Toward Task-Level Multimodal Evaluation of Wearable Robots in Construction Worker Activities

Xinyan Huang and Jingang Yi

***Abstract*— Construction tasks involve complex human motions that challenge the evaluation of wearable robotic systems. Existing approaches typically focus on isolated tasks and single-modal sensing, which do not fully reflect the diverse and dynamic task demands of construction-related activities. In this work, we propose a task-level multimodal evaluation interface for wearable exoskeletons that integrates diverse task primitives (elementary motions or actions), including locomotion, static support, dynamic balance, and precision manipulation, with multimodal sensing such as kinematics, muscle activity, and contact forces. Preliminary results indicate that the proposed framework enables systematic assessment of wearable robotic systems under realistic conditions.**

I. Introduction

Wearable robotic systems are increasingly used in real-world environments such as construction [1]–[3], where human activities are complex, non-periodic, and involve frequent transitions between different task primitives. These characteristics pose significant challenges for evaluating the effectiveness of wearable assistance in realistic scenarios.

However, most existing evaluation approaches are limited to simplified and isolated task primitives [4]–[7], which fail to capture the variability and interaction complexity encountered in practice. In addition, current methods often rely on single-modal measurements or task-specific protocols, making it difficult to systematically assess performance across diverse human-robot interaction scenarios. These limitations highlight a fundamental gap: the lack of a structured evaluation paradigm that can represent heterogeneous tasks while integrating complementary sensing modalities. Rather than treating evaluation as a collection of independent experiments, there is a need to view it as a unified interface that connects task design, sensing, and performance assessment.

In this work, we address this challenge by introducing a task-level multimodal evaluation interface, which functions as a unified framework combining experimental protocol, sensing, and performance evaluation into a structured pipeline, for wearable robotic systems. The proposed interface organizes complex activities into task primitives and integrates multimodal sensing to support systematic and task-aware evaluation across diverse scenarios.

The main contributions of this work are as follows:

- We introduce a task-level evaluation paradigm that represents complex, non-periodic human activities through a structured hierarchy of task primitives.
- We develop a multimodal evaluation interface that integrates complementary sensing modalities for analyzing human-robot interaction across diverse task conditions.
- We provide a scalable and reproducible framework for task-oriented evaluation of wearable robotic systems in realistic environments.

X. Huang and J. Yi are with the Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, NJ 08854 USA (email: xh301@scarletmail.rutgers.edu, jgyi@rutgers.edu).

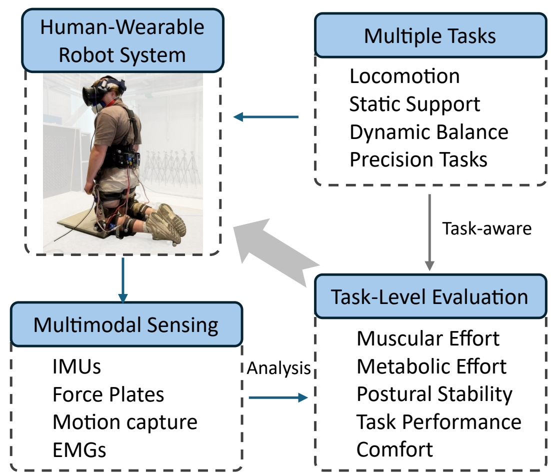


Fig. 1. A task-level multimodal evaluation interface, consisting of a task protocol, human–robot interaction loop, and multimodal sensing and evaluation pipeline.

II. Task-Level Multimodal Evaluation Interface

To enable structured evaluation of wearable robotic systems in complex environments, we propose a task-level multimodal evaluation interface, defined as a structured evaluation framework that integrates (i) a task protocol for generating diverse human activities, (ii) a multimodal sensing pipeline for capturing human–robot interaction, and (iii) a set of task-aware metrics for performance assessment. As illustrated in Fig. 1, this interface connects task design, human–robot interaction, and measurement into a unified evaluation pipeline.

