MASS: Empowering Wi-Fi Human Sensing with Metasurface-Assisted Sample Synthesis

Jiaming Gu[1], Shaonan Chen[2], Yimiao Sun[1], Yadong Xie[1], Rui Xi[3], Qiang Cheng[2], Yuan He[1]

[1] Tsinghua University
[2] Southeast University
[3] University of Electronic Science and Technology of China

Background

➢ High-performance models demand rich datasets

Dataset collection → Model training → Testing

Insufficient data may lead to critical failures:

- Poor generalization
- Cross-domain failure

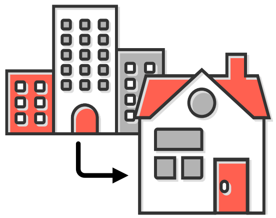
➢ Yet, acquiring such datasets is cost-prohibitive

Labor cost

Hours of repetitive human actions

Cross-domain cost

Re-collection for every new environment

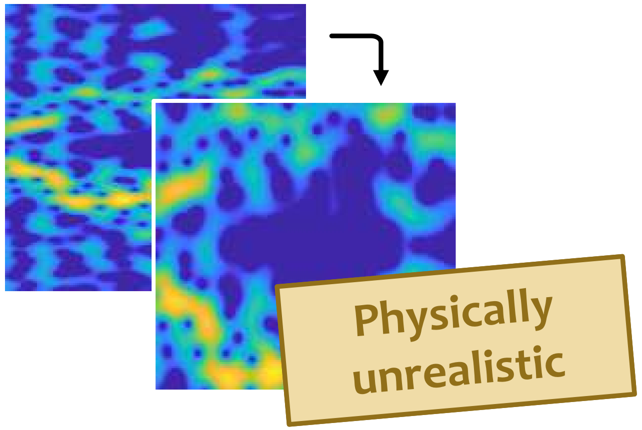


Ethical cost

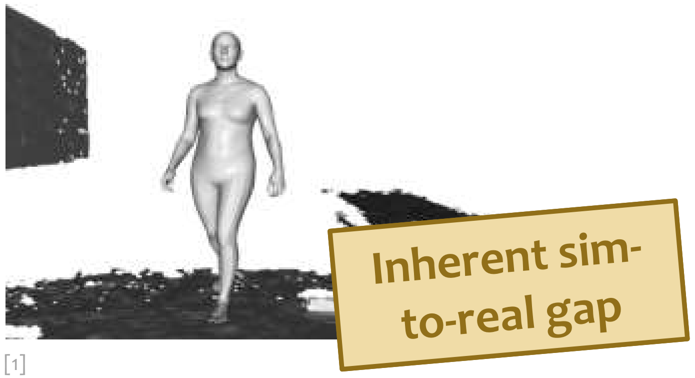
Significant ethical concerns in health monitoring

Background

➢ Existing methods



Augmentation



[1]

Simulation

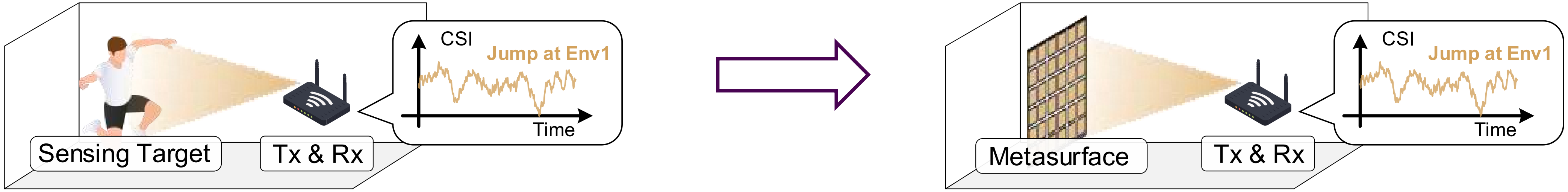
Generative AI

How can we synthesize samples that are both physically-realistic AND environment-specific, without requiring massive initial data?

