Research Statement

Tara Boroushaki

My research interests lie in sensing and mobile technologies with applications in wireless networking, cyber-physical systems (including robotics), and human-computer interaction (HCI). I develop algorithms and build systems for multi-modal sensing to connect, perceive, and interact with the environment in novel ways to enable more efficient, more robust, and more capable mobile, cyber-physical, and cyber-human systems.

Mobile and wireless sensing technologies play a critical role in Internet of Things, Augmented Reality (AR), robotics, environmental monitoring, and human health monitoring. Despite their ubiquity, each of todays mobile sensing technologies is limited in a fundamental way: cameras can capture high-resolution images but are limited to the line of sight and cannot sense behind occlusions; wireless sensing (e.g., using WiFi or Bluetooth) can traverse occlusions but has limited sensing resolution; inertial sensors can track with high accuracy and speed but suffer from spatiotemporal drift.

My research introduces innovative algorithms, learning models, and systems to fuse different sensing modalities, particularly radio frequency (RF) and cameras. The goal of my work is to unlock unprecedented capabilities for IoT-connected and mobile systems, such as AR headsets and robots, and to enable new perception, interaction, and manipulation tasks. To realize this goal, my approach leverages and develops advanced signal processing techniques and mathematical modeling to combine different sensing modalities. In building my systems, I not only use available sensors (such as RGBD cameras and consumer RF), but I also develop next-generation sensing technologies (such as millimeter-waves) in settings where such technologies are needed to deliver transformative capabilities.

Contributions and Impact

My research has appeared in top venues of sensor networking [SenSys'21 Best Paper Finalist, NSDI'23, SIGCOMM'23, IEEE RFID'23 Best Paper] and robotics [ICRA'21, RSS'22, Autonomous Robots'23].

1. **Multi-Modal localization:** Most prior work in RF localization is limited to a single sensing modality, namely RF signals. In contrast, my research introduces novel multimodal localization primitives. I developed geometric optimization methods to fuse time-of-flight RF measurements with RGBD point clouds for accurate and efficient item localization. My research has also extended classical RF localization models like Synthetic Aperture Radar (SAR) with visual inertial odometry (VIO) to deliver high-accuracy item localization on AR headsets. The success of these techniques has enabled my systems to deliver first-of-their-kind capabilities (e.g., centimeter-scale RF localization with AR headsets) as well as two-fold efficiency improvements over past robotics-based localization methods.
2. **Non-line-of-sight perception:** My work uses RF sensing to extend the perception of robots and AR headsets beyond their line of sight. To go from localization to perception, I developed new RF-visual uncertainty methods, including Gaussian belief models, probabilistic occupancy distributions, and expected information gain. Using these models, I built systems that can reason about the environment to extend classical vision tasks (such as object identification, orientation estimation, and picking verification) to fully-occluded settings, enabling cyber-physical and cyber-human systems to operate efficiently in these environments.
3. **Mobile manipulation:** My research introduces first-of-their-kind RF-visual primitives for robotics. These include optimization methods for RF-visual servoing, path planning (including obstacle avoidance), mechanical search, and grasping. They also include novel RF-visual reinforcement learning (RL) networks that enable robots to efficiently perceive and manipulate their environment. These techniques have significantly boosted the success rate of robotics problems like mechanical search (from 75-85% to 90-95%), led to more robust robotic grasping solutions due to multi-modal sensor fusion, and made these tasks more efficient even in unoccluded settings.

My research has been recognized by the Microsoft Research PhD fellowship, the Best Paper Award in IEEE RFID, and the Best Paper Finalist in ACM SenSys. It was chosen as one of the Top 10 AI trends to watch for by *The Wall Street Journal* and as one of "103 ways MIT is making the world better" by MIT. My research has been covered by multiple news outlets including the BBC, the Boston Globe, and the World Economic Forum. I have built live real-time demos of my systems and demonstrated them to over 100 companies. My research has led to multiple patents that have been licensed to a start-up, Cartesian Systems, which has already deployed them to solve problems in retail and supply chain.

Next-Generation Mobile Perception

With the success of smartphones and wearables, AR headsets are emerging as the most exciting next generation of mobile devices. A key challenge with these headsets is that they rely on vision to perceive their environment, which limits their performance in highly cluttered, low-lit, and/or occluded settings.

