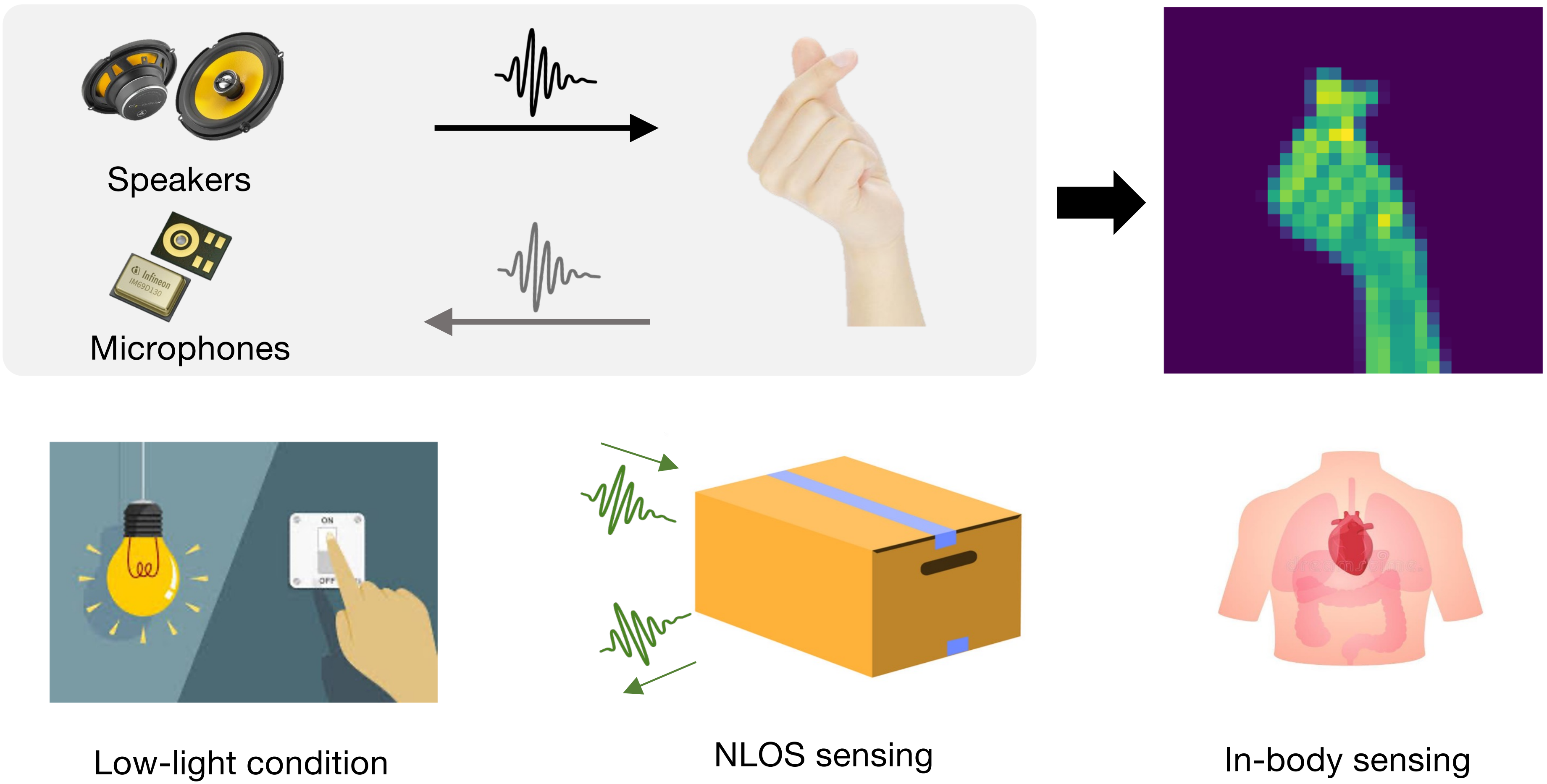
Adaptive Metasurface-based Acoustic Imaging using Joint Optimization

Yongjian Fu, Yongzhao Zhang, Yu Lu, Lili Qiu, Yi-Chao Chen, Yezhou Wang, Mei Wang, Yijie Li, Ju Ren, Yaoxue Zhang