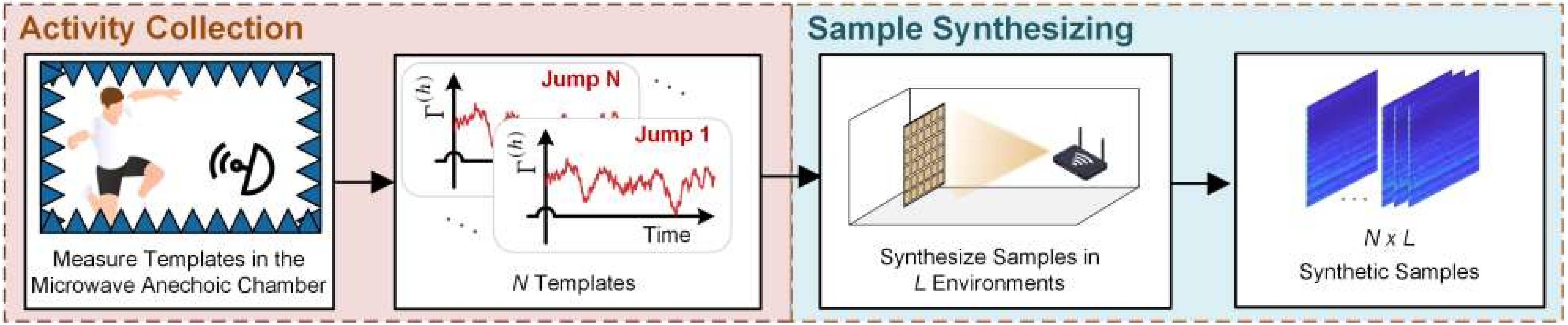
[1] This picture is taken from Huang, Teng, et al. "One Snapshot is All You Need: A Generalized Method for mmWave Signal Generation."

Our approach: metasurface-assisted sample synthesis

- **Replace the human with a programmable "actor"**

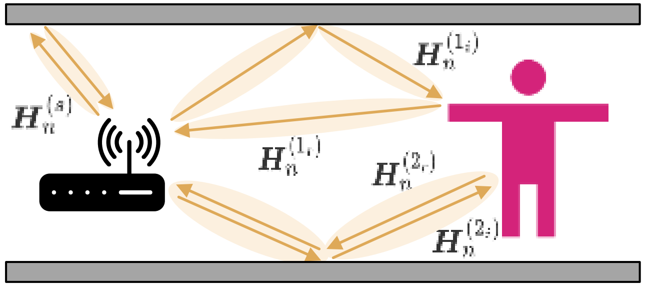


- **One template, synthesize everywhere**



Key issues

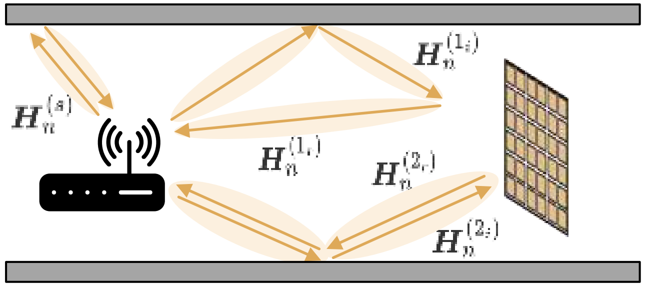
➢ Issue I: theory of the MASS



$$\boldsymbol{H}_n = \boldsymbol{H}_n^{(s)} + \boldsymbol{H}_n^{(1r)}\,\Gamma_n^{(h)}\,\boldsymbol{H}_n^{(1i)} + \boldsymbol{H}_n^{(2r)}\,\Gamma_n^{(h)}\,\boldsymbol{H}_n^{(2i)} + \ldots$$

