Yuheng Wu

+1 (608) 957-5384 | yuheng.wu@wisc.edu | iamwyh.com | LinkedIn: yuhengwu-wisc | Google Scholar

SUMMARY

I work on human-centered AI and human-computer interaction, building context-aware multimodal systems that perceive and understand complex, dynamic real-world environments and help people act in them. I design multi-agent LLM backends, fine-tune vision models, and engineer low-latency pipelines for phones, AR headsets, and smart glasses. I deploy and evaluate these systems with real users in the wild. I focus on accessibility, with applications in education and other everyday tasks.

EDUCATION

University of Wisconsin–Madison — **Madison, WI, USA**
Ph.D., Computer Sciences (HCI). Advisor: Prof. Yuhang Zhao — Sep. 2022 – May 2027 (expected)

- Research focus: Human-Centered AI, Intelligent Interactive Systems, Augmented Reality, Assistive Technology
- David G. Walsh Research Travel Award, Fall 2025

Peking University — **Beijing, China**
B.S., Computer Science, Turing Class (elite CS program) — Sep. 2018 – Jul. 2022

CONFERENCE AND JOURNAL PUBLICATIONS

[5] NavSight in the Wild: Understanding Real-World Use of a Mobile Augmented Reality Application for People with Low Vision in Outdoor Navigation — IMWUT 2026 [arXiv]
Yuheng Wu, Kexin Zhang, Ben Kosa, Ru Wang, Sanbrita Mondal, Yuhang Zhao

[4] What to Distinguish and How? Opportunities and Challenges of Augmenting Multiple, Cluttered Objects in Complex Scenes for People with Low Vision — ASSETS 2026 [DOI] [GitHub]
Yuheng Wu, Ruijia Chen, Jaewook Lee, Jia Li, Kexin Zhang, Meng Fong Lio, Weibing Wang, Sanbrita Mondal, Jon E. Froehlich, Yapeng Tian, Yuhang Zhao

[3] NaviNote: Enabling In-situ Spatial Annotation Authoring to Support Exploration and Navigation for Blind and Low Vision People — CHI 2026 [DOI]
Ruijia Chen*, **Yuheng Wu***, Charlie Houseago, Filipe Gaspar, Filippo Aleotti, Dorian Gálvez-López, Oliver Johnston, Diego Mazala, Guillermo Garcia-Hernando, Maryam Bandukda, Gabriel Brostow, Jessica Van Brummelen (* co-first authors)
🏆 **Honorable Mention | U.S. Patent Pending**

[2] AskNow: An LLM-powered Interactive System for Real-Time Question Answering in Large-Scale Classrooms — CHI 2026 [DOI]
Ziqi Liu, Yuankun Wang, Hui-Ru Ho, **Yuheng Wu**, Yuhang Zhao, Bilge Mutlu

[1] TreeVisual: Design and Evaluation of a Web-Based Visualization Tool for Teaching and Learning Tree Visualization — ASEE 2022 [DOI]
Brendan J. O'Handley, **Yuheng Wu**, Haobin Duan, Chaoli Wang

DOCTORAL SYMPOSIUM

Scene-Aware Augmented Reality for People with Low Vision in Complex, Dynamic Environments — UIST Adjunct 2026
Yuheng Wu

WORKSHOP PAPERS

Preliminary Exploration on Intelligent Social Cues in Wearable Augmented Reality to Support Passing Encounters for People with Low Vision CHI 2026
Ben Kosa, Kexin Zhang, **Yuheng Wu**, Sanbrita Mondal, Yuhang Zhao
Workshop on Next Steps for Augmented Reality On-the-Move: Challenges & Opportunities

PROFESSIONAL EXPERIENCE

Niantic Spatial Inc., Research and Development Intern **London, UK**
Advisors: Dr. Jessica Van Brummelen, Prof. Gabriel Brostow May 2025 – Sep. 2025

- Co-led NaviNote, a voice-first AR system helping blind and low-vision users explore public spaces, author and share in-situ spatial notes, and navigate the last few meters, built with Niantic VPS/Lightship for sub-meter localization.
- Restructured the agentic multimodal LLM backend, an orchestrator routing requests to six tool-calling agents (visual and web search, scene understanding, maps, annotation, customization), which correctly answered 82.8% of 605 live queries. Analyzed the errors and traced a major source to orchestrator-agent miscommunication.
- Took NaviNote from concept to an 18-participant evaluation in four months (participants recruited through a UK sight-loss charity), significantly improving last-few-meters navigation success (14/16 vs. 6/16) over TapTapSee, a commercial AI camera app. CHI 2026 Honorable Mention, U.S. patent pending.

