User Intent-Driven and Context-Aware Personalization for Assistive Exoskeletons

Vaidehi Prakash Wagh

CMU-RI-TR-26-56

June 4, 2026

The Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA

Thesis Committee:
Dr. Inseung Kang, *chair*
Dr. Hartmut Geyer
Michelle Zhao

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Robotics.

To my 9-year-old self who wanted to be a scientist.

Abstract

Personalizing exoskeleton control to individual users remains a central challenge to real world deployment across healthy and clinical populations. Existing approaches optimize personalization parameters for biomechanical objectives such as metabolic cost or joint loading, requiring resource and time-intensive, user-specific data collection, without accounting for user comfort, intent, or task context. Prior work shows that users can perceive and report their preferred assistance, yet no lightweight method maps this expressed intent to quantitative control parameters in real time.

We present a vision-language model (VLM) guided human-exoskeleton interface that translates natural user feedback and egocentric visual context into parameter updates for a hip exoskeleton controller, and treats user feedback as reward to train a contextual bandit on user preferences. The framework separates high-level intent parsing performed by the VLM, from low-level action selection and exploration performed by a user-specific contextual bandit that refines its policy with each interaction, requiring no lab setup or biomechanical instrumentation. This system runs on a lower-limb hip exoskeleton, in real time, tested with users in a pilot study.

We evaluated the framework on a five-task locomotion track across 2 laps, spanning level ground, ramps, and stairs, against a single-shot interpreter and a few-shot recommendation system as personalization baselines. Our contextual bandit framework reduced no improvement personalization instances by 31% and raised terminal user satisfaction by the second lap (5.15 to 5.67), with mean within-episode gains consistent with maximum learning across the first lap (+3.71 gain) and sustained learning by the second (+2.75). Our pilot results validate our framework, supporting the use of VLMs as context-aware interpreters for real-time, user-intent driven assistance personalization for lower limb exoskeletons, and contextual bandits as lightweight online learners for modeling user preferences. This framework highlights the potential of incorporating user preferences through natural-language interfaces for personalizing exoskeleton assistance to ensure user comfort and support long-term adoption of these devices.

Acknowledgments

To Dr. Kang, thank you for guiding me through research in the past 2 years. I've grown a lot from the version of myself who joined the lab in Fall 2024, and I attribute it to my time in the lab, your guidance and mentorship, your trust in my ability to figure it out, and the lab environment you've fostered that is so conducive to collective growth. I have grown to deeply respect the strong perspectives you share on many things, and am grateful for the advice you've shared throughout these two years.

To my committee, Dr. Geyer and Michelle, thank you for your time, encouragement and honest feedback that shaped this work into a thesis. Your insights made me reason and challenge the perspectives I held about my work, which led me to a stronger narrative and a thesis I felt proud of. I'm truly grateful for our discussions.

A big thank you to Changseob for helping me take this project to the finish line. To Maria, who was the most supportive friend I could ask for, and Rajiv, the most entertaining and annoying person I've met, as well as Jimin, Nate, Ilseung, Crystal, Amy, Oscar, Adrian, Jinwoo - MeMo was truly the best lab to be a part of because of you all.

To Sarah, my first research advisor, thank you for taking me on during MITACS four years ago. Two years of working with you, across continents and timezones, only reinforced my love for research, and your unwavering belief and guidance brought me into the school and program of my dreams, which is why I am here today.

To my parents, who believed in the little girl who wanted to be a scientist and made sure that I got every bit of encouragement along the way. Thank you Aai and Papa, for trusting my instincts and decisions, with a firm foundation of your support that made me feel invincible in front of the toughest challenges. I hold immeasurable gratitude in my heart for both of you.

To my sister, Aditi, who was always armed with love, Doordash and retail therapy when I called her up crying at midnight - you always eased the pressure I put on myself. Thank you for grounding me when I flew too close to the sun. To my sister, Vaishnavi, thank you for always being there when I needed a listening ear and ending every call with a laugh. To my family in the states, Anjan and Sakshi, thank you for making your home a place of love and joy, when home truly felt far away.

To the friends I made in CMU, thank you for keeping me sane with every Catan night, Mario's hangout, Treat ice cream run, fall break trips, RI social, coworking sessions, chats in the RI kitchen over lunch, shared block meals, pool and video-game sessions in the RI lounge, or the best chai & bagel dates. You are the reason Pittsburgh feels like home.

To my closest friends from home, thank you for truly being a call or text away. Your support means the world, and it's a joy figuring out life with you.

Contents

When this dissertation is viewed as a PDF, the page header is a link to this Table of Contents.

List of Figures

Chapter 1

Introduction

Lower limb exoskeletons can augment human mobility across healthy and clinical populations [1, 2]. A fundamental tradeoff in building assistive systems lies between the generalization and personalization of robot behavior or control to users. While generalizable policies allow for scaled deployment, user-personalization allows for maximum user benefits, comfort and long-term user adoption [3, 4]. Modern approaches to exoskeleton control, driven by machine learning, have ensured user-generalizable controllers that estimate gait phase, task, joint kinematics and dynamics from user-specific sensor data [5, 6, 7, 8]. While this enables data-driven generalization across populations, personalization is still driven by traditional, in-lab optimization approaches [9]. Conventional approaches like Human-in-the-loop optimization (HILO) [10] use physiological or biomechanical cost functions to personalize control to maximize user benefit. Measuring these objective functions, like metabolic cost reduction, joint load minimization and kinematic symmetry, requires gait-lab instrumentation and long experimentation sessions that are both time and energy intensive for users. Further, traditional approaches rarely account for user perception or preference while tuning parameters. Simpler personalization methods often rely on trial-and-error and manual tuning, leading to static personalization parameters per user that do not adapt to changing task context, user intent or preferences in the real-world. Recent approaches [11, 12, 13, 14] towards personalization have taken the user-preference based approach, using pairwise comparisons between different control parameters to build preference profiles, but they still require in-lab testing, repetitive experimentation and offline analysis. Thus, there is a need for more lightweight approaches to user-preference based personalization in exoskeleton research.

Comfort, confidence, and perceived control are crucial factors in assistive robotics and human-robot interaction studies [15, 16]. The development of an interactive framework between assistive robots and users is a significant avenue in HRI research to build shared autonomy and redirect control back to the user, leading to increased comfort with using the system [17, 18, 19]. This motivates the need for an interactive framework between exoskeletons and their users.

This thesis proposes a user intent-driven and context-aware personalization framework for a lower limb hip exoskeleton, through a multimodal interaction interface guided by VLMs and an online, preference learning algorithm using contextual bandits. Our system combines natural language user feedback with egocentric visual perception, mapping it to quantitative changes in the personalization parameters for the exoskeleton controller. It serves as a real-time, user-driven interface between user intent and exoskeleton controller behavior, enabling a form of shared control in which the user participates meaningfully in the adaptation of their own assistance.

On a technical level, the system integrates Meta Aria glasses for capturing multimodal intent, using its RGB camera for egocentric vision and seven-channel audio microphone for audio feedback capture. An OpenAI Whisper-based speech transcription model works to convert user feedback to text, and a VLM inference pipeline interprets egocentric vision and transcribed user feedback into controller updates. The VLM enables the parsing of direct and potentially ambiguous user feedback, over constrained regex based parsing pipelines, and enables vision-grounded reasoning for controller personalization. We further enable robust task detection through the addition of IMU based pelvis pitch as input to the VLM as trunk tilt markers, to overcome challenges in visual task recognition. From these multimodal inputs, the VLM produces parameter updates directions targeting the two primary axes of the exoskeleton's control spline: assistance magnitude and gait-phase delay, with a context vector encoding multimodal information. A contextual bandit takes in this context vector and chooses between discrete action arms to update personalization parameters for the controller, updated real time.

The novel contribution of our work is as follows.

1. We propose a real-time, lightweight, interactive, exoskeleton personalization

framework that runs without any gait-lab instrumentation.

2. Our learning framework uses contextual bandits to learn user preferences from natural language interactions, without requiring pre-collected, user-specific preference data.
3. Our framework adapts continually to user needs, translating unstructured user feedback to quantitative changes in the exoskeleton controller, and uses a simple rating scale as reward for our preference learner.
4. We propose this system as a proof of concept of foundation-model driven interaction architectures for enabling user-driven personalization in lower-limb exoskeletons.

The structure of this thesis is as follows. Chapter 2 reviews related work in exoskeleton control, human-in-the-loop optimization, and the emerging use of language and vision models in robotics and HRI. Chapter 3 describes the system architecture and hardware integration for our proposed system. Chapter 4 describes the experimental evaluation, while Chapters 5 and 6 talk about pilot results and discussion. Chapter 7 addresses the limitations, future work and final conclusions.

Chapter 2

Background

2.1 Exoskeletons

Robotic exoskeletons have advanced rapidly over the past decade, going from tethered, laboratory prototypes to autonomous wearable devices for everyday use [1, 2]. Consequently, exoskeleton controllers evolved from task- and user-specific tuning toward generalizable, subject-independent control. Modern deep learning approaches have enabled data-driven controllers that generalize across users and locomotion tasks [6] without requiring user-specific data collection. These controllers often estimate gait phase [20], biological joint torque [6], or similar parameters to modulate assistance during the gait cycle. Previous studies have shown metabolic cost benefits across different tasks using ML-driven controllers [21], further strengthening their use in state-of-the-art exoskeletons.

2.2 Assistance Personalization

While ML-driven generalization enables broad deployment of exoskeletons across diverse populations, personalization is a key challenge to ensure user adoption of these devices [9]. Per-user tuning of control parameters for generalizable controllers ensures maximum user comfort and benefit. [3]. Conventional personalization approaches primarily focus on finding optimal control parameters using human-in-the-loop optimization (HILO) approaches, while optimizing for biomechanical or physiological cost functions to maximise user benefit [10, 22, 23]. HILO measures a user's metabolic cost under different control parameters using gait-lab instrumentation, iterating until it converges on user-optimal settings. This process is both time and energy intensive, which precludes online, out-of-lab adaptation and limits its scalability for real-world deployment.

2.3 User Preference guided Personalization

A more recent avenue addresses a key driver of user adoption: user preference. Rather than optimizing a measured physiological cost, some recent studies in exoskeletons and prostheses have used user preference as the optimization signal [12, 24], on the premise that users implicitly gauge comfort, effort, and control, and can express this judgment as feedback that serves as a scalable, non-instrumented reward signal for personalization.

Ingraham et al. [12] used a touchscreen self-tuning interface to show that users can reliably and repeatedly select their preferred ankle exoskeleton torque magnitude and timing based solely on perceived assistance. Convergence to preferred parameters was also fast, within 105 seconds of the trial. Lee et al. [24] introduced a sample-efficient active learning framework that used pretrained user preference models and forced-choice user feedback to rapidly identify preferred values for four ankle exoskeleton control parameters: rise time, peak torque timing, fall time, and peak torque magnitude. Tucker et al. [25] proposed COSPAR, a Bayesian preference-based learning algorithm that combined pairwise user comparisons and coactive feedback to optimize exoskeleton gait parameters such as step length, duration and width, achieving convergence in both simulation and real-world testing. In the domain of prostheses, Clites et al. [26] investigated user-preferred prosthetic ankle stiffness across walking speeds, finding relationships between user preference and kinematic symmetry. Thatte et al. [14] used pairwise preference feedback within a sample-efficient Bayesian optimization framework to tune control parameters of a transfemoral prosthesis, avoiding the need to define an explicit objective function while learning from human feedback.