X-AR [NSDI'23, SIGCOMM'23]: My work introduces X-AR, the first system that extends the perception of AR headsets to non-line-of-sight settings through RF sensing. Bringing RF sensing to AR headsets is challenging because it requires operating within the computational, form-factor, and HCI constraints of these mobile headsets. Moreover, it requires designs that do not hinder or impede any of the existing headset sensors or the users productivity. To overcome these limitations, I introduced multiple contributions. I developed an RF-visual synthetic aperture radar (SAR) localization approach: this technique makes connections between the classical SAR formulation and visual-inertial odometry to enable item-localization while leveraging natural human mobility. I also developed reverse-SAR (RSAR)[1], a new method that combines the hand-tracking technology of AR headsets with the measured RF channel of an RF tag to perform reverse localization of the headset; this enabled me to deliver new tasks such as picking verification (i.e., identifying that a user has picked up the correct RF-tagged item). I built an end-to-end prototype of X-AR with a Microsoft Hololens 2 AR headset, a custom-designed conformal wideband antenna, and BladeRFs software radios. Real-world evaluation of X-AR demonstrated that it can locate hidden items with an accuracy of 9.8 cm and perform item verification with 95% accuracy even when the item is fully occluded inside boxes. A demo video of my system is shown here: youtube.com/watch?v=bdUN21ft7G0

Figure 1: X-AR enables the user to find and see occluded objects.

RF-AR [IEEE RFID'23 Best Paper]: Motivated by the success of X-AR, my next question was: can I optimize the end user tasks (finding, navigation, retrieval) by rethinking the end-to-end system design? My RF-AR system exploits synergies between RF and AR to answer this question. Its techniques use RF properties to deliver new interfaces for users to complete their tasks quicker. I devised a path optimization algorithm that reduces the Dilution of Precision (DoP) while enhancing the signal-to-noise ratio (SNR) in RF measurements to enable faster convergence of localization. I also developed a dynamic RF-based user interface (UI) that displays a hologram which adapts based on the localization certainty to give users intuitive feedback on whether they are getting closer to or further away from their target. Building on these techniques, I explored different cues to bias the users trajectories toward the optimum path, including color coding, holograms, and arrows on the AR headset UI. I conducted a user study with 20 participants. My evaluation showed that these RF-based optimization techniques and human-in-the-loop design enable users to achieve 8.6 cm median localization accuracy within 76 seconds, which is 55% faster than my earlier system (X-AR).

Figure 2: AR user's view when item is found

Research Significance: From a technical perspective, this research introduces RF sensing as a new dimension in the design of next-generation AR headsets from both algorithmic and system perspectives. Practically, it opens the door for next-generation AR systems that enable more advanced, seamless, and efficient end-user tasks, particularly in complex or cluttered environments, including warehouses, retail, manufacturing, etc.

Networked Sensing for Robotics & Cyber-Physical Systems

The past few years have seen a rapid growth in battery-free networked sensors, such as RFIDs, which are attached to over 50 billion items today. My second line of work shows how we can use these existing networks to enable robotic tasks such that were not feasible before, with important applications to commercial, industrial, and smart home environments.

RF-Grasp [ICRA'21]: Most robotic systems rely on computer vision to localize and retrieve items, but vision is inherently limited to the line of sight. Thus robots cannot efficiently find (or retrieve) fully occluded

---

[1] Note that RSAR is different from the well-known inverse SAR which does not have mechanisms to locate the moving target a priori (vs our approach which exploits hand-tracking).

items. My work introduces the first robotic system that can accurately locate a fully occluded item and guide a robotic arm to efficiently retrieve the target while avoiding obstacles. The system, RF-Grasp, uses RF localization to pinpoint the location of RFID-tagged items. However, relying solely on RF localization poses two challenges. First, it can localize items but not recover their shape, which is necessary for object manipulation. Second, because obstacles (such as boxes) are transparent to RF sensing, the robot would collide with these obstacles if it simply reaches to grasp an occluded target upon localization. My work introduces an RF-visual servoing primitive that performs RF-guided active exploration and trajectory optimization by incorporating RF information into Gaussian Belief Space Planning. I also developed an RF-visual RL grasping network that outputs optimal grasping affordances, by folding an RF mask as an attention mechanism into the cameras RGBD information. I implemented RF-Grasp using a UR5 robotic arm, wrist-mounted camera, three Software Defined Radios and antennas positioned in its environment. My evaluation showed that RF-Grasp improves the exploration and retrieval success rate and efficiency by up to 40-50% over a vision-based state-of-the-art baseline. You can watch a demo video at youtu.be/ZAzeYPcTM78?si=85WsH7ca03PAaPYr

