisensory 3D posture estimation approach for teleoperation, which is low-cost and robust to occlusion and achieves high-quality redundancy resolution. In our approach, we used the observations from 2D postures predicted by OpenPose over the video frames of a single RGB camera and the trajectory of the leader robot while performing a teleoperation task. We modeled the problem as a partially observable dynamical system and inferred the 3D posture using a particle filter. Through a human-subject study, we compared our results with the estimated postures solely from the OpenPose, the trajectory of the leader robot, and their ergonomic postural assessment performance. Our results show that the estimated postures from our multisensory approach have less deviation from the gold-standard MoCap postures. Moreover, the observations from two sensors

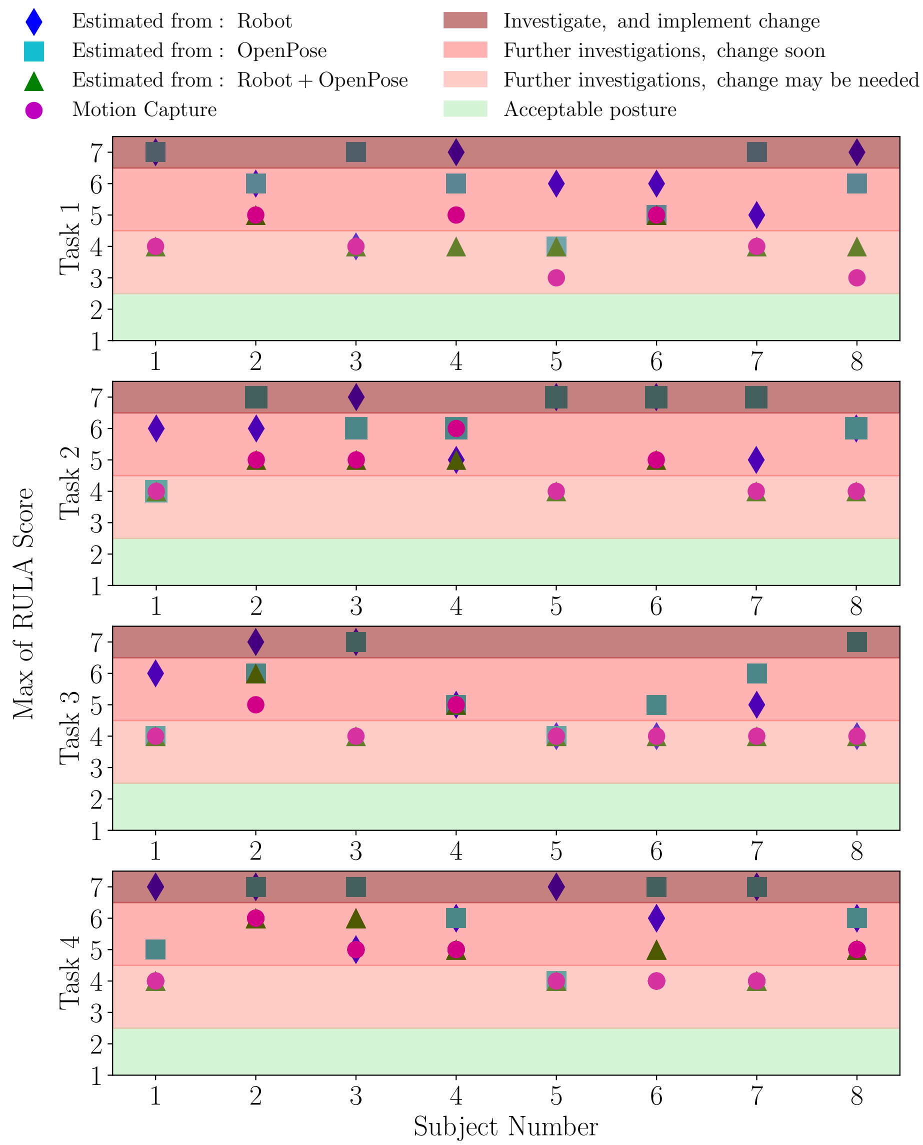


Figure 3.14: Comparison of the maximum values of RULA scores for tasks using the estimated postures from the three approaches and MoCap postures for all the participants.

Table 3.1: Summary of the deviation of the maximum RULA score for all tasks and subjects.

Posture estimation approach	Same max (RULA score)	Same interpretation of max (RULA score)
From Robot	18.75%	31.25%
From OpenPose	21.88%	40.62%
From Robot+OpenPose	75%	87.5%
From Robot (with torso limitation) [29]	65.63%	84.37%

improve the redundancy resolution and make it robust to segment occlusions.

In this chapter, we focused on teleoperation applications. However, we can use our approach for other physical human-robot applications with minor modifications. We would like to extend our approach to be used in the virtual reality (VR) industry to estimate the entire posture of the human based on the sensory information available from a VR headset and controllers and a simple camera that could be a webcam or a phone camera. In addition, we want to investigate the hypothesis that including the DULA and DEBA models for human comfort from [32] in our postural estimation as a weighting function for particles can improve the redundancy resolution and accuracy of the estimated postures. Moreover, the probabilistic distribution of estimated postures and their corresponding ergonomic risk scores can be used in human-aware planning for shared autonomy and other pHRI applications.

3.7 References

[1] S. Corazza, L. Muendermann, A. Chaudhari, T. Demattio, C. Cobelli, and T. P. Andriacchi, "A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach," *Ann. Biomed. Eng.*, vol. 34, no. 6, pp. 1019–1029, 2006.

[2] L. Needham, M. Evans, D. P. Cosker, L. Wade, P. M. McGuigan, J. L. Bilzon, and S. L. Colyer, "The accuracy of several pose estimation methods for 3D joint centre localisation," *Scientific Rep.*, vol. 11, no. 1, pp. 1–11, 2021.

[3] M. Topley and J. G. Richards, "A comparison of currently available optoelectronic motion capture systems," *J. Biomech.*, vol. 106, p. 109820, 2020.

[4] N. Seethapathi, S. Wang, R. Saluja, G. Blohm, and K. P. Kording, "Movement science needs different pose tracking algorithms," *arXiv Preprint arXiv:1907.10226*, 2019.

[5] S. L. Colyer, M. Evans, D. P. Cosker, and A. I. Salo, "A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system," *Sports Med. Open*, vol. 4, no. 1, pp. 1–15, 2018.

[6] S. E. Kessler, M. J. Rainbow, G. A. Lichtwark, A. G. Cresswell, S. E. D'Andrea, N. Konow, and L. A. Kelly, "A direct comparison of biplanar videoradiography and optical motion capture for foot and ankle kinematics," *Frontiers Bioeng. Biotech.*, vol. 7, p. 199, 2019.

[7] H. Tsushima, M. E. Morris, and J. McGinley, "Test-retest reliability and inter-tester reliability of kinematic data from a three-dimensional gait analysis system," *J. Jpn. Physical Therapy Assos.*, vol. 6, no. 1, pp. 9–17, 2003.

[8] C. Metcalf, C. Phillips, A. Forrester, J. Glodowski, K. Simpson, C. Everitt, A. Darekar, L. King, D. Warwick, and A. Dickinson, "Quantifying soft tissue artefacts and imaging variability in motion capture of the fingers," *Ann. Biomed. Eng.*, pp. 1–11, 2020.

[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proc. of the IEEE Conf. on Comput. Vision Pattern Recognit.*, 2017, pp. 7291–7299.

[10] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, "Rmpe: Regional multi-person pose estimation," in *Proc. of the IEEE Int. Conf. on Comput. Vision*, 2017, pp. 2334–2343.

[11] A. Mathis, P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, and M. Bethge, "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," Nature Publishing Group, Tech. Rep., 2018.

[12] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: realtime multi-person 2d pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 43, no. 1, pp. 172–186, 2019.

[13] X. Xiao and S. Zarar, "A wearable system for articulated human pose tracking under uncertainty of sensor placement," in *IEEE Int. Conf. on Biomed. Robot. and Biomechatronics*, 2018, pp. 1144–1150.

[14] X. Chen and J. Davis, "Camera placement considering occlusion for robust motion capture," Computer Graphics Laboratory, Stanford University, Tech. Rep, Tech. Rep., 2000.

[15] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, and X. Twombly, "Vision-based hand pose estimation: A review," *Comput. Vision Image Understanding*, vol. 108, no. 1-2, pp. 52–73, 2007.

[16] T. Steinebach, E. H. Grosse, C. H. Glock, J. Wakula, and A. Lunin, "Accuracy evaluation of two markerless motion capture systems for measurement of upper extremities: Kinect v2 and captiv," *Human Factors Ergonom. Manuf. Service Industries*, vol. 30, no. 4, pp. 291–302, 2020.

[17] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, "Postural optimization for an ergonomic human-robot interaction," in *2017 IEEE/RSJ Int. Conf. on Intell. Robots and Syst. (IROS)*. IEEE, 2017, pp. 2778–2785.

[18] M. Slembrouck, H. Luong, J. Gerlo, K. Schütte, D. V. Cauwelaert, D. D. Clercq, B. Vanwanseele, P. Veelaert, and W. Philips, "Multiview 3d markerless human pose estimation from openpose skeletons," in *Int. Conf. on Adv. Concepts for Intell. Vision Syst.* Springer, 2020, pp. 166–178.

[19] S. Bériault, "Multi-camera system design, calibration and three-dimensional reconstruction for markerless motion capture," Ph.D. dissertation, University of Ottawa (Canada), 2008.

[20] L. Punnett and D. H. Wegman, "Work-related musculoskeletal disorders: the epidemiologic evidence and the debate," *J. Electromyography Kinesiology*, vol. 14, no. 1, pp. 13–23, 2004.

[21] O. Erdinc and O. Vayvay, "Ergonomics interventions improve quality in manufacturing: a case study," *Int. J. Ind. Syst. Eng.*, vol. 3, no. 6, pp. 727–745, 2008.

[22] T. Vos, R. M. Barber, B. Bell, A. Bertozzi-Villa, S. Biryukov, I. Bolliger, F. Charlson, A. Davis, L. Degenhardt, D. Dicker *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013," *The Lancet*, vol. 386, no. 9995, pp. 743–800, 2015.

[23] S. P. Schneider, "Musculoskeletal injuries in construction: a review of the literature," *Appl. Occupat. Envirn. Hygiene*, vol. 16, no. 11, pp. 1056–1064, 2001.

[24] B. R. da Costa and E. R. Vieira, "Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies," *Amer. J. Ind. Med.*, vol. 53, no. 3, pp. 285–323, 2010.

[25] W. Yu, T. Ignatius, Z. Li, X. Wang, T. Sun, H. Lin, S. Wan, H. Qiu, and S. Xie, "Work-related injuries and musculoskeletal disorders among factory workers in a major city of china," *Accident Anal. Prevention*, vol. 48, pp. 457–463, 2012.

[26] P. G. Dempsey, L. M. Kocher, M. F. Nasarwanji, J. P. Pollard, and A. E. Whitson, "Emerging ergonomics issues and opportunities in mining," *Int. J. Environ. Res. Public Health*, vol. 15, no. 11, p. 2449, 2018.

[27] L. Peternel, C. Fang, M. Laghi, A. Bicchi, N. Tsagarakis, and A. Ajoudani, "Human arm posture optimisation in bilateral teleoperation through interface reconfiguration," in *IEEE RAS/EMBS Int. Conf. for Biomed. Robot. and Biomechatronics*, 2020.

[28] S. Yu, J. Lee, B. Park, K. Kim, and I. Cho, "Ergonomic analysis of a telemanipulation technique for a pyroprocess demonstration facility," *Nucl. Eng. Technol.*, vol. 46, no. 4, pp. 489–500, 2014.

[29] A. Yazdani, R. S. Novin, A. Merryweather, and T. Hermans, "Is the leader robot an adequate sensor for posture estimation and ergonomic assessment of a human teleoperator?" in *2021 IEEE 17th Int. Conf. on Automat. Sci. and Eng. (CASE)*. IEEE, 2021, pp. 1946–1952.

[30] A. Yazdani and R. Sabbagh Novin, "Posture estimation and optimization in ergonomically intelligent teleoperation systems," in *Companion of the 2021 ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, 2021, pp. 604–606.

[31] A. Yazdani, R. S. Novin, A. Merryweather, and T. Hermans, "Ergonomically intelligent physical human-robot interaction: Postural estimation, assessment, and optimization," *arXiv Preprint arXiv:2108.05971*, 2021.

[32] A. Yazdani, R. Sabbagh Novin, A. Merryweather, and T. Hermans, "DULA and DEBA: Differentiable Ergonomic Risk Models for Postural Assessment and Optimization in Ergonomically Intelligent pHRI," in *IEEE/RSJ Int. Conf. on Intell. Robots and Syst. (IROS)*, 2022.

[33] E. D'Antonio, J. Taborri, I. Mileti, S. Rossi, and F. Patané, "Validation of a 3d markerless system for gait analysis based on openpose and two rgb webcams," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17 064–17 075, 2021.

[34] Y. Jiang and C. K. Liu, "Data-driven approach to simulating realistic human joint constraints," in *2018 IEEE Int. Conf. on Robot. and Automat. (ICRA)*. IEEE, 2018, pp. 1098–1103.

[35] L. McAtamney and E. N. Corlett, "RULA: a survey method for the investigation of work-related upper limb disorders," *Appl. Ergonom.*, vol. 24, no. 2, pp. 91–99, 1993.

[36] Z. Shi, T. R. Groechel, S. Jain, K. Chima, O. Rudovic, and M. J. Matarić, "Toward personalized affect-aware socially assistive robot tutors for long-term interventions with children with autism," *ACM Trans. Human-Robot Interaction (THRI)*, 2021.

[37] M. Spitale, S. Okamoto, M. Gupta, H. Xi, and M. J. Matarić, "Socially assistive robots as storytellers that elicit empathy," *ACM Trans. Human-Robot Interaction (THRI)*, 2022.

[38] L. F. Figueredo, R. D. C. Aguiar, L. Chen, T. C. Richards, S. Chakrabarty, and M. Dogar, "Planning to minimize the human muscular effort during forceful human-robot collaboration," *ACM Trans. Human-Robot Interaction (THRI)*, vol. 11, no. 1, pp. 1–27, 2021.

[39] L. Vianello, J.-B. Mouret, E. Dalin, A. Aubry, and S. Ivaldi, "Human posture prediction during physical human-robot interaction," *IEEE Robot. Automat. Lett.*, 2021.

[40] T.-C. Lin, A. U. Krishnan, and Z. Li, "Intuitive, efficient and ergonomic tele-nursing robot interfaces: Design evaluation and evolution," *ACM Trans. Human-Robot Interaction (THRI)*, 2022.

[41] T. Zhou, M. E. Cabrera, J. P. Wachs, T. Low, and C. Sundaram, "A comparative study for telerobotic surgery using free hand gestures," *J. Human-Robot Interaction*, vol. 5, no. 2, pp. 1–28, 2016.

[42] A. D. Dragan, S. S. Srinivasa, and K. C. Lee, "Teleoperation with intelligent and customizable interfaces," *J. Human-Robot Interaction*, vol. 2, no. 2, pp. 33–57, 2013.

[43] J. Hadfield, G. Chalvatzaki, P. Koutras, M. Khamassi, C. S. Tzafestas, and P. Maragos, "A deep learning approach for multi-view engagement estimation of children in a child-robot joint attention task," in *2019 IEEE/RSJ Int. Conf. on Intell. Robots and Syst. (IROS)*. IEEE, 2019, pp. 1251–1256.