Together, these works establish two points. First, user preference is reliable and repeatable enough to guide personalization. Second, optimization algorithms can use this perceived feedback as a reward signal to learn user-preferred policies, without an explicit cost function like HILO. However, these works also have some limitations. First, they obtain preference through structured interfaces such as touchscreens or forced-choice comparisons, which yield coarse, low-dimensional preferences that cannot convey richer preference information. Second, these interfaces, while more portable than gait-lab instrumentation, require active physical interaction from the user. Third, they do not account for how preferences vary for the same task across different contexts, such as changing terrain, fatigue, or environment.

These limitations motivate an online, preference learning interface for exoskeleton assistance personalization that is expressive, hands-free, and context-aware.

2.4 Foundation Models

The release of ChatGPT in November 2022 marked a turning point in human-machine interaction. While large language models had been developing for some time [27, 28], GPT made open-ended conversational interfaces practical for the first time[29], having implications well beyond web and voice assistants.

This enabled an era of natural language controlled robots, prompting research on language-conditioned robot control [30, 31]. The subsequent move to multimodal foundation models, which combined vision with language, proved especially powerful for robotics, as natural language could now be grounded in vision. Powerful vision-language backbones enabled the development of Vision-Language-Action models that enabled end-to-end robot control, which became a significant avenue in robotics research. [32, 33]

While these models are powerful and generalizable, they are also data-intensive. Vision-language-action (VLA) models are trained on large-scale, multi-task datasets and can be adapted to downstream robotic tasks by fine-tuning on relatively smaller datasets. However, these approaches rely on structured, high-level action representations and large-scale interaction data that are not currently available in exoskeleton settings. The lack of large, standardized datasets makes direct application of VLA-style policies, for either high-frequency control or low-frequency parameter selection, currently infeasible in exoskeletons.

Nonetheless, the contextual reasoning capacity of vision-language model backbones offers a way to address the research gap in existing preference-based personalization. State-of-the-art VLMs are well-suited to acting as contextual interpreters that parse unstructured, natural-language feedback from users while processing visual information to condition action or personalization parameter selection.

2.4.1 In Assistive Robotics

Recent studies have used large language models to build conversational interfaces enhancing human-robot interaction between users and assistive robots[34, 35]. VoicePilot [18] deployed an LLM (GPT-3.5 Turbo) as a speech interface for a feeding robot, mapping high level user instructions to low-level execution code for feeding robot trajectories. Nwankwo et al. [36] used LLMs and VLMs in a human-robot interaction task, using an LLM to decode natural language commands from users into precise robot actions, and a VLM to extract useful environment information. In preference learning studies, TidyBot [37] used LLM summarizing capabilities to map natural user instructions into known primitives to learn user preferences for a cleaning robot. For users with paralysis, Shankar et al. [38] built a low-burden LLM based preference learning framework that interprets user needs from natural language interaction for a meal preparation task.

2.4.2 In Exoskeletons

Foundation models have begun to enter exoskeleton control, though predominantly at the level of intent interpretation rather than assistance control. Chen et al. [39] used two LLMs and speech recognition to let users command a hand exoskeleton beyond a preprogrammed gesture set, with an incremental learning framework that expands the command set over time. Li et al. [40] extended this to multimodal input, fusing speech and image data through a multimodal LLM to infer hand motion intent and generate motion plans, including for undefined gestures. For upper-limb assistance, Chen et al. [41] integrated an LLM into task planning so that the exoskeleton configures its assistive parameters from the semantics of the manipulated object, with anomaly detection triggering replanning. In occupational settings, [42] provided a voice interface for a back-support exoskeleton, letting wearers adjust controller settings by speech, and reported higher satisfaction and lower cognitive load than a visual interface. Specifically in the field of lower limb exoskeletons, Guo et al. propose a speech based human-exo interface for lower-limb motion planning that achieves a 54% faster task completion time compared to a web interface [43]. Multimodal models also see some application in exoskeletons, primarily for the locomotion-mode recognition relevant to lower-limb devices. Seong et al. [44] applied a VLM to a wearable robot in an industrial setting, estimating the weight of a grasped object from a single camera and adjusting assistance accordingly, without task-specific training. Ahmadi et al. [45] used GPT-4o to predict exoskeleton locomotion modes in construction from field-of-view frames and speech captured by smart glasses, and found that zero-shot GPT-4o (87.9% F1) closely rivaled a fine-tuned CLIP model (90.1%), demonstrating that multimodal foundation models can interpret vision and speech for exoskeleton control without task-specific training.

2.5 Contextual Bandits

Contextual Bandits come from the domain of online decision making. The multi-armed bandit problem formalizes sequential decision-making under uncertainty: an agent repeatedly selects one of K actions, receives a scalar reward, and must balance exploiting high-reward actions discovered so far against exploring undersampled ones to avoid suboptimal convergence. Auer [46] introduced the UCB1 algorithm that minimises regret across time by selecting the action that maximizes an upper confidence bound, pairing the estimated mean reward with an exploration bonus inversely proportional to how often the action has been sampled. This upper-confidence-bound (UCB) principle provides a principled, exploration strategy for complex, multi arm or multi action problems.

Contextual bandits extend the bandit framework by conditioning action selection on an observed context vector. At each round, a feature vector describing the current state is observed, the algorithm selects an action, and it receives a reward for that action-context pair only. Li et al. [47] introduced LinUCB, which models expected reward as a linear function of the context and constructs an upper confidence bound via ridge regression, admitting a closed-form update that is efficient for online deployment. These methods are well-suited to settings where a rich, high-dimensional context is available to condition personalization, such as the scene and activity information a VLM can extract in exoskeleton control.

2.5.1 Bandits in Assistive Robotics

Bandit formulations have begun to see use in assistive robotics, most prominently in robot-assisted feeding, where bite acquisition across diverse, previously unseen food items has been framed as a linear contextual bandit. Gordon et al. [48] showed that LinUCB and ϵ-greedy algorithms could converge quickly to successful manipulation strategies for unseen foods from visual context, and later augmented this with post-hoc haptic context observed after action selection to speed learning and reduce regret [49]. Banerjee et al. [50] extended the feeding setting to a human-in-the-loop contextual bandit, LINUCB-QG, that selectively queries the user for help while explicitly modeling querying workload, balancing task performance against user burden across participants with and without mobility limitations. Contextual bandits have also been used to personalize a robot's interaction with a human over time. Kaushik et al. [51] developed a robot exercise coach that learns, via a contextual bandit, which feedback style best improves a user's performance, and notably found that the style a user performs best with is not always the one they subjectively prefer or report preferring.

2.5.2 Bandits in Exoskeletons

In prostheses and exoskeletons, bandit methods have centered on preference-driven parameter tuning. Thatte et al. [52] framed transfemoral prosthesis tuning as a dueling bandits problem, eliciting pairwise user preferences between control parameter sets from an offline-generated library and selecting among them with Double Thompson Sampling; across five subjects, four different parameter sets were preferred, underscoring the need for per-user personalization. Across these settings, the bandit conditions on either visual or haptic features of an object to manipulate, or selects among a fixed, context-free library of control parameters. To our knowledge, no prior work conditions a contextual bandit on the user's locomotion task and environment for wearable assistance. Our work addresses this gap, using a VLM to supply scene and intent context to a contextual bandit that personalizes exoskeleton assistance per user and per task.

Chapter 3

Proposed Framework

3.1 Overview

We propose a user-intent driven and context-aware human-exoskeleton interface, that uses vision, natural language audio instructions and IMU data to map user instructions to quantitative changes in exoskeleton control. Our action space is defined by two control parameters used in assistance personalization, the magnitude of assistance torque (scale) and the offset timing (delay). We use smart glasses to capture audio-visual input, IMU readings from our hip exoskeleton hardware, and a microphone and speaker setup for the interactive feedback loop to build this system. An off-the-shelf VLM serves as our intent-parser, capable of understanding and interpreting not just direct instructions, but also ambiguous user commands. The VLM outputs an informational context JSON structure that acts as the query to a user-specific contextual bandit, that chooses appropriate action arms for both scale and delay. The contextual bandit is a disjoint LinUCB algorithm, which uses ridge regression to learn a mapping vector between input context and reward. Human feedback reward on a Likert scale rating acts as our reward for training this bandit. The system allows for user-initiated personalization instances, where data is captured for each modality and fed into the VLM-bandit pipeline, updating scale and delay values for the base biotorque controller. Post-update, we capture human feedback ratings on a scale from 1 (strongly dislike) to 7 (strongly like) this update. Satisfaction criteria is counted as Likert ≥ 5, while a rating under 5 triggers a refinement feedback

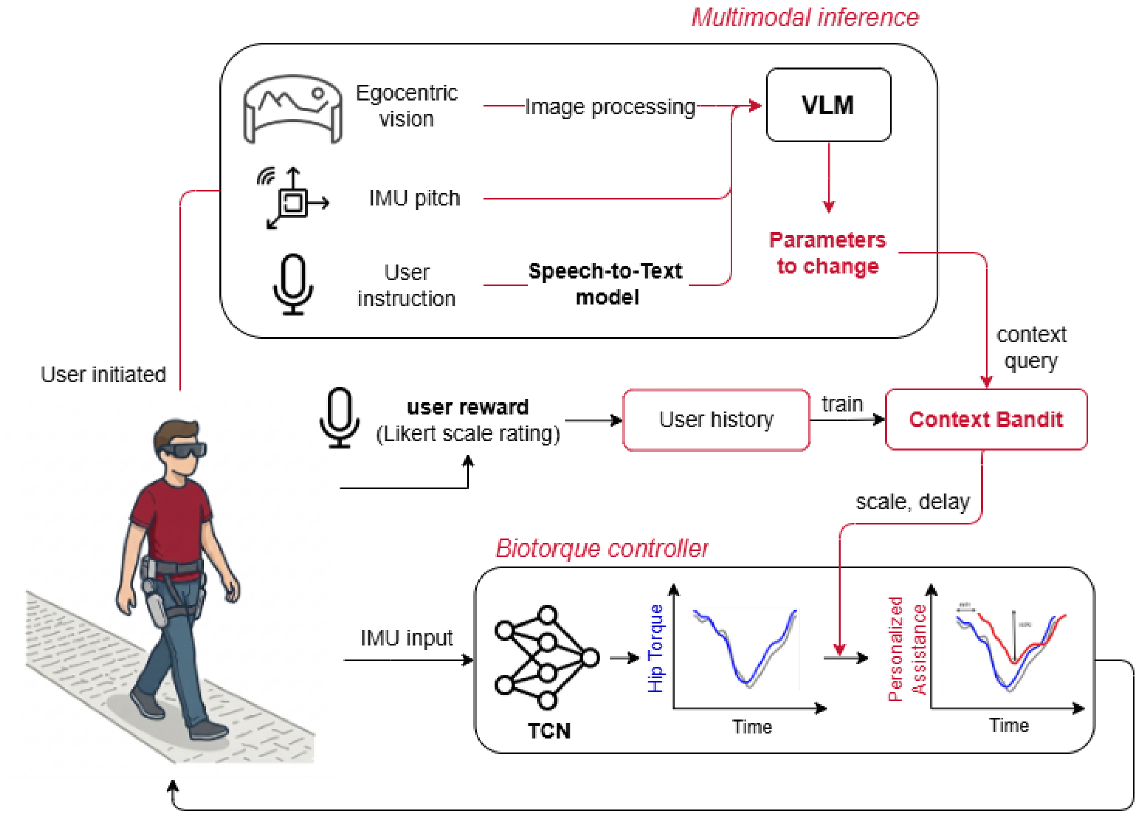

Figure 3.1: Overview of the Proposed Framework

loop that allows for updated audio feedback from the user to adjust the previous update.

