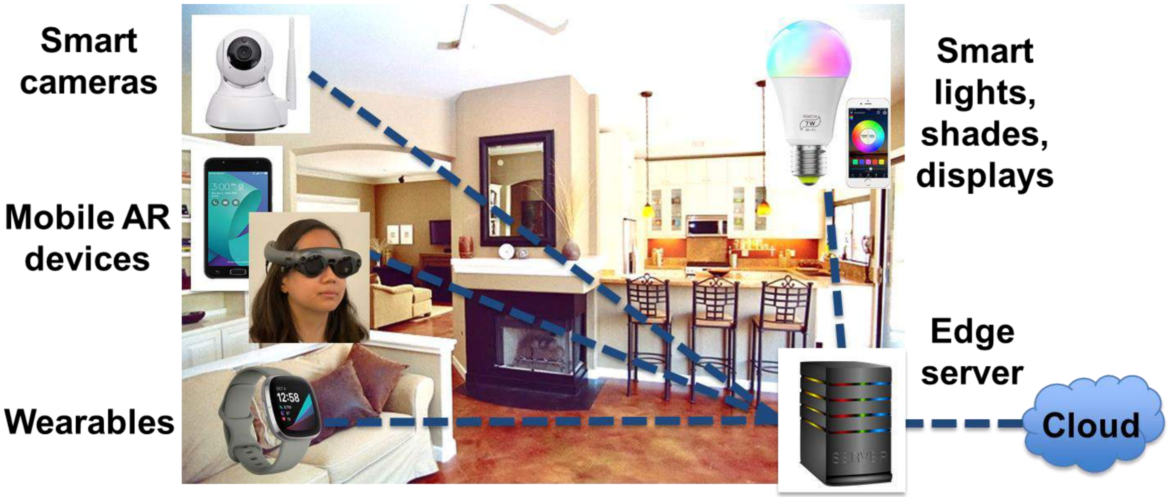
Intelligent Mobile Augmented Reality: Promise, Challenges, and Solutions



Maria Gorlatova

July 1st, 2022

About the Speaker

- Nortel Networks Assistant Professor, ECE/CS, Duke University
- Previously:
  - Associate Research Scholar, Princeton University, Electrical Engineering
  - Ph.D. Columbia University, Electrical Engineering
  - Industry positions:

Telcordia Technologies

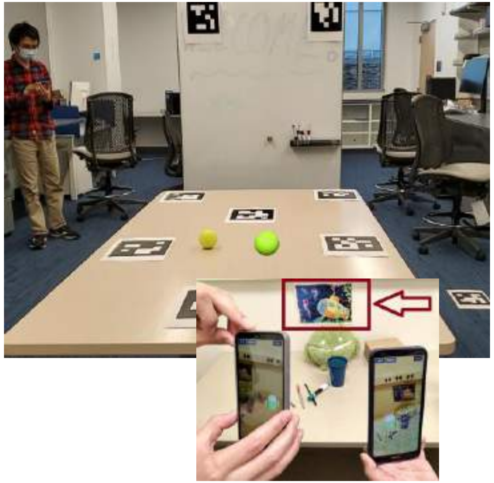
Disney Research

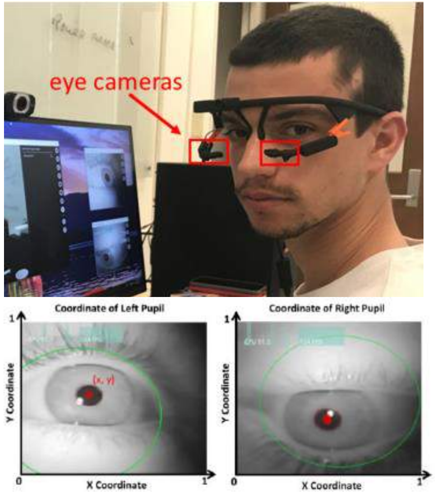
D E Shaw Research

IBM

Duke University Intelligent Interactive Internet of Things (I^3T) Lab

- Core research direction: **reliable context-aware** AR

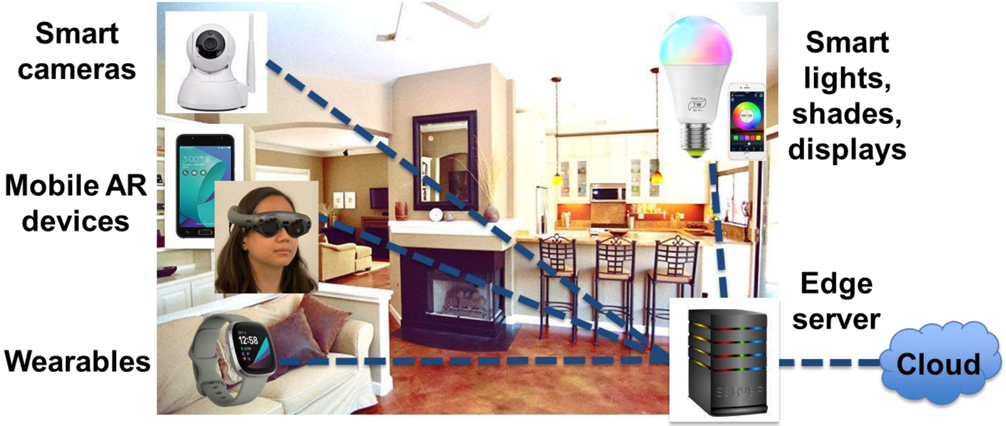




NSF aws DARPA Meta IBM ATHENA

Outline

- **AR: current state and applications**
- Improving AR spatial awareness
- Improving AR semantic awareness
- Improving user context awareness
- Summary and future work



Augmented Reality (AR): A Definition

- The [virtual] content is laid out around a user **in the same spatial coordinates as the physical objects surrounding her/him***

*From: Baldassi et al, Challenges and New Directions in Augmented Reality, Computer Security, and Neuroscience, June 2018.