RFusion [SenSys'21 Best Paper Finalist]: This system takes RF-visual robotic manipulation to another level by co-designing the RF localization with the retrieval task, and integrating RF sensing with the robotic arm. In particular, the retrieval problem suffers from a trade-off between localization accuracy and retrieval efficiency. If the robot moves its antenna on a longer trajectory, it will observe the RF target from more diverse perspectives which improves its DoP (and thus the localization accuracy), but reduces the efficiency (due to longer travel distance). I formulated this as an optimization problem that aims to minimize the overall trajectory for successful target retrieval (which requires accurate localization), and devised the first RF-Visual RL framework for robotic manipulation to solve this problem. The network encodes RF-visual features by geometrically fusing RF-based time of flight estimation within a point cloud based model of unobserved regions. These features along with the DoP from RF measurements are used to represent the RF-visual uncertainty of the target items location for the network. I built RFusion by mounting log-periodic antennas and a camera on the wrist of a UR5 robot, and trained the RL network in simulation. My evaluation demonstrated that RFusion locates fully hidden objects with a median error of 1.1 cm, successfully retrieves the target item with 96% success rate, and adapts to unseen new environments. A demo video is at youtu.be/iqehzw$_aLc0?si = JqGX_YGpkL2I9CG4$

FuseBot [RSS'22, Autonomous Robots'23]: FuseBot extends the benefits of RF perception even to untagged objects. Rather than requiring all items to be tagged, FuseBot operates in typical scenarios where only a subset of items are tagged and efficiently retrieves fully occluded non-tagged target items. To do this, it introduces an RF-visual probabilistic occupancy distribution technique that can predict the positions of fully occluded non-tagged items. This technique models the physical properties of the target items with Gaussian kernels and integrates them into an RF-visual uncertainty map of the environment. The system also introduces a RF-visual extraction policy that chooses optimal next-best-grasp on each visually segmented object on the surface of the pile while taking into account the expected information gain, the expected grasp success, and the occupancy distribution of the target. I evaluated FuseBot over 180 real-world trials and it succeeded in 95% of trials and was 40% more efficient compared to a vision-based state-of-the-art baseline. You can watch a demonstration of the system at youtu.be/TFqz263uPN0?si=Gq1UQLxmTlwFnKs6

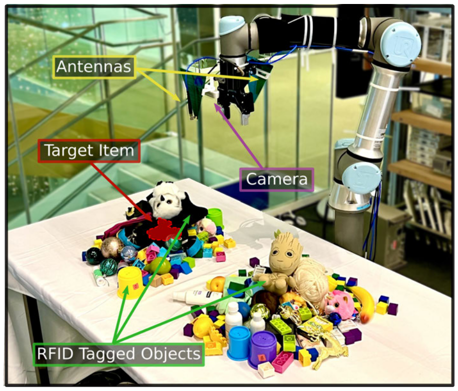

Figure 3: The robotic arm finds fully occluded RFID tagged or non-tagged items.

Research Significance: This research offers a fundamentally new solution for mechanical search, a classical robotics problem, and my solutions are more efficient and robust than past state-of-the-art systems. In addition, it expands the benefits of RF localization. Specifically, in contrast to past work that brings benefits only for RF-tagged objects, my research demonstrates that the mere existence of a single RF-tagged item can significantly boost the (RF-visual) localization accuracy of untagged objects. Finally, this work led to multiple patents that have been licensed and used commercially for indoor mapping and localization.

Future Work

In the future, I plan to continue pushing the boundaries of computational and networked mobile sensing by building systems and developing signal processing and ML models for multi-modal sensing. I will do so via emerging technologies such as mmWave and THz radars, as well as through ubiquitous sensing mechanisms including RF, acoustics, vision, and inertial sensors.

mmWave Perception for next-generation HCI: I am interested in making the leap from non-line-of-sight perception of RF-tagged objects to tagless settings. My FuseBot system took a first step toward this vision. The next step will require bringing the contactless and tagless RF sensing to AR headsets, e.g., using millimeter-wave (mmWave) radars. While mmWave offers higher resolution than sub-GHz RF signals (like the one I have used in X-AR and RF-AR), their sensitivity also makes them susceptible to the slightest movement and head tilts. I plan to develop multi-modal sensor fusion algorithms that combine self-tracking with mmWave sensing to decouple extraneous movements from target tasks and build end-to-end systems to deliver new capabilities for AR/VR headsets, such as through-occlusion imaging, vital sign monitoring, and RF-based affect recognition.