This chapter builds our proposed framework with justifications for each modality and component. Rather than presenting the full multimodal pipeline as a given, each iteration of the design is motivated by limitations of previous approaches.

3.2 System Requirements

We build this framework for a lower-limb hip exoskeleton, running a biotorque controller that maps IMU data into hip joint moment estimates for a given user. Assistance magnitude, scaled as a percentage of user's natural hip joint moment, as well as the timing offset between the start of a user's gait cycle and onset of assistance are two conventional personalization parameters for the biotorque controller. As the biotorque controller is trained on healthy gait, we assume this framework to be used with healthy populations. Based on these assumptions, we establish the following system requirements.

The system must operate in real time. Locomotion tasks are dynamic, and tasks, terrains and environments change quickly. Minimal latency is necessary between user input and controller update.

The system must be portable and usable outside laboratory setups. Hence, all modalities and rewards need to be obtained through non-instrumented, wearable consumer devices.

The system must impose minimal cognitive load on the user. Interaction is thus taken through natural spoken feedback rather than any structured interface (touchscreen, etc), and the user is not trained on valid commands to the exoskeleton. The user is only expected to have a general understanding of assistance magnitude and offset.

The system must initiate without an trained expert in the loop and without prior offline optimization. It should employ strategies to cold-start adaptation and learn with each user interaction, eventually developing a user preference profile with time.

Safety guardrails must be set to ensure smooth transitions between updates and torque cutoff limits to ensure safe behaviour. The system must employ strategies to deal with VLM hallucinations, non-parseable outputs and unsafe suggestions.

Finally, the system must generalize across locomotion contexts, including terrain it was not explicitly designed around, and must behave sensibly when feedback is ambiguous or unrelated to the task.

3.3 Hardware

3.3.1 Hip Exoskeleton

We use a hip exoskeleton developed in the MetaMobility Lab, at Carnegie Mellon University. It is capable of providing bilateral hip flexion/extension assistance by means of two T-Motor AK80-9 quasi-direct-drive actuators aligned with both hip joints. Each motor has an integrated planetary gearhead and encoder, delivering approximately 18.0 Nm peak torque and 9.0 Nm maximum continuous torque, with

Figure 3.2: Hip Exoskeleton worn by a user

a 100° flexion and 30° extension range-of-motion. The system weighs 4.2 kg, with onboard computation and power housed in a backpack unit containing an NVIDIA Jetson Orin Nano and a 22.2 V 4000 mAh 6S LiPo battery. Joint-level state is captured via Invensense ICM-20948 IMUs, and the device runs an ML-based controller that uses IMU data from the pelvis and thighs to estimate the user's biological joint torque and modulate assistance torque using specific personalization parameters (scale and delay).

3.3.2 Wearable sensors

To capture vision and audio modalities, we use Meta's Project Aria glasses (Gen 1), a research device built for egocentric multimodal data collection from the wearer's perspective. The glasses carry five cameras, two monochrome scene cameras, one RGB camera, and two eye-tracking cameras, alongside a non-visual sensor suite of two IMUs, a magnetometer, a barometer, GPS, Wi-Fi and Bluetooth beacons, and microphones. For our purposes, we use the RGB camera for obtaining egocentric frames that capture environmental context, while a seven-channel microphone captures user's audio instructions. The smart glasses weigh about 75 grams, and support easy integration with our exoskeleton system, without requiring integrated camera connections, in line with our requirement for non-instrumented wearable sensing. Eye gaze is available as an input, not used in the current system but holds promise for understanding unspoken user intent.

3.3.3 Orchestration laptop

Due to the nature of the research-grade smart glasses, data collection with the glasses currently requires them to be connected to a laptop for actual processing. Thus, we use a laptop to serve as a communication hub between the Jetson on the exoskeleton running the low-level controller, the VLM and Audio transcription models, and the Aria glasses themselves.

3.3.4 GPU Server

Model inference runs on a dedicated GPU server, an Alienware Area-51 desktop with an NVIDIA RTX 5090 and 32 GB of GPU memory. This capacity lets us host both the audio transcription model (OpenAI's Whisper-medium-en) and the vision language model (Qwen 2.5 VL 7B) locally while maintaining suitable latency for real-time personalization. While the Jetson can run smaller speech-to-text models on its own, hosting transcription on the server lets us use a larger model for better accuracy. Likewise, only a few heavily quantized VLMs can run on the Jetson or on the orchestration laptop's CPU, so the server is what makes it feasible to use a more capable VLM for our task.

3.4 Biotorque Controller

The exoskeleton runs a biotorque controller that estimates the user's biological hip moment in real time. The controller follows the temporal convolutional network (TCN) approach to subject-independent hip moment estimation introduced by Molinaro et al. [6], in which a window of kinematic input is mapped directly to an estimate of the underlying joint moment rather than relying on hand-engineered gait variables such as gait phase and ambulation mode. Our controller uses a similar architecture, with 5 temporal blocks and 2 convolutional layers, and is trained on 11 subjects data collected on the hip exoskeleton, across multi-speed walking tasks.

3.5 Personalization Action Space

The biotorque controller generates assistive torque from the control spline generated by the model across the gait cycle and exposes two adjustable parameters to the personalization layer: scale and delay. Scale defines the assistance magnitude, which controls the amplitude of the assistive torque and ranges from 0 to 30 percent of the user's biological torque, capped at 10 N-m. The second is delay, which controls the timing of assistance relative to the gait cycle and ranges from 0 to 0.3 seconds. Both parameters are split into discrete arms across their ranges, allowing the bandit to choose from a range of options per parameter based on context.

3.6 Input Modalities

3.6.1 Natural Language

Spoken feedback acts as the primary signal in our framework for user intent. The user is asked to describe naturally about how the assistance feels, and we transcribe this recording using OpenAI's Whisper model (medium size, english only). The transcribed text is used as input to the VLM's system prompt.

Spoken user intent is also our strongest cue to map onto the action space directly. A request for more help, or for assistance that is stronger, concerns assistance magnitude and pushes it upward. Feedback that the assistance arrives at the wrong moment, that it feels too late or too early or just "off", concerns delay. Using language as the intent signal keeps the interaction in the user's own terms and supports the shared control that the system is built around, in which the user steers the personalization and the system carries out the adjustment.

3.6.2 Limitations of Language-only approach

While language is sufficient to identify the parameter and the direction, it has a few limitations. Similar words or phrases mean different things across users, and can map to different magnitudes in different situations. A request for more help means one thing during level walking and another during ramp ascent, while a request to make the assistance feel more balanced does not map cleanly onto either parameter at all. Also, expecting rich, task and environment specific context increases the cognitive workload for the user, as they have to describe not only intent, but other factors like task, terrain, and obstacles. A language-only system is therefore context-sparse, and cannot be used to learn context-specific user preferences. It can only map direct instructions to parameter changes, something that can be done using a touchscreen or buttons. Thus, we require a modality that can provide unsaid context that takes away the user-burden of scene description, while allowing the user to focus on expressing intent.

3.6.3 Vision

Vision, as an additional modality, proves to be a denser source of information. From a single egocentric frame, the system can interpret the task the user is performing, the terrain and surface they are on, whether the setting is indoor, outdoor, or transitional, and the presence of obstacles ahead. The smart glasses capture the scene from the user's own perspective, as compared to a camera located on the waist or shoulder, helping provide context that clearly informs user intent. Paired with audio feedback, vision lets us obtain both user intent and rich scene context cleanly, with minimal cognitive load on the user.

3.6.4 Limitations of Vision-Language approach

Nonetheless, either fisheye or rectified image inputs from the smart glasses can struggle with task recognition for visually confusing tasks such as level ground vs gentle ramps. While optical flow based approaches with continual vision-streaming can help with this, they add considerable overhead with the research-grade smart glasses. Thus, there is a need to introduce a simple, third modality that can increase task recognition

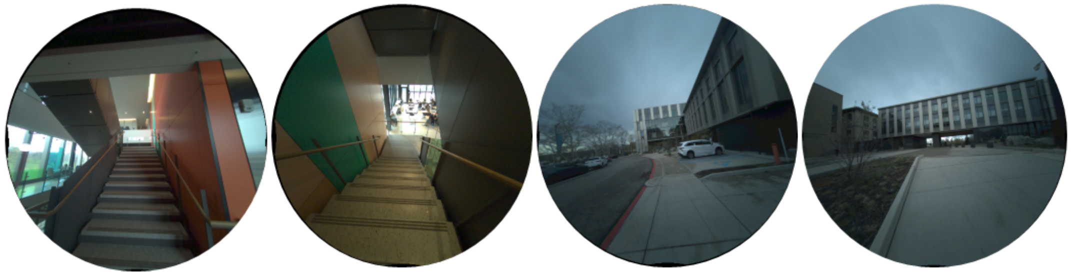

Figure 3.3: Visually confusing frames for task recognition. From left, stair ascent, stair descent, ramp descent, ramp ascent

robustness for the VLM for visually confusing tasks.

3.6.5 IMU

To ground visual task recognition in user orientation, we add a single IMU feature as our third modality. The pelvis IMU onboard the exoskeleton is used to compute pelvis pitch, the arctangent of z-axis over negative x-axis acceleration. This captures the user's trunk tilt, which varies across different inclinations as the user leans to maintain balance. Pelvis pitch therefore serves as a minimal, direct estimate of the gradient of the locomotion task. This signal makes task recognition robust where vision alone fails. The VLM identifies stairs easily, but cannot distinguish ascent from descent due to similar geometric compositions of first-person perspective images. With no visual anchors, it also fails at ramp identification. We therefore prompt the VLM to trust the IMU pitch and corroborate it with vision, using the sign of the gradient to resolve ascent against descent and to separate ramps from level ground.

Because resting trunk posture varies across users, pelvis pitch is compared to a per-user baseline, set by a short calibration step of ten seconds of level-ground walking and reused across sessions.