Microsoft Research Asia, Research Intern **Beijing, China**
Advisor: Dr. Yun Wang Jan. 2022 – Jun. 2022

- Built WonderFlow (published in IEEE TVCG 2024), a web-based authoring tool for data-visualization animation (Vue.js, D3.js, GSAP), including an SVG element parser, timeline control, and 40 composable animations.

RESEARCH PROJECTS

SceneGlance: AI-Powered Wearable AR for Multi-Object Scene Perception Oct. 2023 – Apr. 2026
ASSETS 2026 (first author)

- Led development of a real-time AR system on Microsoft HoloLens 2 that recognizes multiple objects in cluttered egocentric video and highlights them by importance to support complex scene perception for people with low vision.
- Built two ~20K-image importance-labeled datasets by relabeling public data, fine-tuned RTMDet instance-segmentation models to over 2× the pretrained model's mAP, and engineered a streaming pipeline between headset and edge server running at 28.6 FPS with 176 ms latency.
- Evaluated SceneGlance in indoor and outdoor studies with 12 and 13 low-vision participants, showing in the indoor study that importance-based highlighting effectively shifted attention to key objects at the cost of overall scene recall, and identifying design challenges for AR in cluttered, dynamic scenes.

NavSight: Field Deployment of an On-Device Mobile AR Navigation System Feb. 2025 – Aug. 2026
IMWUT 2026 (first author)

- Led development of an iOS AR app (Unity + Core ML) helping low-vision users navigate outdoors through six accessible visual enhancements, with a fine-tuned YOLO11-seg model recognizing 21 object categories across four navigation challenges, fully on-device at 25.6 FPS with 70 ms latency on iPhone 13.
- Shipped to the Apple App Store and ran a longitudinal diary study with 12 participants using it unsupervised in their own environments for 5–12 days. Captured 202 usage sessions (31 h) alongside 87 daily survey entries and 12 exit interviews on subjective experience to understand why users adopted, customized, and abandoned features.
- Characterized real-world model failures, including environmental conditions degrading model performance, asymmetry between error types (false positives vs. false negatives), users' sense-making around errors, and overlap between model and user vision failures. Derived implications for evaluating and deploying real-world AI.

Real-Time AR System for Navigating Surface-Level Changes Sep. 2025 – Present
Manuscript in preparation

- Leading development of a wearable AR system (Quest 3, depth camera, phone) that helps people navigate surface-level changes, fusing depth with headset IMU to localize step edges and landings in real time.
- Built a C++ step detector running on the phone (79 FPS, 6–11 mm median step-height error), and integrated SAM 3 and Depth Anything 3 on a remote GPU for long-range detection.
- Collected IMU data from 7 participants (1,076 steps) and built a CNN-based step detector that detects each step the user takes at 0.994 step-level F1 in leave-one-participant-out testing.
- Designing and evaluating visual and audio feedback for safe navigation of surface-level changes across different stages of traversal.

SERVICE AND TEACHING

Co-organizer: Workshop on Vision Foundation Models and Generative AI for Accessibility: Challenges and Opportunities, ICCV 2025
Reviewer: CHI 2027; UIST 2026; CHI 2026; ICCV 2025; ISMAR 2024; CSCW 2024
Mentoring: Two undergraduates, Weibing Wang and Meng Fong Lio (now at Microsoft), co-authors on ASSETS 2026.
Teaching, University of Wisconsin–Madison

Teaching Assistant, CS320: Data Science Programming II	Fall 2025, Spring 2026
Head Teaching Assistant, CS220: Data Science Programming I	Spring 2024
Teaching Assistant, CS577: Introduction to Algorithms	Spring 2023
Teaching Assistant, CS220: Data Science Programming I	Fall 2022, Fall 2023, Summer 2024

SKILLS

LLMs: Multi-Agent Orchestration, Tool Use, Multimodal LLMs & VLMs, Agent Evaluation, Human-Centered Evaluation
Computer Vision: PyTorch, Object Detection, Instance Segmentation, Fine-Tuning, Core ML, On-Device Inference
AR / XR: Unity, MRTK, Meta XR, Niantic VPS/Lightship, ARKit, Spatial Computing, Sensor Fusion, Real-Time Streaming
Programming: Python, C#, C/C++, Swift, JavaScript, SQL, React, Vue, Git