ACM MobiSys 2024

Enabling Acoustic Imaging on IoT Devices



Enabling Acoustic Imaging on IoT Devices

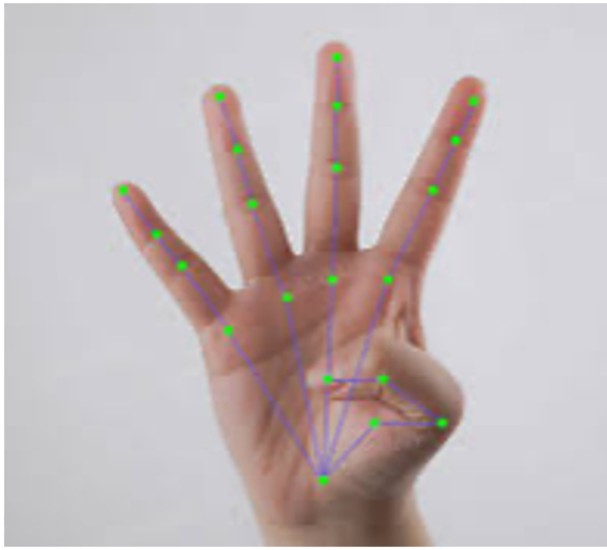
Smartphones

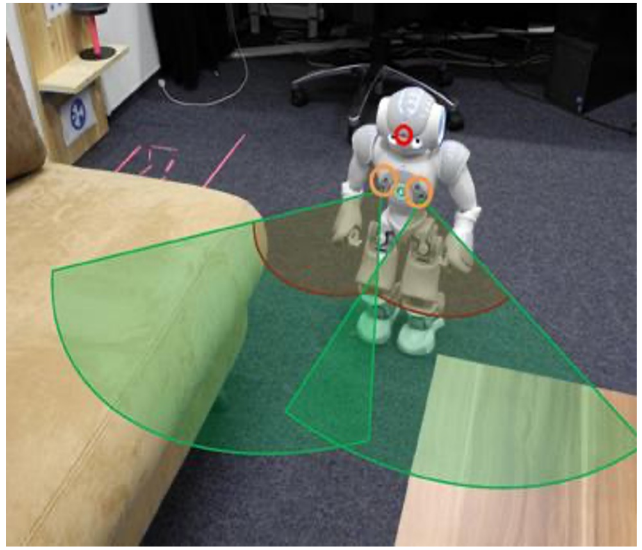
Smart speakers

Mobile robots

Enabling Acoustic Imaging on IoT Devices

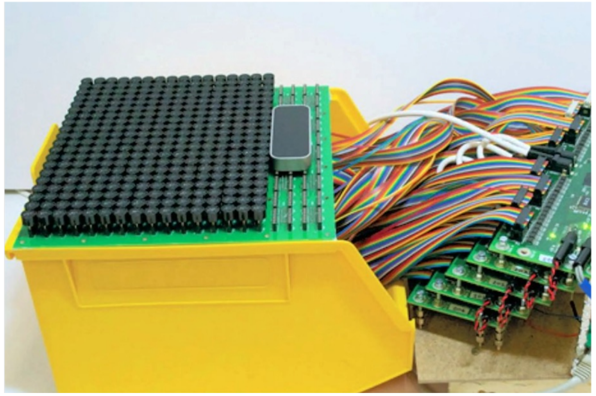
Such imaging functionality can fundamentaly change the way of acoustic sensing.



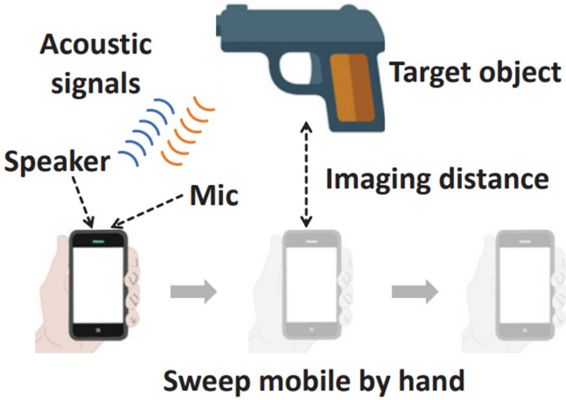


Related Works of Acoustic Imaging

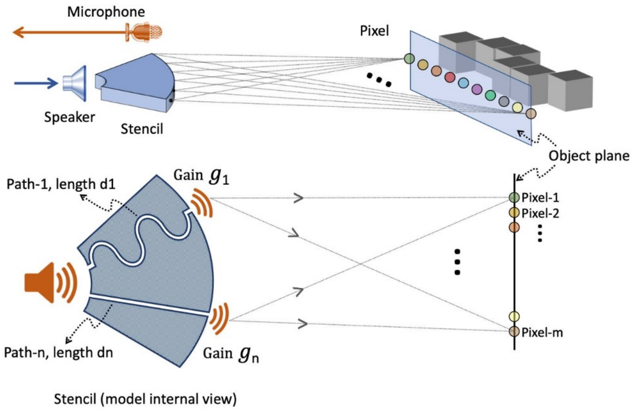
Using large array



Requiring mechanical movement[1]



Limited scalability[2]



[1] Mao, Wenguang, Mei Wang, and Lili Qiu. "Aim: Acoustic imaging on a mobile." ACM MobiSys. 2018.

[2] Bai, Yang, Nakul Garg, and Nirupam Roy. "Spidr: Ultra-low-power acoustic spatial sensing for micro-robot navigation." ACM MobiSys. 2022.

Expensive! Time-consuming! Limited scalability!