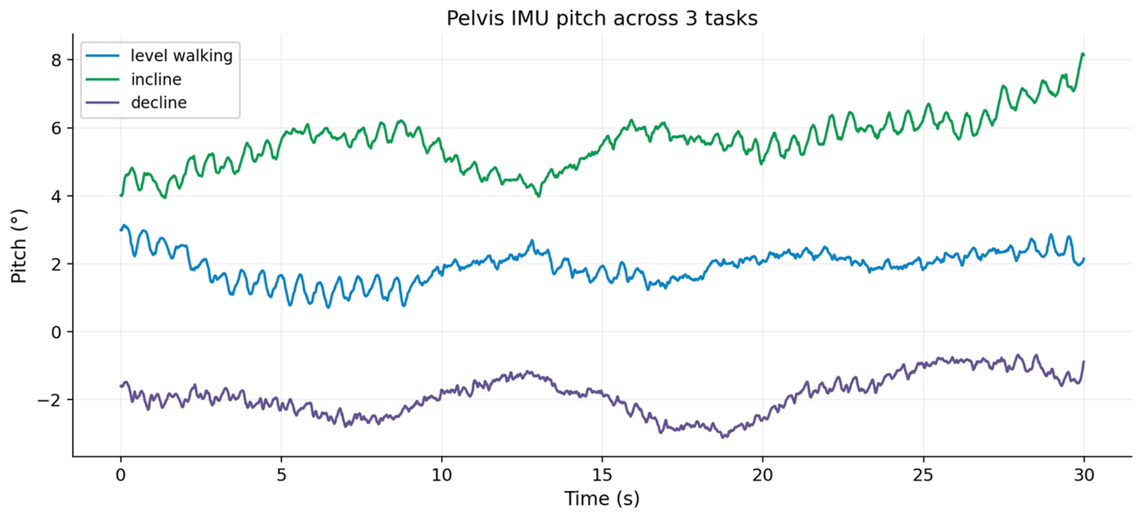

Figure 3.4: Pelvis Pitch for a user measured across level ground, ramp ascent (incline) and ramp descent

3.6.6 Multimodal inputs

We thus justify each modality as an input to our framework, to accurately and robustly capture both user intent and scene context. Our combined modalities address limitations of using individual ones, with language contributing intent, vision contributing environmental context, and IMU based pitch contributing orientation information to accurately guide VLM intent parsing. The remainder of the chapter concerns the model that performs that interpretation and the loop that grounds its output in the user.

3.6.7 VLM

We use Qwen 2.5 VL 7B as our VLM, served locally through Ollama as a persistent service on the GPU server. The multimodal inputs from the user are wrapped in a system prompt that informs the VLM about the baseline controller, its personalization parameters, and the current parameter values, gives general guidance on what to extract from each modality, and specifies a structured JSON output indicating which parameter to vary and how. The structured output expects the following fields. Parameters to change is one of scale, delay, both, or none. Direction of change is one of increase, decrease, or none. Locomotion task is one of level ground walking, ramp ascent, ramp descent, stair ascent, stair descent, or unknown. Terrain grade is reported as none, gentle, moderate, or steep. Exertion demand is reported as minimal, low, moderate, high, or very high. Environmental context is reported as indoor, outdoor, or transitional. Together these fields form a rich context vector that is passed to the contextual bandit. A natural alternative would be to have the VLM map this context vector directly to values for scale and delay, rather than passing it to a separate learning layer. We deliberately separate interpretation from magnitude selection, for the following reasons: While VLMs can act as interpreters for contextual instruction, it is complex to ground them in user preferences. Building a user-preference aware, continual-learning prompt for a VLM requires considerable prompt iterations, complex design and safety guardrails, long context and agent-like wrappers. Within the latency and compute constraints of this framework, we believe that a VLM-guided approach to personalization and user-preference learning is beyond the scope of this thesis. Nonetheless, a simpler approach would be to offload learning to a separate layer, letting VLM predict magnitudes of personalization parameters. While this serves as our baseline comparison, we suggest that the structured and principled approach to user preference learning enabled by a context bandit would be more difficult and intensive to deploy using a VLM in real time.

3.6.8 Context Bandit

The VLM determines which parameter to change and in which direction, but the magnitude is learned separately by a contextual bandit. We maintain two disjoint LinUCB bandits per user, one for scale and one for delay, each learning a preferred absolute value per task context. Each bandit operates over a fixed, discrete set of arms spanning the parameter's safe range. The context is a one-hot encoding of the structured scene, a bias term together with the locomotion task, terrain grade, and exertion demand reported by the VLM. For each arm, ridge regression estimates the expected reward as a linear function of this context, and the bandit selects the arm maximizing an upper confidence bound, which trades off the arm with the highest estimated reward against arms whose value under the current context is still uncertain. Because the bandits are disjoint, each arm maintains its own estimate, and the context conditioning keeps what is learned in one setting, such as an incline, from overwriting what is learned in another. The reward is the user's reported satisfaction, a Likert rating mapped to the interval $[-1, 1]$.

A key design choice is that the VLM's intent enters only as a constraint at selection time, not as a training signal. Thus, a request to increase a parameter restricts the bandit to arms at or above the current value, leaving the choice of magnitude to the learned reward estimates. This keeps the two roles cleanly separated, the VLM interprets intent and scene, and the bandit alone grounds magnitude in the user's feedback.

Each user's bandits are retrained from their full interaction history after every interaction, so the learned preferences persist across sessions and accumulate as a per-user preference model rather than resetting each time.

3.6.9 Human Reward

The signal that drives the bandit's learning is the user's own reported satisfaction. After an adjustment is applied, the user rates how the assistance feels on a Likert scale from 1 (strongly dislike) to 7 (strongly like). When satisfaction is below threshold (<5), the loop records audio instructions from the user again for further refinement, reusing the same context image to readjust personalization parameters, continuing until the user is satisfied, and the bandit updates from the outcome.

This places a human judgment at the center of the loop, allowing for the bandit to learn online from instant human feedback, rather than post-hoc offline. The reward is what corrects an overshoot or undershoot against the user's felt experience, and it is the mechanism by which the system's notion of the right magnitude is defined by the user rather than asserted by the model.

3.7 Proposed System

The final proposed system is described in Figure 3.5. The model's recommendation closes a loop in which the user rates the result, the loop iterates while satisfaction is below threshold, and the contextual bandit learns from each outcome so that recommendations improve across interactions. Where the first condition interprets once and the second corrects once, the proposed system continues until the user is satisfied and carries what it learns forward. This is the condition the system is built around and the one the evaluation examines most closely. A simpler, more intuitive interaction flow diagram is shown below in Fig 3.6. A system implementation diagram that shows the distribution of components across hardware is also shown in 3.7.

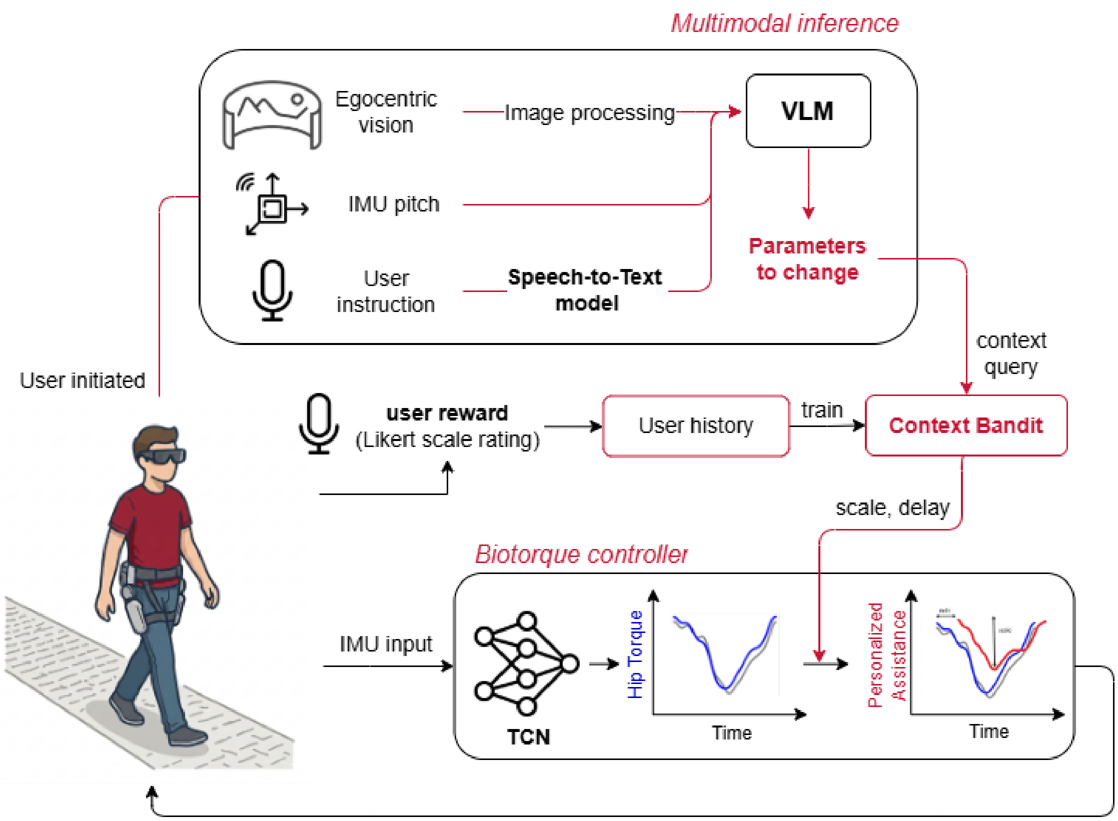

Figure 3.5: Overview of the Proposed Framework

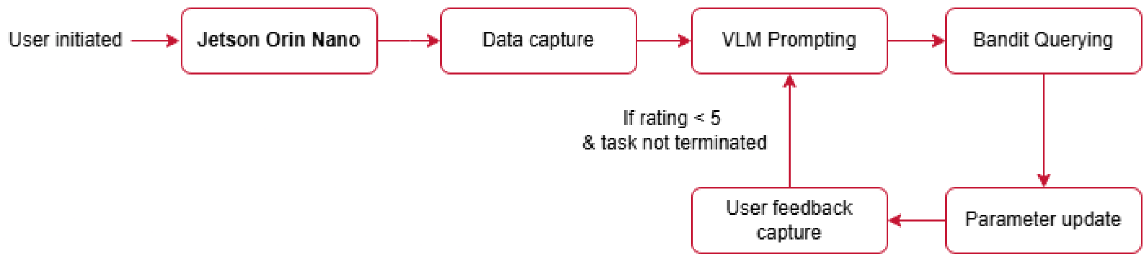

Figure 3.6: Interaction Flow

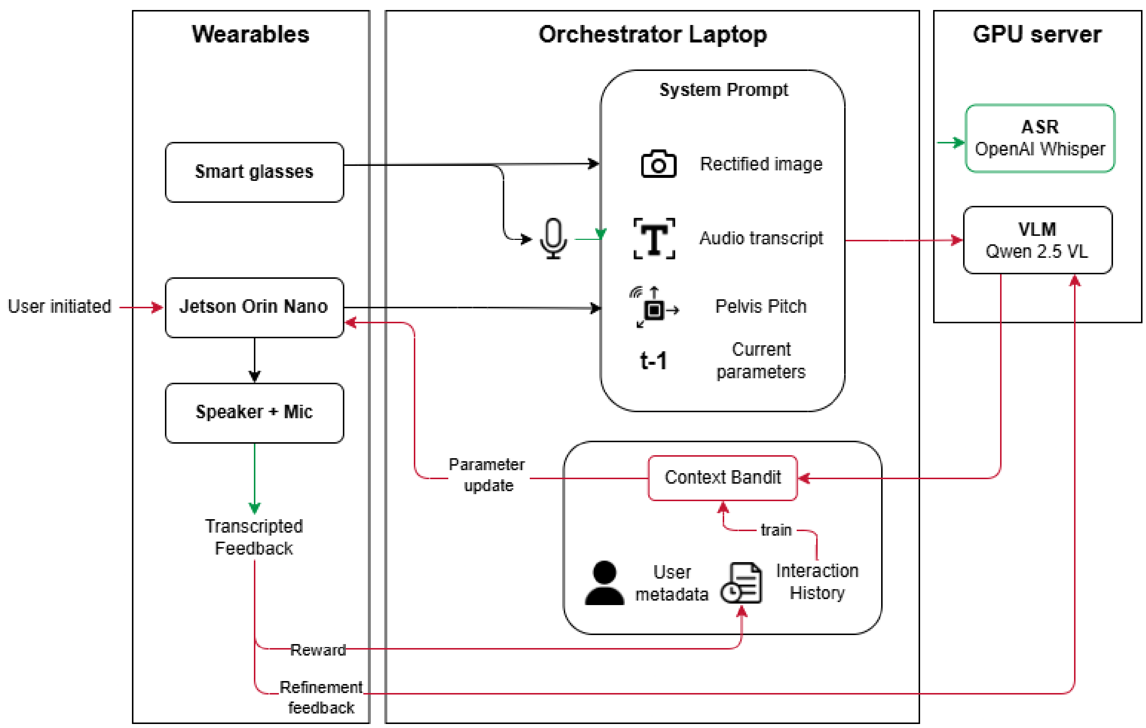

Figure 3.7: System Implementation

3.8 Baselines

3.8.1 Baseline 1: Single Shot Interpreter

The first baseline is a single VLM pass that lets the VLM decide magnitudes for scale and delay. The model receives the multimodal input, interprets the user's feedback, and emits one parameter adjustment. There is no reward, feedback or learning across interactions. This condition isolates the model's interpretation on its own, but is incapable of truly personalizing to a given user.