At the task level, human activities are represented as compositions of task primitives derived from a 30-minute experimental protocol, as illustrated in Fig. 2. The protocol is designed to capture common interaction patterns in construction-related activities, including locomotion, static support, dynamic balance, and precision tasks, under varying workload conditions. Specifically, the trial is organized into three stages with increasing load levels. In the no-load

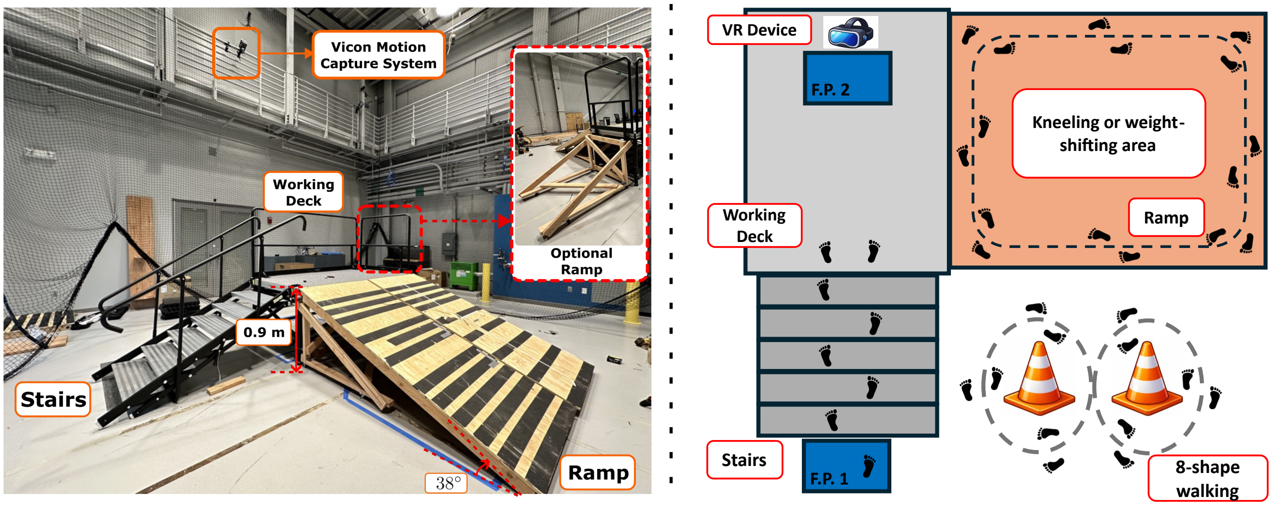


Fig. 2. The experimental platform and task protocol for evaluating the knee exoskeleton in construction environments. (Left) A photograph of the custom-built platform, highlighting the main components: a six-step staircase, a ramp (38° incline) leading to an elevated working deck (0.9 m height), and the Vicon motion capture system that is used for kinematic tracking. (Right) A corresponding schematic diagram illustrating the layout and the multiple tasks performed by participants. These tasks include straight-line stair ascent/descent, walking along the perimeter of the ramp, kneeling or weight-shifting on the ramp, virtual welding with VR device, and navigating a figure-8 course around cone obstacles. The positions of the two force plates (F.P. 1 and F.P. 2), embedded within the platform to measure GRFs, are explicitly marked.

TABLE I
SUMMARY OF PRIMARY WEARABLE DEVICES, SENSORS, AND SAMPLING RATES.

Modality	*Sensor/Device*	*Sampling Rate*
Motion capture	Vicon Vantage	100 Hz
Kinetics	AMTI ACG-O	1000 Hz
Muscle activity	Noraxon DTS EMG	1500 Hz
Torso/limb movement	LP-RESEARCH IMU	100 Hz
Metabolic cost	VO2 Master	-
VR environment	HTC Vive Pro	-

stage, participants perform locomotion and navigation task primitives, including stair climbing and descending, 8-shape walking, ramp walking, and precision task primitives such as VR-assisted welding. In the medium-load stage, the protocol introduces interaction-intensive activities, such as prolonged kneeling for simulated flooring tasks and dynamic weight-shifting on the ramp, alongside locomotion tasks under 7-kg load. In the heavy-load stage, participants perform similar task compositions, including ramp walking, kneeling on inclined surfaces, obstacle navigation, and stair climbing while carrying 9-kg loads.

Within this context, wearable robotic assistance is incorporated as part of the human–robot system, where the device interacts with the user during task execution. This interaction serves as the central element linking task conditions to measurable outcomes, enabling the analysis of how assistance influences human motion and effort across different activities.

