Bidirectional Human-AI Learning in Real-Time Disoriented Balancing

Sheikh Mannan, Nikhil Krishnaswamy

Situated Grounding and Natural Language (SIGNAL) Lab, Colorado State University
Fort Collins, CO 80523 USA
{sheikh.mannan,nkrishna}@colostate.edu

Abstract

We present a real-time system that enables bidirectional human-AI learning and teaching in a balancing task that is a realistic analogue of disorientation during piloting and spaceflight. A human subject and autonomous AI model of choice guide each other in maintaining balance using a visual inverted pendulum (VIP) display. We show how AI assistance changes human performance and vice versa.

Introduction

Spatial disorientation remains a leading causes of fatal aircraft accidents (Braithwaite et al. 1998; Gibb, Ercoline, and Scharff 2011). An automated system, without the physical characteristics that lead to the inducement of disorientation, can potentially serve as a countermeasure (Wang et al. 2022). However, the automated system may execute strategies counter to the human's own preference, and too many such instance may result in the human losing trust in the AI (Nikolaidis et al. 2017) even when AI assistance makes overall human task performance better. One way to mitigate such conflict is *dyadic interaction*, wherein the process of interacting with a partner changes one's own underlying behavior (Park and Kim 2012; Roy, Singhal, and Srivastava 2017).

In this system, we demonstrate how mutual human-AI adaptation may be conducted in real-time using a realistic simulation of balancing under disorientation, where human participants use a joystick to stabilize themselves about the direction of balance (DOB). See Fig. 1. A human user may be paired with different types of AI models which themselves display a variety of proficiency and performance strategies in this action-learning task with well-controlled parameters that is nonetheless challenging for humans. Human and AI mutually correct each other's actions with either visualized suggestions (AI-to-human) or direct numerical joystick input (human-to-AI). The differences in both human and AI performance before and after dyadic learning can be clearly visualized in intuitive phase portraits.

Task Setting

The **virtual inverted pendulum (VIP)** task is a documented, realistic simulation of upright balance control in a disorienting condition analogous to spaceflight or piloting in the absence of strong gravitational cues. The IP is represented as a visually simulated circular array of dots (random dot kinematogram, RDK) which rolls in the plane of the display screen governed by $\ddot{\theta} = k_P sin\theta$, where θ is degrees deviation from the DOB and pendulum constant $k_P = 600°/s^2$ (Vimal, DiZio, and Lackner 2017). If the IP drifts beyond $\pm 60°$ from the DOB, it has "crashed". To induce disorientation, the VIP is rendered at 50% *coherence*: between every two consecutive frames, half the dots displace coherently while the other half jump randomly. This eliminates configural displacement cues relative to the DOB while providing low-level retinal motion cues. The result is that humans watching the VIP in motion can tell how fast they are moving, but have difficulty telling how far from the upright they have fallen, as if they had been denied gravitational cues. Thus, in this high-throughput, portable setting, humans experience performance degradations in maintaining balance that strongly correlate with those induced by placement in a physical apparatus that disrupts signals from the vestibular system (Panic et al. 2015, 2017; Vimal, Lackner, and DiZio 2016, 2018; Vimal, DiZio, and Lackner 2017, 2019, 2021; Vimal et al. 2020; Wang et al. 2022; Mannan and Krish-

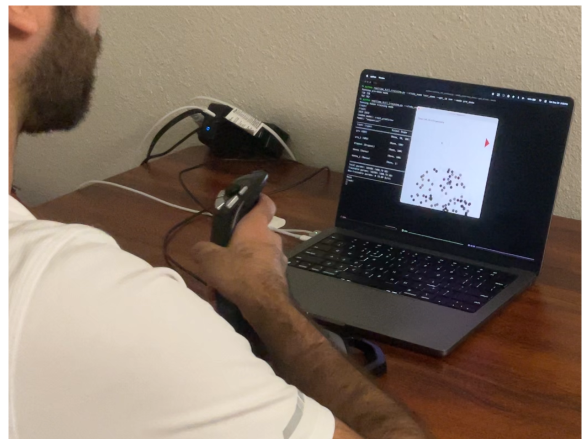

Figure 1: A human engaging in VIP balancing with cues (arrows) rendered by an AI assistant. A demonstration video can be viewed at https://youtu.be/coJdj0LIYa4.

naswamy 2022; DiZio et al. 2023; Mannan et al. 2024a,b).