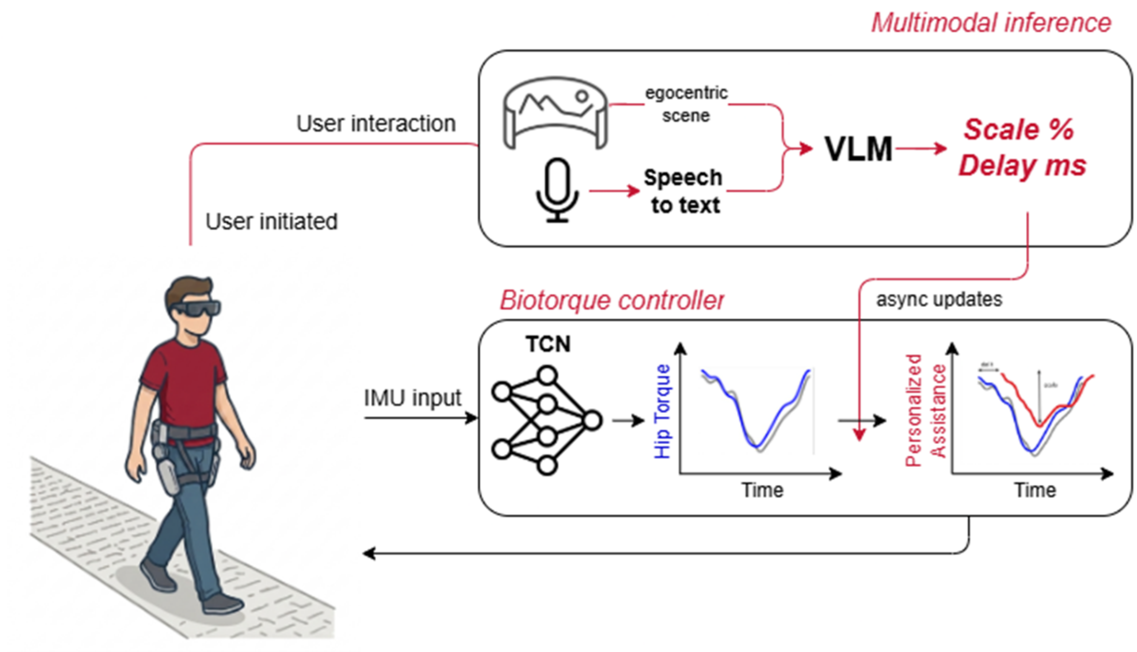

Figure 3.8: Overview of Baseline 1

3.8.2 Baseline 2: Few-shot recommendation system

The second baseline introduces personalization without a dedicated learning layer. Each VLM recommendation is logged in the user's history with a user-assigned reward. At inference, the highest-rated prior interaction for each task is retrieved and provided to the VLM as a few-shot exemplar, conditioning its output on previously preferred values. While this yields recommendations that are personalized to the individual, the resulting learning is implicit and unstructured, with no explicit model of preferred values and no reward-guided exploration of the action space. This baseline exhibits two characteristic limitations. First, a user may assign high ratings to differing scale values across distinct events, presenting the VLM with conflicting examples and no principled basis for resolving them. Second, the approach affords no mechanism for exploration beyond what has been observed before, so the recommendation remains anchored to past selections even when the user's preference in the current context can differ.

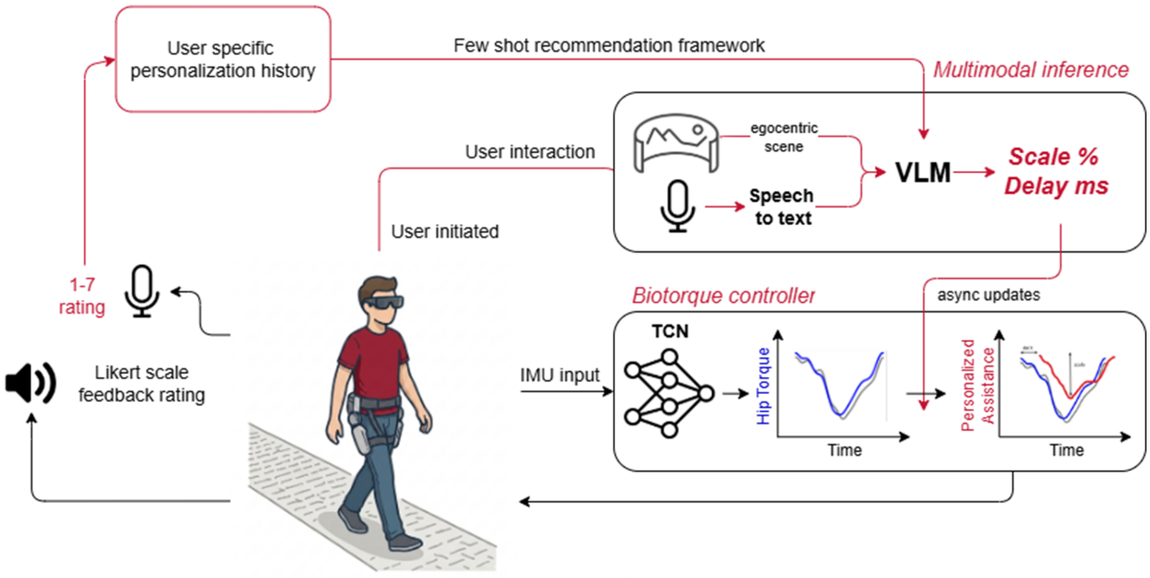

Figure 3.9: Overview of Baseline 2

Chapter 4

Experimentation Protocol

We piloted the proposed framework across a set of participants to assess system feasibility and gather preliminary evidence of learning across pipeline conditions. Participants wore the hip exoskeleton and the Aria glasses, and a short calibration was performed at the start to set each user's pelvis-pitch baseline. The calibration involved the user walking on level ground for 10 seconds, while the pelvis pitch was calculated for the user across time and averaged to established per user pelvis pitch.

The actual experiment consisted of 5 tasks performed by the user across the CMU campus, from level ground walking, to ramp and stairs ascent and descent. The stair ascent descent tasks were performed on the 4 flights of stairs in Scaife hall, the ramp ascent and descent on the spiral slope in Gates-Hillman centre, and the level ground walking task was conducted in Wean & Scott Hall. The three conditions were the single-shot interpreter (F1), the few-shot recommendation system (F2), and the proposed closed-loop feedback system (F3), as described in Chapter 3. Each framework was tried across two laps, leading to 30 total laps across 5 tasks x 3 frameworks x 2 laps. Personalization for each task was user-initiated, rather than at fixed points during each session. Every interaction was logged per participant, capturing the parameters before and after each update, the spoken-feedback transcript, the Likert rating, the VLM's structured scene reasoning, and the adaptation iteration. These logs form the basis of the analyses in Chapter 5.

Chapter 5

Pilot Results

We compare three levels of personalization: F1 (no personalization), F2 (naive, few shot personalization) and F3 (bandit based personalization) through our experimentation protocol described in Chapter 4.

5.1 Subjects

We test across n=3 subjects (age: 23 ± 2.65) for our pilot. The subjects were familiar with lower-limb exoskeletons, but not with the proposed frameworks. Data was also collected for 2 more subjects (n=5), but excluded from analysis due to transcription issues with the ASR model leading to incorrect bandit training, discussed later in 6.

5.2 Framework 3

Framework 3 is the contribution of this work, and we evaluate it on the following 3 questions: whether it brings user ratings to satisfaction within the limitations of the experimentation protocol, whether it shows learning across user interactions, and whether we observe personalization to individual users. Comparisons to Framework 2 are noted where possible, but a controlled isolation of the loop and the bandit is left to future work, as the frameworks differ in more than one component.

Within a user-initiated personalization instance, henceforth referred to as an episode, Framework 3 improved outcomes across laps through two complementary mechanisms. Within an episode, the feedback loop recovered from unsatisfactory initial settings: among loop-engaged episodes (those reaching satisfaction through iteration), every episode improved, with mean within-episode gains (terminal minus first Likert rating) of +3.71 and +2.75 across the two laps (Figure 5.1). Mean terminal Likert rose from 5.15 to 5.67, with all second-lap episodes at or above the satisfaction threshold of 5. As the bandit accumulated per-user preferences, fewer episodes required the loop, which is why within-episode gain declines as terminal satisfaction rises.

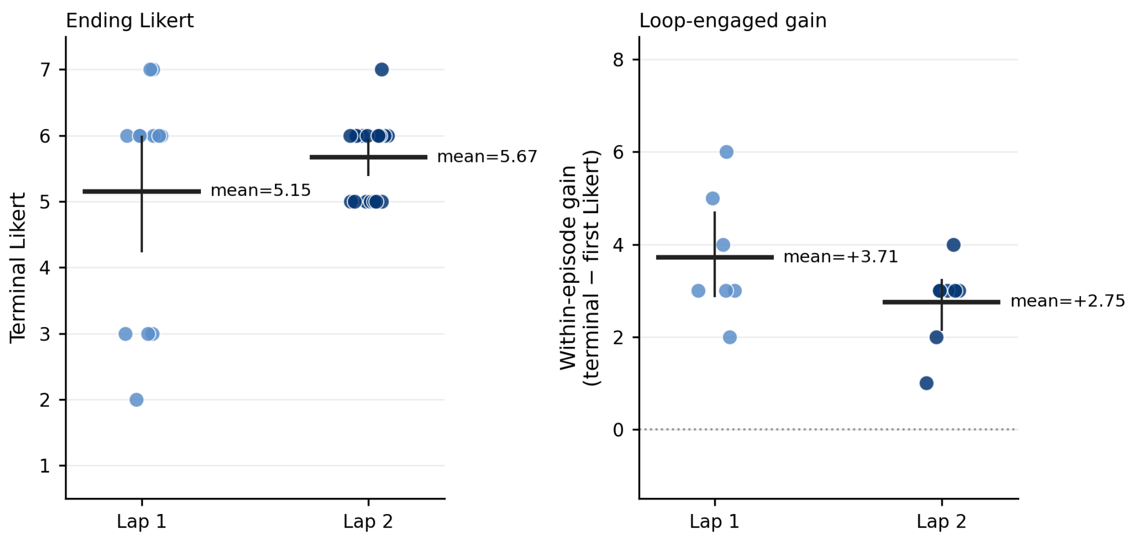

Figure 5.1: Framework 3 satisfaction and within-episode recovery

Across episodes and the cohort, the system increasingly suggested preferred settings for users on the first attempt, referred to as one-shot success where episodes satisfied without iteration. The proportion of one-shot success rose from 15% to 56% across laps, and no-improvement episodes fell from 31% to zero by the second lap (Figure 5.2), consistent with our expectation that the bandit learnt user preferences.

