Research Statement

Tengxiang Zhang

Digital content is breaking out of the screens into the physical world with the help of augmented reality and IoT technologies. How do we transition into the new computing paradigm and interact with digital content and physical objects in a natural, consistent, and spontaneous way? I build stretchable skin sensors and smart wearables (*e.g.*, glasses, rings) to monitor human states and behaviors, develop wireless IoT sensing systems (*e.g.*, millimeter wave tags) to identify and understand the environment, then design interaction interfaces powered by AI agents to merge the tagged physical world, the digital metaverse, and the humans. Finally, I apply the research outputs in areas like health and education. Thus, my previous work falls into three categories: **Wearable Sensors, IoT Tags and Localization Systems**, and **AI-agent-based Interaction Interfaces** (Figure 1). Such efforts have led to many publications at top venues, including SIGCHI and IMWUT, with two paper awards.

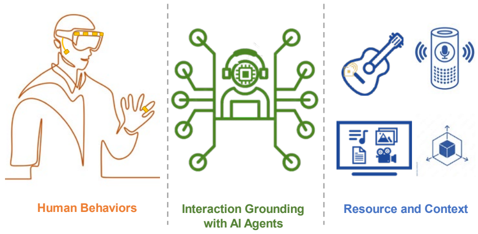

Figure 1: Research methodology: design wearables to sense human behaviors, develop IoT tags to sense the environment and context, use AI agents to associate the two for user understanding.

1. Head, Wrist and Finger Mounted Wearable Devices

I focus on smart rings and smart glasses to sense both explicit and implicit interaction intent. Our smart glasses prototype (Figure 2 left) deploys electrooculography (EOG) electrodes on the nose pads and between the eyebrows to sense upper facial activity. It also places two inertial measurement units (IMUs) close to the skin below the ears to sense lower facial movement. The project was demonstrated at the MobileHCI 2022 Student Design Competition and later invited for video demonstration at the Huawei Developer Conference 2022. I have also designed on-skin sensors that can monitor the superficial vascular activities for long-term cognitive load detection (Figure 2 right). Such sensors attach to the temporal area of the user's face and uses infrared sensors to monitor blood component changes.

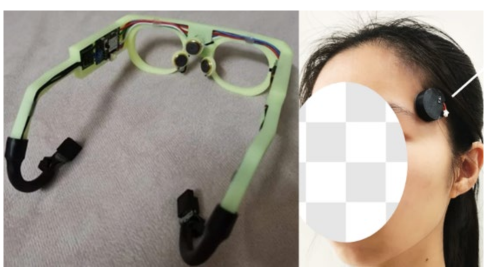

Figure 2: Smart glasses support novel interaction via EOG and IMU sensing (left); a long-term vascular activity sensor for cognitive load detection(right).

Finger wearables can support accessible and subtle gesture inputs. However, finger wearables are challenging to design due to the inherent size constraints which limit the available computing, I/O, and power resources. To overcome these limitations, I designed a modular smart ring (Figure 3). Instead of integrating all functionality into a single piece of hardware, the novel design of **ModularRing** [1] separates the I/O module from the wireless MCU and battery. Users can then switch the module and combine multiple rings with different I/O modules to support the desired interaction.

Figure 3: ModularRing adopts a modular design for smart rings.

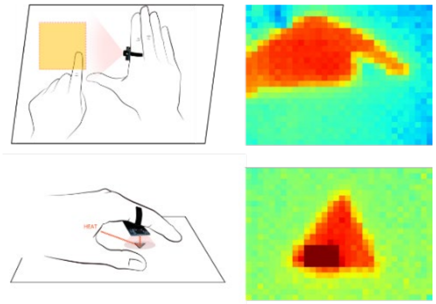

Figure 4: Thermal imaging of a hand and of a passive tag (which reflects the hand's thermal emissions)..

In **ThermalRing** [2], I use a low-power, low-resolution thermal camera to support gestural and tag-based input. ThermalRing analyzes the heat silhouette of the hand to recognize drawing gestures on flat surfaces (Figure 4). A bag-of-words model is trained to recognize gestures based on movement distance and the angle of the hand. I also proposed ThermalTags, passive tags that can be easily made using materials with high heat reflectivity, such as copper tape. When covered by a hand, ThermalTags reflect the heat radiated from the hand and thus can be imaged. This demonstrates how the on-body wearables can work together with off-body tags, enabling more interaction possibilities.

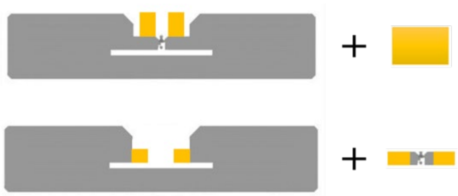

Figure 5: BitID tags work by selectively shorting and separating the IC with the antenna (top and bottom, respectively).