[44] G. Chalvatzaki, X. S. Papageorgiou, P. Maragos, and C. S. Tzafestas, "Learn to adapt to human walking: A model-based reinforcement learning approach for a robotic assistant rollator," *IEEE Robot. Automat. Lett.*, vol. 4, no. 4, pp. 3774–3781, 2019.

[45] A. Dini, C. Murko, S. Yahyanejad, U. Augsdörfer, M. Hofbaur, and L. Paletta, "Measurement and prediction of situation awareness in human-robot interaction based on a framework of probabilistic attention," in *2017 IEEE/RSJ Int. Conf. on Intell. Robots and Syst. (IROS)*. IEEE, 2017, pp. 4354–4361.

[46] W. Kim, M. Lorenzini, K. Kapıcıoğlu, and A. Ajoudani, "Ergotac: A tactile feedback interface for improving human ergonomics in workplaces," *IEEE Robot. Automat. Lett.*, vol. 3, no. 4, pp. 4179–4186, 2018.

[47] C. Fang, A. Ajoudani, A. Bicchi, and N. G. Tsagarakis, "A real-time identification and tracking method for the musculoskeletal model of human arm," in *IEEE Int. Conf. on Syst., Man, and Cybernetics*, 2018, pp. 3472–3479.

[48] A. M. Zanchettin, N. M. Ceriani, P. Rocco, H. Ding, and B. Matthias, "Safety in human-robot collaborative manufacturing environments: Metrics and control," *IEEE Trans. Automat. Sci. Eng.*, vol. 13, no. 2, pp. 882–893, 2016.

[49] L. Peternel, C. Fang, N. Tsagarakis, and A. Ajoudani, "Online human muscle force estimation for fatigue management in human-robot co-manipulation," in *IEEE/RSJ Int. Conf. on Intell. Robots and Syst. (IROS*, 2018, pp. 1340–1346.

[50] P. Vartholomeos, N. Katevas, A. Papadakis, and L. Sarakis, “Design of motion-tracking device for intuitive and safe human-robot physical interaction,” in *IEEE Int. Conf. on Telecommun. and Multimedia*, 2016, pp. 1–6.

[51] J. Buzzi, E. De Momi, and I. Nisky, “An uncontrolled manifold analysis of arm joint variability in virtual planar position and orientation tele-manipulation,” *IEEE Trans. Biomed. Eng.*, 2018.

[52] G. Martinez, I. A. Kakadiaris, D. Magruder, and I. Magruder, “Teleoperating ROBONAUT: A case study,” in *British Machine Vision Conf.*, 2002, pp. 1–10.

[53] R. Rahal, G. Matarese, M. Gabiccini, A. Artoni, D. Prattichizzo, P. R. Giordano, and C. Pacchierotti, “Caring about the human operator: haptic shared control for enhanced user comfort in robotic telemanipulation,” *IEEE Trans. Haptics*, vol. 13, no. 1, pp. 197–203, 2020.

[54] K. M. Lynch and F. C. Park, *Modern Robotics*. Cambridge University Press, 2017.

[55] H. Kim, L. M. Miller, N. Byl, G. M. Abrams, and J. Rosen, “Redundancy resolution of the human arm and an upper limb exoskeleton,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1770–1779, 2012.

[56] M. Li, W. Guo, R. Lin, C. Wu, and L. Han, “An efficient motion generation method for redundant humanoid robot arm based on the intrinsic principles of human arm motion,” *Int. J. Humanoid Robots*, vol. 15, no. 06, p. 1850026, 2018.

[57] A. Yang, Y. Chen, W. Naeem, M. Fei, and L. Chen, “Humanoid motion planning of robotic arm based on human arm action feature and reinforcement learning,” *Mechatronics*, vol. 78, p. 102630, 2021.

[58] S. Kim, E. Gribovskaya, and A. Billard, “Learning motion dynamics to catch a moving object,” in *2010 10th IEEE-RAS Int. Conf. on Humanoid Robots*. IEEE, 2010, pp. 106–111.

[59] C. Yang, C. Zeng, C. Fang, W. He, and Z. Li, “A dmps-based framework for robot learning and generalization of humanlike variable impedance skills,” *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 3, pp. 1193–1203, 2018.

[60] Y.-H. Deng and J.-Y. J. Chang, “Human-like posture correction for seven-degree-of-freedom robotic arm,” *J. Mechanisms Robot.*, vol. 14, no. 2, 2022.

[61] X. Wang, M. Maurin, F. Mazet, N. D. C. Maia, K. Voinot, J. P. Verriest, and M. Fayet, “Three-dimensional modelling of the motion range of axial rotation of the upper arm,” *J. Biomech.*, vol. 31, no. 10, pp. 899–908, 1998.

[62] I. Akhter and M. J. Black, “Pose-conditioned joint angle limits for 3d human pose reconstruction,” in *Proc. of the IEEE Conf. on Comput. Vision and Pattern Recognit.*, 2015, pp. 1446–1455.

[63] P. K. Levangie and C. C. Norkin, *Joint Structure and Function: A Comprehensive Analysis*. FA Davis, 2011.

[64] *NASA-STD-3001 Space Flight Human-System Standard Volumes 1 (Crew Health)*, Available at https://msis.jsc.nasa.gov/sections/section03.htm, Std., 1995.

[65] S. A. Ismail, S. B. M. Tamrin, M. R. Baharudin, M. A. M. Noor, M. H. Juni, J. Jalaludin, and Z. Hashim, “Evaluation of two ergonomics intervention programs in reducing ergonomic risk factors of musculoskeletal disorder among school children,” *Res. J. Med. Sci.*, vol. 4, no. 1, pp. 1–10, 2010.

[66] Z. Khodabakhshi, S. A. Saadatmand, M. Anbarian, and R. Heydari Moghadam, “An ergonomic assessment of musculoskeletal disorders risk among the computer users by rula technique and effects of an eight-week corrective exercises program on reduction of musculoskeletal pain,” *Iranian J. Ergonom.*, vol. 2, no. 3, pp. 44–56, 2014.

[67] S. G. Hart and L. E. Staveland, “Development of nasa-tlx (task load index): Results of empirical and theoretical research,” in *Advances in Psychol.*, vol. 52, 1988, pp. 139–183.

[68] S. Hignett and L. McAtamney, “Rapid entire body assessment (REBA),” *Appl. Ergonom.*, vol. 31, no. 2, pp. 201–205, 2000.

[69] J. Steven Moore and A. Garg, “The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders,” *Amer. Ind. Hygiene Assos. J.*, vol. 56, no. 5, pp. 443–458, 1995.

[70] J. M. Kapellusch, A. Garg, K. T. Hegmann, M. S. Thiese, and E. J. Malloy, “The strain index and acgih tlv for hal: risk of trigger digit in the wistah prospective cohort,” *Human Factors*, vol. 56, no. 1, pp. 98–111, 2014.

[71] L. Chen, L. Figueredo, and M. Dogar, “Planning for muscular and peripersonal-space comfort during human-robot forceful collaboration,” in *IEEE-RAS Int. Conf. on Humanoid Robots*, 2018, pp. 1–8.

[72] L. Peternel, W. Kim, J. Babič, and A. Ajoudani, “Towards ergonomic control of human-robot co-manipulation and handover,” in *IEEE-RAS Int. Conf. on Humanoid Robots*, 2017, pp. 55–60.

[73] W. Kim, J. Lee, L. Peternel, N. Tsagarakis, and A. Ajoudani, “Anticipatory robot assistance for the prevention of human static joint overloading in human–robot collaboration,” *Robot. Automat. Lett.*, vol. 3, no. 1, pp. 68–75, 2017.

[74] L. F. Figueredo, R. C. Aguiar, L. Chen, S. Chakrabarty, M. R. Dogar, and A. G. Cohn, “Human comfortability: Integrating ergonomics and muscular-informed metrics for manipulability analysis during human-robot collaboration,” *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 351–358, 2020.

[75] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT press, 2005.

[76] Openpose documentation. [Online]. Available: https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/md_doc_02_output.html

[77] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer Science & Business Media, 2010.

CHAPTER 4

ERGONOMICALLY INTELLIGENT PHYSICAL HUMAN-ROBOT INTERACTION: POSTURAL ESTIMATION, ASSESSMENT, AND OPTIMIZATION

In this chapter, we leverage the posture estimation method discussed in the second chapter and introduce a continuous and differentiable version of the RULA risk assessment tool (DULA) for ergonomically intelligent teleoperation and pHRI. We also investigate and compare gradient-based and gradient-free approaches for posture optimization in teleoperation in a simulated environment. We presented this part of the work at AAAI Fall Symposium Series 2021: Artificial Intelligence for Human-Robot Interaction held as an online event.

I primarily completed this work under the guidance of Prof. Tucker Hermans and Prof. Andrew Merryweather. Roya Sabbagh Novin assisted in system design and human-subject studies.

Ergonomically Intelligent Physical Human-Robot Interaction: Postural Estimation, Assessment, and Optimization*

Amir Yazdani,[1] **Roya Sabbagh Novin,**[1†] **Andrew Merryweather,**[1] **Tucker Hermans**[1,2]

[1]University of Utah Robotics Center, SLC, UT; [2]NVIDIA Corporation, Seattle, WA;
amir.yazdani@utah.edu

Abstract

Ergonomics and human comfort are essential concerns in physical human-robot interaction. Common practical methods in the area either fail in estimating the correct posture due to occlusion or suffer from inaccurate ergonomics models in performing postural optimization. We propose a novel alternative framework for posture estimation, assessment, and optimization for ergonomically intelligent physical human-robot interaction. We show that we can estimate human posture solely from the trajectory of the interacting robot with median deviation of 5 deg from motion capture. We propose DULA, a differentiable ergonomics assessment tool with 99.73% accuracy comparing to RULA. We use DULA in postural optimization for physical human-robot interaction tasks such as co-manipulation and teleoperation. We evaluate our framework through human and simulation experiments.

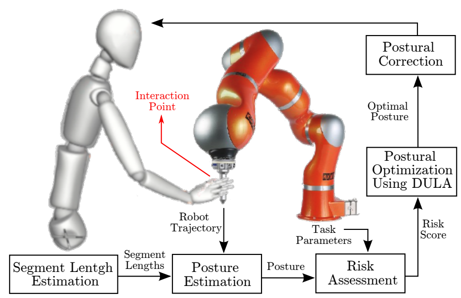


Figure 1: The framework for ergonomically intelligent p-HRI.

Introduction

Improving workplace ergonomics and reducing the risk of work-related musculoskeletal disorders (WMSDs) has been the focus of researchers for many years (Punnett and Wegman 2004; Erdinc and Vayvay 2008). High rate of injuries in industry (Schneider 2001; Yu et al. 2012) has further motivated these studies. WMSDs are the second-largest source of disabilities worldwide (Vos et al. 2015), and awkward postures are the main contributor to the work-related injuries (da Costa and Vieira 2010). Industry 4.0 initiative (Gorecky et al. 2014) and the development of collaborative workplaces where smart agents (e.g., collaborative robots (Sherwani, Asad, and Ibrahim 2020)) physically collaborate with humans to perform different types of tasks, highlights the importance of human comfort and ergonomics in physical human-robot interaction (p-HRI). In this paper, we introduce a framework for ergonomically intelligent p-HRI including posture estimation, assessment, optimization and correction. Figure 1 shows our proposed framework.

p-HRI covers a vast number of applications such as generic p-HRI which includes direct physical interaction with the robot (e.g., robotics physical therapy (Guerrero et al. 2013)), co-manipulation (Kim et al. 2017), assistive holding (Chen, Figueredo, and Dogar 2018), handover (Busch, Toussaint, and Lopes 2018), and teleoperation (Rahal et al. 2020). We categorize teleoperation into three main types: *goal-constrained teleoperation* (e.g., pick and place with a desired pose for the object), *path-constrained teleoperation* (e.g., turning a valve where the path to follow is constrained based on the diameter of the valve), and *trajectory-constrained teleoperation* (e.g., arc welding where the operator should follow a velocity profile in addition to the welding path). Figure 2 visualizes the different types of p-HRI applications.

A common and practical solution to improve workplace ergonomics is to perform risk assessments and analyze human comfort during the task. Ergonomists usually perform this process for the worst posture during the task by taking measurements of human posture onsite or from recorded images or videos. They normally use risk assessment tools (Ramaganesh et al. 2021) to assess the task and provide inter-

*Research reported in this publication was supported in part by DARPA under grant N66001-19-2-4035 and the National Institute for Occupational Safety and Health under award number T420H008414-10.

†Roya Sabbagh Novin is currently with Embark Trucks Inc.

vention and improvement in forms of user training (Bazazan et al. 2019), workstation modification (Shikdar and Al-Hadhrami 2007) or task rotation (Motabar and Nimbarte 2021). However, in most cases, it is not possible to have an ergonomist present during the entire task to perform a comprehensive risk assessment, which results in missing events or only partial understanding of awkward postures. In addition, these sampling strategies lead to failures in detecting the most hazardous combinations of force and posture, increasing the inaccuracy of an assessment. Finally, although postural correction by ergonomists generally helps reduce the risk of injuries, in many cases, they are not optimal or tailored to specific task requirements.

Early improvements on continuous monitoring of posture, such as using motion capture (MoCap) systems to estimate the posture and task parameters (e.g., frequency and duration of the task), still require significant time and effort for set up (Alvarez et al. 2017). Moreover, putting markers on human operators can be inconvenient. Alternative markerless techniques are more adaptable, minimally intrusive, and less expensive. However, they need calibration to deal with errors and uncertainties from the sensors (Xiao and Zarar 2018; Ganesh et al. 2021). Vision-based markerless methods can also be perturbed by the lighting, background color, and even the user's clothing (Colyer et al. 2018). The proximity between the human and robot in p-HRI increases occlusion and reduces the accuracy in estimate posture as mentioned in (Busch et al. 2017).

Postural optimization in p-HRI has recently received substantial attention in research (see Table 1) and developing a model of human comfort lies at the heart of effective postural optimization. Researchers either use risk assessment tools from ergonomics or propose computational models for human comfort and ergonomics. Among all risk assessment tools, RULA (McAtamney and Corlett 1993) and REBA (Hignett and McAtamney 2000) depend most on human posture and provide quantitative scores, making them good choices for postural optimization applications. However, the discrete scores and the presence of plateaus in RULA and REBA (Busch et al. 2017) create challenges when using them in gradient-based postural optimization. Based on our experience, using the risk assessment models directly in gradient-free optimization is time-expensive, and the plateaus often prevent progress toward the optimal global solution in postural optimization. Thus, researchers in p-HRI often use computational models based on approximations of ergonomic assessment models in gradient-based postural optimizations; quadratic approximations (Rahal et al. 2020; Busch et al. 2017; Busch, Toussaint, and Lopes 2018) being the standard approach in the literature. However, these approximations deviate far from the scientifically validated assessments, causing doubt that they can reliably provide the same level of ergonomic benefit.

Assuming perfect ergonomics assessment and optimization, the postural correction process is also a challenging part in improving ergonomics in p-HRI. Manual correction by ergonomists is tedious and cannot be applied continuously for the human operator. More systematic approaches such as continuous visual (Lorenzini et al. 2018) and vibro-

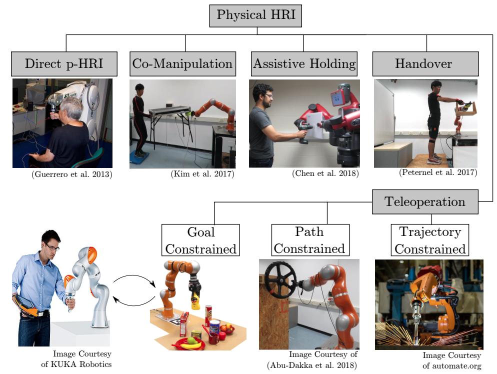


Figure 2: Categories of physical HRI applications.

tactile (Kim et al. 2018) feedback to the user perform well in correcting the posture. However, they need different devices to provide feedback, require the user to accept and apply the correction. Guiding the human toward the optimal posture using the force from the robot (Rahal et al. 2020) also may cause extra force on the human and increase the risk of injuries, as well as affecting the realistic haptic feeling.