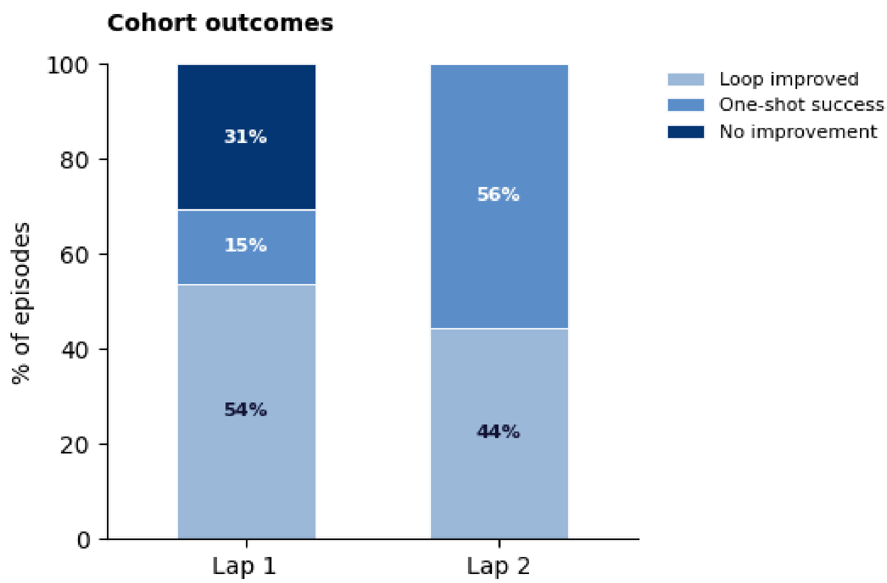

Figure 5.2: Framework 3 cohort outcomes across laps

5.2.1 User specific personalization

Under the same experimentation protocol, the three participants converged to distinct regions of the controller parameter space (Figure 5.3) based on the available interaction data. One participant settled at high assistance with short delay, another at low assistance with longer delay, and the third varied by task demand. These operating points are separable between users even with limited data, which supports the potential of our system adapting and personalizing to the individual rather than to a population average.

5.3 Qualitative Evaluations

We do not evaluate Framework 2 or 1 on quantitative metrics for this sample size, and we restrict its analysis to qualitative failure patterns that motivate and justify the design of Framework 3, and discuss these in Chapter 6.

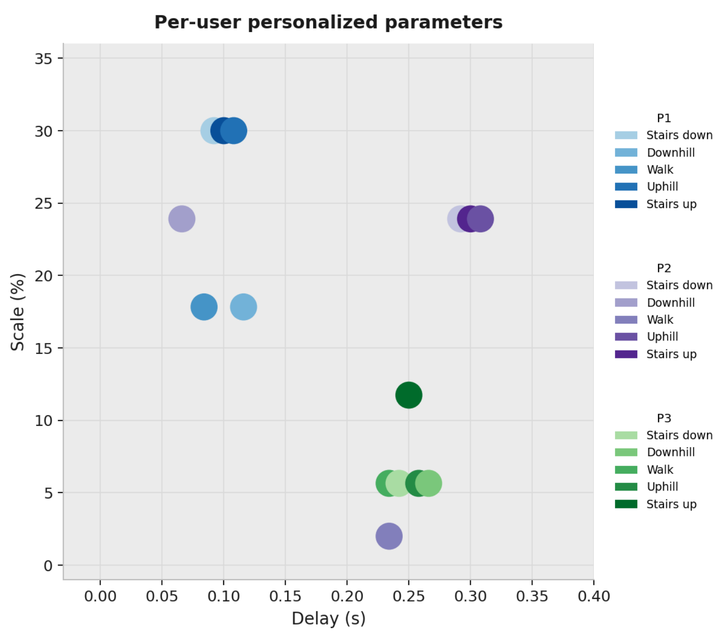

Figure 5.3: Settled scale and delay values by task context from user-specific contextual bandits.

Chapter 6

Discussion

6.1 VLM as an interpreter

We hypothesized that off-the-shelf VLMs could be used as contextual interpreters of user intent and environmental information in a personalization task for lower limb exoskeleton control. We also hypothesized that our contextual-bandit guided learning framework could learn user preferences across real-world interactions. We structure the following discussion evaluating these hypotheses both quantitatively and qualitatively.

6.1.1 Framework 1: Interpretation without exploration

We use Framework 1 to evaluate the hypothesis that VLMs can act as interpreters of user intent and scene context, and can map these to concrete, quantitative updates in exoskeleton control. Framework 1, through prompt iterations during design, achieves robust mapping capabilities for unambiguous user intent. Nonetheless, we observe a particular failure case during actual experimentation, where the VLM correctly reasons about the expected mapping, but fails to suggest updated parameters. We note this as a freeze response in 13.8% interactions of Framework 1 across our cohort. Further, the lack of a feedback response means that Framework 1 cannot finetune to recover from freeze responses or get human reward for learning user preferences. This motivates the design of our feedback based system that uses human reward to "grade" VLM suggestions, and the feedback loop that allows for fine-adjustment.

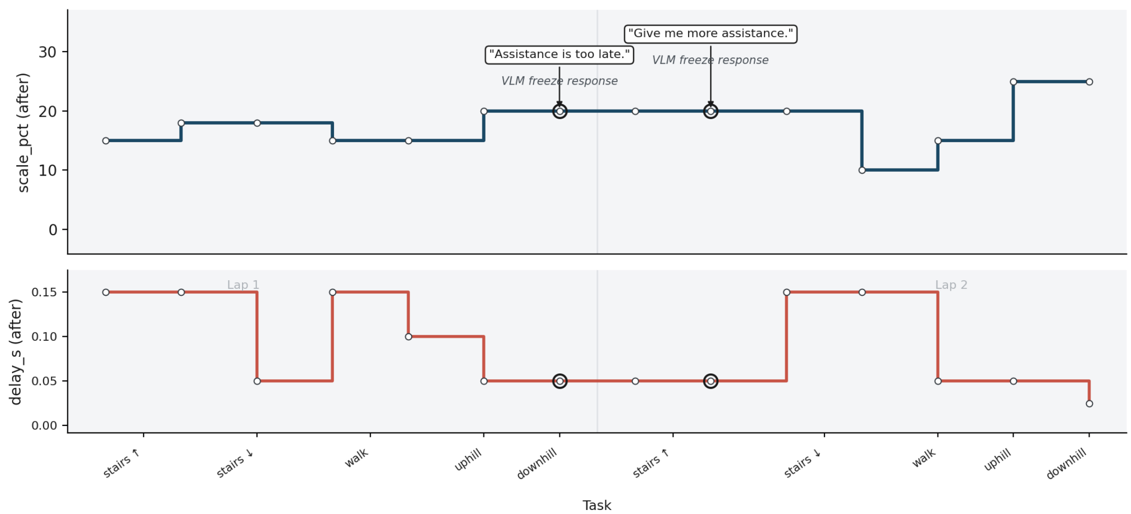

Figure 6.1: Two user prompts causing VLM freeze response

6.1.2 Framework 2: Anchoring without structured exploration

To build up on Framework 1's failures, we add human reward to "grade" VLM suggestions and use best-rated exemplars as few-shot, user-specific conditioning, to create a naive personalization approach. In practice, we see a specific failure attached to this. While few-shotting with exemplars helps VLM predict more-personalized parameters, the VLM tends to anchor itself in these exemplars and struggles with any exploration beyond those values. While this anchoring behaviour doesn't necessarily point to a VLM failure, and can be associated with model capabilities or the system prompt, it does motivate the need for a structured approach towards preference-learning. The lack of a lightweight, principled approach to exploration and exploitation of our action space motivates the addition of our bandit-based learning pipeline.

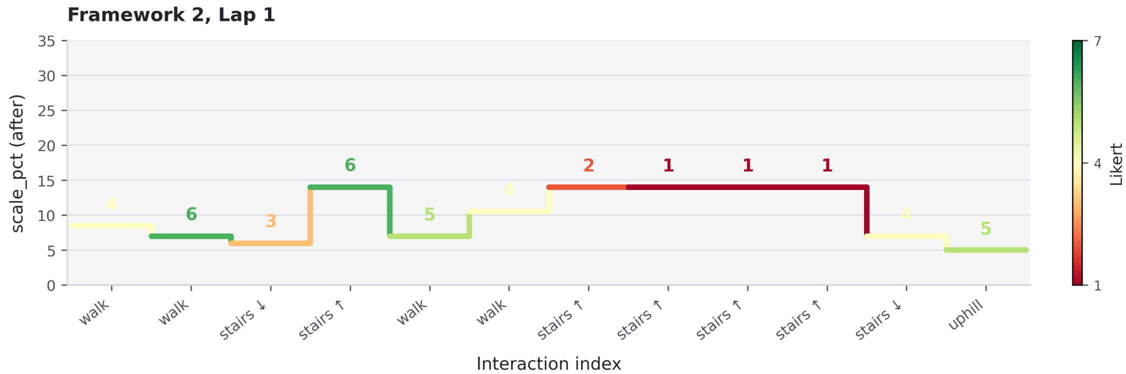

Figure 6.2: Framework 2's few shot approach anchors itself on exemplar values and struggles with exploration.

6.2 Contextual bandit guided personalization

Framework 3 settles on distinct user-specific preferences for within our experimentation framework, validating the hypothesis that the context bandit can model and learn user preferences across interactions. Given that diversity in interactions leads to maximum learning for the context bandit, it raises an important question about the required number of interactions for a bandit to "converge" to user-specific parameters. In our experiments, across the cohort, across 182 user prompts, 54.8% were regarding scale, 33.7% focused on delay, 8.4% discussed both and the rest 3% were ambiguous. In terms of user perception, the effect of varying scale was more obvious than delay, and users could clearly request for changes in scale, but not in delay. This imbalance in the diversity of user interactions needs to be accounted for while training the bandit effectively, in minimum interactions. As the action space gets more diverse in future work, this imbalance can lead to poor learning behaviour. This was also reflected in some HRI surveys we took post-experiment, which are discussed in section 6.3. Every interaction datapoint was used to train the bandit, but this gave rise to a very interesting problem.

User 3 initially requested for a later timing, to which the bandit responded appropriately by increasing the delay value. User perception of this change was negative, and was recorded in the feedback as a Likert rating of 1. The context bandit learnt that the user disliked that particular delay value, which was incorrect as the user prompt was wrong. This is a limitation of the current framework. It is pertinent to have a more robust approach to validating the quality of each interaction before using it to train the bandit, so it does not learn from corrupted inputs.