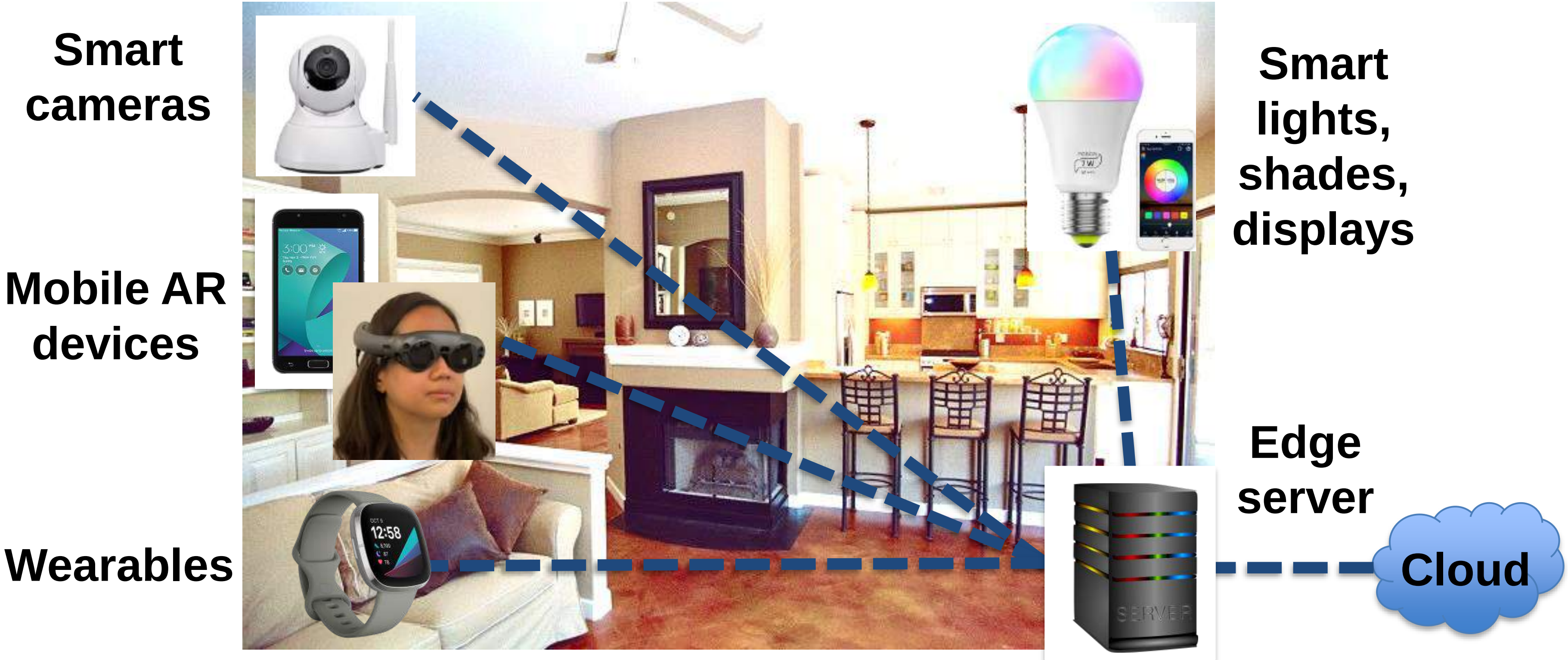
Modern AR: Multiple Device Options

- Google ARCore (2018), Apple ARKit (2017)
  - ➢ Vast majority of modern phone models support it
- Constantly expanding in devices and capabilities

Microsoft HoloLens (2016)

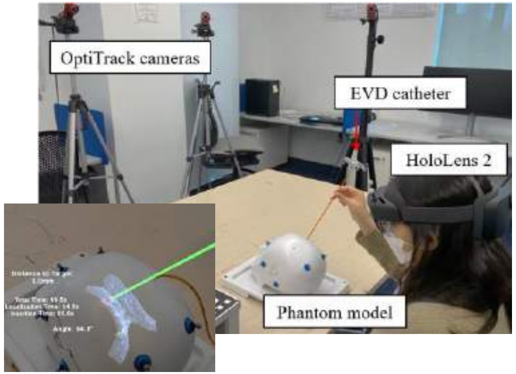
Magic Leap One (2018)

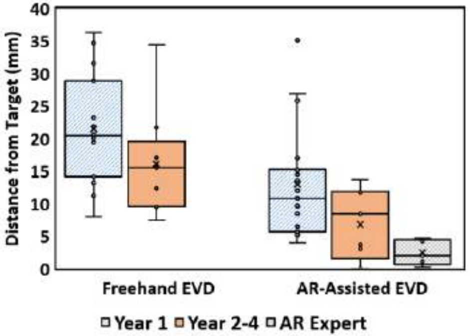
Our Vision: Multi-Device Mobile AR Architectures



Augmented Reality: Applications (1/2)

- **Seeing the invisible**
  - Navigation, education
  - AR-based **task guidance** has been demonstrated to out-perform paper-based and other guidance



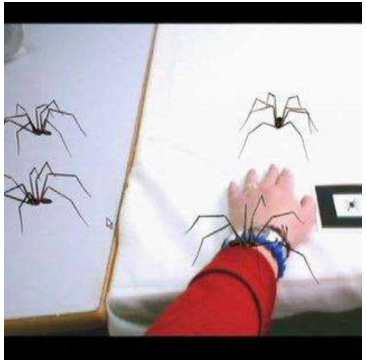


Accuracy improvements in AR-aided neurosurgery

- S. Eom, D. Sykes, S. Rahimpour, **M. Gorlatova**, under review. Work in progress version: AR-Assisted Surgical Guidance System for Ventriculostomy, in Proc. *IEEE XR for Healthcare and Wellbeing Workshop*, Mar. 2022 (co-located with *IEEE VR'22*).

Augmented Reality: Applications (2/2)

- **Influences users deeply and profoundly**
  - ➢ Studied in context of treating phobias
  - ➢ Interesting temperature perception studies
  - ➢ Several projects on AR for substance use disorders



❑ T. Scargill, Y. Chen, S. Eom, J. Dunn, **M. Gorlatova**, Environmental, User, and Social Context-Aware Augmented Reality for Supporting Personal Development and Change, in *Proc. IEEE Workshop for Building the Foundations of the Metaverse*, Mar. 2022 (co-located with *IEEE VR'22*).

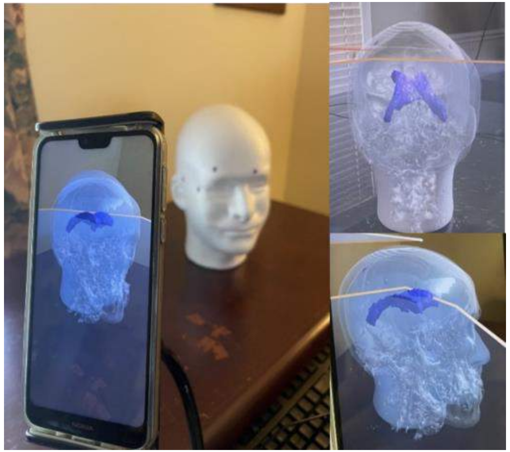
AR: Core Mobile Technology of the Future

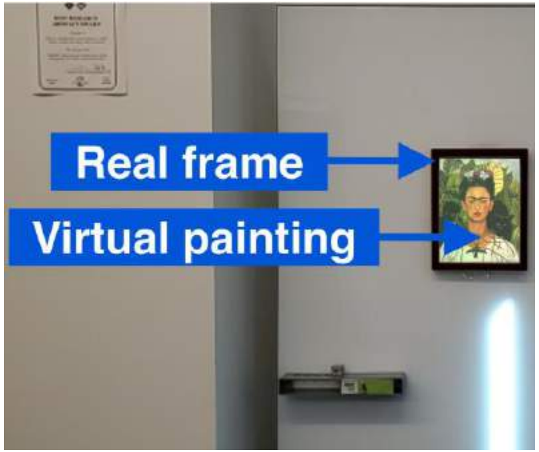
"AR will redefine our relationship with technology"

"It is the next big thing, and will pervade our entire lives"