To capture different aspects of human-robot interaction, we employ a multimodal sensing setup that combines motion capture, force measurement, physiological monitoring, and wearable sensing, as summarized in Table I. These measurement modalities are widely used in human movement and ergonomics studies and together provide complementary information for evaluating human-exoskeleton interaction. Based on these measurements, task-level evaluation is performed through a set of metrics, including muscle and metabolic effort, postural stability, precision task performance, and feedback from subjects. Muscle and metabolic effort are

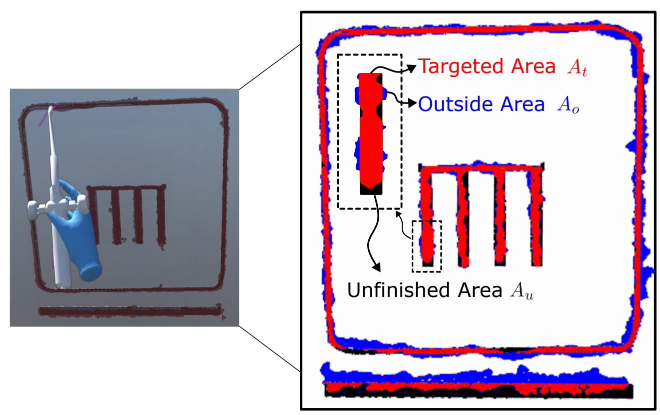


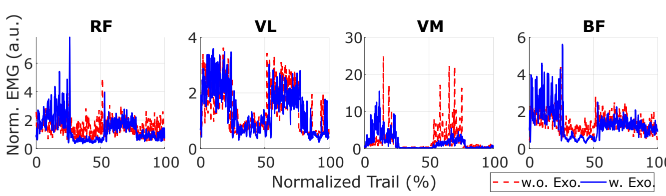
Fig. 3. An illustrative schematic for the image processing method for welding performance assessment on a virtual welding piece.

measured by electromyography (EMG) and metabolic analyzer (VO2 Master) to evaluate muscle activation and rate of oxygen consumption, respectively. Postural stability is derived from trunk IMU and force plate data: the 95% confidence ellipse areas for trunk acceleration (A^{Acc}) and center of pressure (A^{Sway}). These ellipses are computed from the covariance of bivariate signals and represent the area enclosing approximately 95% of the data samples. A^{Acc} reflects dynamic movement variability and control, and A^{Sway} quantifies postural sway and indicates balance stability. Task performance is evaluated in a representative precision task using a VR-based welding scenario (see Fig. 3). Similar to our previous work [8], performance metrics are obtained through image-based analysis of the virtual workpiece, where segmented regions corresponding to correctly welded, excess, and unfinished areas are used to compute accuracy, precision, and completion rate. Subjective feedback is collected via a post-experiment questionnaire to assess usability, comfort, fatigue, and overall experience, providing a complementary perspective to the objective measurements.

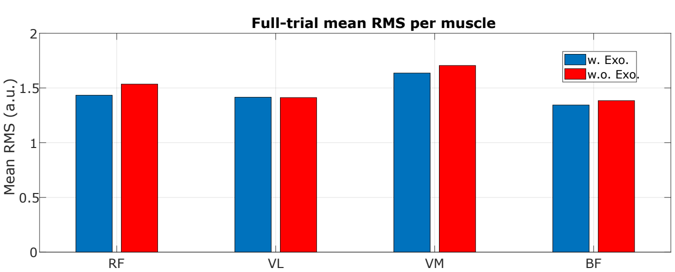
By comparing these metrics with and without exoskeleton assistance across task primitives, the proposed framework enables structured and task-aware analysis of wearable robot performance. Rather than relying on a single unified metric, the proposed interface supports structured analysis across task primitives by associating multimodal measurements with their corresponding task contexts. This enables task-aware interpretation of performance and facilitates systematic comparison of wearable robotic assistance under diverse scenarios.

III. Preliminary Insights

The proposed evaluation framework is demonstrated through representative multimodal measurements across diverse task scenarios, as shown in Fig. 4–7. These results illustrate how the interface enables task-aware analysis by linking sensing data to task contexts and performance indicators.



(a)



(b)

Fig. 4. (a) Comparison of the normalized EMG time-series data for RF, VL, VM, and BF muscles from a representative subject under mid-load conditions. The dashed red and solid blue lines denote the "w.o. Exo." and "w. Exo." trials, respectively. Each activity segment corresponds to stair climbing, kneeling, ramp walking, and dynamic weight shifting. Across all activities, muscle activation amplitudes were reduced with exoskeleton assistance, particularly in the quadriceps group (RF and VM), indicating decreased knee extensor effort during both dynamic locomotion and static tasks. (b) Root-mean-square (RMS) EMG amplitudes throughout the selected trials.