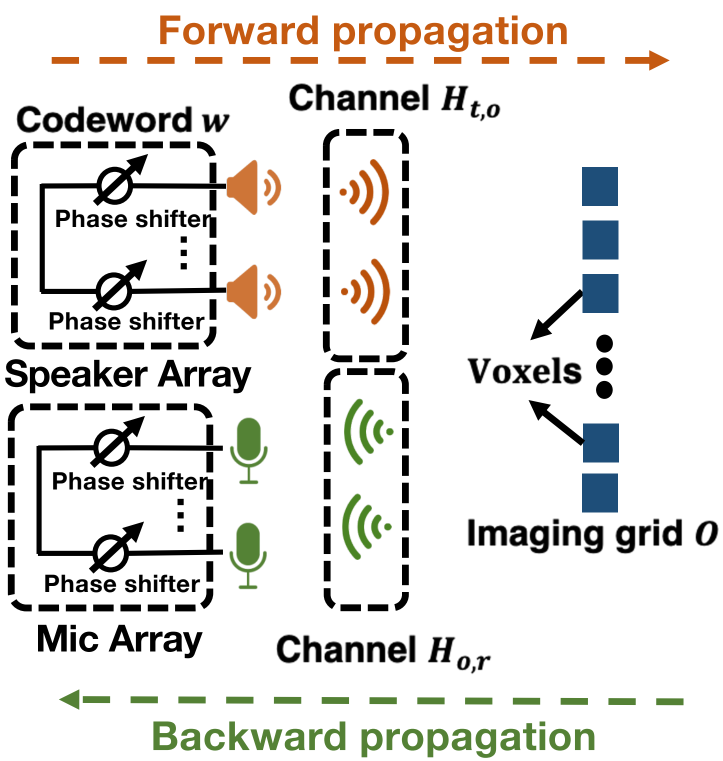
Limitations of Current Acoustic Imaging on IoT Devices

≤ 6 Microphones
≤ 6 Speakers

≤ 3 Microphones
≤ 3 Speakers

Can we achieve accurate acoustic imaging on IoT devices?

Effective Rank Defects for Imaging



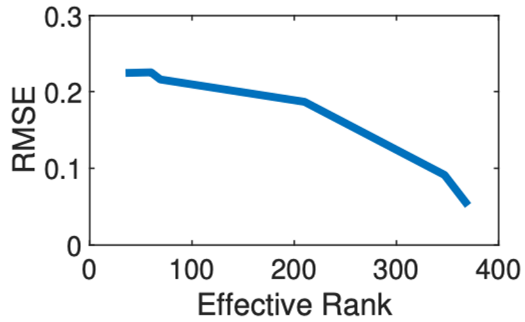
The received signal of each microphone:

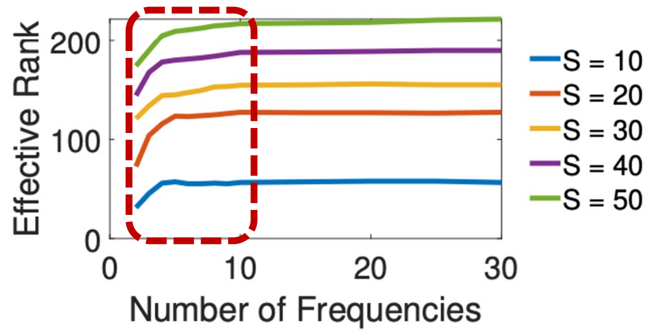
$$R_m = H_{o,r}\, O \cdot H_{t,o} w + \epsilon$$
$$= A(w) O + \epsilon$$

$$\mathrm{y = Ax + e, A \in \mathbb{E}^{M\times N}} \xrightarrow{\text{Rank-sufficient}} x \approx A^{-1} y$$

We can infer the image O by optimizing (rank-deficient):