6.3 HRI metrics: Cognitive workload

In the design of an interactive framework, it is crucial for the users to know and effectively convey their preferences to maximize system performance and benefits. This is trivial in the case of intuitive parameters like scale. Delay is the timing offset between the onset of a gait cycle and the assistance provided to the user. Due to the cyclical nature of locomotion tasks, this is complicated to isolate and perceive for users, as offset that is late for one cycle is inherently early for the next cycle. We recorded the NASA-TLX cognitive workload survey and an anecdotal response from all users.

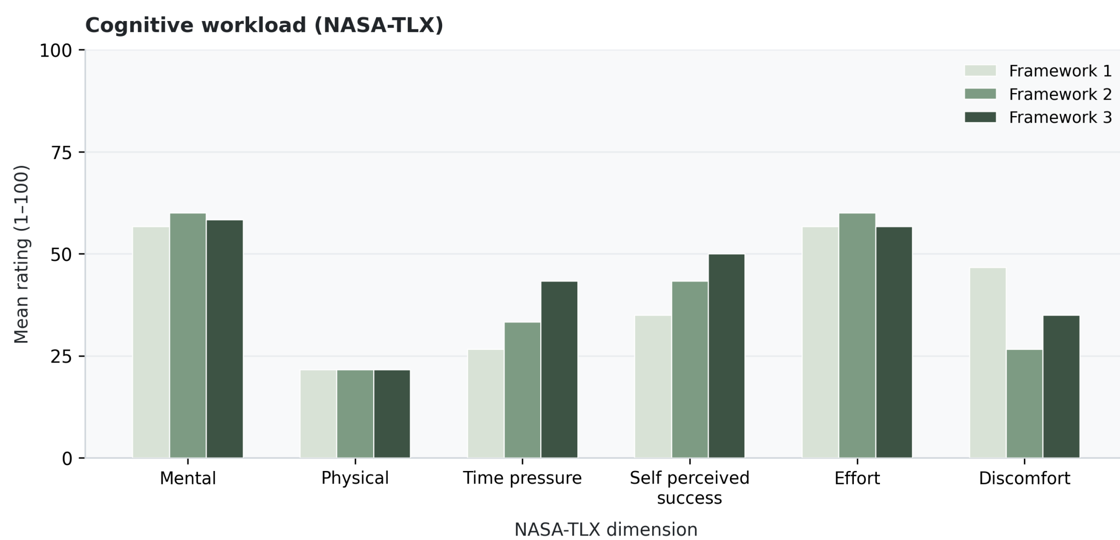

Figure 6.3: NASA-TLX Cognitive Workload survey across frameworks

In the workload response across the cohort, we see similar user perception of mental activity and effort required across frameworks. Time pressure increased across frameworks, largely due to the repetitive feedback loop of the context bandit that triggered a few seconds after each adaptation. In line with our hypothesis, we see user perception of self-perceived success increase across frameworks. Discomfort or stress varies across the frameworks, with framework 1 having the highest discomfort, followed by framework 3. This could potentially be attributed to the lack of personalization in framework 1 that requires constant user-initialization of personalization from users to get to their preferred values, and the increased time pressure for framework 3 adding to user-perceived discomfort.

In anecdotal analysis, we share three primary comments from each participant.

> "For someone new to exoskeletons, there might be some struggle in learning how to adjust the scale and delay, like knowing what to ask the system for compared to a person that has a better understanding of how to adjust the controller for better feeling results."

> "Sometimes it was hard to determine how to adjust the assistance to improve it, more with timing than scaling as timing is hard to feel"

> "I had a really hard time determining which direction to move the control in. It was easy to say something feels off, but it was very hard for me to figure out should the timing be sooner or later even if it is very pronounced."

These reflect user concerns regarding the learning curve associated with the adapting the exoskeleton assistance, specifically with the timing of assistance. An important note here is that each framework was tested for about $\tilde{4}0$ minutes in total per user, with the entire experiment requiring about 2.5-3 hours from setup to completion. User comments thus, also reflect the perception of initial adjustment to a new framework and limited adjustment time for the user. A more concrete experiment could test workload across sessions and days to isolate such effects.

Chapter 7

Future Work

We presented a VLM-guided framework that maps user-intent and contextual scene information to quantitative changes in exoskeleton assistance, with an online preference learning layer using a contextual bandit. We pilot this framework across an n=3 cohort to validate it across real-world experimentation and validate our hypotheses for it.

In terms of the experimentation protocol, future work should test this framework across a larger cohort for results that can be tested for statistical significance. Baseline comparisons can include concrete ablations of system components to isolate and evaluate each contribution effectively, over comparisons of pure personalization approaches. An ablation across VLM and speech transcription models can also be insightful regarding VLM limitations for this task. We observed issues with speech transcription models when working with different accents, leading to the exclusion of 2 participants' data. Robust speech transcription is crucial to this framework, and future work should ablate through state-of-the-art models to ensure the same.

Iterations of this framework can test richer action spaces beyond scale and delay, potentially building custom, parametrized splines, user-specific spline profiles, or adapt other standard controllers using a similar approach. While our baseline controller enabled user-intent based personalization due to relatively intuitive nature of scaling and delaying assistance, it is important to consider where user-preferences are informative enough to guide parameter selection in spline design or alternative controllers. We believe there is potential in using a user-intent guided framework in tandem with other conventional approaches to personalize assistance, as it addresses the shortcoming where users might prefer settings which are uncomfortable but biomechanically accurate.

Another observation regarding personalization lies in the varying nature of user preferences across different sessions. Our framework enables a static bandit that learns one preference per context, while user preferences act as moving targets across sessions. While the bandit can be retrained using a sliding window of user interaction history, we can also try alternative approaches or dynamic bandit implementations to address this aspect of varying user-preference learning.

This framework has promising implications for future work in posing this as an assistance planning problem. An important use of the vision modality is its ability to give context for future tasks and states, and use this information to pre-emptively personalize, ensuring seamless transitions between overground tasks and a reduced cognitive workload for the users. Specifically for clinical populations, who face difficulties in initiating challenging tasks like stair ascent, a vision-based planning framework holds a lot of potential.

Chapter 8

Conclusions

We presented a VLM-guided, real-time personalization framework that maps user-intent and contextual scene information to quantitative changes in exoskeleton assistance across two parameters, scale (assistance magnitude) and delay (timing offset), with an online preference learning layer guided by a contextual bandit. The framework used wearable smart glasses to provide multimodal input - transcribed user speech and visual scene frame - to a vision language model (Qwen 2.5 VL: 7B), to obtain a context vector. This vector was used to query a context bandit (disjoint LinUCB) to select action arms across scale and delay according to user intent as described in the context vector. Human feedback ratings on a Likert scale from 1-7 acted as rewards for training the bandit with each user interaction, with a feedback loop that allowed for refinement of bandit suggestions.

We piloted this framework across a cohort of 3 participants on a real-world experiment track spanning 2 laps of 5 locomotion tasks, obtaining the following results. We saw an elimination of no-improvement feedback loops by the second lap of the proposed framework, with an increase in terminal user satisfaction. Within-episode gain across laps decreased, consistent with our expectation of maximum bandit learning in lap 1 and sustained learning in lap 2. Compared to Framework 2, Framework 3 showed higher Likert ratings in the scope of the pilot.

Through our work, we hope to motivate further research in preference-guided user personalization for lower limb exoskeletons, in line with previous studies in this area [11, 12, 14, 24, 25, 26]. While user-preference cannot replace conventional approaches that evaluate physiological benefit, we believe that preference can act as a valid input in such optimization studies to account for user comfort, ensuring better user perception and long-term adaptation of lower-limb exoskeletons.

Bibliography

[1] Shuo Gao et al. Wearable technologies for assisted mobility in the real world. *Nature Communications*, 16(1), 2025. doi: 10.1038/s41467-025-67126-4. 1, 2.1

[2] Gregory S. Sawicki, Owen N. Beck, Inseung Kang, and Aaron J. Young. The exoskeleton expansion: improving walking and running economy. *Journal of NeuroEngineering and Rehabilitation*, 17(1):25, 2020. doi: 10.1186/s12984-020-00663-9. 1, 2.1

[3] Patrick Slade, Mykel J. Kochenderfer, Scott L. Delp, and Steven H. Collins. Personalizing exoskeleton assistance while walking in the real world. *Nature*, 610 (7931):277–282, 2022. doi: 10.1038/s41586-022-05191-1. 1, 2.2

[4] Inseung Kang, Dean D. Molinaro, Dongho Park, Dawit Lee, Pratik Kunapuli, Kinsey R. Herrin, and Aaron J. Young. Online adaptation framework enables personalization of exoskeleton assistance during locomotion in patients affected by stroke. *IEEE Transactions on Robotics*, 41:4941–4959, 2025. doi: 10.1109/TRO.2025.3595701. 1

[5] Inseung Kang, Pratik Kunapuli, and Aaron J. Young. Real-time neural network-based gait phase estimation using a robotic hip exoskeleton. *IEEE Transactions on Medical Robotics and Bionics*, 2(1):28–37, 2020. doi: 10.1109/TMRB.2019.2961749. 1

[6] Dean D. Molinaro, Inseung Kang, Jonathan Camargo, Matthew C. Gombolay, and Aaron J. Young. Subject-independent, biological hip moment estimation during multimodal overground ambulation using deep learning. *IEEE Transactions on Medical Robotics and Bionics*, 4(1):219–229, 2022. doi: 10.1109/TMRB.2021.3130979. 1, 2.1, 3.4

[7] Clément Lhoste, Emek Barış Küçüktabak, Lorenzo Vianello, Lorenzo Amato, Matthew R. Short, Kevin M. Lynch, and Jose L. Pons. Deep-learning estimation of weight distribution using joint kinematics for lower-limb exoskeleton control. *IEEE Transactions on Medical Robotics and Bionics*, 7(1):20–26, 2025. doi: 10.1109/TMRB.2024.3503922. 1

[8] Longbin Zhang, Davit Soselia, Ruoli Wang, and Elena M. Gutierrez-Farewik.