Mobile AR: Current State and Limitations

Already: impressive in **merging physical and digital worlds**

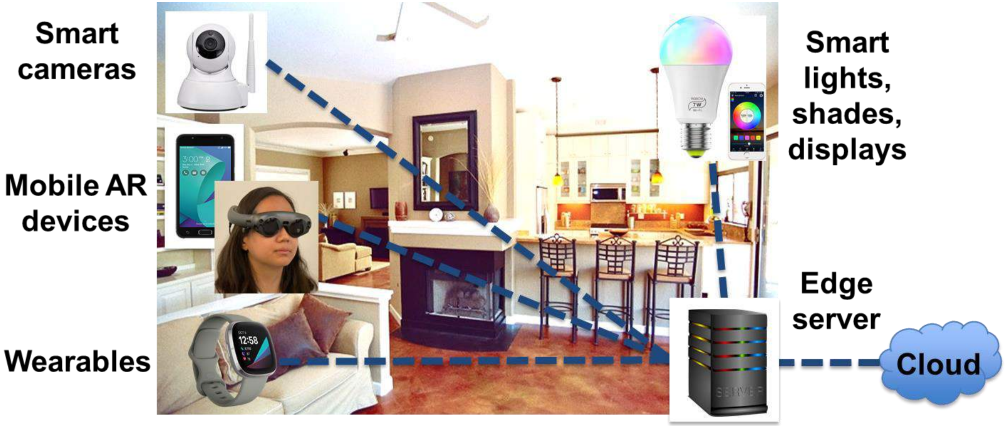




- Among key limitations: resource consumption, headset form factor, security and privacy
- **Context awareness**

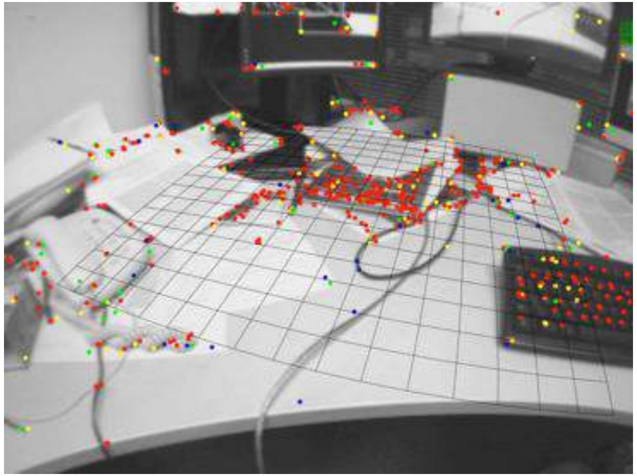
Outline

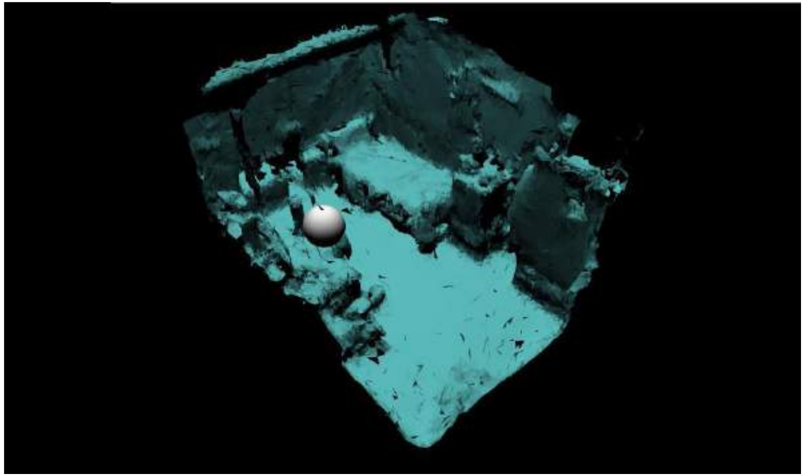
- AR: current state and applications
- **Improving AR spatial awareness**
- Improving AR semantic awareness
- Improving user context awareness
- Summary and future work



AR Spatial Awareness

- Essential for integrating virtual objects with the real world
  - **Sub-cm**-level localization
- Without markers, achieved via **Visual-Inertial Simultaneous Localization and Mapping (VI-SLAM)**





Errors in AR Spatial Awareness

- Virtual objects rendered in the wrong place

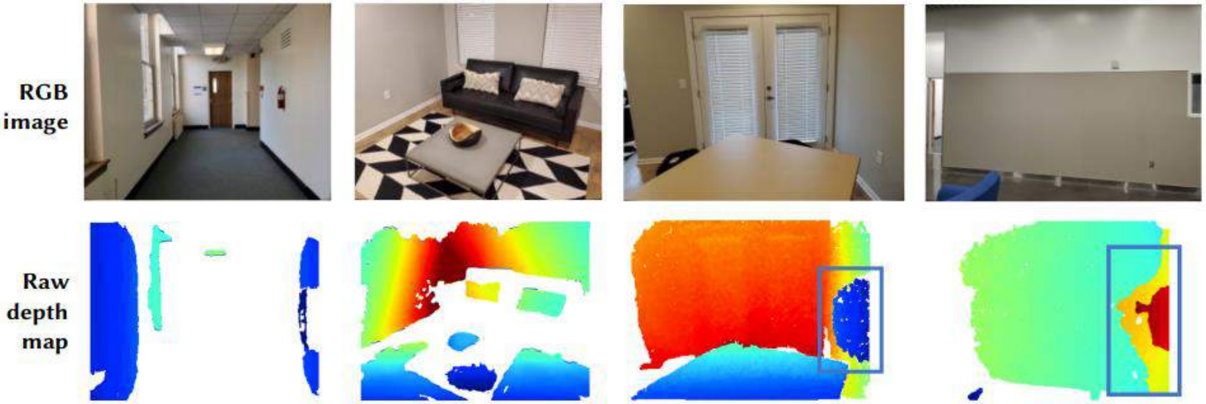
Unintended motion (*drift*)

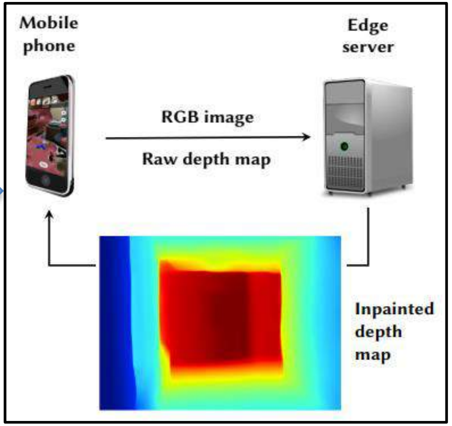
Incorrect scale

- T. Scargill, G. Premsankar, J. Chen, **M. Gorlatova**, Here To Stay: A Quantitative Comparison of Virtual Object Stability in Markerless Mobile AR, in *Proc. IEEE/ACM CPHS Workshop*, 2022 (co-located with CPS-IoT Week).
- Y. Zhang, T. Scargill, A. Vaishnav, G. Premsankar, M. Di Francesco, **M. Gorlatova**, InDepth: Real-time Depth Inpainting for Mobile Augmented Reality, to appear in *Proc. ACM IMWUT*, 2022.