CSI

- Static component
- Compound of the multipath and the human's impact

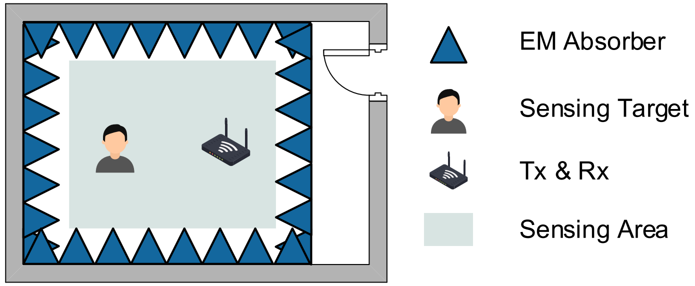


$$\boldsymbol{H}_n{}' = \boldsymbol{H}_n^{(s)} + \boldsymbol{H}_n^{(1r)}\,\Gamma_n^{(m)}\,\boldsymbol{H}_n^{(1i)} + \boldsymbol{H}_n^{(2r)}\,\Gamma_n^{(m)}\,\boldsymbol{H}_n^{(2i)} + \ldots$$

To synthesize a sensing sample, we simply need to make the metasurface's impact mimic the human's: If $\Gamma_n^{(m)} \approx \Gamma_n^{(h)}$, then $\boldsymbol{H}'_n \approx \boldsymbol{H}_n$.

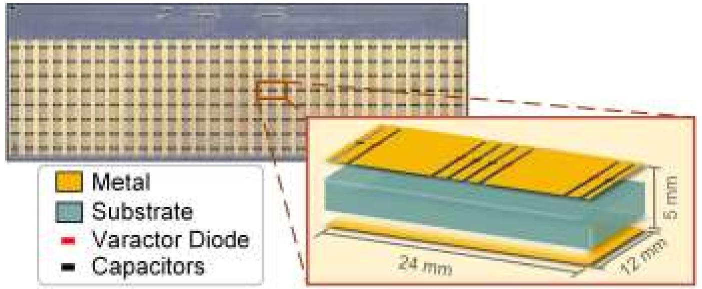
Key issues

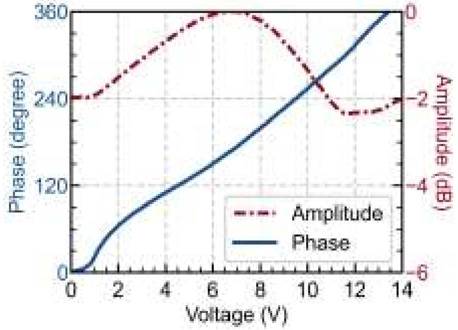
➢ Issue II: activity template collection



- Multipath effects are minimized.
- The echo signal is collected and bandpass filtered to be the activity template.