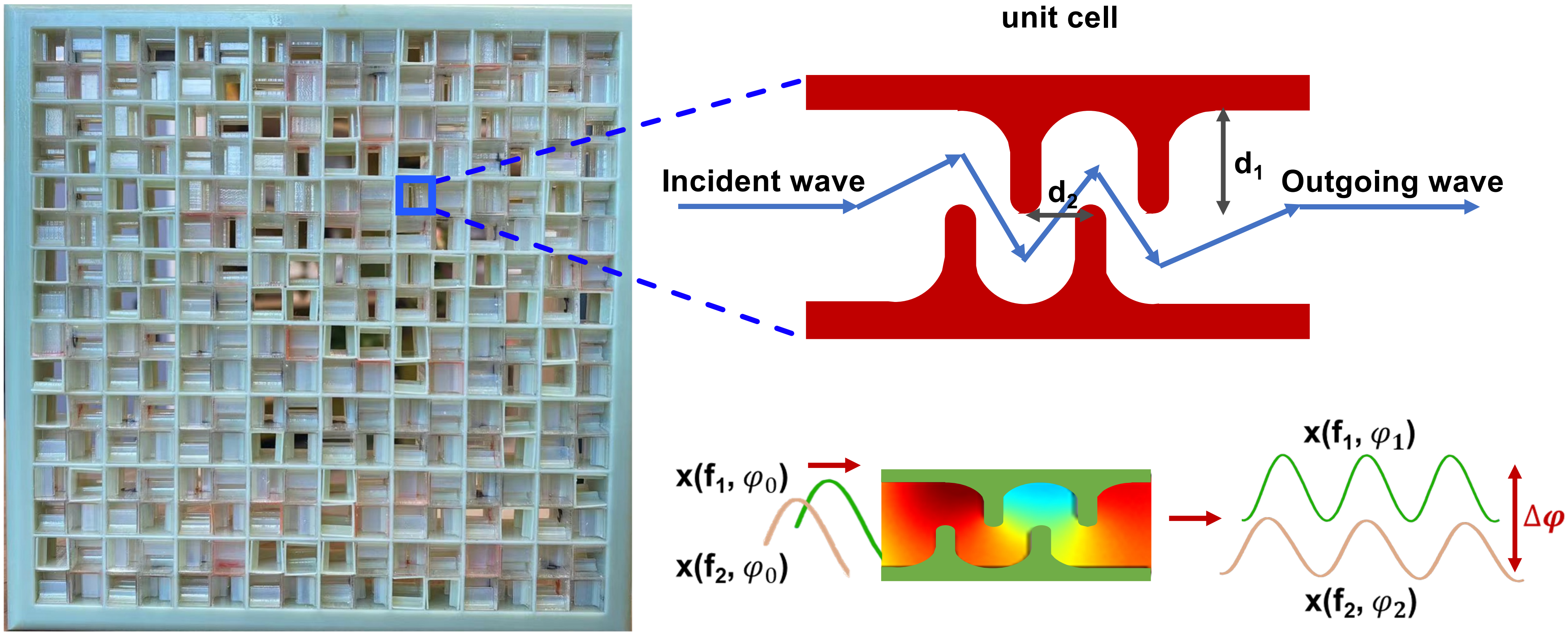
$$\arg\min \| Ax - y \|_2$$



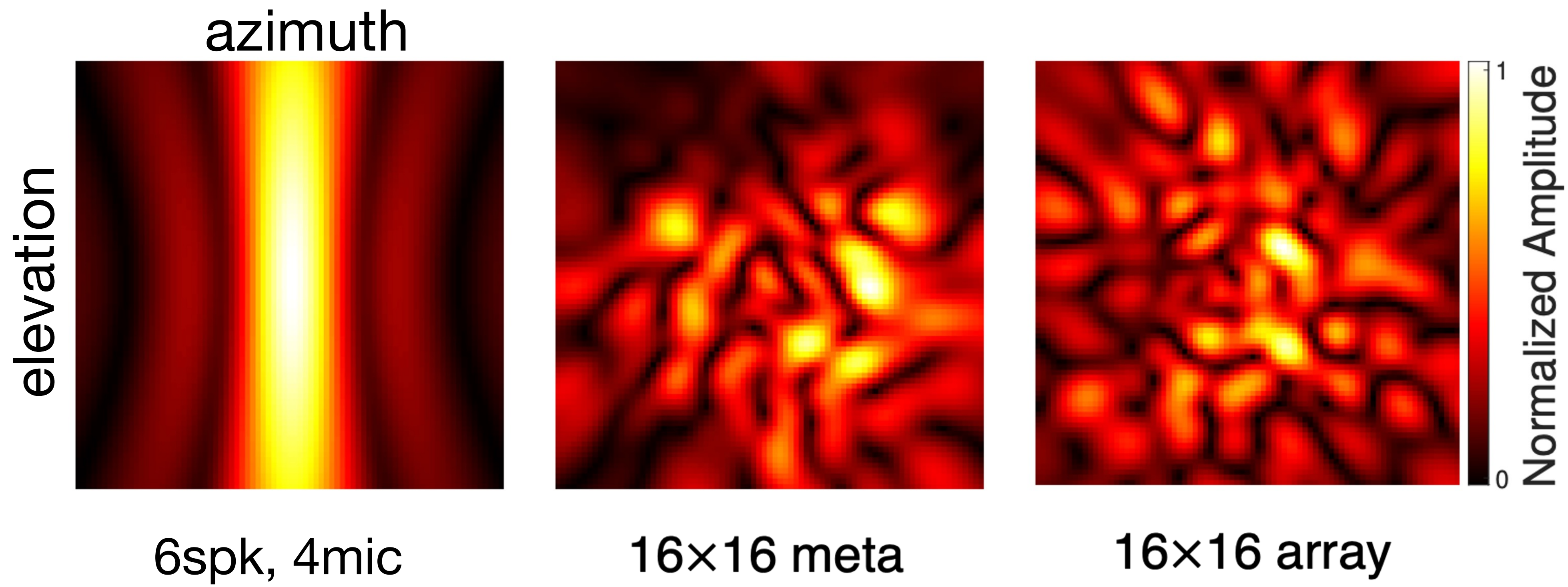


Acoustic Imaging is limited by the effective rank of measurement matrix!