Improving Depth Perception with Depth Inpainting

- Depth data artifacts lead to incorrect mapping of the space



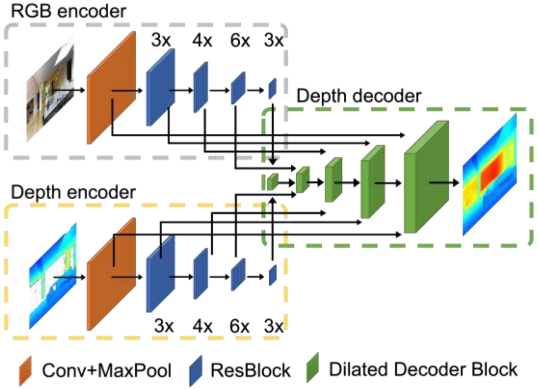


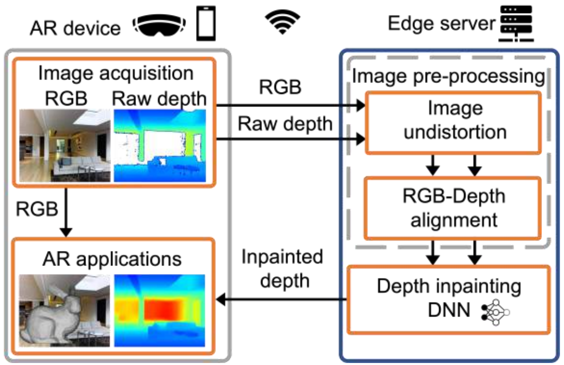
- Our approach:
  - ➢ New DNN and data augmentation for depth data *inpainting* and correction
  - ➢ The only depth inpainting solution suitable for real-time operation in mobile AR

❑ Y. Zhang, T. Scargill, A. Vaishnav, G. Premsankar, M. Di Francesco, **M. Gorlatova**, InDepth: Real-time Depth Inpainting for Mobile Augmented Reality, to appear in *Proc. ACM IMWUT*, 2022.

InDepth: Design and Performance

- Encoder-decoder architecture with two independent encoder branches
  - ➢ Trained on Matterport 3D
  - ➢ Reduces mean average error from 0.34m to 0.29m and inference latency **from 70.2ms to 8.7ms** (edge server execution)

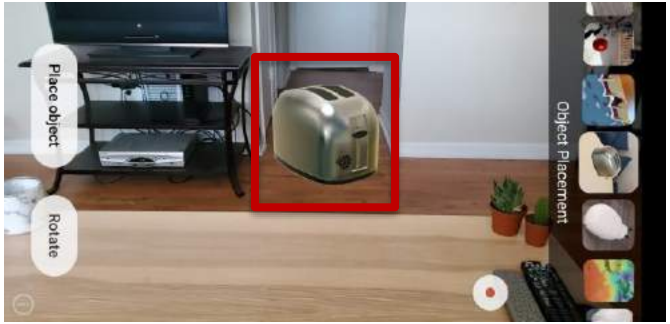




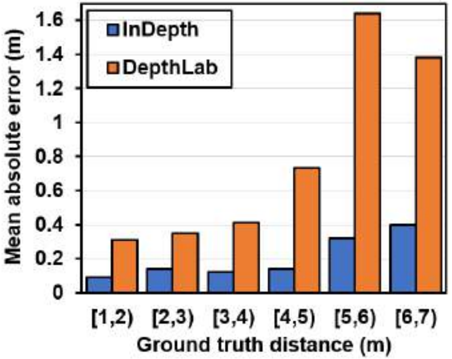
	Latency (ms)	
Edge testbed	**W-class**	**E-class**
End-to-end (i.e., total)	26.3	36.5
Communication overhead	13.1	9.2
Image pre-processing	4.5	10.8
DNN inference	8.7	16.5
Mobile device-only	**GPU**	**CPU**
DNN inference	489.2	646.5

❑ Y. Zhang, T. Scargill, A. Vaishnav, G. Premsankar, M. Di Francesco, **M. Gorlatova**, InDepth: Real-time Depth Inpainting for Mobile Augmented Reality, to appear in *Proc. ACM IMWUT*, 2022.

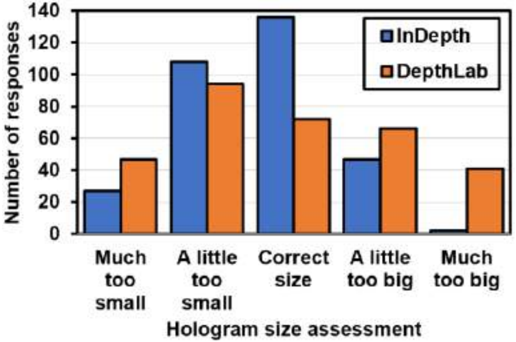
InDepth vs. Google DepthLab



DepthLab error example



Error at different distances



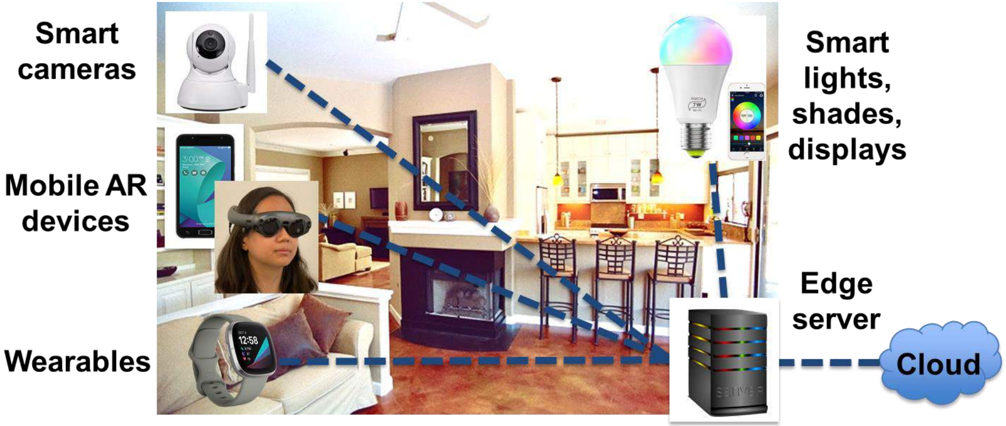
User perception

- Reduce mean absolute error from **78cm to 20cm** on average
  - From **1.51m to 36cm** at distances over 5m
- Reduce the percentage of holograms seen as “much too small/big” from **28% to 9%**

❑ Y. Zhang, T. Scargill, A. Vaishnav, G. Premsankar, M. Di Francesco, **M. Gorlatova**, InDepth: Real-time Depth Inpainting for Mobile Augmented Reality, to appear in *Proc. ACM IMWUT*, 2022.