To overcome these issues in postural estimation and optimization in p-HRI, we propose ergonomically intelligent physical human-robot interaction (see Fig. 1). It includes four main parts, targeting the drawbacks of current methods:

- *Posture estimation in p-HRI solely from the trajectory of the interacting robot*: We show that the interacting robot is an adequate sensor for continuous posture estimation of the human and for monitoring the risk of injuries. We model the posture estimation problem as a partially observable dynamical system where the robot's trajectory at the interaction point is the only observation. We provide an approach for estimating human segment lengths using interactive motion routines and a particle filter to infer the human joint angles.
- *Differentiable ergonomics assessment model*: We learn an accurate, continuous, and differentiable model called *DULA* (Differentiable Upper Limb Assessment) for upper body ergonomics using a standard neural network from a dataset of exhaustively sampled postures, labeled by the standard RULA risk assessment tool.
- *Gradient-based postural optimization*: We use DULA in a gradient-based optimization and compare it with a gradient-free optimization directly using RULA.
- *Automatic postural correction*: We provide a new approach to correct human posture in teleoperation through the velocity scale between the leader and follower robots.

In this paper, we focus on the first three parts of our framework. We introduce our approach for general p-HRI scenarios. However, we implement and evaluate them in a simple teleoperation as an example. We initially proposed the

<table>
<tr><td colspan="2">Ergonomics Model</td><td>Analytical Models</td><td>Learned Models</td><td colspan="2">Risk Assessment Tools</td></tr>
<tr><td colspan="2">Approximation</td><td></td><td></td><td>Quadratic</td><td>No-Approximation</td></tr>
<tr><td rowspan="2">Optimization</td><td>Gradient Based</td><td>■(Peternel et al. 2018)
▩■(Peternel et al. 2017)
□(Chen, Figueredo, and Dogar 2018)
◆(Peternel et al. 2020)
■(Kim et al. 2017)</td><td>This work</td><td>◇(Rahal et al. 2020)
▩(Busch et al. 2017)
▩(Busch, Toussaint, and Lopes 2018)</td><td></td></tr>
<tr><td>Gradient Free</td><td></td><td>□(Marin et al. 2018)</td><td></td><td>■(Van der Spaa et al. 2020)</td></tr>
<tr><td colspan="2">No Optimization</td><td></td><td></td><td></td><td>□(Shafti et al. 2019)</td></tr>
<tr><td>Legend</td><td colspan="5">□ Physical-HRI: Assistive Holding, ▩ Physical-HRI: Handover, ■ Physical-HRI: Co-Manipulation
◆ Teleop: Goal-Constrained with Postural Optimization By Interface Reconfiguration
◇ Teleop: Goal-Constrained with Online Postural Optimization</td></tr>
</table>

Table 1: State of the art of postural optimization in p-HRI and teleoperation.

idea of ergonomically intelligent teleoperation systems in the HRI Pioneers workshop (Yazdani and Sabbagh Novin 2021). Our posture estimation for teleoperation approach was previously published at IEEE CASE conference (Yazdani et al. 2021b), and we introduced the differentiable ergonomics model for postural optimization in the RSS Robotics for People workshop (Yazdani et al. 2021a). In this paper, we do not add any novel experimental evolution. Instead, we formalize the framework for general p-HRI and more tightly link the previous contributions. We believe this paper brings all the previous work together, improving clarity and advancing the discussion of ergonomics in p-HRI.

Related Work

In this section we provide the state of the art in posture estimation, risk assessment and postural optimization in p-HRI.

Posture Estimation in p-HRI

For many years, MoCap systems using reflective markers ((Das et al. 2011)) were the dominant approach for posture estimation. However, advances in computer vision have led to increased popularity in markerless techniques (Saremi et al. 2018; Mathis et al. 2018). In p-HRI and teleoperation, researchers mainly have used external sensors (i.e., a vision system or IMU) to estimate a user's posture, especially for hand gestures (Vartholomeos et al. 2016).

The idea of solely using the interacting robot's trajectory (leader robot in teleoperation) for posture estimation in p-HRI has been introduced concurrently with this research by (Rahal et al. 2020), where they solved the inverse kinematics of the human arm. Unlike our probabilistic approach, their heuristic does not always hold across different tasks. In comparison, our work provides a probabilistic distribution over the human posture and ergonomic risk score. Moreover, our human model covers the entire upper body, including the arm and torso. Later, (Vianello et al. 2021) proposed probabilistic prediction of human postures for a given robot trajectory executed in a collaborative scenario. They learned the distribution of the null space of the human arm's Jacobian and the weights of the weighted pseudo-inverse from demonstrations. Their approach depends highly on the task and demonstrations. In comparison, our approach is task-independent and does not require any demonstration.

Defining a model for human joint limits is challenging. Studies show that the range of motion for a joint varies depending on the positions of other joints (inter-joint dependency) or other degrees-of-freedom in the same joint (intra-joint dependency) (Wang et al. 1998; Jiang and Liu 2018) and vary by gender and person. (Akhter and Black 2015) used a dataset of recorded MoCap of human motion to develop a discontinuous mathematical model for a posture-dependant range of motion and check the validity of a full-body posture. (Jiang and Liu 2018) used the above model to label the validity of a set of randomly generated postures and learned a differentiable neural network based on the generated data, and used it as a constraint in the inverse kinematics optimization. Their arm model only includes the shoulder and elbow and not the wrist. We use this learned network for checking the validity of the arm posture.

Ergonomics Models and Postural Optimizations in p-HRI

The literature in postural optimization in p-HRI and teleoperation is minimal. Table 1 summarizes the relevant literature for in both areas.

Researchers have examined ergonomics and postural optimization in three different types of p-HRI: (1) *Assistive Holding* (e.g., (Marin et al. 2018; Shafti et al. 2019; Chen, Figueredo, and Dogar 2018)), (2) *Object handover* (e.g., (Busch et al. 2017; Busch, Toussaint, and Lopes 2018; Peternel et al. 2017)), and (3) *Co-Manipulation* (e.g., (Van der Spaa et al. 2020; Peternel et al. 2018, 2017; Kim et al. 2017)). In the first two tasks the postural optimization provides optimized posture for the human and joint configurations for the robot while maintaining contact with the object at the interaction interface. In co-manipulation, the postural optimization outputs a trajectory of optimal postures for the human and an optimal joint-space trajectory for the robot to perform the task. In teleoperation, postural optimization has only been developed for goal-constrained tasks.

Ergonomic assessment models provide the primary cost in postural optimization objectives. p-HRI researchers have proposed many computational models to assess ergonomics and comfort of users, including peripersonal space (Chen, Figueredo, and Dogar 2018), muscle fatigue (Peternel et al. 2017), and joint overloading (Kim et al. 2017). To make the optimization simpler, (Marin et al. 2018) suggested a contextual ergonomics model, which is a set of Gaussian process models including joint angles, moments, reaction, load, and muscle activation trained with the musculoskeletal simulation task contexts. Using Gaussian process models enables a 2D latent space to search while their cost function is defined in the high-dimensional musculoskeletal space.

Some works use approximations of the ergonomics risk assessment tools. (Busch, Toussaint, and Lopes 2018) proposed a differentiable surrogate of the REBA score by fitting a weighted combination of quadratic functions plus a constant for the task payload. (Rahal et al. 2020) suggested a quadratic approximation for RULA, which is the summation of the quadratic norm of the deviation from the human neutral posture for shoulder, elbow, and wrist joints angles. Their approximation conceptually agrees with the qualitative idea behind RULA, in which the risk score goes higher when the human deviates more from the neutral posture.

(Van der Spaa et al. 2020) provided the only study of directly using risk assessment tools in derivative-free postural optimization. They added the REBA score to the transition cost function in an A* optimization for task and motion planning of a robot in co-manipulation. Finally, (Shafti et al. 2019) identified six main causes of low ergonomic states, leading to six universal robot responses to allow the human to return to an optimal ergonomics state. Their framework identifies the causes and controls the robot to constantly adapt to the environment in an ergonomic way.

The table shows that most of the publications use analytical models and quadratic approximation to risk assessment tools in their approaches. Due to the differentiability of the models, they can be used in gradient-based postural optimization, which is faster and easier to implement for postural optimization problems. However, the reliability and optimality of the level of ergonomic benefit they provide are questionable comparing to standard risk assessment tools in ergonomics. Table 1 also shows an excellent opportunity for researchers to develop accurate ergonomics models learned from standard risk assessment tools and leverage them in gradient-based optimization. We target this area of research and propose DULA, a differentiable and continuous ergonomics model that is learned from the standard assessment tool, RULA.

Problem Statement

In this section, we introduce the problem of posture estimation and optimization in p-HRI. Among the different p-HRI applications, we focus on co-manipulation and teleoperation since they have the most range of continuous motions. However, the same problem statement can be extended to other applications with minor modifications.

Posture Estimation via the Robot Proprioceptive Sensing in p-HRI

In p-HRI, the physical interaction occurs at the *interaction point*. This is the contact point either between the human body and the robot in direct p-HRI and teleoperation or between the human body and the shared object in co-manipulation, assistive holding, and handover. In the latter case, we assume the robot firmly grasps the shared object, adding it to the robot's kinematic chain.

We seek to solve the problem of estimating the human joint-space trajectory $\boldsymbol{\tau} = [\mathbf{q}_t; \dot{\mathbf{q}}_t]_{t=1:T}$ using only the observed task-space poses and velocities of the interaction point $\mathcal{Z} = [\mathbf{z}_t; \dot{\mathbf{z}}_t]_{t=1:T}$ from the robot. The mapping from human joint angles into the pose of the interaction point is through the kinematics of the human model parameterized by the segment length ψ. We estimate ψ independently, prior to posture estimation. Hence,

$$[\mathbf{z}_t; \dot{\mathbf{z}}_t] = h(\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi)) \tag{1}$$

where h is the observation function, and ϕ is the forward kinematics of the human model. This defines only a partial observation of the human posture because of redundancy in the human kinematics and a noisy measurement at the interaction point, which may change slightly during a task.

We can write the postural estimation as an optimization:

$$\begin{aligned}\tau^* = \arg\min_{\tau} \sum_{t=1}^{T} & ||\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi) - [\mathbf{z}_t; \dot{\mathbf{z}}_t]||^2_{\Sigma_1} + \\ & ||[\mathbf{q}_t, \dot{\mathbf{q}}_t] - f([\mathbf{q}_{t-1}, \dot{\mathbf{q}}_{t-1}])||^2_{\Sigma_2} \\ \text{s.t.} \quad & \mathbf{q}_{\min} \leq \mathbf{q} \leq \mathbf{q}_{\max}\end{aligned} \tag{2}$$

where $\mathbf{q}_{\min}$ and $\mathbf{q}_{\max}$ are the joint limits, Σ is the weight vector for position and orientation elements, and f is the human motion model defined by:

$$\begin{bmatrix}\mathbf{q}_t \\ \dot{\mathbf{q}}_t\end{bmatrix} = \begin{bmatrix}1 & \Delta t \\ 0 & 1\end{bmatrix}\begin{bmatrix}\mathbf{q}_{t-1} \\ \dot{\mathbf{q}}_{t-1}\end{bmatrix} + \begin{bmatrix}0 \\ \ddot{\mathbf{q}}_{t-1}\Delta t\end{bmatrix} \tag{3}$$

The high degree-of-freedom in human kinematics makes this problem *redundant* with an infinite number of solutions. We seek the solution closest to the actual human's posture.

Postural Optimization in Co-Manipulation

In co-manipulation, the robot helps a human to move an object (usually a heavy one) from an initial pose to the desired pose (Peternel et al. 2017). There is no specific trajectory for the object to follow; the human plans and leads the motion, and the robot follows and provides help carrying the object. Meanwhile, the ergonomically intelligent p-HRI seeks to find the optimal posture for the human at each instance while holding the pose and velocities at the interaction point close to the current state. Once found, the system will suggest the solution to the human. We call this *online postural optimization*. Using the same notation of our posture estimation approach, we can formalize finding the ergonomically optimal posture with the following optimization problem:

$$\begin{aligned}\mathbf{q}^*_t = \arg\max_{\mathbf{q}_t} \quad & \text{Ergonomics}(\mathbf{q}_t) \\ \text{s.t.} \quad & ||\phi([\mathbf{q}_t; \dot{\mathbf{q}}_t], \psi) - [\mathbf{z}_t; \dot{\mathbf{z}}_t]||^2_{\Sigma_1} < \epsilon \\ & \mathbf{q}_t \in \text{Range of Motion}\end{aligned} \tag{4}$$

where $\mathbf{q}^*_t$ is the optimal posture at time t. It is important to note that the optimal posture $\mathbf{q}^*_t$ from Eq. (4) for each time step is then suggested to the human to move towards. The human can refuse or accept and try to apply the correction as much as possible while completing the task.

Postural optimization in Teleoperation

The main distinction of teleoperation from other p-HRI applications is the motion (and force) coupling between the leader and follower robots. The motion coupling between the two robots is not necessarily a one-to-one scale, and the coupling can be paused and resumed. This enables the teleoperator to disengage the leader robot in the middle of a task in many types of application, reposition their postures (and the leader robot as the result), and resume the teleoperation from a a more comfortable posture (Peternel et al. 2020). The relative position trajectory of the follower robot remains unchanged; however, its velocity trajectory changes due to the pause. The motion scaling also helps keep the teleoperator's posture in a smaller comfort zone while the follower robot operates in a much wider zone.

To use this feature of teleoperation in postural optimization, we define three types of postural optimization, customized for different types of teleoperation from Fig. 2.

Online Postural Optimization Online postural optimization in teleoperation is very similar to online postural optimization in co-manipulation. The difference is that instead of having initial and goal poses for the shared object in co-manipulation, in teleoperation, we have initial and goal poses for the end-effector of the follower robot (or the object it manipulates). Hence, the ergonomically intelligent system should find the optimal posture of the human at each instance, in which the pose and velocities of the follower robot's end-effector are not far from the current states, and suggest it to the human. The formulation in Eq. 4 works here too, where $[\mathbf{z}_t; \dot{\mathbf{z}}_t]$ are observed from the leader robot.

Initial Postural optimization The remote connection between the leader and follower robots makes it possible to start the teleoperation task from any initial human posture and the corresponding leader robot's initial posture and perform the same task with the follower robot, considering the human and leader robot's workspace. We use this feature to propose *initial postural optimization* for path-constrained and trajectory-constrained teleoperation tasks. In this type of postural optimization, the ergonomically intelligent system observes the human doing the task once, then calculates the optimal initial posture for the human to start *the same task* with the same task-space motion for the follower robot. It also requires recording the entire motion of the robots and the human during the task.

We form initial postural optimization as the following:

$$\mathbf{q}_0^{h^*} = \arg\max_{\mathbf{q}_0^h} \sum_{t=0}^{T} \text{Ergonomics}(\mathbf{q}_t^h) \quad (5)$$
$$\text{s.t.} \quad ||\hat{\dot{\mathbf{x}}}_t^h - J^h(\mathbf{q}_t^h, \psi)\dot{\mathbf{q}}_t||_\Sigma^2 < \epsilon \quad \forall t \geq 0$$

where $\hat{\dot{\mathbf{x}}}_t^h$ is the recorded velocity of the interaction point during the performed task, and $J^h(\mathbf{q}_t^h, \psi)$ is the Jacobian of the human kinematic chain. Here, the superscripts h, l, f refer to human, leader robot, and follower robot, respectively.