Enhancing the Rank using Acoustic Metasurface

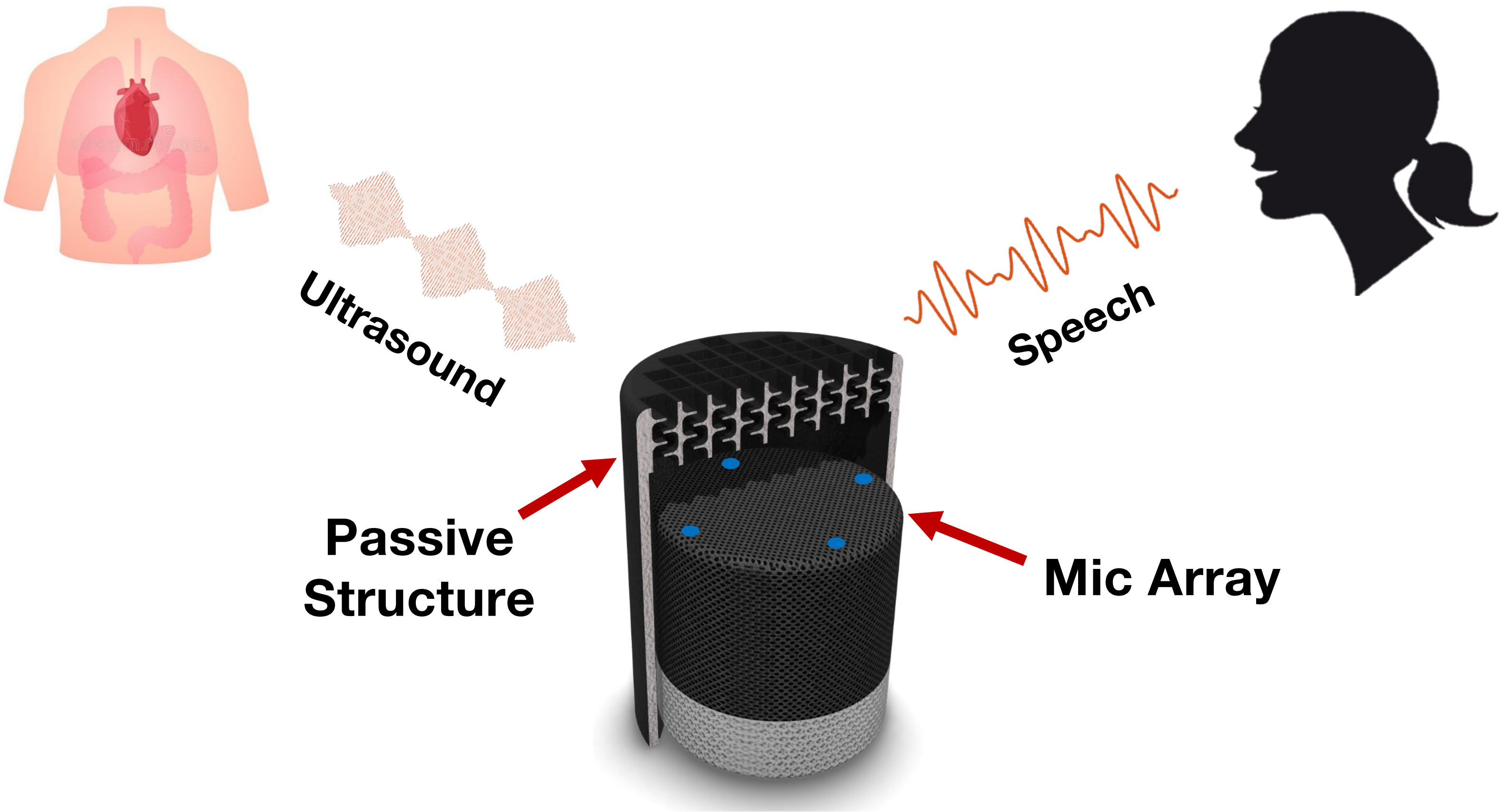


More Powerful and Diverse Beamforming Capabilities

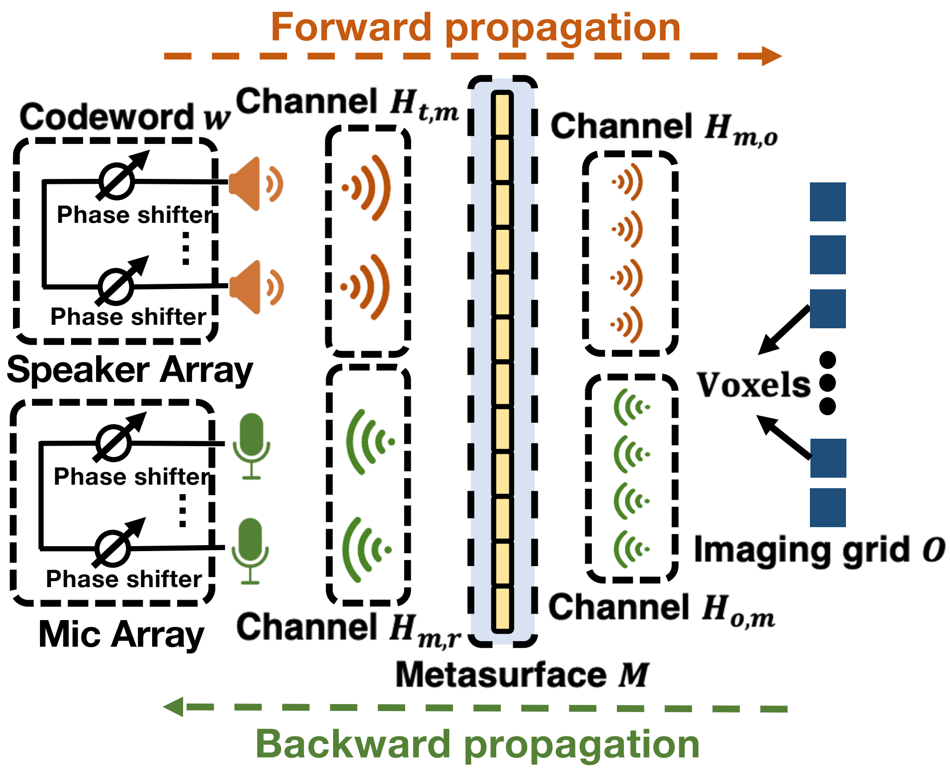


Sound field distribution in space

Integrating Acoustic Metasurface as Device Shell



Enhanced Measurement Matrix using Metasurface

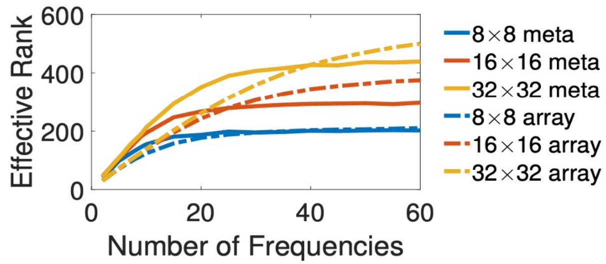


- **The received signal of each microphone:**

$$
\begin{aligned}
R &= H_{m,r}M \cdot H_{o,m}O \cdot H_{m,o}M \cdot H_{t,m}w \\
&= H_{m,r}M \cdot H_{o,m}(H_{m,o}M \cdot H_{t,m}w) \cdot O \\
&= H_{m,r}M \cdot H_{o,m}diag(H_{m,o}M \cdot H_{t,m}w)O \\
&= H_{m,r}diag(M)H_{o,m}diag(H_{m,o}M \cdot H_{t,m}w)O \\
&= A(M, w)O
\end{aligned}
$$

- **We can infer the image *O* by optimizing:**

$$\min_{O} \|A(M, w)O - R_m\|_2^2$$



Enhancing the rank like lager arrays!

Design Challenges

- How to design an imaging algorithm achieve high imaging quality?