Fig. 4 presents EMG signals across multiple task segments (i.e., temporally segmented executions of task primitives), including stair climbing, kneeling, ramp walking, and dynamic weight shifting. Muscle activation was recorded from key lower-limb muscle groups, including knee extensors [rectus femoris (RF), vastus lateralis (VL), vastus medialis (VM)], knee flexors [biceps femoris long head (BF), semitendinosus (ST)], shank extensor [lateral gastrocnemius (LG)], and hip extensors/abductors [gluteus maximus (MG), gluteus medius (GM)]. Both the time-series data and aggregated RMS values show consistent reductions in muscle activation under wearable assistance, particularly in the quadriceps group. These results demonstrate the capability of the framework to quantify muscle effort and reveal task-dependent variations in physical demand.

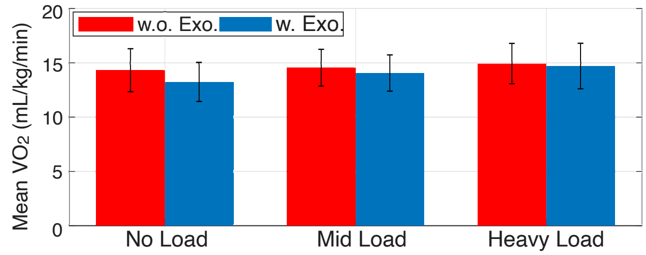


Fig. 5. Mean oxygen consumption rate ($\dot{V}O_2$) across load conditions with and without exoskeleton assistance. Bars represent group means and error bars indicate the one standard deviation.

Fig. 5 shows the mean oxygen consumption rate (VO_2) across different load conditions. Mean $\dot{V}O_2$ was slightly lower with assistance by approximately -1.1, -0.5, and -0.2 mL·kg^{-1}·min^{-1} under the "No Load", "Mid Load", and "Heavy Load" conditions, respectively. Paired comparisons indicate that these differences were not statistically significant ($p > 0.05$ in all cases). This finding implies that although a mild reduction trend was observed (most significant at lighter loads), the overall metabolic cost of performing mixed construction tasks was not materially changed by exoskeleton use. This suggests that the assistance reduces local muscular effort without substantially affecting whole-body energy expenditure across short-duration task blocks.

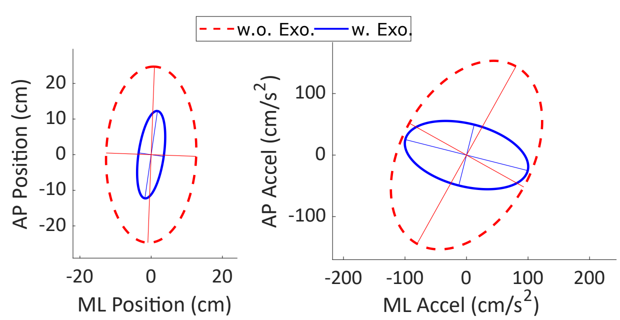


Fig. 6. Representative 95% COP sway ellipses (A^{Sway}) and 95% IMU acceleration ellipses (A^{Acc}) from a randomly selected subject under mid-load conditions. (Left) During static kneeling in the VR environment. (Right) During dynamic weight shifting on ramp. Blue solid lines represent the "w. Exo." condition, and red dashed lines denote the "w.o. Exo." condition.

Fig. 6 illustrates postural stability metrics derived from

COP sway and IMU acceleration measures. The coordinate system was established with the anterior-posterior (A-P) and medial-lateral (M-L) directions corresponding to the x- and y-axes, respectively. Both the COP sway area (A^{Sway}) and IMU acceleration area (A^{Acc}) were smaller with the exoskeleton assistance. The reduction in sway area indicates improved balance control with wearable assistance during static and dynamic tasks.

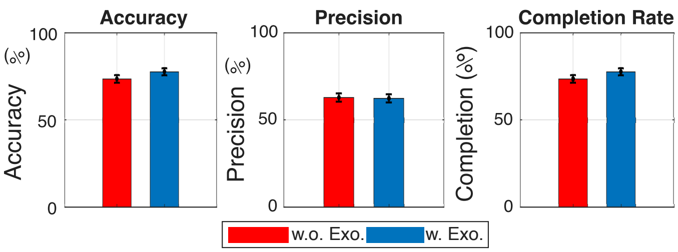


Fig. 7. Performance metrics for the virtual welding task under the mid-load condition. Error bars represent ± 1 standard error.