Postural Optimization by Interface Reconfiguration An inherent feature of path-constrained teleoperation is the ability to pause the teleoperation by disengaging the leader and follower robots from each other, move the leader robot to a new position, and resume the teleoperation. This is usually done by a clutch switch placed under the user's foot. Benefiting from this feature, we propose *postural optimization by interface reconfiguration*, in which, when the ergonomically intelligent system detects high-risk postures during a previously recorded task, suggests to pause the teleoperation. When teleoperation is paused, the system calculates a new optimal posture for the human to resume the teleoperation from it and continue the task.

The formulation is similar to the initial postural optimization but covers the time after the pause:

$$\mathbf{q}_{t_p+1}^{h^*} = \arg\max_{\mathbf{q}_{t_p+1}^h} \sum_{t=t_p+1}^{T} \text{Ergonomics}(\mathbf{q}_t^h) \quad (6)$$
$$\text{s.t.} \quad ||\hat{\dot{\mathbf{x}}}_t^h - J^h(\mathbf{q}_t^h, \psi)\dot{\mathbf{q}}_t||_\Sigma^2 < \epsilon \quad \forall t > t_p$$

where t_p is the time at pause in teleoperation.

Approach and Implementation

This section proposes our approach to solving the problems mentioned in the previous section and details their implementations.

Human Kinematics Model

We use a 10-DOF kinematics model (see Fig. 3) to analyze the upper-body motion of a human sitting on a chair and interacting with a robot. We assume that the chair is stationary with a known position w.r.t. the robot, and the human sits on the center of the chair. In (Yazdani et al. 2021b), we proposed an approach based on circle point analysis (Mooring, Roth, and Driels 1991) for estimating human segment lengths in p-HRI using interactive motion routines and we use it here. We encode human motion limits based on both fixed limits and the posture-dependant joint limit model from (Jiang and Liu 2018).

Particle Filter for Posture Estimation

We approximate the solution for the partially-observable problem of posture estimation in Eq. 2 by a particle filter (Thrun, Burgard, and Fox 2005). As the estimation of the human model from the trajectory of the leader robot has high ambiguity due to the redundancy, we add the joint angular velocities to our state variables and use the velocity of the robot's end-effector as a part of the observation. As a prior, we encode that the human *starts* the task in a static and neutral posture. We initialize M particles using a truncated normal distribution with the mean at the neutral posture $\mathbf{q}_{neutral}$ and set the initial angular velocities to zero.

DULA: Differentiable Upper Limb Assessment Model for Postural Optimization

Solving Eqs. 4, 5, 6 using a gradient-based optimization requires a differentiable and continuous ergonomics model as objective functions. Instead of using analytical models or approximations, we introduce *DULA* (Differentiable Upper

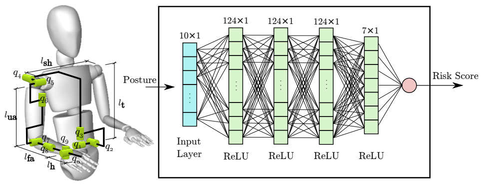


Figure 3: The structure of the DULA neural network.

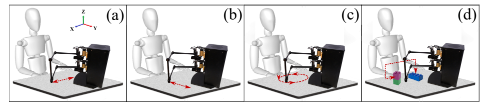


Figure 4: The tasks for human subject study: repetitive motions following a straight line in the X direction (a), a straight line in the Y direction (b), a circular path (c), and repetitive motions between random sides of two blocks positioned at different heights, with an *unprescribed* motion and high range of hand rotation (d).

Limb Assessment) which is an accurate, continuous, and differentiable model for upper body ergonomics. It is learned using a standard neural network from a dataset of exhaustively sampled postures, labeled by RULA.

To learn a continuous and differentiable model for RULA, we designed a fully connected regression-based neural network. While RULA provides discrete integer scores from 1 to 7, we choose to predict continuous labels. If we had instead performed multi-class classification based on the discrete labels, the resulting model would be less useful in optimization as there is no natural choice of a smooth objective to minimize or constrain ergonomic cost. This would negate our desire for a computationally helpful model. Hence, we perform regression to the multi-class labels.

The structure of the neural network is shown in Fig. 3. It includes 4 hidden layers with ReLU activation function. We found that a network with 124 units for the first three layers and 7 units for the last hidden layer worked best. We train the network using the standard mean squared error loss function for 2000 epochs using a learning rate of 0.001. We used 5-fold cross-validation to find the optimal network parameters.

We developed a dataset of 7.5 million upper body postures of a human. We additionally defined task parameters based on the RULA worksheet—the frequency of the arm and body motions; type and maximum load on the arm and body; neck angle; and whether any legs, feet, or arms are supported. We developed a script for automatic RULA assessment based on the posture and tasks parameters and verified it with several ergonomists. We used this script to label the posture dataset. Since postures with labels 1, 2, 6, and 7 are not frequent in the full range of human motion, we balanced the dataset by forcing the data generation scripts to generate enough data points with those labels. We split the dataset into 80% training and 20% testing sets.

Training of the neural networks takes about 60 hours on a Core i9 CPU with GPU support. We implemented and trained our network using PyTorch (Paszke et al. 2019).

Postural optimization

We explore two methods for postural optimization. Initially, we use the cross-entropy method (CEM) (Kobilarov 2012; De Boer et al. 2005) as a sample-based and gradient-free optimization. In our implementation, we use a Gaussian probability distribution N=10000 samples.

As an alternative method, we utilize a derivative-based optimization using our learned DULA model. We solve the nonlinear and constrained least-square postural optimization using sequential least squares programming (SLSQP) (Nocedal and Wright 2006) solver in SciPy. We calculated DULA gradients using automatic differentiation in PyTorch.

The following section examines variations of postural correction optimizations using these two methods and compares the resulting RULA score distributions.

Evaluation & Experimental Protocol

We evaluated our proposed postural estimation and optimization approaches through human subject studies and simulation experiments, respectively.

Human Subject Study for Posture Estimation

We conducted a human subject experiment in which participants interact with a 6-DOF Quanser HD^2 haptic interface as the leader robot (see Fig. 7) in teleoperation, and we recorded their upper body motion using a 12-camera Optitrack MoCap system for comparison. We recruited 8 participants (4 females, 4 males) with ages ranging from 25 to 33 years and heights in the range of 171 $\pm$ 21cm. Participants were graduate students from various programs and did not have any experience with teleoperationing robots and received 15 minutes of training with the robot. Each participant performed 4 tasks visualized in Fig 4. We provided a printed visual guide on the table for the first three tasks. However, the participants were not required to follow the path accurately. The participants were not told what posture to initialize the task from and how high they should be above the table to do the task. The robot collected data from the participant's motion.

To compare with other well-known approaches, we solved the postural estimation problem using two other deterministic methods: (1) boosted and bounded online least-squares IK optimization (*Online-IK*) in which we simply solve the inverse kinematics optimization independently at each time step by initializing it with the solution from the previous step, and (2) boosted and bounded offline least-squares trajectory IK optimization (*Offline-TrajIK*) in which we solve the inverse kinematics problem for the whole trajectory by initializing it with the solution trajectory from Online-IK. We used the dogleg algorithm with rectangular trust regions (Nocedal and Wright 2006) in SciPy as the solver.

Simulated Environment For Postural Optimization

We developed an open-source simulator for postural correction using ROS. It includes a human seated on a stool and

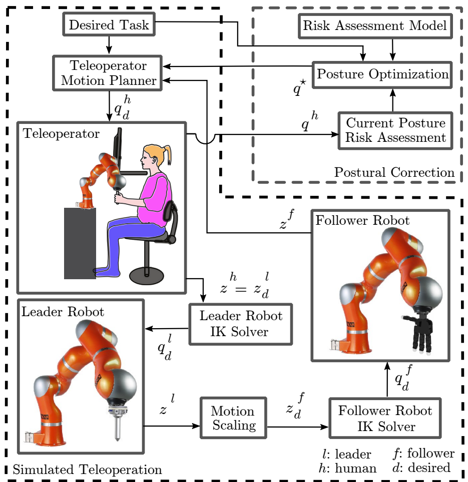


Figure 5: Overview of teleoperation simulation.

a KUKA LWR-4 robot (two robots for teleoperation). The simulated human behaves like a human in two ways: (1) physically controlling the interactive task and (2) accepting or rejecting the recommended optimal postural corrections.

We model this in an optimal motion planning framework with re-planning that finds a human joint trajectory that interacts with the robot to do the task while moving toward the optimal ergonomic posture. For example, it is formulated as the following for a teleoperation task:

$$\tau^{h^*}_{t\to H} = \arg\min_{\tau^{h}_{t\to H}} \sum_{t=t}^{H} ||\mathbf{x}^f_g - \mathbf{x}^f_t||^2_\Sigma + \alpha||\mathbf{q}^h_t - \mathbf{q}^{h^*}_t||^2_2 \quad (7)$$

where $\tau^f_{t\to H}$ is the trajectory of human posture from time t to the time of the horizon H, $\mathbf{x}^f_g$ is the goal pose of the follower robot, and ${}^f\mathbf{x}_t$ is the pose of the follower robot at time t. Here, $0 \leq \alpha \leq 1$ is a scalar number that models the postural correction acceptance and the effort of the human operator towards applying the postural correction. Details of the motion planning for the humans and the robots are presented in Fig. 5. We note that this model is likely a simplification of true human interactive behavior. We do not advocate for its use over human subject studies. Instead, we propose it as a useful tool for systematically exploring new human safety assessment and improvement algorithms.

Results

This section provides results for our ergonomically intelligent p-HRI system. We discuss the proposed posture estimation approach's performance compared with MoCap, and their corresponding estimated risk assessment results. Moreover, we showcase the accuracy of the learned model for DULA and its effect on postural optimization.

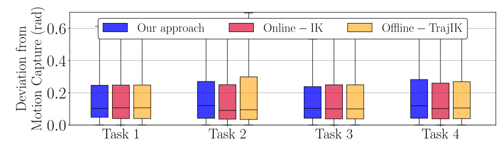


Figure 6: Deviation of the posture estimated by our approach vs Online-IK and Offline-TrajIK methods among all the tasks.

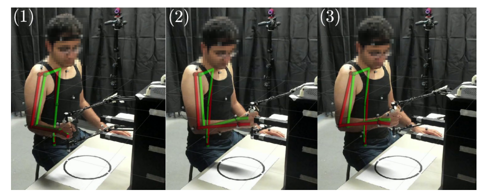


Figure 7: Video-overlaid skeletons show the posture estimated from our proposed approach (red) and MoCap (green).

Posture Estimation

Fig. 6 compares our posture estimation approach with two other least-squares IK solutions for redundant robots for all participants among 4 tasks. From the statistical analysis of the results, we can conclude that our probabilistic *particle filter* approach performance is not significantly different from the other two methods. Overall, the approach generally agrees with MoCap with a median deviation less than 0.09rad (less than 5deg) and upper quartile less than 0.25rad (less than 15deg) considering the observation solely from the stylus trajectory and having no extra sensors.

Since there is no ground truth for human posture to compare with, we provide a qualitative evaluation using video frames of a participant during task (3) with overlaid reconstructed skeletons from MoCap (green) and our approach (red) in Fig. 7. The estimated postures align well with the MoCap postures during the entire task.

Fig. 8 shows the maximum RULA score during a task for the postures from our approach and MoCap postures. Our approach successfully identified all instances where the score was higher than 2 (i.e., future investigation or change may be or is needed) in all 32 trials. It also resulted in the same interpretation of the RULA in 27 trials (84.37%) and the same RULA score in 21 trials (65.63%).

Overall, the results showed that our posture estimation solely from the robot has the potential to be used for continuous monitoring of ergonomics in p-HRI and rising alerts when further investigation or change in the task is required.

Postural Optimization

The trained model for DULA performs well on the test dataset and results in 99.73% accuracy. In the accuracy test, we round our continuous output to the nearest integer when reporting accuracy. The lowest diagonal element in the confusion matrix for DULA is 99.38%, which shows the high

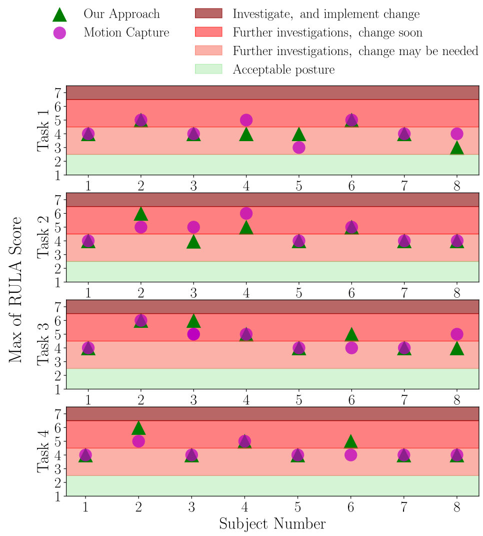


Figure 8: Comparison of the maximum value of RULA scores of a task using estimated posture from our approach and MoCap for all the participants and trials.

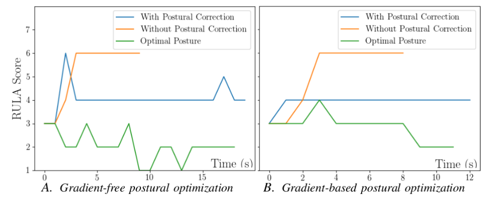


Figure 9: Comparison between gradient-free and gradient-based postural correction on a teleoperation task. Lower scores are better.

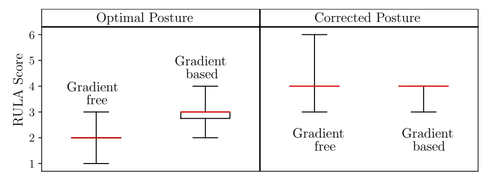


Figure 10: Comparison in optimal posture and corrected posture between gradient-free and gradient-based optimization.

accuracy of the learned model across all ranges.

Figure 9 shows the postural correction results for gradient-free (left) and gradient-based (right) optimization in a simple teleoperation task. It presents the RULA scores for calculated optimal posture (green), human posture without postural correction (orange), and the human posture after applying the correction (blue) according to the human control and acceptance model during the task using $\alpha \sim \{0, 0.75\}$. The plot shows that the risk is reduced after the simulated human applies the suggested optimal postural correction. Also, the task completion time increases without a significant decrease in risk. We can also see that the gradient-based approach provides a smoother motion concerning the RULA risk score and avoids going through postures with a risk higher than 4. It also results in a shorter task completion time.

Figure 10 provides more information on comparing gradient-free and gradient-based postural optimization. It shows that the median risk scores of target optimal postures from the gradient-free approach is lower. However, the postures of the simulated human are more comfortable after applying the optimal posture calculated from the gradient-based optimization. We believe this is due to the smoother optimal postures that has been suggested to the simulated human from the gradient-based method. Moreover, the gradient-based approach is much faster. Each iteration of the gradient-free approach using 10000 samples takes 2.7 minutes to solve, while the gradient-based approach takes only tenths of a second.

Conclusion

This paper proposed a framework for ergonomically intelligent physical human-robot interaction that includes posture estimation, risk assessment, and postural optimization. We showed that the interacting robot is an adequate sensor to continuously monitor posture to assess the ergonomics of an interactive task. We described a probabilistic approach based solely on the trajectory of the robot, which is already necessary to perform the p-HRI task. Additionally, we introduced DULA, a differentiable and continuous ergonomics model to assess human upper body posture. DULA is learned to replicate the non-differentiable RULA assessment tool using a neural network, and it is 99.73% accurate and computationally designed for efficient postural optimization for p-HRI and other related applications. We also introduced a framework for postural optimization in p-HRI using DULA and presented a demo task. The results reveal that postural optimization using DULA lowers the risk of injuries.