- How to jointly optimize beamforming, metasurface, and imaging algorithms?
- How to maintain high imaging quality across a wide range of distances?

Design Challenges

- ☐ How to design an imaging algorithm achieve high imaging quality?
  - ➢ A new imaging algorithm that uses unrolled ADMM + refinement network
- ☐ How to jointly optimize beamforming, metasurface, and imaging algorithms?
- ☐ How to maintain high imaging quality across a wide range of distances?

Solution 1: Neural-enhanced Imaging Algorithm

- **Utilizing sparsity to simplify the optimization:**

$$\min \|A(M,W,D)O - R_m\|_2^2 + \underline{\alpha\|\mathcal{N}(z)\|_1}$$

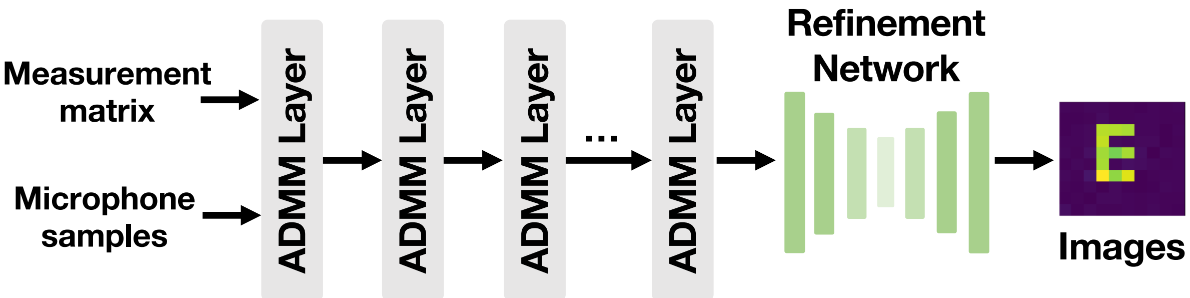
$$s.t.\ O - z = 0 \quad \text{Neural constrains}$$

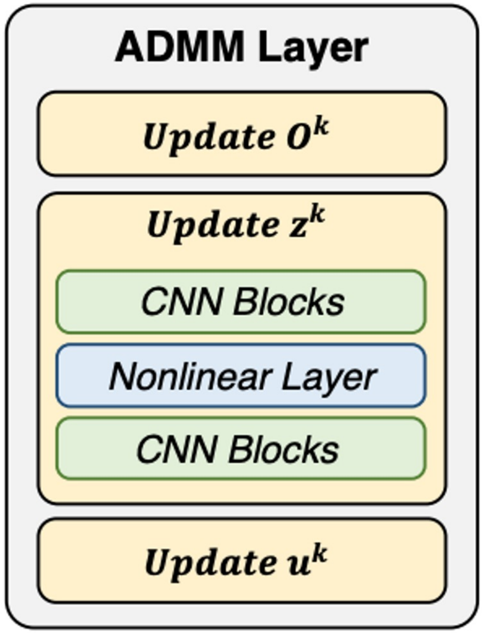
- **Using unrolled ADMM to solve this problem:**

$$O^{k+1} = \underset{O}{argmin}\{\|A(M,W,D)O - R_m\|_2^2 + \frac{\rho}{2}\|O - z^k + u^k\|_2^2\}$$