Outline

- AR: current state and applications
- Improving AR spatial awareness
- **Improving AR semantic awareness**
- Improving user context awareness
- Summary and future work



AR Semantic Awareness

- Understanding the nature of objects and surfaces in the space around the user
  - ➢ 2D and 3D object detection, segmentation, semantic mapping, …
- Will enable many important **context-aware** AR applications
  - ➢ Safety guidance, mobile health, diminished reality

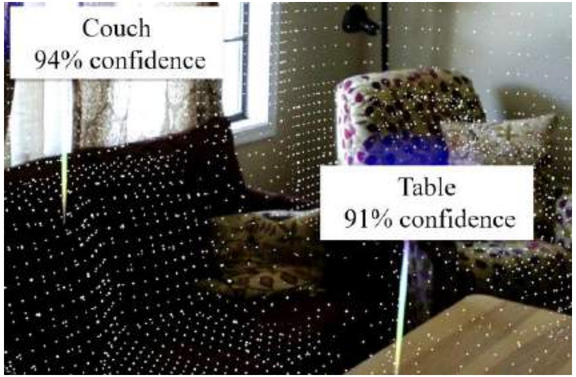
Avatar behavior adapted to conditions

Tahara et al, *IEEE ISMAR-adjunct'20*

Cars replaced with virtual objects

Kari et al, *IEEE ISMAR'21*

Semantic Awareness in AR: Challenges



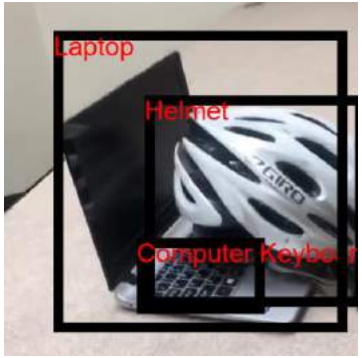
"*Even though object recognition rates can exceed 90% on many datasets, in the real world the results from these algorithms are* ***almost unusable***" – Huynh et al, *IEEE VR'19*

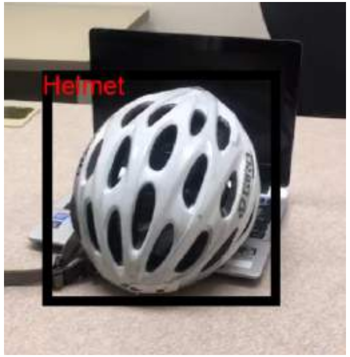
- Resource limitations, latency constraints
  - Edge computing: de-facto standard approach to solving these
- Multiple types of domain mismatches
  - Image quality, object pose, camera pose
- Open-set conditions

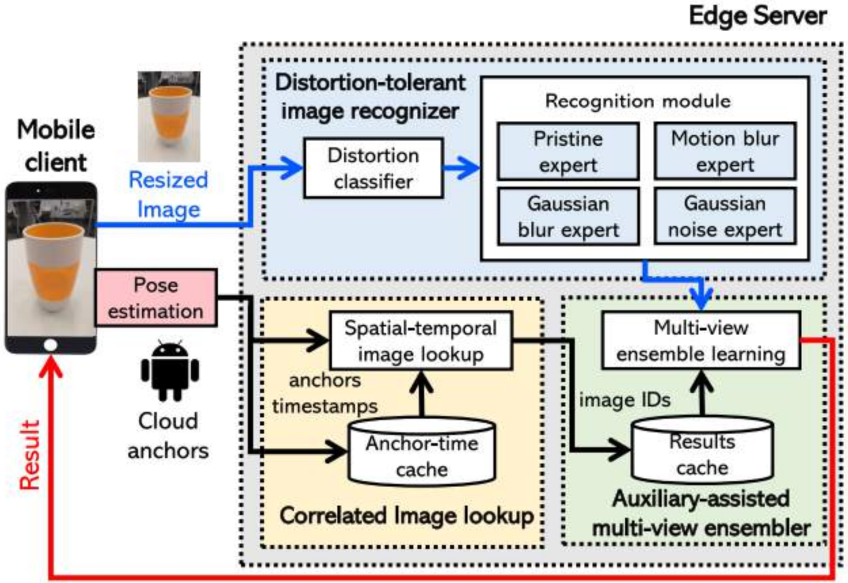
❑ Z. Liu, G. Lan, J. Stojkovic, Y. Zhang, C. Joe-Wong, **M. Gorlatova**, CollabAR: Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality, in Proc. *IEEE/ACM IPSN*, Apr. 2020. **IEEE/ACM IPSN Best Research Artifact Award.**

Multiple Views of the Same Scene to Improve Semantic Awareness

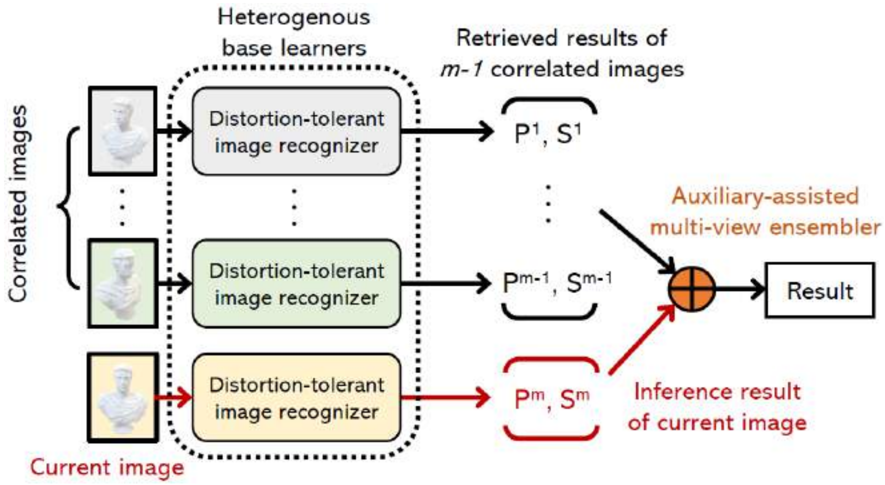
- DNN-based image recognition enables seamless contextual AR
- Challenge: DNNs make mistakes
  - ➢ Distortions, non-standard object poses, occlusions
- Multiple views of a scene can help







CollabAR: Collaborative Image Recognition for AR



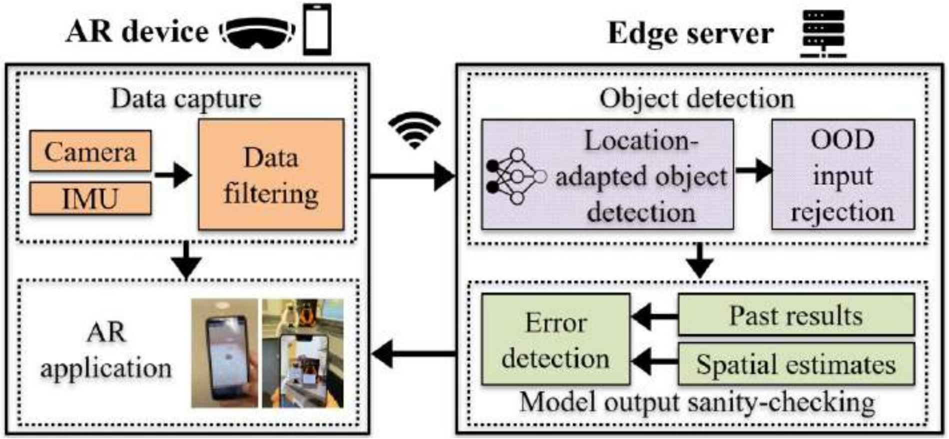
Auxiliary-assisted Multiview Ensemble Learning