There are several directions for future work. For the posture estimation part, Our approach can be used for full arm representation in virtual reality systems based on the pose of controllers and headsets. When available, it would be straightforward to combine our proposed approach with other sensing modalities (e.g., vision or MoCap). We are examining the combination of our approach with OpenPose (Cao et al. 2019) using incremental smoothing (Kaess et al. 2012) in order to improve estimation results and decrease runtime. Our risk assessment work can be immediately extended by following the same procedure for REBA, a whole-body assessment tool. Additionally, we intend to conduct a human subject study to evaluate our postural optimization and correction approach. We wish to compare different means of feedback for the posture correction to the human, including visual, auditory, and haptic feedback.

References

Akhter, I.; and Black, M. J. 2015. Pose-conditioned joint angle limits for 3D human pose reconstruction. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1446–1455.

Alvarez, M.; Torricelli, D.; del Ama, A. J.; Pinto, D.; Gonzalez-Vargas, J.; Moreno, J. C.; Gil-Agudo, A.; and Pons, J. L. 2017. Simultaneous estimation of human and exoskeleton motion: a simplified protocol. In *IEEE Int. Conference on Rehabilitation Robotics*, 1431–1436.

Bazazan, A.; Dianat, I.; Feizollahi, N.; Mombeini, Z.; Shirazi, A. M.; and Castellucci, H. I. 2019. Effect of a posture correction–based intervention on musculoskeletal symptoms and fatigue among control room operators. *Applied ergonomics* 76: 12–19.

Busch, B.; Maeda, G.; Mollard, Y.; Demangeat, M.; and Lopes, M. 2017. Postural optimization for an ergonomic human-robot interaction. In *International Conference on Intelligent Robots and Systems*, 2778–2785.

Busch, B.; Toussaint, M.; and Lopes, M. 2018. Planning ergonomic sequences of actions in human-robot interaction. In *International Conference on Robotics and Automation*, 1916–1923.

Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.-E.; and Sheikh, Y. 2019. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence* 43(1): 172–186.

Chen, L.; Figueredo, L. F.; and Dogar, M. R. 2018. Planning for muscular and peripersonal-space comfort during human-robot forceful collaboration. In *IEEE-RAS International Conference on Humanoid Robotics*, 1–8.

Colyer, S. L.; Evans, M.; Cosker, D. P.; and Salo, A. I. 2018. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open* 4(1): 1–15.

da Costa, B. R.; and Vieira, E. R. 2010. Risk factors for work-related musculoskeletal disorders: a systematic review of recent longitudinal studies. *American journal of industrial medicine* 53(3): 285–323.

Das, S.; Trutoiu, L.; Murai, A.; Alcindor, D.; Oh, M.; De la Torre, F.; and Hodgins, J. 2011. Quantitative measurement of motor symptoms in Parkinson's disease: A study with full-body motion capture data. In *IEEE Int. Conf. of Engineering in Medicine and Biology Society*, 6789–6792.

De Boer, P.-T.; Kroese, D. P.; Mannor, S.; and Rubinstein, R. Y. 2005. A tutorial on the cross-entropy method. *Annals of operations research* 134(1): 19–67.

Erdinc, O.; and Vayvay, O. 2008. Ergonomics interventions improve quality in manufacturing: a case study. *International Journal of Industrial and Systems Engineering* 3(6): 727–745.

Ganesh, P.; Volle, K.; Buzaud, P.; Brink, K.; and Willis, A. 2021. Extrinsic calibration of camera and motion capture systems. In *SoutheastCon 2021*, 01–08. IEEE.

Gorecky, D.; Schmitt, M.; Loskyll, M.; and Zühlke, D. 2014. Human-machine-interaction in the industry 4.0 era. In *2014 12th IEEE international conference on industrial informatics (INDIN)*, 289–294. Ieee.

Guerrero, C. R.; Marinero, J. C. F.; Turiel, J. P.; and Muñoz, V. 2013. Using "human state aware" robots to enhance physical human–robot interaction in a cooperative scenario. *Computer methods and programs in biomedicine* 112(2): 250–259.

Hignett, S.; and McAtamney, L. 2000. Rapid entire body assessment (REBA). *Applied ergonomics* 31(2): 201–205.

Jiang, Y.; and Liu, C. K. 2018. Data-driven approach to simulating realistic human joint constraints. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 1098–1103.

Kaess, M.; Johannsson, H.; Roberts, R.; Ila, V.; Leonard, J. J.; and Dellaert, F. 2012. iSAM2: Incremental smoothing and mapping using the Bayes tree. *International Journal of Robotics Research* 31(2): 216–235.

Kim, W.; Lee, J.; Peternel, L.; Tsagarakis, N.; and Ajoudani, A. 2017. Anticipatory robot assistance for the prevention of human static joint overloading in human–robot collaboration. *IEEE Robotics and Automation Letters* 3(1): 68–75.

Kim, W.; Lorenzini, M.; Kapıcıoğlu, K.; and Ajoudani, A. 2018. Ergotac: A tactile feedback interface for improving human ergonomics in workplaces. *IEEE Robotics and Automation Letters* 3(4): 4179–4186.

Kobilarov, M. 2012. Cross-entropy motion planning. *International Journal of Robotics Research* 31(7): 855–871.

Lorenzini, M.; Kim, W.; De Momi, E.; and Ajoudani, A. 2018. A real-time graphic interface for the monitoring of the human joint overloadings with application to assistive exoskeletons. In *International Symposium on Wearable Robotics*, 281–285. Springer.

Marin, A. G.; Shourijeh, M. S.; Galibarov, P. E.; Damsgaard, M.; Fritzsch, L.; and Stulp, F. 2018. Optimizing contextual ergonomics models in human-robot interaction. In *International Conference on Intelligent Robots and Systems*, 1–9.

Mathis, A.; Mamidanna, P.; Cury, K. M.; Abe, T.; Murthy, V. N.; Mathis, M. W.; and Bethge, M. 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Technical report, Nature Publishing Group.

McAtamney, L.; and Corlett, E. N. 1993. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied ergonomics* 24(2): 91–99.

Mooring, B. W.; Roth, Z. S.; and Driels, M. R. 1991. *Fundamentals of manipulator calibration*. Wiley New York.

Motabar, H.; and Nimbarte, A. D. 2021. The effect of task rotation on activation and fatigue response of rotator cuff muscles during overhead work. *Applied Ergonomics* 97: 103461.

Nocedal, J.; and Wright, S. 2006. *Numerical optimization*. Springer.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; and Chintala, S. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Neural Info. Proc. Sys.*

Peternel, L.; Fang, C.; Laghi, M.; Bicchi, A.; Tsagarakis, N.; and Ajoudani, A. 2020. Human Arm Posture Optimisation in Bilateral Teleoperation Through Interface Reconfiguration. In *IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics*.

Peternel, L.; Kim, W.; Babič, J.; and Ajoudani, A. 2017. Towards ergonomic control of human-robot co-manipulation and handover. In *IEEE-RAS International Conference on Humanoid Robotics*, 55–60.

Peternel, L.; Tsagarakis, N.; Caldwell, D.; and Ajoudani, A. 2018. Robot adaptation to human physical fatigue in human-robot co-manipulation. *Autonomous Robots* 1–11.

Punnett, L.; and Wegman, D. H. 2004. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *Journal of electromyography and kinesiology* 14(1): 13–23.

Rahal, R.; Matarese, G.; Gabiccini, M.; Artoni, A.; Prattichizzo, D.; Giordano, P. R.; and Pacchierotti, C. 2020. Caring about the human operator: haptic shared control for enhanced user comfort in robotic telemanipulation. *IEEE Tran. on Haptics* 13(1): 197–203.

Ramaganesh, M.; Jayasuriyan, R.; Rajpradeesh, T.; Bathrinath, S.; and Manikandan, R. 2021. Ergonomics hazard analysis techniques-A technical review. *Materials Today: Proceedings* .

Saremi, S.; Mirjalili, S.; Lewis, A.; Liew, A. W. C.; and Dong, J. S. 2018. Enhanced multi-objective particle swarm optimisation for estimating hand postures. *Knowledge-Based Systems* .

Schneider, S. P. 2001. Musculoskeletal injuries in construction: a review of the literature. *Applied occupational and environmental hygiene* 16(11): 1056–1064.

Shafti, A.; Ataka, A.; Lazpita, B. U.; Shiva, A.; Wurdemann, H. A.; and Althoefer, K. 2019. Real-time robot-assisted ergonomics. In *International Conference on Robotics and Automation*, 1975–1981.

Sherwani, F.; Asad, M. M.; and Ibrahim, B. 2020. Collaborative robots and industrial revolution 4.0 (ir 4.0). In *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*, 1–5. IEEE.

Shikdar, A. A.; and Al-Hadhrami, M. A. 2007. Smart workstation design: an ergonomics and methods engineering approach. *International Journal of Industrial and Systems Engineering* 2(4): 363–374.

Thrun, S.; Burgard, W.; and Fox, D. 2005. *Probabilistic Robotics*. MIT press.

Van der Spaa, L.; Gienger, M.; Bates, T.; and Kober, J. 2020. Predicting and Optimizing Ergonomics in Physical Human-Robot Cooperation Tasks. In *International Conference on Robotics and Automation*, 1799–1805. IEEE.

Vartholomeos, P.; Katevas, N.; Papadakis, A.; and Sarakis, L. 2016. Design of motion-tracking device for intuitive and safe human-robot physical interaction. In *IEEE Int. Conference on Telecommunications and Multimedia*, 1–6.

Vianello, L.; Mouret, J.-B.; Dalin, E.; Aubry, A.; and Ivaldi, S. 2021. Human Posture Prediction during Physical Human-Robot Interaction. *IEEE Robotics and Automation Letters* .

Vos, T.; Barber, R. M.; Bell, B.; Bertozzi-Villa, A.; Biryukov, S.; Bolliger, I.; Charlson, F.; Davis, A.; Degenhardt, L.; Dicker, D.; et al. 2015. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 386(9995): 743–800.

Wang, X.; Maurin, M.; Mazet, F.; Maia, N. D. C.; Voinot, K.; Verriest, J. P.; and Fayet, M. 1998. Three-dimensional modelling of the motion range of axial rotation of the upper arm. *Journal of biomechanics* 31(10): 899–908.

Xiao, X.; and Zarar, S. 2018. A wearable system for articulated human pose tracking under uncertainty of sensor placement. In *IEEE Int. Conference on Biomedical Robotics and Biomechatronics*, 1144–1150.

Yazdani, A.; Novin, R. S.; Merryweather, A.; and Hermans, T. 2021a. DULA: A Differentiable Ergonomics Model for Postural Optimization in Physical HRI. *arXiv preprint arXiv:2107.06875* .

Yazdani, A.; Novin, R. S.; Merryweather, A.; and Hermans, T. 2021b. Is The Leader Robot an Adequate Sensor for Posture Estimation and Ergonomic Assessment of A Human Teleoperator? In *IEEE International Conference on Automation Science and Engineering*.

Yazdani, A.; and Sabbagh Novin, R. 2021. Posture Estimation and Optimization in Ergonomically Intelligent Teleoperation Systems. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 604–606.

Yu, W.; Ignatius, T.; Li, Z.; Wang, X.; Sun, T.; Lin, H.; Wan, S.; Qiu, H.; and Xie, S. 2012. Work-related injuries and musculoskeletal disorders among factory workers in a major city of China. *Accident Analysis & Prevention* 48: 457–463.

CHAPTER 5

DULA AND DEBA: DIFFERENTIABLE ERGONOMIC RISK MODELS FOR POSTURAL ASSESSMENT AND OPTIMIZATION IN ERGONOMICALLY INTELLIGENT pHRI

This chapter focuses on the second part of the ergonomically intelligent teleoperation and physical human-robot interaction framework. Assuming the known posture from Chapters 2 and 3, we investigate the postural assessment and optimization using different gradient-free and gradient-based approaches. We propose two differentiable and continuous ergonomics models that are easy to use in gradient-based optimization and as informative as well-studied risk-assessment tools in ergonomics. We also use learned and generic risk-assessment models to assess the risk of injuries for two datasets of human motions, including solo and interactive tasks.

I primarily completed this work under the guidance of Prof. Tucker Hermans and Prof. Andrew Merryweather. Roya Sabbagh Novin assisted in system design and learning the ergonomic models. Adam Conkey helped model the robots in Gazebo and the solver for cross-entropy optimization. Katie Allen helped with one of the human-subject studies. Lorenzo Vianello and Prof. Serena Ivaldi from INRIA Nancy, France, helped by preparing and sharing the dataset of their human-subject study.

2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2022)
October 23-27, 2022, Kyoto, Japan

DULA and DEBA: Differentiable Ergonomic Risk Models for Postural Assessment and Optimization in Ergonomically Intelligent pHRI

Amir Yazdani[1], Roya Sabbagh Novin[1], Andrew Merryweather[1], and Tucker Hermans[2]

Abstract— **Ergonomics and human comfort are essential concerns in physical human-robot interaction applications. Defining an accurate and easy-to-use ergonomic assessment model stands as an important step in providing feedback for postural correction to improve operator health and comfort. Common practical methods in the area suffer from inaccurate ergonomics models in performing postural optimization. In order to retain assessment quality, while improving computational considerations, we propose a novel framework for postural assessment and optimization for ergonomically intelligent physical human-robot interaction. We introduce DULA and DEBA, differentiable and continuous ergonomics models learned to replicate the popular and scientifically validated RULA and REBA assessments with more than 99% accuracy. We show that DULA and DEBA provide assessment comparable to RULA and REBA while providing computational benefits when being used in postural optimization. We evaluate our framework through human and simulation experiments. We highlight DULA and DEBA's strength in a demonstration of postural optimization for a simulated pHRI task.**

I. Introduction and Motivation

Improving workplace ergonomics and reducing the risk of work-related musculoskeletal disorders (WMSDs) has been the focus of researchers for many years [1,2]. High rates of injuries in industry [3] further motivates these studies. The industry 4.0 and 5.0 initiatives [4,5] and the development of collaborative workplaces where smart agents (e.g., cobots [6]) physically collaborate with humans to perform different types of tasks, highlights the growing importance of human comfort and ergonomics in physical human-robot interaction (pHRI). Moreover, the fact that smart agents share a physical interaction point with the operator provides the opportunity to develop novel technologies for those agents to consider human comfort in their behaviors.

A common and practical solution to improve workplace ergonomics is to perform risk assessments and analyze human comfort during task execution. Ergonomists usually focus their analysis on the worst posture achieved during the task by taking measurements of the worker's posture onsite or from recorded images or videos. They normally use risk assessment tools [7] to assess the task and provide intervention and improvement in forms of user training [8], workstation modification [9] or task rotation [10]. However, in most cases, it is not possible to have an ergonomist

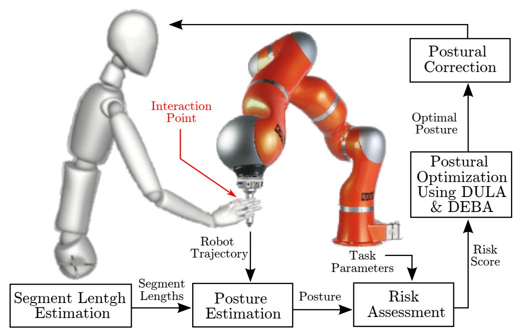


Fig. 1: Ergonomically intelligent pHRI, adapted from [11].

present during the task to perform a comprehensive risk assessment, which results in missing events or only partial understanding of awkward postures. In addition, these sampling strategies lead to failures in detecting the most hazardous combinations of force and posture, increasing the inaccuracy of an assessment. Finally, although postural correction by ergonomists generally helps reduce the risk of injuries, in many cases, the corrections are sub-optimal or tailored to specific task requirements, which might not generalize across task variability.