Neural Radiance Fields (NeRFs) & Gaussian Splatting for Wireless Sensing: Recent advancements in NeRFs (Neural Radiance Fields) and Gaussian splatting in computer vision have significantly enhanced 3D modeling and rendering in both resolution and efficiency. NeRFs enable detailed 3D reconstruction of objects and scenes from limited 2D image data. This stems from their ability to predict how light interacts with and reflects off the objects. I am interested in bridging NeRFs with the field of wireless sensing, especially mmWave and THz imaging. To do so, I intend to develop imaging techniques that can utilize mmWave and THz measurements (including phase information, which is not traditionally available in vision) collected from a few vantage points with limited aperture sizes to produce high-resolution 3D models of the environment and objects, even where fully occluded. This future generation of mmWave and THz sensing will enable robots and cyber-physical systems to perceive occluded items with higher efficiency and in more realistic scenarios since it will require collecting fewer measurements and within smaller apertures. Moreover, existing mmWave radars on cars are still limited to recognizing several large objects categories (cars, pedestrians, bikes). This line of work will enable autonomous vehicles to create 3D models of the environment based on a few mmWave radars on the car, allowing them to perceive the environment regardless of weather and lighting conditions, and improve their spatial awareness.

Joint Sensing and Networking: With the future adoption of AR/VR technologies and their need for higher bandwidth communication, they will be equipped with 6G communication technologies. I am interested in enabling simultaneous high-speed wireless communication and sensing using mmWave and THz frequencies. My future work will explore new algorithms and protocols that can seamlessly switch between communication and sensing modes or share spectrum.

Privacy and Security: The advancement in mmWave imaging and sensing on mobile devices, particularly in AR/VR headsets, brings with it privacy and security concerns. I aim to develop algorithms and systems to detect and safeguard against unauthorized mmWave imaging and sensing of private spaces or personal belongings. One potential approach is to identify the transmission of unauthorized waveforms, and employ countermeasures such as jamming. This is a first line of defense for certain types of waveforms (e.g., FMCW) but is challenging for wireless transmissions used for joint communication and sensing. In such cases, jamming could inadvertently impair legitimate communication channels. I envision that future counter-measures could be built into the devices themselves (similar to how we cannot use WiFi to simply jam the medium), for example, by integrating the human user into the security framework. One approach is to require the users of AR headsets to demonstrate physical access to specific areas if they wish to image or sense with RF and mmWave signals. This approach not only adds a layer of security but also ensures that the sensing technology is used within authorized boundaries, potentially reducing the risk of misuse.

Multi-modal sensing for Climate & Ecosystem Monitoring: I am interested in applying my skills in multi-modal sensor fusion for climate and sustainability applications. One area I am interested in is subsea monitoring. Since RF signals die exponentially underwater, I intend to investigate adapting my RF-visual sensing techniques to acoustic-visual sensing. This enhanced sensing approach can support underwater robots in mapping the ocean and extend their perception and manipulation capabilities. Another area of my research interest is RF-visual sensing for satellite-based Earth monitoring. The emergence of nanosatellites offers a new opportunity in Earth observation. However, in contrast to traditional satellites, nanosatellites face limitations in computational power and memory capacity. I am interested in developing efficient multi-modal sensor fusion mechanisms using thousands of distributed nanosats that can unlock new pathways for Earth monitoring and environmental sensing.

More generally, I remain broadly interested in all aspects of wireless and mobile computing systems - from multi-modal sensing and perception to IoT connectivity and edge processing and learning - and their applications to robotics, AR, and other emerging cyber-physical systems.

References

[NSDI'23] **Boroushaki, Tara**, et al. "Augmenting Augmented Reality with Non-Line-of-Sight Perception." 20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23). 2023.

[SIGCOMM'23] **Boroushaki, Tara**, et al. "Real-time X-Ray Vision via Augmented Reality with RF Sensing." Proceedings of the ACM SIGCOMM 2023 Conference. 2023.

[IEEE RFID'23] **Boroushaki, Tara**, et al. "Exploiting Synergies between Augmented Reality and RFIDs for Item Localization and Retrieval." 2023 IEEE International Conference on RFID (RFID). IEEE, 2023, Best Paper Award

[RSS'22] **Boroushaki, Tara**, et al. "Fusebot: Rf-visual mechanical search." Robotics: Science and Systems. Vol. 2022. 2022.

[Autonomous Robots'23] **Boroushaki, Tara**, et al. "FuseBot: mechanical search of rigid and deformable objects via multi-modal perception." Autonomous Robots (2023): 1-18

[SenSys'21] **Boroushaki, Tara**, et al. "Rfusion: Robotic grasping via rf-visual sensing and learning." Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, 2021, Best Paper Finalist

[ICRA'21] **Boroushaki, Tara**, et al. "Robotic grasping of fully-occluded objects using rf perception." 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021.