2. Ultra-low-power IoT Tags and Localization Systems

The digital ID and spatial coordinates of physical objects are essential for a consistent interaction experience in the merged world. Wireless tags can assign ID and enable localization of non-smart objects, while wireless signals can augment the human senses since they can 'see' through and further. My initial attempt involves modifying UHF RFID tags so that they can broadcast the ID and sense the tagged object at the same time [3]. Such maintenance-free backscatter wireless tags offload many hardware components (e.g., power source, RF oscillator) to resource-abundant edge devices. This minimizes communication power consumption and makes it possible to power such tags with harvested energy. The tag's readability is modulated by externally switching the antenna matching impedance (Figure 5). Users can easily make and deploy BitID sensors to detect the binary states of various objects. Codes for a complete BitID system, including sensor registration, definition, and event recognition, are open-sourced [5].

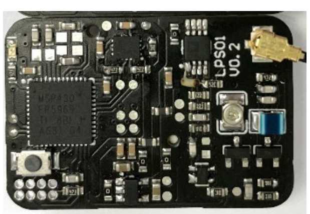

Figure 6: Bluetooth-compatible backscatter sensing tag (33x22mm, 800uA).

To reduce transceiver size and leverage existing computing devices, I built a Bluetooth-compatible backscatter tag (Figure 6) that broadcasts the IMU data as Bluetooth advertisements, which commercial devices like smartphones and laptops can receive. The average broadcasting current is only 800uA, a four-fold improvement on commercial Bluetooth chips (3.5mA). **BLEselect** [4] (Figure 7) localizes such Bluetooth backscatter tags with a head-mounted receiver. A compact antenna array is fitted on a pair of smart glasses to estimate the angle of arrival (AoA) of advertising signals from wrist-worn devices and Bluetooth tags. A sensing pipeline supports three natural selection gestures—nodding, pointing, and encircling—with lightweight machine-learning models trained in real-time for each performed gesture. Extensive evaluations show our system is accurate, low-power, and privacy-preserving despite the small-size antenna array.

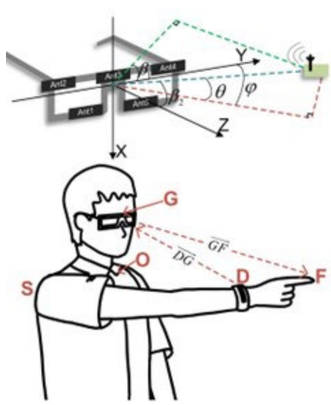

Figure 7: BLEselect deploys an antenna array on a pair of smart glasses to detect AoA from wearables and tag.

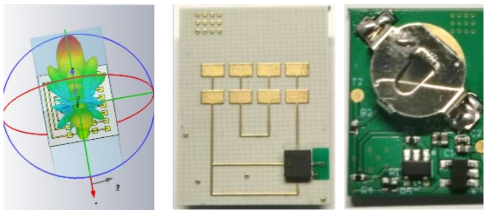

Figure 8: The full-wave simulation result of the mmWave tag (left); Front (middle) and back (right) of the mmWave tag.

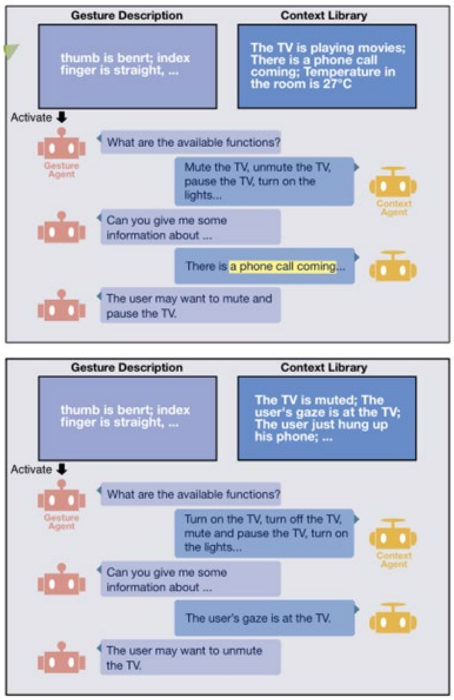

Figure 9: The two-agent framework of GestureGPT. The user mutes the TV to answer a phone call (top), then unmutes the TV when the call ended (bottom).

To further improve the localization accuracy, I turned to millimeter wave FMCW radar (Figure 8). I design 60Ghz backscatter tags based on Van Atta array, which retroreflects signal back to the radar. The tag only consumes 940uW power. A head-worn 60Ghz FMCW radar localizes the tags with a 3D error distance of 3.26cm. I have also developed power[5][6] and interaction[7][8] techniques for such tags.