Estimation of joint torque by emg-driven neuromusculoskeletal models and lstm networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:3722–3731, 2023. doi: 10.1109/TNSRE.2023.3315373. 1

[9] Patrick Slade, Christopher Atkeson, J. Maxwell Donelan, Han Houdijk, Kimberly A. Ingraham, Myunghee Kim, Kyu-Jin Kong, Katherine L. Poggensee, Robert Riener, Martin Steinert, Juanjuan Zhang, and Steven H. Collins. On human-in-the-loop optimization of human–robot interaction. *Nature*, 633(8031): 779–788, 2024. doi: 10.1038/s41586-024-07697-2. 1, 2.2

[10] Juanjuan Zhang, Pieter Fiers, Kirby A. Witte, Rachel W. Jackson, Katherine L. Poggensee, Christopher G. Atkeson, and Steven H. Collins. Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, 356(6344): 1280–1284, 2017. doi: 10.1126/science.aal5054. 1, 2.2

[11] Philipp Arens et al. Preference-based assistance optimization for lifting and lowering with a soft back exosuit. *Science Advances*, 11:eadu2099, 2025. doi: 10.1126/sciadv.adu2099. 1, 8

[12] Kimberly A. Ingraham, C. David Remy, and Elliott J. Rouse. The role of user preference in the customized control of robotic exoskeletons. *Science Robotics*, 7 (64):eabj3487, 2022. doi: 10.1126/scirobotics.abj3487. 1, 2.3, 8

[13] Kimberly A Ingraham, Maegan Tucker, Aaron D Ames, Elliott J Rouse, and Max K Shepherd. Leveraging user preference in the design and evaluation of lower-limb exoskeletons and prostheses. *Current Opinion in Biomedical Engineering*, 28:100487, 2023. 1

[14] Nitish Thatte, Haoyong Duan, and Hartmut Geyer. A sample-efficient black-box optimizer to train policies for human-in-the-loop systems with user preferences. *IEEE Robotics and Automation Letters*, 2(2):993–1000, 2017. doi: 10.1109/LRA.2017.2654358. 1, 2.3, 8

[15] Maggie A Collier, Rithika Narayan, and Henny Admoni. The sense of agency in assistive robotics using shared autonomy. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 880–888. IEEE, 2025. 1

[16] Sumukha Udupa, Vineet R Kamat, and Carol C Menassa. Shared autonomy in assistive mobile robots: a review. *Disability and Rehabilitation: Assistive Technology*, 18(6):827–848, 2023. 1

[17] Duosi Dai, Yuchong Zhang, Yong Ma, and Danica Kragic. Ai that moves with you: A review of interactive technologies powered by large foundation models for mobility impairment. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, CHI '26, New York, NY, USA, 2026. Association for Computing Machinery. ISBN 9798400722783. doi: 10.1145/3772318.3791239. URL `https://doi.org/10.1145/3772318.3791239`. 1

[18] Akhil Padmanabha, Jessie Yuan, Janavi Gupta, Zulekha Karachiwalla, Carmel Majidi, Henny Admoni, and Zackory Erickson. Voicepilot: Harnessing llms as speech interfaces for physically assistive robots. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, UIST '24, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400706288. doi: 10.1145/3654777.3676401. URL `https://doi.org/10.1145/3654777.3676401`. 1, 2.4.1

[19] T. Pulikottil, E. Biffi, E. Diella, et al. Testing the usability of a voice control system for assistive robotic arms in people with neurological conditions. *Journal of NeuroEngineering and Rehabilitation*, 23:90, 2026. doi: 10.1186/s12984-026-01902-1. URL `https://doi.org/10.1186/s12984-026-01902-1`. 1

[20] Inseung Kang, Dean D. Molinaro, Srijan Duggal, Yue Chen, Pratik Kunapuli, and Aaron J. Young. Real-time gait phase estimation for robotic hip exoskeleton control during multimodal locomotion. *IEEE Robotics and Automation Letters*, 6(2):3491–3497, 2021. doi: 10.1109/LRA.2021.3062562. 2.1

[21] Dean D. Molinaro, Inseung Kang, and Aaron J. Young. Estimating human joint moments unifies exoskeleton control, reducing user effort. *Science Robotics*, 9 (88):eadi8852, 2024. doi: 10.1126/scirobotics.adi8852. 2.1

[22] S. Song and Steven H. Collins. Optimizing exoskeleton assistance for faster self-selected walking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:786–795, 2021. doi: 10.1109/TNSRE.2021.3074154. 2.2

[23] Myunghee Kim, Charles Liu, Jinsoo Kim, Sangjun Lee, Adham Meguid, Conor Walsh, and Scott Kuindersma. Bayesian optimization of soft exosuits using a metabolic estimator stopping process. In *2019 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9173–9179, 2019. doi: 10.1109/ICRA.2019.8793817. 2.2

[24] Ung Hee Lee, Varun S. Shetty, Patrick W. Franks, Jie Tan, Georgios Evangelopoulos, Sehoon Ha, and Elliott J. Rouse. User preference optimization for control of ankle exoskeletons using sample efficient active learning. *Science Robotics*, 8(83):eadg3705, 2023. doi: 10.1126/scirobotics.adg3705. 2.3, 8

[25] Maegan Tucker, Ellen Novoseller, Claudia Kann, Yanan Sui, Yisong Yue, Joel W. Burdick, and Aaron D. Ames. Preference-based learning for exoskeleton gait optimization. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2351–2357, 2020. doi: 10.1109/ICRA40945.2020.9197356. 2.3, 8

[26] Tyler R. Clites, Max K. Shepherd, Kimberly A. Ingraham, Leslie Wontorcik, and Elliott J. Rouse. Understanding patient preference in prosthetic ankle stiffness. *Journal of NeuroEngineering and Rehabilitation*, 18(1):128, 2021. doi:

10.1186/s12984-021-00916-1. 2.3, 8

[27] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 30, 2017. URL https://arxiv.org/abs/1706.03762. 2.4

[28] Tom B. Brown et al. Language models are few-shot learners. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, 2020. URL https://arxiv.org/abs/2005.14165. 2.4

[29] Long Ouyang et al. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, 2022. URL https://arxiv.org/abs/2203.02155. 2.4

[30] Mohit Shridhar, Lucas Manuelli, and Dieter Fox. CLIPort: What and where pathways for robotic manipulation. In *Proceedings of the 5th Conference on Robot Learning (CoRL)*, volume 164 of *PMLR*, 2022. URL https://arxiv.org/abs/2109.12098. 2.4

[31] Michael Ahn et al. Do as I can, not as I say: Grounding language in robotic affordances. In *Proceedings of the 6th Conference on Robot Learning (CoRL)*, volume 205 of *PMLR*, 2022. URL https://arxiv.org/abs/2204.01691. 2.4

[32] Anthony Brohan et al. RT-2: Vision-language-action models transfer web knowledge to robotic control. arXiv:2307.15818, 2023. URL https://arxiv.org/abs/2307.15818. 2.4

[33] Moo Jin Kim et al. OpenVLA: An open-source vision-language-action model. In *Proceedings of the 8th Conference on Robot Learning (CoRL)*, volume 270 of *PMLR*, 2024. URL https://arxiv.org/abs/2406.09246. 2.4

[34] Maria R. Lima, Amy O'Connell, Feiyang Zhou, Alethea Nagahara, Avni Hulyalkar, Anura Deshpande, Jesse Thomason, Ravi Vaidyanathan, and Maja Matarić. Promoting cognitive health in elder care with large language model-powered socially assistive robots. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400713941. doi: 10.1145/3706598.3713582. URL https://doi.org/10.1145/3706598.3713582. 2.4.1

[35] Chao Wang, Stephan Hasler, Daniel Tanneberg, Felix Ocker, Frank Joublin, Antonello Ceravola, Joerg Deigmoeller, and Michael Gienger. Lami: Large language models for multi-modal human-robot interaction. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2024. 2.4.1

[36] Linus Nwankwo and Elmar Rueckert. The conversation is the command: Interact-

ing with real-world autonomous robots through natural language. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 808–812, 2024. 2.4.1

[37] Jimmy Wu, Rika Antonova, Adam Kan, Marion Lepert, Andy Zeng, Shuran Song, Jeannette Bohg, Szymon Rusinkiewicz, and Thomas Funkhouser. Tidybot: Personalized robot assistance with large language models. *Autonomous Robots*, 47(8):1087–1102, 2023. 2.4.1

[38] Keshav Shankar, Dan Ding, and Wei Gao. Low-burden llm-based preference learning: Personalizing assistive robots from natural language feedback for users with paralysis. *arXiv preprint arXiv:2604.01463*, 2026. 2.4.1

[39] Wenyuan Chen, Guangyong Li, Mingwei Li, Wenxue Wang, Peng Li, Xiujuan Xue, Xingang Zhao, and Lianqing Liu. Llm-enabled incremental learning framework for hand exoskeleton control. *IEEE Transactions on Automation Science and Engineering*, 22:2617–2626, 2025. doi: 10.1109/TASE.2024.3382679. 2.4.2

[40] Houcheng Li, Zhenchan Su, Honglei Guo, Yifan Wang, Zeyu Liu, and Long Cheng. Eic framework for hand exoskeletons based on a multimodal large language model. In *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10063–10068. IEEE, 2025. 2.4.2

[41] Yu Chen, Shu Miao, Chunyu Wu, Jingsong Mu, Bo Ouyang, and Xiang Li. A semantic-aware framework for safe and intent-integrative assistance in upper-limb exoskeletons. *IEEE Robotics and Automation Letters*, 11(5):5938–5945, 2026. 2.4.2

[42] Raajshekhar Parameswari, Christian Di Natali, Jesus Ortiz, Darwin G Caldwell, et al. Task assessment of xonli: A natural language interface for occupational exoskeletons. In *2024 10th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, pages 634–640. IEEE, 2024. 2.4.2

[43] Eddie Guo, Christopher Perlette, Mojtaba Sharifi, Lukas Grasse, Matthew Tata, Vivian K. Mushahwar, and Mahdi Tavakoli. Speech-based human-exoskeleton interaction for lower limb motion planning. In *2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS)*, pages 1–6, 2024. doi: 10.1109/ICHMS59971.2024.10555587. 2.4.2

[44] Hyeon-Seok Seong, Seif Farag, Sang-Wook Lee, Igor Gaponov, and Jee-Hwan Ryu. Wearable robot control method based on vision-language models. In *First Workshop on Vision-Language Models for Navigation and Manipulation at ICRA 2024*, 2024. 2.4.2

[45] Ehsan Ahmadi, Asif Faisal Chowdhury, Chao Wang, and Mahmood Jasim. Exoskeleton locomotion mode prediction in construction using gpt-4o: Zero-shot

learning from vision and speech. *Journal of Intelligent Construction*, 2026. 2.4.2

[46] P. Auer. Using upper confidence bounds for online learning. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 270–279, 2000. doi: 10.1109/SFCS.2000.892116. 2.5

[47] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, page 661–670, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605587998. doi: 10.1145/1772690.1772758. URL `https://doi.org/10.1145/1772690.1772758`. 2.5

[48] Ethan K. Gordon, Xiang Meng, Tapomayukh Bhattacharjee, Matt Barnes, and Siddhartha S. Srinivasa. Adaptive robot-assisted feeding: An online learning framework for acquiring previously unseen food items. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9659–9666, 2020. doi: 10.1109/IROS45743.2020.9341359. 2.5.1

[49] Ethan K. Gordon, Sumegh Roychowdhury, Tapomayukh Bhattacharjee, Kevin Jamieson, and Siddhartha S. Srinivasa. Leveraging post hoc context for faster learning in bandit settings with applications in robot-assisted feeding. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10528–10535, 2021. doi: 10.1109/ICRA48506.2021.9561520. 2.5.1

[50] Rohan Banerjee, Rajat Kumar Jenamani, Sidharth Vasudev, Amal Nanavati, Katherine Dimitropoulou, Sarah Dean, and Tapomayukh Bhattacharjee. To ask or not to ask: Human-in-the-loop contextual bandits with applications in robot-assisted feeding. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1378–1384, 2025. doi: 10.1109/ICRA55743.2025.11127795. 2.5.1

[51] Roshni Kaushik, Rayna Hata, Aaron Steinfeld, and Reid Simmons. Choosing robot feedback style to optimize human exercise performance. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 658–666, 2025. doi: 10.1109/HRI61500.2025.10974072. 2.5.1

[52] Nitish Thatte, Haoyong Duan, and Hartmut Geyer. A method for online optimization of lower limb assistive devices with high dimensional parameter spaces. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5380–5385, 2018. doi: 10.1109/ICRA.2018.8461196. 2.5.2