To provide a solution for the above issues, we previously introduced a framework for an *ergonomically intelligent teleoperation system* in [11], which includes smart and autonomous posture estimation, postural ergonomics assessment, and postural optimization. Fig. 1 visualizes our proposed framework. As a first step to realizing this system, we investigated a new approach for posture estimation in teleoperation relying solely on sensing from the leader robot [12]. Although, our implementation focused on teleoperation, it can be extended to other pHRI tasks with minor modifications. In this paper, we focus on ergonomics assessment and postural optimization in the context of pHRI.

pHRI covers a vast number of applications. To help present our framework, we categorize pHRI tasks into *direct physical interaction* with the robot (e.g., robotic physical therapy [13]), *co-manipulation* [14], *assistive holding* [15], *handover* [16], and *teleoperation* [17]. We split teleoperation into three main types: *goal-constrained teleoperation* (e.g., pick and place with a desired pose for the object), *path-constrained teleoperation* (e.g., turning a valve where the path to follow is constrained based on the diameter of the valve), and *trajectory-constrained teleoperation* (e.g., welding or sealing where the operator should follow a reference

[1]Department of Mechanical Eng. and Robotics Center, University of Utah, Salt Lake City, UT, USA, amir.yazdani@utah.edu.
[2]School of Computing and Robotics Center, University of Utah, Salt Lake City, UT, USA, and NVIDIA, Seattle, WA, USA. Research reported in this publication was supported in part by DARPA under grant N66001-19-2-4035 and the National Institute for Occupational Safety and Health under award number T420H008414-10.

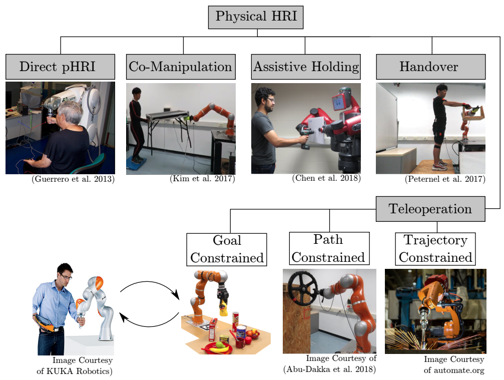


Fig. 2: Categories of physical HRI applications.

travel velocity in addition to a defined path to perform a uniform and good quality weld or seal). Fig. 2 visualizes the different types of pHRI applications.

Autonomous Postural Optimization has received substantial research attention using various technologies around humans such as collaborative robots [18], smart personal trainers [19], and VR systems [20]. In these systems, the interacting agent should consider human comfort and ergonomics in its behaviour and motion planning [11]. For example, in co-manipulation the collaborative robot should account for satisfying ergonomics and comfort of the human during the collaboration when generating controls. Ergonomics and comfort can be included in postural optimization in two ways: (1) as a constraint, and (2) as a part of the objective function. In the above example, the robot can behave in such a way that avoids high-risk and awkward postures (comfort as a constraint) or minimizes the risk of musculoskeletal disorders for the whole duration of the task (comfort as a part of the objective function). The literature in postural optimization in pHRI and teleoperation is minimal. Table I summarizes the relevant literature in both areas. Most works use gradient-based optimization techniques for postural optimization [14–18,21–23], while a couple opt for gradient-free methods [24,25]. Although gradient-free approaches are more time and resource consuming, they can work with non-differentiable objective functions. In contrast, gradient-based approaches are usually faster, but require the costs and constraints to be continuous and differentiable.

Developing a model of human comfort lies at the heart of effective postural optimization. Ergonomists have provided simple models which are easier for human experts to calculate by hand and as such are common in practice. These models include the NASA TLX [26], RULA [27], REBA [28], strain index [29], and ACGIH TLV [30]. Importantly, these models are supported by extensive human subject studies that validate their effectiveness on reducing ergonomic risk factors [31]. pHRI researchers have also proposed computational models for human comfort, including peripersonal space [15], muscle fatigue [21], joint overloading [14], and muscular comfort [15].

Among all ergonomics risk assessment tools, RULA (*Rapid Upper Limb Assessment*) [27] and REBA (*Rapid Entire Body Assessment*) [28] depend most directly on human posture and provide quantitative scores, making them good choices for postural optimization applications. RULA targets the upper body, while REBA assesses the whole body posture of the human. Both models take as input the joint angles of the human and output discrete, integer scores where the higher number indicates the higher risk of WMSDs, i.e. lower human comfort [32]. The interpretation of each score can be found in [27,28]. Van der Spaa et al. provided the only study of directly using risk assessment tools in derivative-free postural optimization [25].

The discrete, step-function like output scores in RULA and REBA [23] prevent its use in gradient-based postural optimization, as the gradient is either 0 or undefined. Based on our experience, using the risk assessment models directly in gradient-free optimization is also time-expensive, as the plateaus often prevent progress toward the optimal global solution. Thus, researchers in pHRI often use computational approximations of ergonomic assessment models to enable gradient-based postural optimizations. As Table I shows, quadratic approximations [16,17,23] make up the standard approach in the literature. However, these approximations deviate far from the scientifically validated assessments, causing uncertainty about their reliability in providing ergonomic assessment and improvement.

In contrast, this paper presents a continuous and differentiable, tight approximation of the fully validated RULA and REBA assessments tools. This enables us to improve both gradient-based and gradient-free optimization using the continuous models. We propose two continuous and differentiable ergonomics models for postural optimization, *Differentiable Upper Limb Assessment* (DULA) and *Differentiable Entire Body Assessment* (DEBA) approximating the RULA and REBA survey tools respectively. We build DULA and DEBA as neural network regression models, which we train to predict the same output scores as RULA and REBA, from associated human postures. However, DULA and DEBA output continuous, not discrete, values across the assessment range enabling smoother optimization landscapes compared to the non-differentiable models. Furthermore, they provide the gradient of the risk with respect to each joint enabling efficient use in optimization. We show that both models achieve over 98% agreement with RULA and REBA on large scale synthetic datasets. We then investigate the utility of DULA and DEBA to improve both gradient-based and gradient-free postural optimization in pHRI and teleoperation. We provide our models for others use at https://sites.google.com/view/differentiable-ergo.

We organize the remainder of the paper as follows. Section II introduces our framework for postural optimization in pHRI. Section III defines the structure, training, and data generation for DULA and DEBA. We define our ergonomic assessment and optimization experiments in Sec IV and analyze our results in Sec V. We conclude in Sec. VI.

Ergonomics Model		Analytical Models	Learned Models	Risk Assessment Tools	
				Quadratic Approximation	No Approximation
Optimization	Gradient-Based	■ Peternel et al. [18] ▪■ Peternel et al. [21] □ Chen et al. [15] ◆ Peternel et al. [22] ■ Kim et al. [14]	This work	◇Rahal et al. [17] ▪ Busch et al. [23] ▪ Busch et al. [16]	Infeasible for current assessment tools
Optimization	Gradient-Free				■ Van der Spaa et al. [25]
Non-Optimization					□ Shafti et al. [33]
Legend		□ pHRI: Assistive Holding, ▪ pHRI: Handover, ■ pHRI: Co-Manipulation ◆ Teleop: Goal-Constrained with Postural Optimization By Interface Reconfiguration ◇ Teleop: Goal-Constrained with Online Postural Optimization			

TABLE I: State of the art of postural optimization in pHRI and teleoperation.

II. Postural Optimization in pHRI and Teleoperation

In this section, we introduce the problem of postural optimization in pHRI. We focus our presentation on the applications of co-manipulation and teleoperation since they contend with the largest ranges of continuous motions. However, the same problem statement can be extended to other applications with minor modifications. We present the problem formulation with a generic "comfort" objective. We formally define our proposed metrics in Section. III.

A. Postural Optimization in Co-Manipulation

In co-manipulation, the robot helps a human to move an object (usually a heavy one) from an initial pose to a desired pose [21]. There is no specific trajectory for the object to follow; the human plans and leads the motion, and the robot follows and provides help carrying the object. Meanwhile, ergonomically intelligent pHRI seeks to find the optimal posture for the human at each instance while holding the pose and velocities at the interaction point close to the current state. Once found, the system will suggest the solution to the human. We call this *online postural optimization*. We can formalize finding the ergonomically optimal posture with the following optimization problem:

$$\begin{aligned} \mathbf{q}_t^* = \arg\max_{\mathbf{q}_t} & \quad \mathrm{Comfort}(\mathbf{q}_t) \\ \text{s.t.} & \quad ||\phi([\mathbf{q}_t;\dot{\mathbf{q}}_t],\psi) - [\mathbf{z}_t;\dot{\mathbf{z}}_t]||_\Sigma^2 < \epsilon \\ & \quad \mathbf{q}_t \in \text{Range of Motion} \end{aligned} \tag{1}$$

where $\mathbf{q}_t^*$ and $\mathbf{q}_t$ are the optimal posture and current posture at time t, respectively, ϕ is the forward kinematics of the human, and ψ is the human segment lengths. It is important to note that the optimal posture $\mathbf{q}_t^*$ from (1) for each time step is then suggested to the human to move towards. The human can refuse or accept the correction as much as they desire while completing the task.

B. Postural Optimization in Teleoperation

The main distinction of teleoperation from other pHRI applications is the motion (and force) coupling between the leader and follower robots. The coupling between the two need not be a one-to-one scale and can be paused and resumed during task execution. This enables the teleoperator to disengage the leader robot in the middle of a task in many types of application, reposition their postures (and the leader robot as the result), and resume the teleoperation from a more comfortable posture [22]. The relative position trajectory of the follower robot remains unchanged; however, its velocity trajectory changes due to the pause. The motion scaling also helps keep the teleoperator's posture in a smaller comfort zone while the follower robot operates in a much wider zone. To use this feature of teleoperation in postural optimization, we customize three postural optimization problems for the three teleoperation problems shown in Fig. 2.

1) Online Postural Optimization: Online postural optimization in teleoperation is very similar to online postural optimization in co-manipulation. The difference is that instead of having initial and goal poses for the shared object in co-manipulation, in teleoperation, we have initial and goal poses for the end-effector of the follower robot (or the object it manipulates). Hence, the ergonomically intelligent system should find the optimal posture of the human at each instance, in which the pose and velocities of the follower robot's end-effector are not far from the current states, and suggest it to the human. The formulation in (1) works here too, where $[\mathbf{z}_t;\dot{\mathbf{z}}_t]$ are observed from the leader robot.

2) Initial Postural Optimization: The remote connection between the leader and follower robots makes it possible to start the teleoperation task from any initial human posture and corresponding initial posture of the leader robot without changing the follower robot. We use this feature to propose *initial postural optimization* for path-constrained and trajectory-constrained teleoperation tasks. In this type of postural optimization, the ergonomically intelligent system can get info about the path (or trajectory) of the task either by observing the human doing the task once, or using inference approaches based on priors on the task objectives (e.g. using object pose estimation techniques to understand the diameter, center, and the rotation axis of a valve to rotate [34]), then calculates the optimal initial posture for the human to start *the same task* with the same task-space motion for the follower robot. Formally we define the problem

$$\begin{aligned} \mathbf{q}_0^{h^*} = \arg\max_{\mathbf{q}_0^h} & \quad \sum_{t=0}^{T} \mathrm{Comfort}(\mathbf{q}_t^h) \\ \text{s.t.} & \quad ||\hat{\dot{\mathbf{x}}}_t^h - J^h(\mathbf{q}_t^h,\psi)\dot{\mathbf{q}}_t^h||_\Sigma^2 < \epsilon \quad \forall t \geq 0 \end{aligned} \tag{2}$$

where $\hat{\dot{\mathbf{x}}}_t^h$ is the recorded or inferred velocity profile of the interaction point during the task, and $J^h(\mathbf{q}_t^h, \psi)$ is the Jacobian of the human kinematic chain. Here, the superscripts h, l, f refer to human, leader robot, and follower robot, respectively.

3) Postural Optimization by Interface Reconfiguration: An inherent feature of path-constrained teleoperation is the ability to pause the teleoperation by disengaging the leader and follower robots from each other, move the leader robot to a new position, and resume the teleoperation. This is usually done by a clutch switch placed under the user's foot. To benefit from this feature, we propose *postural optimization by interface reconfiguration*. In this case the ergonomically intelligent system suggests to pause the teleoperation and reconfigure the human posture when it detects high-risk postures during a task. When teleoperation is paused, the system calculates a new optimal posture for the human to resume the teleoperation from and continue the task.

The formulation is the same as that of initial postural optimization in (2), but only covers the time after the pause. We define t_p as the time at pause in teleoperation. Then we simply change the indices for the summation in the objective and enumeration of constraints to start at t_p instead of 0. Similar to the initial posture optimization, $\hat{\dot{\mathbf{x}}}_t^h$ can come from a variety of sources. This could be derived from a history of a human performing the task or inferred from task objectives (e.g. using object tracking to locate the welding seam [35] and use the prior knowledge on proper reference travel velocity to follow [36]).

III. Differentiable Ergonomics Models

This section proposes our approach to develop differentiable comfort models for use in the objective functions in Sec. II. We achieve this by learning neural network models that replicate the output of the RULA and REBA models. By using neural networks we produce continuous, differentiable models to replace the discrete, non-differentiable RULA and REBA. We name the differentiable variants of these models *DULA* (Differentiable Upper Limb Assessment) and *DEBA* (Differentiable Entire Body Assessment). We train these models using a standard supervised learning approach. In this remainder of this section we give the details of the DULA and DEBA architectures, training the models and generating the training data.

A. DULA and DEBA

To learn continuous and differentiable models approximating RULA and REBA, we designed fully connected neural networks for regressions. While RULA provide discrete integer scores from 1 (lowest risk of WMSDs) to 7 (highest risk of WMSDs), and REBA provide discrete integer scores from 1 (lowest risk of WMSDs) to 15 (highest risk of WMSDs), we choose to predict continuous labels for those models. If we had instead performed multi-class classification based on the discrete labels, the resulting model would be less useful for optimization as there is no natural choice of a smooth function to minimize or constrain ergonomic cost. This would negate our desire for a computationally helpful model. Hence, we perform continuous regression to the multi-class integer labels.

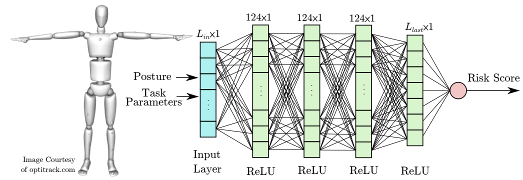


Fig. 3: The structures of the neural networks for learning DULA and DEBA. For DULA L_{in} = 10+task parameters and L_{last} = 7, where for DEBA L_{in} = 11 + task parameters and L_{last} = 12.

We calculate the loss for each forward path of the training based on the difference between the continuous score from the forward path and the integer label for the data point

$$\frac{1}{N}\sum_{i=1}^{N}||y_i - f(\mathbf{q}_i, \theta)||_2^2 \qquad (3)$$

where y_i is the label, $\mathbf{q}_i$ is the input posture, θ is the task parameters, and N is the number of data points.

However, in calculation of the models' accuracy, we round the continuous scores to their nearest integers, and then compare them with the integer labels from RULA and REBA.