➢ Issue III: metasurface design and control

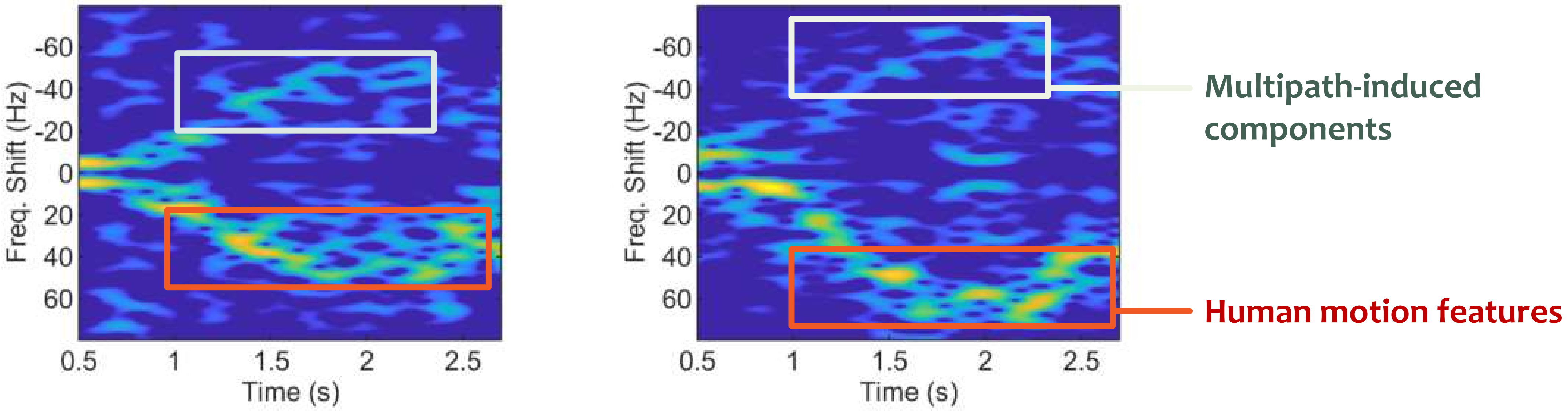




- Metasurface
  - Continuous phase modulation capability.
  - 16×8 atoms
- Phase prioritized modulation
  - Only $\angle\Gamma_n^{(m)} = \angle\Gamma_n^{(h)}$ is ensured

Fidelity of synthesized samples

➢ Comparison at the Micro-Doppler signatures

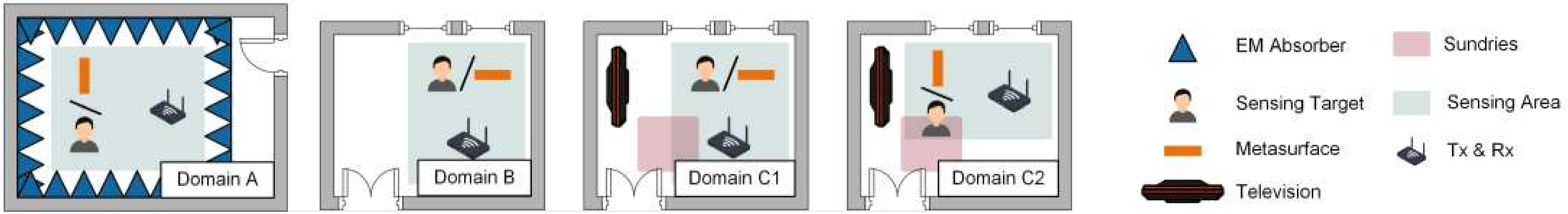


(a) The Human-Based Sample

(b) The Synthetic Sample

Experimental setup

➢ Environment



➢ End-to-end evaluation

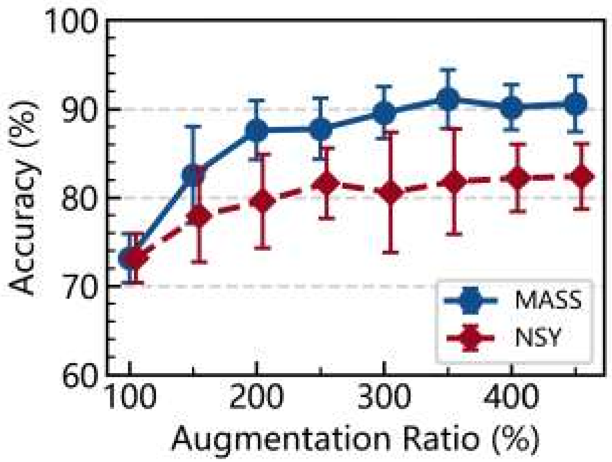
- Metric: activity recognition accuracy
- Classifier: LeNet
- Preprocessing: CFO + static component removal

➢ Dataset

- 7 daily activities (Walk, Run, Sit, etc.)
- 140 activity templates
- 659 human-based samples
- 1520 synthesized samples

Experimental result

➢ Significant accuracy improvement with MASS



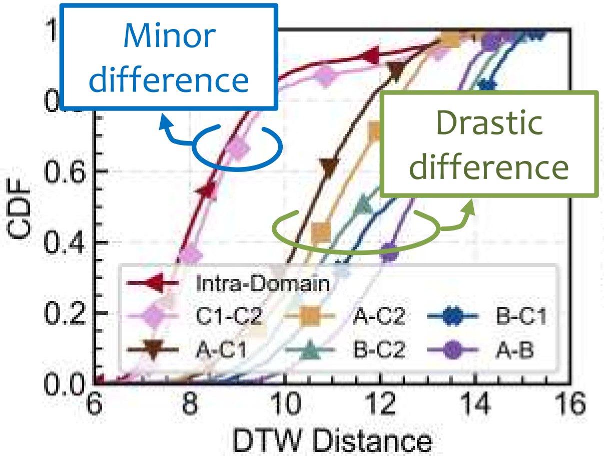
- Setup
  - Augment training set with different methods, and test on a common independent set of real samples.
  - Use k-fold cross-validation for reliability.
- Methods compared
  - MASS: synthesizing samples from our method
  - NSY: duplicating samples by adding Gaussian noise