Fig. 7 presents task performance metrics in a representative precision task (VR-based welding), including accuracy, precision, and completion rate. The results indicate that the use of the exoskeleton leads to improved accuracy and completion rate. These findings suggest that wearable assistance can enhance task execution quality, highlighting its potential to improve performance in precision-demanding tasks.

Subjective feedback further complements these objective measurements. Survey results from 21 participants indicate that the system is generally easy to use, with moderate ratings in ease of use and learnability. Perceived usefulness is more task-dependent, with ratings centered near neutral and some participants reporting reduced efficiency in low-load conditions, while noting benefits during physically demanding tasks such as kneeling.

Based on these measurements, the evaluation interface incorporates a set of task-level performance indicators derived from multimodal sensing, including: (i) muscle effort (EMG), (ii) metabolic effort (VO_2), (iii) postural stability (COP and IMU acceleration metrics), (iv) task performance (accuracy, precision, completion), and (v) subjective feedback capturing user comfort and experience.

Rather than relying on a single unified metric, the proposed interface enables structured, task-aware analysis by associating multimodal measurements with their corresponding task contexts. This supports systematic comparison of wearable robotic assistance across heterogeneous activities.

While the current study establishes a controlled and reproducible evaluation framework under laboratory conditions, several directions can further extend its applicability. Future work will focus on validating the proposed framework in real-world construction environments, where irregular terrain, environmental variability, and prolonged task duration may influence performance. In addition, expanding the participant pool to include experienced construction and industrial workers would provide further insights into usability and task-specific performance in practical settings. Finally, incorporating adaptive stiffness control and intent-aware mode switching represents a promising direction to enable smoother transitions between stabilization and free motion across diverse task scenarios.

IV. Conclusion

This work presents a task-level multimodal evaluation interface for wearable robotic systems, addressing the limitations of existing approaches that rely on isolated tasks and single-modal measurements. By integrating task primitives, human-robot interaction, and multimodal sensing within a unified framework, the proposed interface enables structured and task-aware analysis of performance in complex, non-periodic environments.

The presented results demonstrate how the interface links multimodal measurements to task contexts, allowing consistent interpretation of muscle effort, postural stability, and task execution across diverse activities. This provides a more comprehensive perspective on wearable robotic assistance compared to conventional evaluation protocols.

Importantly, the proposed approach does not rely on a single unified metric, but instead supports task-specific and multimodal analysis within a consistent evaluation structure. This makes the framework flexible and adaptable to different application domains and task scenarios.

Future work will focus on extending this interface toward more standardized evaluation protocols and real-time integration with adaptive and task-aware wearable robotic control strategies.

References

[1] D.-H. Moon, D. Kim, and Y.-D. Hong, "Development of a single leg knee exoskeleton and sensing knee center of rotation change for intention detection," *Sensors*, vol. 19, no. 18, p. 3960, 2019.

[2] L. Chen, H. Xie, Z. Liu, B. Li, H. Cheng, *et al.*, "Exploring challenges and opportunities of wearable robots: A comprehensive review of design, human-robot interaction and control strategy," *APSIPA Trans. Signal Inform. Processing*, vol. 12, no. 1, 2023.

[3] S. Wang, C. Feng, X. Chen, M. Shan, and W. Niu, "A biomechanical evaluation of firefighters' musculoskeletal loads when carrying self-contained breathing apparatus in walking and running," *J. Safety Res.*, vol. 87, pp. 1–14, 2023.

[4] S. Chen, D. T. Stevenson, S. Yu, M. Mioskowska, J. Yi, H. Su, and M. Trkov, "Wearable knee assistive devices for kneeling tasks in construction," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 4, pp. 1989–1996, 2021.

[5] S. Chen, Y. Yu, C. Di, M. Trkov, J. Gong, and J. Yi, "Postural balance of kneeling gaits on inclined and elevated surface for construction workers," in *Proc. IEEE Conf. Automat. Sci. Eng.*, Lyon, France, 2021, pp. 753–758.

[6] G. Sreenivasan, C. Zhu, and J. Yi, "Neural balance control of human quiet stance in construction tasks," in *Proc. IEEE Conf. Automat. Sci. Eng.*, Auckland, New Zealand, 2023.

[7] J. Camargo, A. Ramanathan, W. Flanagan, and A. Young, "A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions," *J. Biomech.*, vol. 119, p. 110320, 2021.

[8] G. Sreenivasan, C. Zhu, and J. Yi, "Exoskeleton-assisted stance and kneeling balance and work task evaluation in construction," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 22 356–22 370, 2025.