3. Novel Interfaces with AI Agents

Current interaction systems usually stops at successful recognition of user interaction behaviors (*e.g.* the category of the gesture). However, the system is yet to figure out how to meet the user's need. For example, a right swipe can mean "next song" or "next page" when a user is reading online. I develop a new interaction framework that grounds the interaction behavior with the system function in a zero-shot manner. As an example, in GestureGPT[9], I designed a novel zero-shot gesture understanding and grounding framework leveraging large language models (LLMs). Gesture descriptions are formulated based on hand landmark coordinates from gesture videos and fed into our dual-agent dialogue system. A gesture agent deciphers these descriptions and queries about the interaction context (e.g., interface, history, gaze data), which a context agent organizes and provides. Following iterative exchanges, the gesture agent discerns user intent, grounding it to an interactive function. Such agent-based interfaces shed light on the paradigm shift of future interaction interface design.

Future Research Directions

I am eager to collaborate with colleagues with various backgrounds to apply my interdisciplinary knowledge to new application domains, including but not limited to education, healthcare, and accessibility. Topics that I am actively working towards include:

1. Ultra-thin Skin Sensors for Long-term Monitoring

The sensing positions of wearable devices are limited. Taking advantage of my flexible electronics fabrication skills, I develop ultra-thin skin sensors that can be placed at most places on-body for skin characteristic monitoring. This enables multi-spot long-term physiological and psychological signal monitoring, which leads to a new set of possibilities for health (*e.g.* cough detection) and education (*e.g.* cognitive load detection) applications.

2. Multi-agent-based User Understanding and Interaction Grounding

I plan to expand the two-agent model of GestureGPT to include more interaction modalities like facial expression and speech. I will explore how to moderate the dialogs between agents for efficient user understanding and decision making.

REFERENCES

1. *Yuanchun Shi, Yinshuai Zhang, and Tengxiang Zhang. 2018. CN Patent 108926081.A Smart Ring and Its Wearing Method.* ***Codes****: https://github.com/saintnever/ModularRing*

2. *Tengxiang Zhang, Xin Zeng, Yinshuai Zhang, Ke Sun, Yuntao Wang, and Yiqiang Chen. 2020. ThermalRing: Gesture and Tag Inputs Enabled by a Thermal Imaging Smart Ring. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20), 13.* ***Codes****: https://github.com/saintnever/thermalring*

3. *Tengxiang Zhang, Nicholas Becker, Yuntao Wang, Yuan Zhou, and Yuanchun Shi. 2017. BitID: Easily Add Battery-Free Wireless Sensors to Everyday Objects. In 2017 IEEE International Conference on Smart Computing (SMARTCOMP), 1–8.* ***Codes****: https://github.com/AlexFxw/BitID*

4. *Tengxiang Zhang, Zitong Lan, Chenren Xu, Yanrong Li, and Yiqiang Chen. 2022.BLEselect: Gestural IoT Device Selection via BLuetooth Angle of Arrival Estimation from Smart Glasses. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 6, 4: 198:1–198:20.*

5. *Yuntao Wang, Jianyu Zhou, Hanchuan Li, Tengxiang Zhang, Minxuan Gao, Zhuolin Cheng, Chun Yu, Shwetak Patel, and Yuanchun Shi. 2019. FlexTouch: Enabling Large-Scale Interaction Sensing Beyond Touchscreens Using Flexible and Conductive Materials. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 3, 3: 109:1–109:20.*

6. *Tengxiang Zhang, Xin Yi, Chun Yu, Yuntao Wang, Nicholas Becker, and Yuanchun Shi. 2017. TouchPower: Interaction-based Power Transfer for Power-as-needed Devices. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 1, 3: 121:1–121:20.*

7. *Tengxiang Zhang, Xin Yi, Ruolin Wang, Yuntao Wang, Chun Yu, Yiqin Lu, and Yuanchun Shi. 2018. Tap-to-Pair: Associating Wireless Devices with Synchronous Tapping. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 4: 201:1–201:21.* ***Codes****: https://github.com/starsjf/Tap2pair*

8. *Tengxiang Zhang, Xin Yi, Ruolin Wang, Jiayuan Gao, Yuntao Wang, Chun Yu, Simin Li, and Yuanchun Shi. 2019. Facilitating Temporal Synchronous Target Selection through User Behavior Modeling. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 3, 4: 1–24.*

9. *Xin Zeng, Xiaoyu Wang, Tengxiang Zhang, Chun Yu, Shengdong Zhao, and Yiqiang Chen. 2023. GestureGPT: Zero-shot Interactive Gesture Understanding and Grounding with Large Language Model Agents. http://arxiv.org/abs/2310.12821*