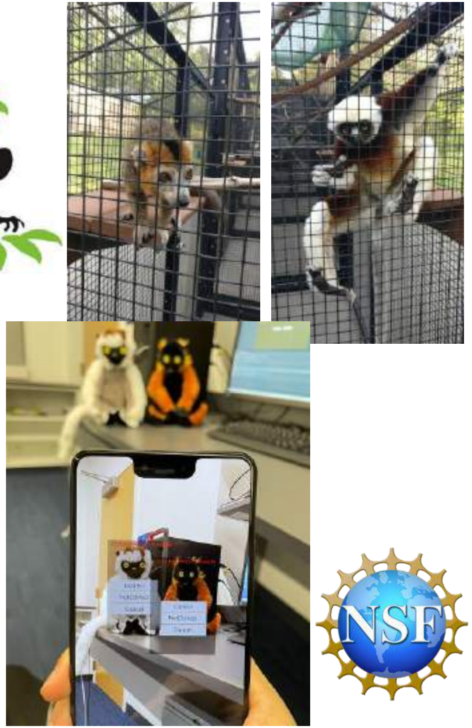
- Improves recognition accuracy by 16% to 20% in severe distortions scenario

	Nokia 7.1	Pixel 2 XL
On-device	67.2 ms	34.7 ms
Edge	19.9 ms	19.8 ms

- G. Lan, Z. Liu, Y. Zhang, T. Scargill, J. Stojkovic, C. Joe-Wong, **M. Gorlatova**, Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality, *ACM Transactions on Sensor Networks*, Vol. 18, No 1, Feb. 2022.
- Z. Liu, G. Lan, J. Stojkovic, Y. Zhang, C. Joe-Wong, **M. Gorlatova**, CollabAR: Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality, in Proc. *IEEE/ACM IPSN*, Apr. 2020. **IEEE/ACM IPSN Best Research Artifact Award.**

Ongoing Work: Robust AR Object Detection



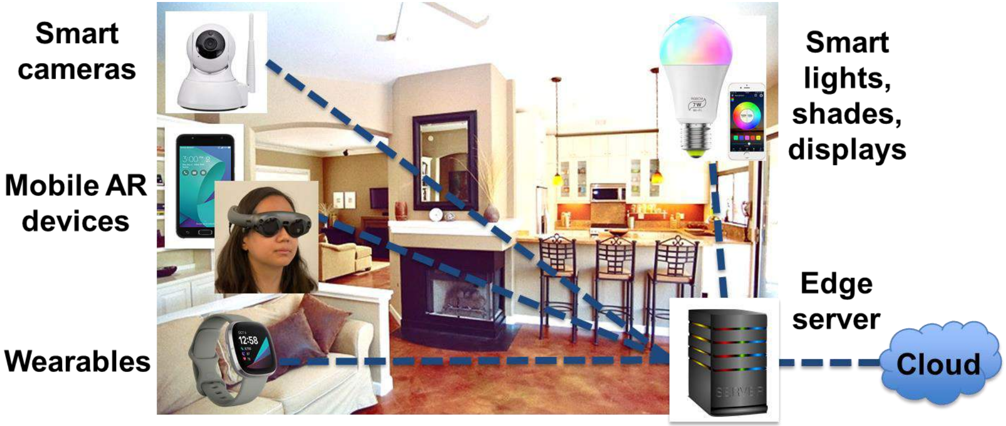


- First application: AR enhancing experience of Duke Lemur Center visitors
  - ➢ Traditional object detectors' results are ≈ random guess

❑ Students Code Alternate Realities, Pratt School of Engineering, Aug. 2020, https://pratt.duke.edu/about/news/students-code-alternate-realities

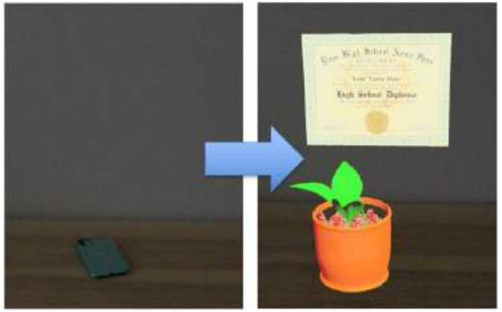
Outline

- AR: current state and applications
- Improving AR spatial awareness
- Improving AR semantic awareness
- **Improving AR user context awareness**
- Summary and future work



AR User Context Awareness

- User context awareness:
  - ➢ Understanding what the user is doing and how the user is feeling
  - ➢ Automatically assessing user’s level of expertise, concentration, mental and physical fatigue
- AR: immersive experiences tailored to the state of the user



❑ T. Scargill, Y. Chen, S. Eom, J. Dunn, **M. Gorlatova**, Environmental, User, and Social Context-Aware Augmented Reality for Supporting Personal Development and Change, in *Proc. IEEE Workshop for Building the Foundations of the Metaverse*, Mar. 2022 (co-located with *IEEE VR’22*).

User Context Awareness: Opportunities and Challenges

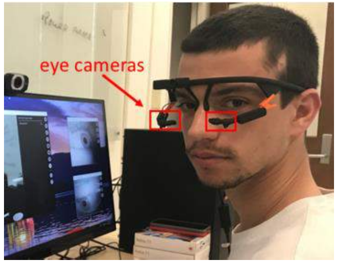
- Intimately connected with the user
  - AR **sees what you see and hears what you hear**
  - Eye tracking: looks deep inside your soul
- Wearables add new opportunities

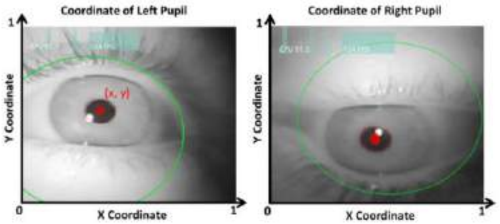
- Challenges: training **personalized** algorithms
  - Population-level models do not work well on individual users
  - Can’t collect too much data → Fine-tuning and few-shot learning
  - Personal data can be highly sensitive → Edge computing for enhanced privacy controls

❑ Y. Jiang, W. Wang, T. Scargill, M. Rothman, **M. Gorlatova**, J. Dunn, Digital biomarkers reflect stress reduction after Augmented Reality guided meditation: a feasibility study, in *Proc. ACM Workshop on Emerging Devices for Digital Biomarkers*, May 2022 (co-located with ACM MobiSys 2022).