$$z^{k+1} = \mathcal{N}(O^{k+1}, u^k)$$

$$u^{k+1} = u^k + O^{k+1} - z^{k+1}$$



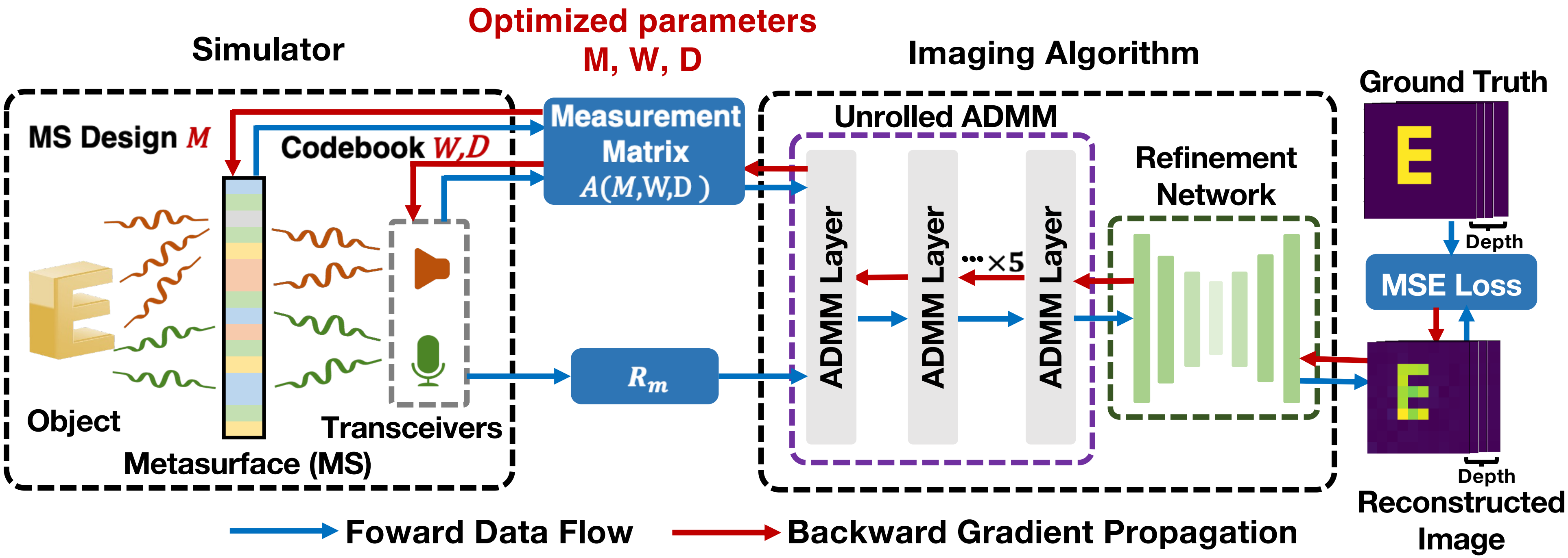


Structure of ADMM layer

Design Challenges

- ☐ **How to design an imaging algorithm achieve high imaging quality?**
  - ➢ **A new imaging algorithm that uses unrolled ADMM + refinement network**
- ☐ **How to jointly optimize metasurface and beamforming algorithms?**
  - ➢ **A novel joint optimization framework**
- ☐ **How to maintain high imaging quality across a wide range of distances?**