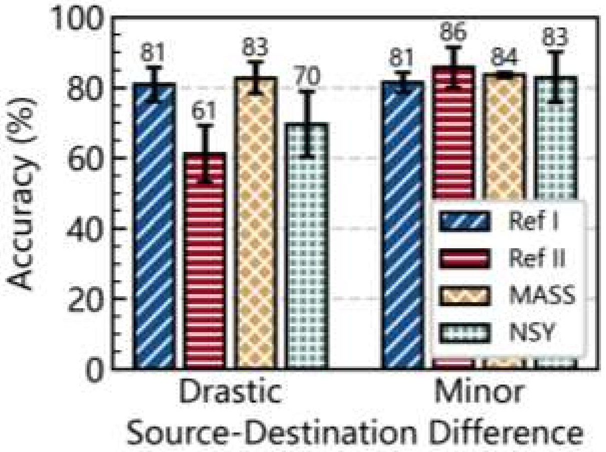
- Significant accuracy boost up to **18%** (from 73% to 91% at 350% augmentation)
- Clear superiority over NSY (**9%** higher)
- Higher stability (smaller error bars compared to NSY)

MASS captures essential motion features, offering superior training benefits over noise perturbation.

Experimental result

➢ Cross-domain sensing accuracy improvement with MASS





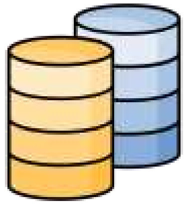
- Ref I (Ideal but unrealistic upper-bound)
  Train with samples from both the destination and the source domain.
- Ref II (Realistic Baseline)
  Train with samples from the source domain only.

- MASS outperforms the Realistic Baseline
  - Boosts accuracy by **22%** over the real-world baseline (Re
- MASS closes the cross-domain gap
  - MASS (**83%**) reaches the level of the ideal upper-bound (Ref I at **81%**).

MASS successfully captures domain-specific multipath features, validating it as a powerful solution for cross-domain sensing.

Future directions

➢ Several directions that address the limitations and unlock its full potential

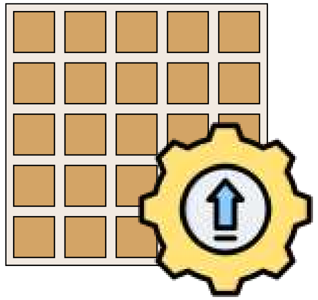


Comprehensive evaluation

- Our initial evaluation covers a limited set of activities.
- => Richer activities & more sensing domains

Human-free template collection

- Templates still rely on manual collection.
- => Fully automated, human-free pipeline.



Next-Gen Metasurface

- Phase-only control, relatively-high cost
- => More control capabilities & low-cost deployment.

Conclusion

- **Data scarcity remains a critical bottleneck for practical Wi-Fi sensing.**
- **Metasurface offers a new, powerful paradigm.**
  - It synthesizes samples that capture both motion dynamics and environment-specific features, offering a novel solution to the data scarcity problem.
- **We contribute MASS, a complete framework to realize this paradigm.**
- **Our results validate the feasibility of MASS.**
  - MASS boosts in-domain accuracy **18%.**
  - More crucially, MASS enhances cross-domain accuracy by **22%**, nearly matching the ideal upper-bound.

By fundamentally changing how data is acquired, MASS enables the creation of large-scale, diverse datasets at a significantly lower cost.

MASS

Thank You!

Please find more details in:

http://tns.thss.tsinghua.edu.cn/sun/