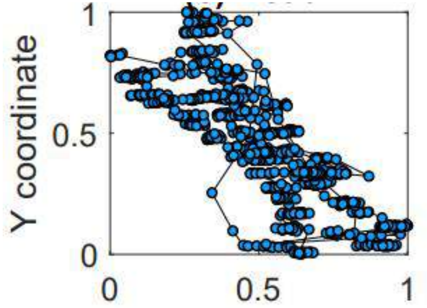
Eye Tracking-based Cognitive Context Sensing

- Important part of AR and VR hardware
- Powerful source of information about the user
  - ➢ Both ***where*** and ***how*** the user looks

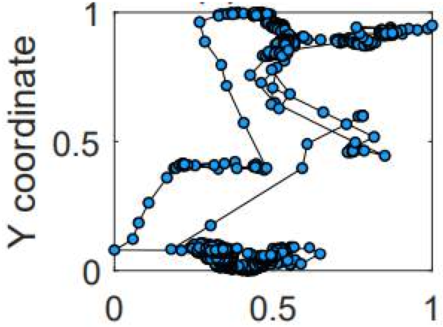




- Challenges:
  - ➢ Need to work in few-shot conditions
  - ➢ Highly heterogenous across subjects, stimuli, eye tracking devices

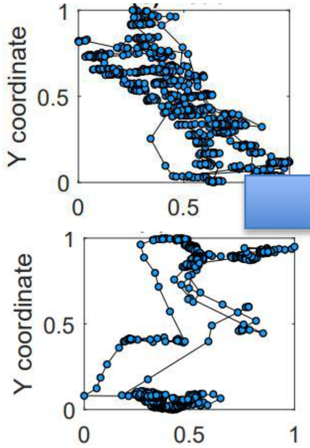


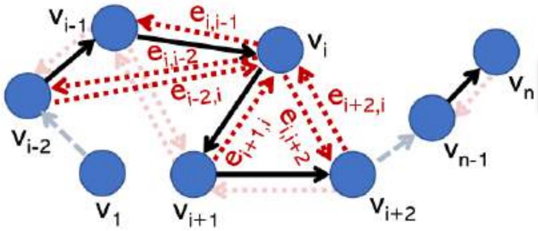
Reading text

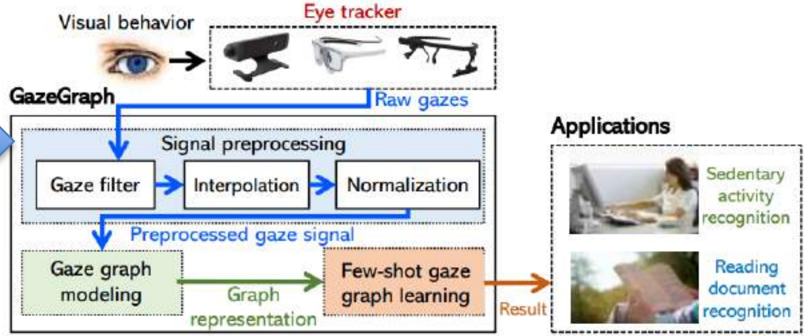


Watching a movie with subtitles

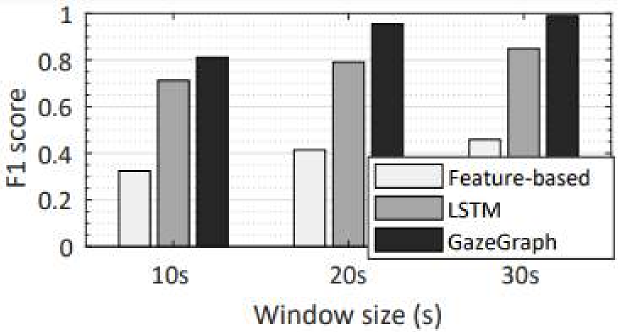
Cognitive Context Detection with GazeGraph







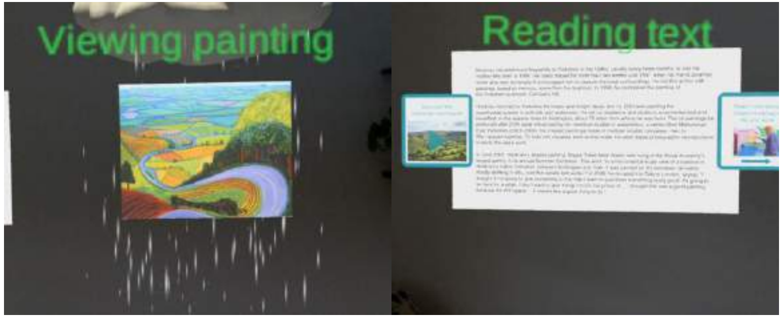
- First gaze-based cognitive context sensing method that *models human visual behavior as graphs*
- Outperforms feature-based methods by 37-54%, LSTM-based methods by 10-23%

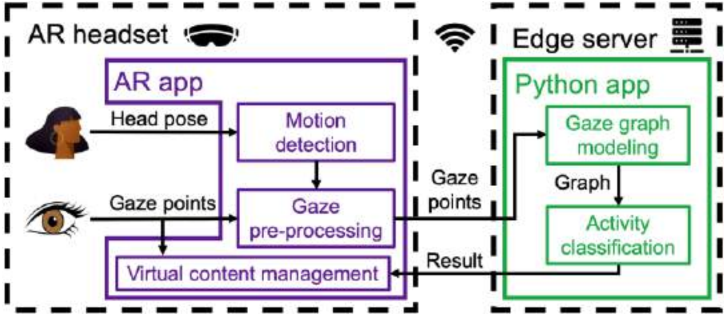


❑ G. Lan, B. Heit, T. Scargill, **M. Gorlatova**, GazeGraph: Graph-based Few-Shot Cognitive Context Sensing from Human Visual Behavior, in *Proc. ACM SenSys*, Nov. 2020.

GazeGraph in an AR Application

- Demonstrated *first gaze-based activity recognition in an AR app*





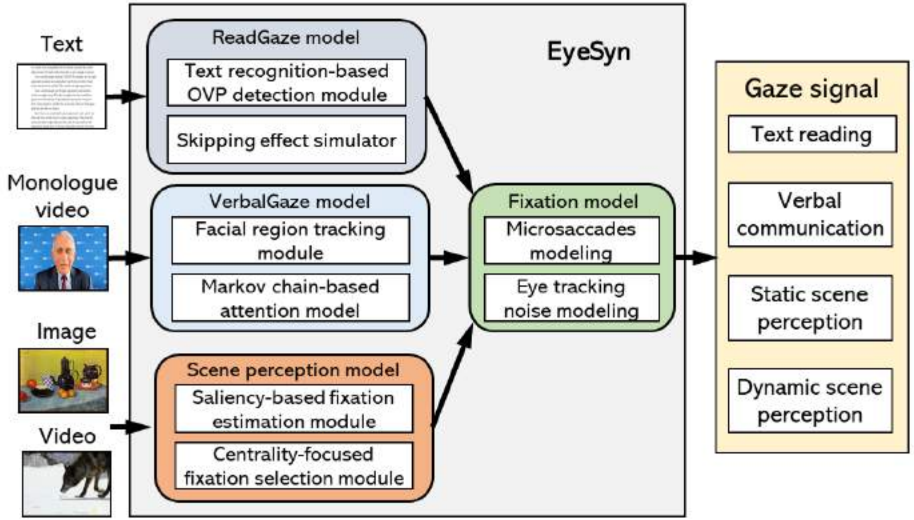
- Pre-processing, graph construction, and DNN inference time: under 15 ms
- Overall round-trip latency: under 100 ms

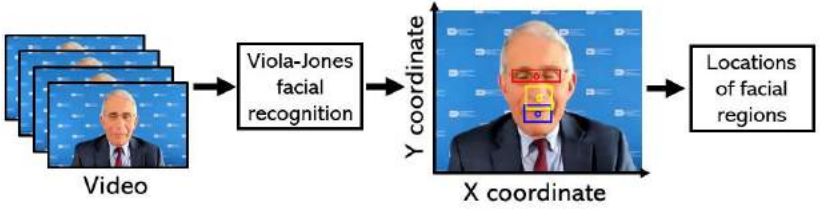
❑ T. Scargill, G. Lan, **M. Gorlatova**, Demo: Catch My Eye: Gaze-Based Activity Recognition in an Augmented Reality Art Gallery, in *Proc. ACM/IEEE IPSN*, May 2022.