Looking at the REBA literature [28], we observe that the *activity score* parameter is added directly to the output of Table C to calculate the final REBA score. As the information for computing the activity score is available to us a priori and remains constant during a task, we do not want to re-learn it. We thus exclude it from the learning process and design the output of the network to represent scores from 1 to 12.

The structure of the neural networks for DULA and DEBA are shown in Fig. 3. We performed 5-fold cross-validations to find the optimal network parameters, which revealed that the networks with 4 hidden layers with ReLU activation function, with 124 units for the first three layers and 7 units for DULA and 12 units for DEBA at the last hidden layer worked the best. We trained the networks using the loss function in (3) for 2000 epochs using a learning rate of 0.001. Training of the neural networks are done on a NVIDIA 1070-Ti GPU. We implemented and trained our network using a mixture of tools from PyTorch [37], scikit-learn [38], and skorch [39].

B. Dataset for Learning DULA and DEBA

We generated two datasets of 6 million upper and full body postures each for DULA and DEBA respectively. We additionally defined task parameters based on the RULA and REBA worksheets—the frequency of the arm and body motions; type and maximum load on the arm and body; neck angle; the activity score; coupling score; and whether any legs, feet, or arms are supported. We developed scripts for automatic RULA and REBA assessment based on the posture and tasks parameters and verified it with several ergonomists. We used these scripts to label the posture datasets. Since postures with RULA labels 1, 2, 6, and 7 and DEBA labels of 1, 2, 3 are not frequent in the full range of human motion,

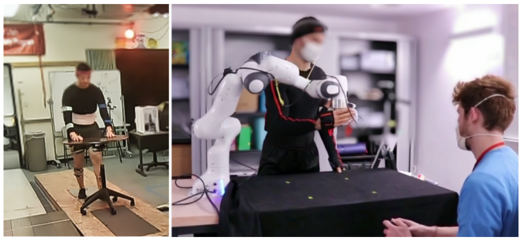


Fig. 4: Examples of human subject studies: (left): pushing a table, and (right) interacting with a remotely-controlled robot from [40].

we balanced the dataset by setting the data generation scripts to generate sample data points around the neutral posture until we have an equal number of data with each label. We split the dataset into 80% training and 20% testing sets.

IV. Human Subject and Simulation Experiments

In this section, we share the details of the postural optimization algorithms and the simulation environment we developed to evaluate the effect of postural optimization in reducing the risk of WMSDs.

A. Postural Assessment Using DULA and DEBA

To evaluate our ergonomic models, we compare the ergonomic risk scores from DULA/DEBA to scores from RULA/REBA over two dataset of postures from human subject studies. We conducted a human subject study[1] in which participants perform four common solo tasks in industrial environments including lifting and twisting, pulling, pushing, and reaching. We recorded their postures using an OptiTrack motion capture system and assessed them using RULA/REBA and DULA/DEBA models. We recruited 4 participants, 22-31 years old from university students.

We additionally evaluate the dataset of human motions during a pHRI subject study from Vianello et al. [40] using RULA/REBA and DULA/DEBA. In this human subject study, the human performs a 3 type of co-manipulation tasks with a Franka Emika Panda robot. The subject's posture is recorded using a Xsens MVN suit. Fig. 4 shows snapshots of these two studies corresponding to trials in Fig 7.

B. Simulated Environment For Postural Optimization

We developed an open-source simulator for postural correction using ROS. It includes a human and a KUKA LWR-4 robot (two KUKA robots for teleoperation). The simulated human behaves like a human in two ways: (1) physically controlling the interactive task, and (2) accepting or rejecting the recommended optimal postural corrections. We model this in an optimal motion planning framework with re-planning that finds a human joint trajectory that interacts with the robot to do the task while moving toward the optimal ergonomic posture. For example, it is formulated as the following for a teleoperation task:

$$\tau^{h^*}_{t\to H} = \arg\min_{\tau^h_{t\to H}} \sum_{t=t}^{H} ||\mathbf{x}^f_g - \mathbf{x}^f_t||^2_\Sigma + \alpha||\mathbf{q}^h_t - \mathbf{q}^{h^*}_t||^2_2 \quad (4)$$

[1] IRB number: IRB_0040280

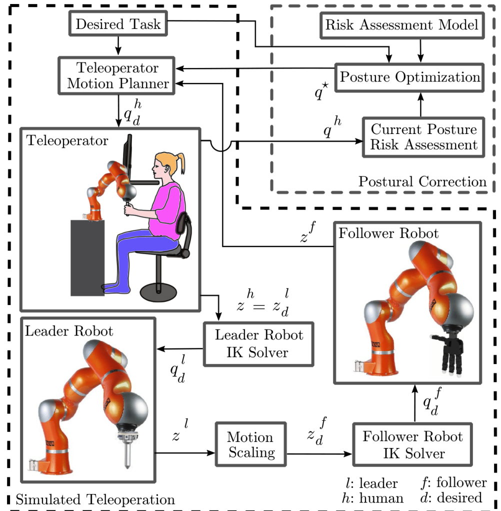


Fig. 5: Overview of teleoperation simulation.

where $\tau^f_{t\to H}$ is the trajectory of human posture from time t to the time of the horizon H, $\mathbf{x}^f_g$ is the goal pose of the follower robot, and ${}^f\mathbf{x}_t$ is the pose of the follower robot at time t. Here, $0 \leq \alpha \leq 1$ is a scalar number that models the postural correction acceptance and the effort of the human operator towards applying the postural correction. Details of the motion planning for the humans and the robots are presented in Fig. 5. We note that this model is likely a simplification of true human interactive behavior. We do not advocate for its use over human subject studies. Instead, we propose it as a useful tool for systematically exploring new human safety assessment and improvement algorithms.

C. Methods for Postural Optimization

We explore two different solvers for postural optimization, one gradient-based (SQP [41]) and one gradient-free (CEM [42]). We investigate the use of DULA/DEBA with both solvers by replacing the objective of maximizing human comfort in the problems from Sec. II with minimizing the ergonomic risk encoded by e.g. RULA or DEBA.

$$\mathbf{q}^* = \arg\min_{\mathbf{q}} \quad \mathrm{RULA/REBA/DULA/DEBA}(\mathbf{q}) \quad (5)$$

Our primary interest lies in examining to what extent the use of derivative computation in DULA/DEBA improves performance over the gradient-free optimization required in using RULA/REBA as an objective. However, as a secondary question we seek to examine if using the continuous DULA/DEBA models improves gradient-free optimization performance over the discrete RULA/REBA models.

We use the SLSQP solver from SciPy [43] for SQP. We calculate DULA and DEBA gradients using automatic differentiation in PyTorch. For CEM, we use a Gaussian sampling distribution with 10000 samples.

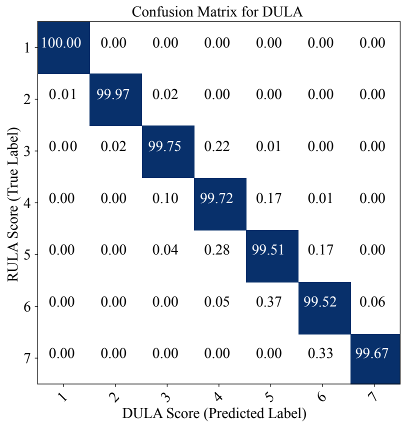


Fig. 6: Confusion matrix (accuracy %) for DULA vs RULA.

V. Results and Discussion

This section showcases the accuracy results for DULA and DEBA and the effect of their use in postural optimization.

A. DULA and DEBA Accuracy

Fig. 6 represents the confusion matrix for DULA. The trained model for DULA performs well on the test dataset and results in 99.73% accuracy. The lowest diagonal element in the confusion matrix shown in Fig. 6 for DULA is 99.38%, which shows the high accuracy of the learned model across all ranges. DEBA also has a high accuracy of 99.16% and the lowest diagonal element of the confusion matrix is 98.5%.

Training DULA took 61 hours, while DEBA took 64 hours. Note, however, that this training only has to take place once as the models can be used for all relevant ergonomic assessment tasks.

B. Postural Assessment

Fig. 7 compares the ergonomics assessment scores during representative trials from the pushing (top), and pHRI co-manipulation tasks (bottom). We compare RULA and DULA for the upper body; and REBA and DEBA scores for full body assessments. The results show that the continuous DULA score follows the RULA score very well in both tasks. The same is true with DEBA and REBA scores. The plots also show the smoothness of the continuous DULA/DEBA scores in comparison to the discrete RULA/REBA. These plots highlight that the majority of DULA/DEBA model errors happen between scores 3 and 4. Fig. 8 shows the deviation of the DULA/DEBA scores from RULA/REBA scores for all subjects and tasks. The plot reveals the high accuracy of our learned models. The above results show that DULA and DEBA can be used instead of RULA and REBA in autonomous ergonomics risk assessments.

C. Postural Optimization

There are three types of information that can be used to correct the human posture: (1) the general rule of thumb of moving toward neutral posture, (2) moving toward the optimal posture calculated from our postural optimization, and (3) use the gradient of DULA/DEBA. The first one is a naive idea that cannot be used in postural optimization in intelligent systems. The optimal posture from our approach is the best one for ergonomists and ergonomically intelligent systems, however it requires solving the postural optimization. In addition to the benefit of gradient of DULA/DEBA in gradient-based postural optimization, it can also provide insight on how to move the posture to increase the human comfort on the fly, since it is fast and does not require solving the postural optimization problem. Fig. 9 visualizes the gradient of DULA as arrow markers on the joints and

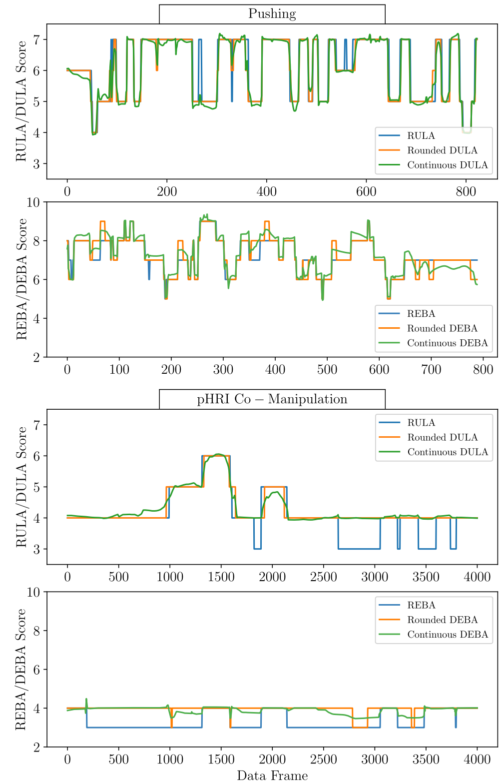


Fig. 7: Comparison of (RULA, continuous DULA, and rounded DULA) scores, and (REBA, continuous DEBA, and rounded DEBA) scores for a pushing task (top) and a pHRI co-manipulation with a teleoperated robot from [40] (bottom).

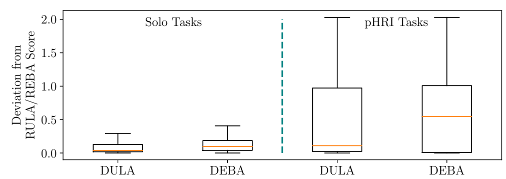


Fig. 8: Deviation of rounded DULA and DEBA from RULA and REBA scores among all the subjects and tasks. Lower is better.

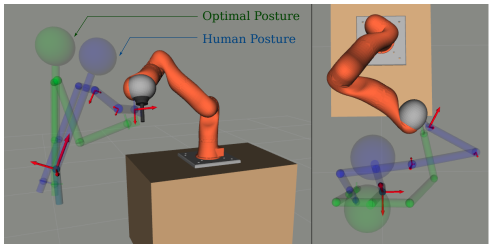


Fig. 9: Two views of the optimal posture, and the gradient of DULA at an example posture. We show the scaled gradient of DULA over the human joints using arrow markers. Larger markers correspond to relatively higher values of gradient; the positive angle of joint follows the right hand rule around the markers.

the calculated optimal posture at a the given posture. Larger arrows correspond to relatively higher negative gradient (larger arrow means more increase in human comfort by a change in the joint angle). It shows that at this posture, changes in the trunk and wrist joints have the highest effects on lowering the risk score.

Fig. 10 shows the RULA scores for the the postural correction results for gradient-free and gradient-based optimizations in 4 simulated teleoperation tasks with correction acceptance $\alpha \sim \{0, 0.75\}$. The plots show that the risk is reduced after the human applies the suggested optimal postural correction, however, task completion time increases, which is reasonable.

Fig. 11 provides more information on comparing the RULA score of resultant optimal postures and the corrected postures (the output of human-like motion generation for a simulated human) using 3 approaches: (a) gradient-free postural optimization with RULA, (b) gradient-free postural optimization with DULA, and (c) gradient-based optimization with DULA for 4 simulated tasks. The whiskers of the box plots are at min and max of the data series. From the left side of the figure, we can see that although all three methods have similar median risk scores for target optimal postures across all tasks, the gradient-free approach using RULA provides optimal postures with lowest risk scores. From the right side, initially we see that the corrected postures of the simulated human are more comfortable after applying the optimal posture from all three methods Moreover, the gradient-free RULA and gradient-based DULA approaches perform similarly and better than the other method. However, the gradient-based approach outperforms the others in terms of much faster computation. Each iteration of the gradient-free approach with RULA using 10000 samples takes ≈ 162s to solve, versus ≈ 0.19s for the gradient-based approach. This makes the gradient-based postural optimization using DULA an ideal approach for postural optimization, especially for online applications. Furthermore, the plot shows that using DULA instead of RULA in a gradient-free optimization degrades the overall performance of postural optimization.

The results from our simple simulated experiments reveal the effect of intelligent postural optimization on lowering the risk of WMSDs in pHRI, and opens an area for future

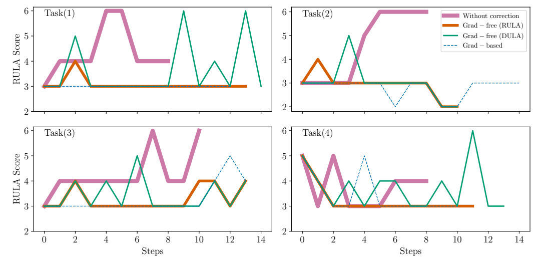


Fig. 10: Comparison between the effect of gradient-free and gradient-based postural optimization on 2 teleoperation tasks.

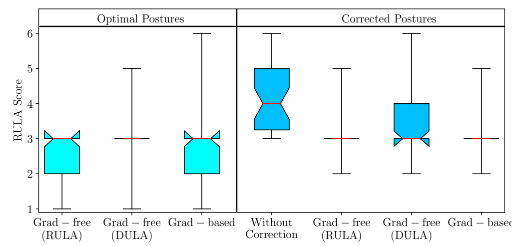


Fig. 11: Comparing the RULA score of the optimal postures from gradient-free and gradient-based postural optimizations and their effect on correcting postures in all tasks. Whiskers are at min and max of each data series, and lower score are better.

research on performing extensive human subject studies to evaluate postural optimization using DULA and DEBA.

VI. Conclusion

This paper proposes a framework for postural optimization in ergonomically intelligent physical human-robot interaction that includes and introduces two differentiable and continuous ergonomics risk assessment models to assess human upper body and full body posture. The DULA and DEBA ergonomics models are learned to replicate the non-differentiable RULA and REBA assessment tool using neural networks. DULA achieves 99.73% and DEBA 99.16% accuracy. The models are designed for computationally efficient postural optimization for pHRI and other related applications. We highlighted the benefits of the differentiable models in gradient-based postural optimization. The results reveal that postural optimizations using DULA and DEBA lower the ergonomics risk score across the tasks, with improved computational performance to using RULA and REBA.