A perfect performer in this task would rotate to the DOB and remains there with no drift or oscillation. Thus a proficient performer minimizes, e.g., distance from the DOB, angular velocity, and magnitude of actions. This means we can demonstrate bidirectional learning in real-time as humans or AIs display characteristic behaviors that may be more idealized (as above) and/or more human-like (such as small intermittent deflections, as in Vimal et al. (2020)) and these characteristics may converge as the two learn from each other.

System Functionality

An AI "assistant" can be instantiated as any of a variety of reinforcement learning (RL) or supervised deep learning models. Examples include SAC or DDPG instances trained in an environment programmed with the IP physics, or MLP, RNN, LSTM, or GRU models trained on data from humans performing disoriented IP balancing tasks using the VIP or an analogous physical apparatus (Panic et al. 2015; Vimal, DiZio, and Lackner 2017; Mannan et al. 2024a,b). There are 26 available assistants, detailed in Mannan et al. (2024a). All AI models predict what the next joystick deflection would be given the current angular position and velocity, and so depending on the model type and training data, the AI model may display different levels of native proficiency at performing the task.

In our system, after a tutorial to help the user acclimate to the joystick, controls, pendulum, and RDK movement, bidirectional human-AI learning proceeds in two phases:

1. Human training: First, the user performs the task alone to determine baseline performance. Then, the human is assisted by an AI, which provides suggestions when deemed necessary, rendered as arrows on screen.
2. AI training:
   a. An AI performs the task alone. During solo performance, the AI receives only numerical input, but depending on model and training data may display any level of proficiency at the task.
   b. The human then assists the AI in a second run by deflecting the joystick to keep the VIP balanced. This is done in the 50% coherent VIP condition to ensure that signals the AI receives are those from a human who is experiencing disorientation. All episodes where the human and AI disagreed on the direction of the movement are recorded. After the run, a brief finetuning is performed to update the assistant.
   c. The updated AI performs the task again where the human can determine whether the AI has improved or requires further corrections and updates. If further corrections are required step (b) can be repeated until the AI achieves acceptable performance.

After each phase, the baseline and assisted performances of the human or AI are shown in the form of phase portraits of angular velocity vs. angular position (e.g., Fig. 2).

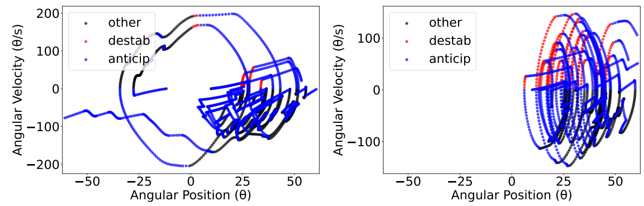

Figure 2: Phase portraits of sample human VIP performance without [L] and with [R] AI assistance. With AI assistance, this human subject decreased their oscillation and maintained stability even while offset from the DOB.

Technical Specifics

RL assistant models learned directly from exposure to environmental physics using a custom variation of Gymnasium's classic-control Pendulum environment, modified to reflect the dynamics of the VIP task. The deep learning models have been pretrained on human subject data gathered from trials performed in a physical Multi-Axis Rotation System (MARS) apparatus, programmed with identical dynamics, where subjects likewise use a joystick to balance the device while deprived of orientational cues (details in Wang et al. (2022)). Additional data was gathered from subject trials in the VIP setting (details in Mannan et al. (2024a)).

We also incorporate a *crash predictor*, which is a stacked GRU model as reported in Wang et al. (2022), which predicts the likelihood of a crash occurring. AI cueing is provided in cases of imminent danger (crash is $\geq$80% likely) where angular distance from the DOB exceeds 12°.

For fine-tuning models during the AI training phase, the actor networks of the SAC and DDPG are fine-tuned using behavior cloning over the trial data, the SAC-AIRL model is updated using AIRL over the trial data, and the deep learning models undergo standard fine-tuning. The RL models are fine-tuned for 100 epochs with a learning rate of $1e-5$ and a batch size of 64. The deep learning models are fine-tuned for 20 epochs with a learning rate of $1e-7$, a 9:1 train-test ratio, and a batch size of 16. All model fine-tuning can be conducted on a consumer laptop and takes approximately 30 seconds for a single training run.

Conclusion

Our system concisely demonstrates how human-AI mutual adaptation manifests through dyadic bidirectional learning, evident in changes in both human and AI behavior due to the interaction. Our demonstration is lightweight and showcases mutual human-AI learning in near-real-time on equipment as common as a consumer laptop, making it an accessible public-facing demonstration of AI. Our task setting is directly applicable to problems of spatial disorientation as in piloting or spaceflight, but our demonstration of human-AI bidirectional learning has broader applicability to the study of shared autonomy and problems in human-AI trust and allows for rapid parameterization of multiple experiments to test hypotheses in this area. Our code is available at https://github.com/csu-signal/HITL-VIP/releases/tag/v1.0.