❑ G. Lan, B. Heit, T. Scargill, **M. Gorlatova**, GazeGraph: Graph-based Few-Shot Cognitive Context Sensing from Human Visual Behavior, in *Proc. ACM SenSys*, Nov. 2020.

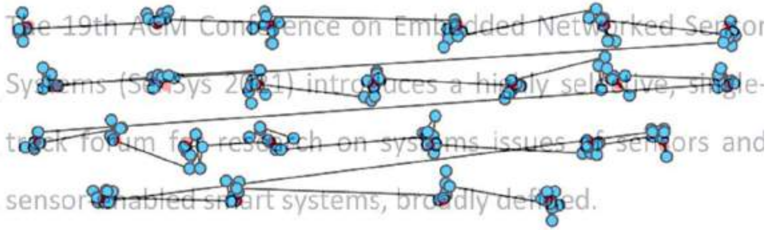
Eye Movement Synthesis for Context Detection

- Psychology-inspired generative models for *synthesizing* arbitrarily large realistic eye movement datasets





Human visual attention in a conversation

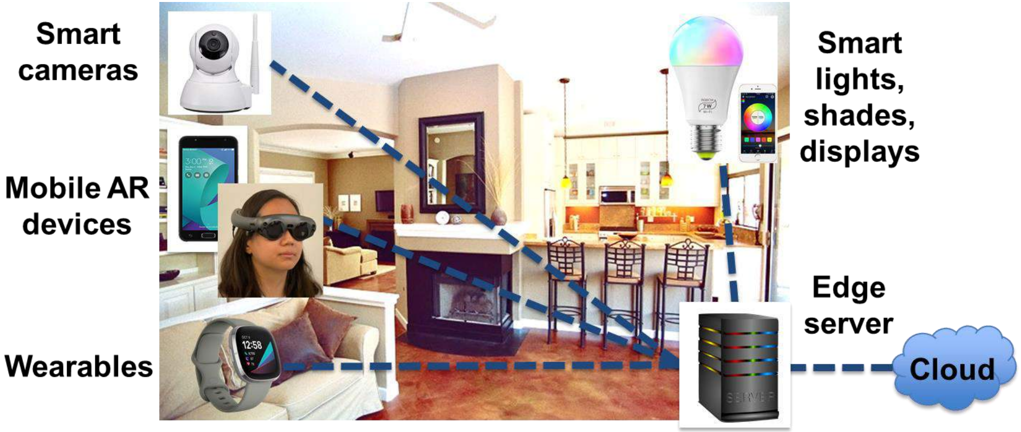


Human visual attention when reading

❑ G. Lan, T. Scargill, **M. Gorlatova**, EyeSyn: Psychology-inspired Eye Movement Synthesis for Gaze-based Activity Recognition, in *Proc. IEEE/ACM IPSN*, May 2022. Code & data available via GitHub at https://github.com/EyeSyn/EyeSynResource. **Selected media coverage: vice.com, hackster.io. Highlighted in the NSF-wide Discoveries newsletter.**

Summary and Future Work

- **Spatial**, **semantic**, and **user** context awareness
- **Edge computing**: key enabler of advanced capabilities
- Future work: wearables and the IoT: additional sources of information



- Need privacy-aware solutions
  - ➢ But cannot compromise on performance

❑ T. Scargill, Y. Chen, S. Eom, J. Dunn, **M. Gorlatova**, Environmental, User, and Social Context-Aware Augmented Reality for Supporting Personal Development and Change, in *Proc. IEEE Workshop for Building the Foundations of the Metaverse*, Mar. 2022 (co-located with *IEEE VR'22*).

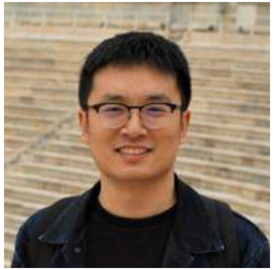
Acknowledgements



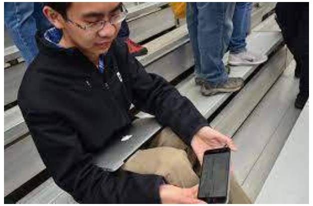










NSF aws DARPA Meta IBM ATHENA

Selected Publications

- Y. Zhang, T. Scargill, A. Vaishnav, G. Premsankar, M. Di Francesco, **M. Gorlatova**, InDepth: Real-time Depth Inpainting for Mobile Augmented Reality, to appear in *Proc. ACM IMWUT*, 2022.
- Y. Jiang, W. Wang, T. Scargill, M. Rothman, **M. Gorlatova**, J. Dunn, Digital biomarkers reflect stress reduction after Augmented Reality guided meditation: a feasibility study, in *Proc. ACM Workshop on Emerging Devices for Digital Biomarkers*, May 2022 (co-located with *ACM MobiSys'22*).
- G. Lan, T. Scargill, **M. Gorlatova**, EyeSyn: Psychology-inspired Eye Movement Synthesis for Gaze-based Activity Recognition, in *Proc. IEEE/ACM IPSN*, May 2022. **Selected media coverage: vice.com, hackster.io. Highlighted in the university-wide Duke Daily and in the NSF-wide Discoveries newsletters.**
- G. Lan, Z. Liu, Y. Zhang, T. Scargill, J. Stojkovic, C. Joe-Wong, **M. Gorlatova**, Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality, *ACM Transactions on Sensor Networks*, Vol. 18, No 1, Feb. 2022.
- G. Lan, B. Heit, T. Scargill, **M. Gorlatova**, GazeGraph: Graph-based Few-Shot Cognitive Context Sensing from Human Visual Behavior, in *Proc. ACM SenSys*, Nov. 2020.
- Z. Liu, G. Lan, J. Stojkovic, Y. Zhang, C. Joe-Wong, **M. Gorlatova**, CollabAR: Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality, in Proc. *IEEE/ACM IPSN*, Apr. 2020. **IEEE/ACM IPSN Best Research Artifact Award.**
- X. Ran, C. Slocum, Y.-Z. Tsai, K. Apicharttrisorn, **M. Gorlatova**, J. Chen, Multi-User Augmented Reality with Communication Efficient and Spatially Consistent Virtual Objects, in *Proc. ACM CoNEXT*, Dec. 2020.

Questions?

- E-mail the speaker: maria.gorlatova@duke.edu