There are several directions for future work. We wish to improve our training using hard negative mining techniques around the region of data with worst performance to further improve the accuracy of our models. The differentiable ergonomics models can be used in human-like motion generation in virtual reality, animation, and games. Our postural optimization work can be immediately extended by conducting a human subject study to evaluate our postural optimization and correction approach. We wish to compare different means of feedback for the posture correction to the human, including visual, auditory, and haptic feedback.

References

[1] L. Punnett and D. H. Wegman, "Work-related musculoskeletal disorders: the epidemiologic evidence and the debate," *Journal of electromyography and kinesiology*, vol. 14, no. 1, pp. 13–23, 2004.

[2] O. Erdinc and O. Vayvay, "Ergonomics interventions improve quality in manufacturing: a case study," *Intl. Journal of Industrial and Systems Engineering*, vol. 3, no. 6, pp. 727–745, 2008.

[3] W. Yu, T. Ignatius, Z. Li, X. Wang, T. Sun, H. Lin, S. Wan, H. Qiu, and S. Xie, "Work-related injuries and musculoskeletal disorders among factory workers in a major city of china," *Accident Analysis & Prevention*, vol. 48, pp. 457–463, 2012.

[4] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zühlke, "Human-machine-interaction in the industry 4.0 era," in *IEEE international conference on industrial informatics*, 2014, pp. 289–294.

[5] P. K. R. Maddikunta, Q.-V. Pham, B. Prabadevi, N. Deepa, K. Dev, T. R. Gadekallu, R. Ruby, and M. Liyanage, "Industry 5.0: A survey on enabling technologies and potential applications," *Journal of Industrial Information Integration*, p. 100257, 2021.

[6] F. Sherwani, M. M. Asad, and B. Ibrahim, "Collaborative robots and industrial revolution 4.0 (ir 4.0)," in *2020 Intl. Conf. on Emerging Trends in Smart Technologies*. IEEE, 2020, pp. 1–5.

[7] M. Ramaganesh, R. Jayasuriyan, T. Rajpradeesh, S. Bathrinath, and R. Manikandan, "Ergonomics hazard analysis techniques-a technical review," *Materials Today: Proceedings*, 2021.

[8] A. Bazazan, I. Dianat, N. Feizollahi, Z. Mombeini, A. M. Shirazi, and H. I. Castellucci, "Effect of a posture correction–based intervention on musculoskeletal symptoms and fatigue among control room operators," *Applied ergonomics*, vol. 76, pp. 12–19, 2019.

[9] A. A. Shikdar and M. A. Al-Hadhrami, "Smart workstation design: an ergonomics and methods engineering approach," *Intl. Journal of Industrial and Systems Engineering*, vol. 2, no. 4, pp. 363–374, 2007.

[10] H. Motabar and A. D. Nimbarte, "The effect of task rotation on activation and fatigue response of rotator cuff muscles during overhead work," *Applied Ergonomics*, vol. 97, p. 103461, 2021.

[11] A. Yazdani and R. Sabbagh Novin, "Posture estimation and optimization in ergonomically intelligent teleoperation systems," in *Companion of the ACM/IEEE Intl. Conf. on Human-Robot Interaction*, 2021.

[12] A. Yazdani, R. S. Novin, A. Merryweather, and T. Hermans, "Is the leader robot an adequate sensor for posture estimation and ergonomic assessment of a human teleoperator?" in *IEEE Intl. Conf. on Automation Science and Engineering*, 2021.

[13] C. R. Guerrero, J. C. F. Marinero, J. P. Turiel, and V. Muñoz, "Using "human state aware" robots to enhance physical human–robot interaction in a cooperative scenario," *Computer methods and programs in biomedicine*, vol. 112, no. 2, pp. 250–259, 2013.

[14] W. Kim, J. Lee, L. Peternel, N. Tsagarakis, and A. Ajoudani, "Anticipatory robot assistance for the prevention of human static joint overloading in human–robot collaboration," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 68–75, 2017.

[15] L. Chen, L. F. Figueredo, and M. R. Dogar, "Planning for muscular and peripersonal-space comfort during human-robot forceful collaboration," in *IEEE-RAS Intl. Conf. on Humanoid Robotics*, 2018.

[16] B. Busch, M. Toussaint, and M. Lopes, "Planning ergonomic sequences of actions in human-robot interaction," in *Intl. Conf. on Robotics and Automation*, 2018, pp. 1916–1923.

[17] R. Rahal, G. Matarese, M. Gabiccini, A. Artoni, D. Prattichizzo, P. R. Giordano, and C. Pacchierotti, "Caring about the human operator: haptic shared control for enhanced user comfort in robotic telemanipulation," *IEEE Tran. on Haptics*, vol. 13, no. 1, pp. 197–203, 2020.

[18] L. Peternel, N. Tsagarakis, D. Caldwell, and A. Ajoudani, "Robot adaptation to human physical fatigue in human-robot co-manipulation," *Autonomous Robots*, pp. 1–11, 2018.

[19] S. Chen and R. R. Yang, "Pose trainer: correcting exercise posture using pose estimation," *arXiv preprint arXiv:2006.11718*, 2020.

[20] T. N. Hoang, M. Reinoso, F. Vetere, and E. Tanin, "Onebody: remote posture guidance system using first person view in virtual environment," in *Proceedings of the 9th Nordic Conf. on Human-Computer Interaction*, 2016, pp. 1–10.

[21] L. Peternel, W. Kim, J. Babič, and A. Ajoudani, "Towards ergonomic control of human-robot co-manipulation and handover," in *IEEE-RAS Intl. Conf. on Humanoid Robotics*, 2017, pp. 55–60.

[22] L. Peternel, C. Fang, M. Laghi, A. Bicchi, N. Tsagarakis, and A. Ajoudani, "Human arm posture optimisation in bilateral teleoperation through interface reconfiguration," in *IEEE RAS/EMBS Intl. Conf. for Biomedical Robotics and Biomechatronics*, 2020.

[23] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, "Postural optimization for an ergonomic human-robot interaction," in *Intl. Conf. on Intelligent Robots and Systems*, 2017, pp. 2778–2785.

[24] A. G. Marin, M. S. Shourijeh, P. E. Galibarov, M. Damsgaard, L. Fritzsch, and F. Stulp, "Optimizing contextual ergonomics models in human-robot interaction," in *Intl. Conf. on Intelligent Robots and Systems*, 2018, pp. 1–9.

[25] L. Van der Spaa, M. Gienger, T. Bates, and J. Kober, "Predicting and optimizing ergonomics in physical human-robot cooperation tasks," in *Intl. Conf. on Robotics and Automation*. IEEE, 2020, pp. 1799–1805.

[26] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*, vol. 52, 1988, pp. 139–183.

[27] L. McAtamney and E. N. Corlett, "Rula: a survey method for the investigation of work-related upper limb disorders," *Applied ergonomics*, vol. 24, no. 2, pp. 91–99, 1993.

[28] S. Hignett and L. McAtamney, "Rapid entire body assessment (reba)," *Applied ergonomics*, vol. 31, no. 2, pp. 201–205, 2000.

[29] J. Steven Moore and A. Garg, "The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders," *American Industrial Hygiene Association Journal*, vol. 56, no. 5, 1995.

[30] J. M. Kapellusch, A. Garg, K. T. Hegmann, M. S. Thiese, and E. J. Malloy, "The strain index and acgih tlv for hal: risk of trigger digit in the wistah prospective cohort," *Human factors*, vol. 56, no. 1, 2014.

[31] Z. Khodabakhshi, S. A. Saadatmand, M. Anbarian, and R. Heydari Moghadam, "An ergonomic assessment of musculoskeletal disorders risk among the computer users by rula technique and effects of an eight-week corrective exercises program on reduction of musculoskeletal pain," *Iranian Journal of Ergonomics*, vol. 2, no. 3, 2014.

[32] L. F. Figueredo, R. C. Aguiar, L. Chen, S. Chakrabarty, M. R. Dogar, and A. G. Cohn, "Human comfortability: Integrating ergonomics and muscular-informed metrics for manipulability analysis during human-robot collaboration," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 351–358, 2020.

[33] A. Shafti, A. Ataka, B. U. Lazpita, A. Shiva, H. A. Wurdemann, and K. Althoefer, "Real-time robot-assisted ergonomics," in *Intl. Conf. on Robotics and Automation*, 2019, pp. 1975–1981.

[34] J. Lim, I. Lee, I. Shim, H. Jung, H. M. Joe, H. Bae, O. Sim, J. Oh, T. Jung, S. Shin, *et al.*, "Robot system of drc-hubo+ and control strategy of team kaist in darpa robotics challenge finals," *Journal of Field Robotics*, vol. 34, no. 4, pp. 802–829, 2017.

[35] A. Rout, B. Deepak, and B. Biswal, "Advances in weld seam tracking techniques for robotic welding: A review," *robotics and computer-integrated manufacturing*, vol. 56, pp. 12–37, 2019.

[36] P. J. D. de Oliveira Evald, R. T. S. da Rosa, J. L. Mór, R. Z. Azzolin, V. M. de Oliveira, and S. S. da Costa Botelho, "Velocity regulation of a linear welding robot by unscented and cubature kalman filter output estimation-based sliding mode control," in *Annual Conf. of the IEEE Industrial Electronics Society*, 2017.

[37] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Neural Info. Proc. Sys.*, 2019.

[38] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[39] M. Tietz, T. J. Fan, D. Nouri, B. Bossan, and skorch Developers, *skorch: A scikit-learn compatible neural network library that wraps PyTorch*, 2017.

[40] L. Vianello, J.-B. Mouret, E. Dalin, A. Aubry, and S. Ivaldi, "Human posture prediction during physical human-robot interaction," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 6046–6053, 2021.

[41] J. Nocedal and S. Wright, *Numerical optimization*. Springer, 2006.

[42] M. Kobilarov, "Cross-entropy motion planning," *Intl. Journal of Robotics Research*, vol. 31, no. 7, pp. 855–871, 2012.

[43] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, and et al, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

CHAPTER 6

DISCUSSION

This dissertation discusses some of the main challenges in improving ergonomics and human comfort in ergonomically intelligent teleoperation and physical human-robot interaction. We provided multiple contributions to solve each problem in each chapter. However, there are some limitations in this work, which are the potential for new research opportunities in the future. We will discuss the main limitations in the next section and follow up with future directions in Section 6.2.

6.1 Limitations

In this dissertation, we focused on formulating and implementing posture estimation and optimization on teleoperation. However, we can use our approach for other physical human-robot applications with minor modifications.

A general limitation in 3D posture estimation is that the ground truth for the human posture is not achievable using standard approaches, and motion capture systems provide only gold-standard estimations of the human posture. Therefore, the deviation between the postures estimated from the proposed approaches in this dissertation and the motion capture postures do not necessarily represent the actual error and accuracy of our approaches. Moreover, when we re-targeted the motion capture with higher DOF into the 10-DOF kinematic model of humans, we did not account for the length change in the virtual link from the hip to the neck. The fix-length assumption for the virtual link causes some errors when we overlay the motion capture skeleton on the video frames.

Moreover, our posture estimation approach using the robot's trajectory assumes that the human hand stays attached to the robot's stylus and that the grasping point and pose are the same during the task. However, this assumption may not be valid during some tasks, and the human operator may slightly slide their attachment point on the robot's stylus. In addition, we assume that the relative position of the center of the chair to the robot is known and fixed. This assumption may also not hold during some tasks, which causes a change in the chair position or the sitting position of the human on the chair.

Although we use the posture validation from [1] to validate the posture of particles, their approach only validates the shoulder and elbow joints and does not include the wrist angles. This limitation may cause non-bio-mimetic estimated postures for the wrist and torso joints.

One issue with using a particle filter for 3D posture estimation is that it is computationally expensive, and the run time for each filtering step is proportional to the number of particles. However, a higher number of particles results in a more accurate estimation. Based on the computation resources available for this research and for the sake of run time, we only used 500 particles in our implementation. Higher computational resources would help improve the accuracy of the posture estimation algorithm in the future.

A significant limitation of using OpenPose in this research is that our OpenPose implementation did not provide predicted keypoints on the hand. This issue limits the sensory information available to predict the joint angles at the human wrist. A better implementation that could include hand keypoint will help improve the accuracy of the estimated 3D posture.

Our human skeletal model does not include the left arm, legs, and neck, and we assume that the torso and the right arm are the moving segments during the tasks. We can resolve this limitation by improving the skeletal model to represent the entire human body and perform whole-body posture estimation, assessment, and optimization.

In the postural correction for the simulated human operator, we modeled the human motion in a receding horizon formulation with re-planning. However, we could use more sophisticated human-like motion generation methods here. In addition, we modeled the acceptance of the suggested postural correction in a simple random-based way. The acceptance model can be further investigated based on the literature or human subject studies.

We evaluated our postural optimization approach only in simulation. Further human subject studies will help evaluate the approach's effectiveness in natural work environments.

6.2 Future directions

One of the main future directions for the work in this dissertation is to apply the postural estimation, assessment, and optimization approaches to different pHRI applications and VR systems. We can benefit from the differentiable ergonomics models in human-like motion generation in virtual reality, animation, and games.

Significantly, we can extend the human skeletal model into the entire body skeleton and perform whole-body posture estimation, assessment, and optimization.

Another exciting possibility for future research in the posture estimation part of this dissertation is to use more sophisticated particle filters or inference approaches such as [2] to improve the estimation accuracy and reduce the processing time.

In addition, we want to investigate the hypothesis that including the DULA and DEBA models for human comfort in our postural estimation as a weighting function for particles can improve the redundancy resolution and accuracy of the estimated postures.

Moreover, the probabilistic distribution of estimated postures and their corresponding ergonomic risk scores can be used in human-aware planning for shared autonomy and other pHRI applications.

We wish to use better designs of neural networks and improve our model training using hard negative mining techniques around the region of data with the worst performance to improve further the accuracy of our differentiable DULA and DEBA ergonomics models in order to improve the accuracy of the model among different risk scores.

Finally, we can immediately extend our postural optimization work by conducting a human subject study to evaluate our postural optimization and correction approach. We wish to compare different means of feedback for the posture correction to the human, including visual, auditory, and haptic feedback.

6.3 References

[1] Y. Jiang and C. K. Liu, "Data-driven approach to simulating realistic human joint constraints," in *2018 IEEE Int. Conf. on Robot. and Automat. (ICRA)*. IEEE, 2018, pp. 1098–1103.

[2] F. Dellaert, "Factor graphs and gtsam: A hands-on introduction," Georgia Institute of Technology, Tech. Rep., 2012.

CHAPTER 7

CONCLUSION

This dissertation presents an artificial intelligence solution for improving the human comfort and ergonomics of the workplace in teleoperation and physical human-robot interaction. We present a complete framework for the ergonomically intelligent system in teleoperation and physical human-robot interaction that includes novel posture estimation, assessment, and optimization approaches.

We propose using the interacting robot as the sensor for 3D segment lengths and posture estimation. Later, we combine it with another markerless posture estimation method to develop an occlusion-robust estimation algorithm.

We also use deep neural networks to learn differentiable and continuous ergonomic models based on popular and well-studied risk assessment tools in the ergonomics community. We use them to assess and optimize the posture in teleoperation and physical human-robot interaction.

Finally, we introduce a framework for postural optimization tailored for different types of teleoperation and physical human-robot interaction tasks. We use gradient-free and gradient-based approaches to solve the optimizations, which benefit from our differentiable ergonomic models.