Acknowledgments

Thanks to Paul DiZio, Vivekanand Pandey Vimal, and James R. Lackner for foundational work on the spatial disorientation problem and collecting and providing the initial training data, to Hannah N. Davies for additional data collection, and to Paige Hansen for help training the candidate assistant models.

References

Braithwaite, M. G.; Durnford, S. J.; Crowley, J. S.; Rosado, N. R.; and Albano, J. P. 1998. Spatial disorientation in US Army rotary-wing operations. *Aviation, space, and environmental medicine*, 69(11): 1031–1037.

DiZio, P.; Krishnaswamy, N.; Mannan, S.; and Hansen, P. 2023. Manual balancing of a visual inverted pendulum by quantized versus proportional joystick commands. In *Neuroscience*.

Gibb, R.; Ercoline, B.; and Scharff, L. 2011. Spatial disorientation: decades of pilot fatalities. *Aviation, space, and environmental medicine*, 82(7): 717–724.

Mannan, S.; Hansen, P.; Vimal, V. P.; Davies, H. N.; DiZio, P.; and Krishnaswamy, N. 2024a. Combating Spatial Disorientation in a Dynamic Self-Stabilization Task Using AI Assistants. In *Proceedings of the 12th International Conference on Human-Agent Interaction*, 113–122.

Mannan, S.; and Krishnaswamy, N. 2022. Where am I and where should I go? Grounding positional and directional labels in a disoriented human balancing task. In *Proceedings of the 2022 CLASP Conference on (Dis) embodiment*, 70–79.

Mannan, S.; Vimal, V. P.; DiZio, P.; and Krishnaswamy, N. 2024b. Embodying Human-Like Modes of Balance Control Through Human-In-the-Loop Dyadic Learning. In *Proceedings of the AAAI Symposium Series*, volume 3, 565–569.

Nikolaidis, S.; Zhu, Y. X.; Hsu, D.; and Srinivasa, S. 2017. Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 294–302.

Panic, A. S.; Panic, H.; DiZio, P.; and Lackner, J. R. 2017. Gravitational and somatosensory influences on control and perception of roll balance. *Aerospace medicine and human performance*, 88(11): 993–999.

Panic, H.; Panic, A. S.; DiZio, P.; and Lackner, J. R. 2015. Direction of balance and perception of the upright are perceptually dissociable. *Journal of neurophysiology*, 113(10): 3600–3609.

Park, W.-W.; and Kim, S. 2012. The need of leader–subordinate reciprocal dyadic trust to build the subordinate's trust in the organization: the case of Korean air pilots. *The International Journal of Aviation Psychology*, 22(2): 97–119.

Roy, A.; Singhal, A.; and Srivastava, J. 2017. Formation and reciprocation of dyadic trust. *ACM Transactions on Internet Technology (TOIT)*, 17(2): 1–24.

Vimal, V. P.; DiZio, P.; and Lackner, J. R. 2017. Learning dynamic balancing in the roll plane with and without gravitational cues. *Experimental brain research*, 235: 3495–3503.

Vimal, V. P.; DiZio, P.; and Lackner, J. R. 2019. Learning and long-term retention of dynamic self-stabilization skills. *Experimental brain research*, 237: 2775–2787.

Vimal, V. P.; DiZio, P.; and Lackner, J. R. 2021. The role of spatial acuity in a dynamic balancing task without gravitational cues. *Experimental brain research*, 1–11.

Vimal, V. P.; Lackner, J. R.; and DiZio, P. 2016. Learning dynamic control of body roll orientation. *Experimental brain research*, 234: 483–492.

Vimal, V. P.; Lackner, J. R.; and DiZio, P. 2018. Learning dynamic control of body yaw orientation. *Experimental brain research*, 236: 1321–1330.

Vimal, V. P.; Zheng, H.; Hong, P.; Fakharzadeh, L. N.; Lackner, J. R.; and DiZio, P. 2020. Characterizing individual differences in a dynamic stabilization task using machine learning. *Aerospace medicine and human performance*, 91(6): 479–488.

Wang, Y.; Tang, J.; Vimal, V. P.; Lackner, J. R.; DiZio, P.; and Hong, P. 2022. Crash prediction using deep learning in a disorienting spaceflight analog balancing task. *Frontiers in physiology*, 13: 51.