Solution 2: MAJIC: A Joint Optimization Framework



Optimizing System Configuration Using Mean Square Error

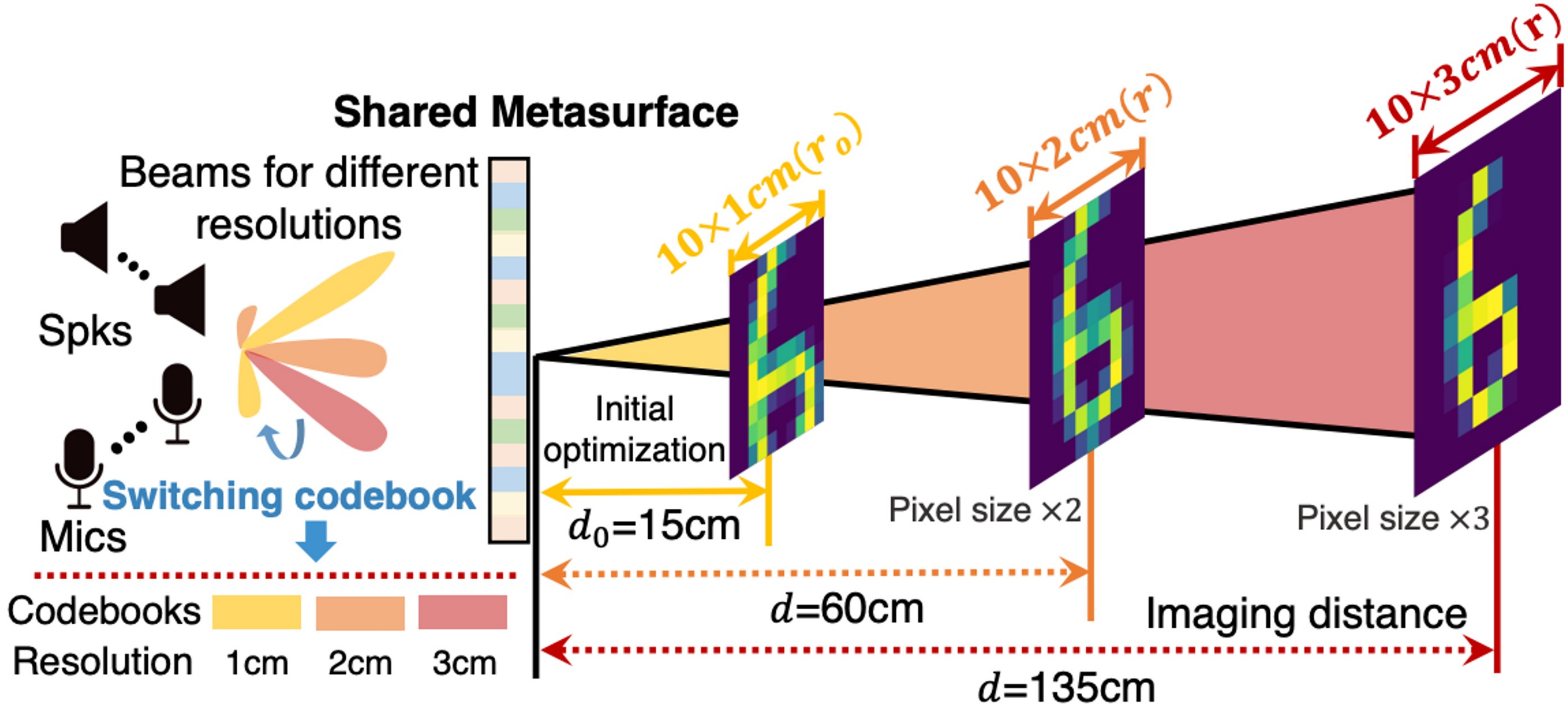
Design Challenges

- ☐ **How to design an imaging algorithm achieve high imaging quality?**
  - ➢ **A new imaging algorithm that uses unrolled ADMM + refinement network**

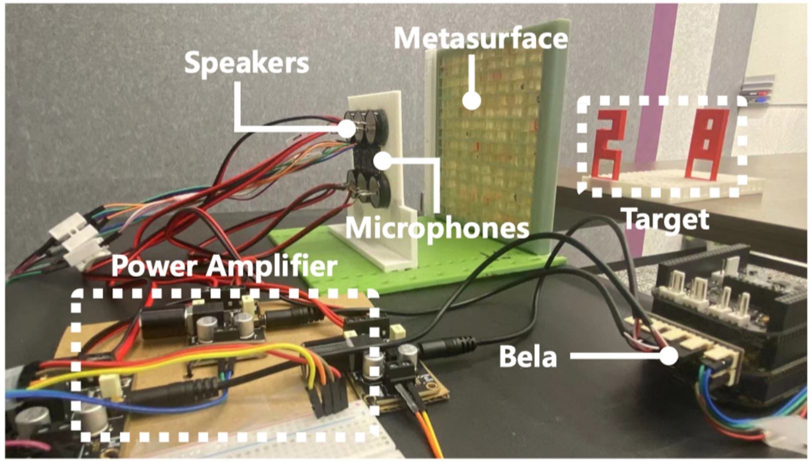
- ☐ **How to jointly optimize beamforming, metasurface, and imaging algorithms?**
  - ➢ **A novel joint optimization framework**

- ☐ **How to maintain high imaging quality across a wide range of distances?**
  - ➢ **Adapt imaging resolution and beamforming according to the distance**

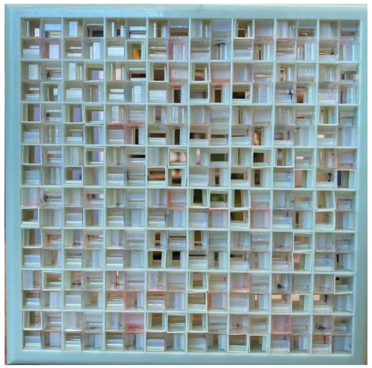
Solution 3: Adaptive Scheme for Varying Distance



Experimental Setup



a small phased array + a passive metasurface



16x16 metasurface

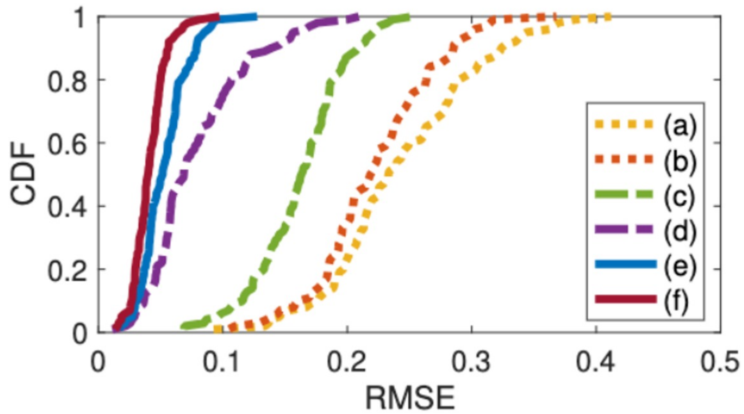
2x3 commodity speakers
1x4 commodity microphones

- Imaging resolution: 1cm
- Imaging area: $10\times10\times10$
- Open-sources Fashion-MNIST as training dataset
- Implement on the Bela board
- Metasurface to imaging region distance: 30cm
- 18-20kHz Chirp signal, 20ms for each sweep
- Metric: Root Mean Square Error (RMSE)

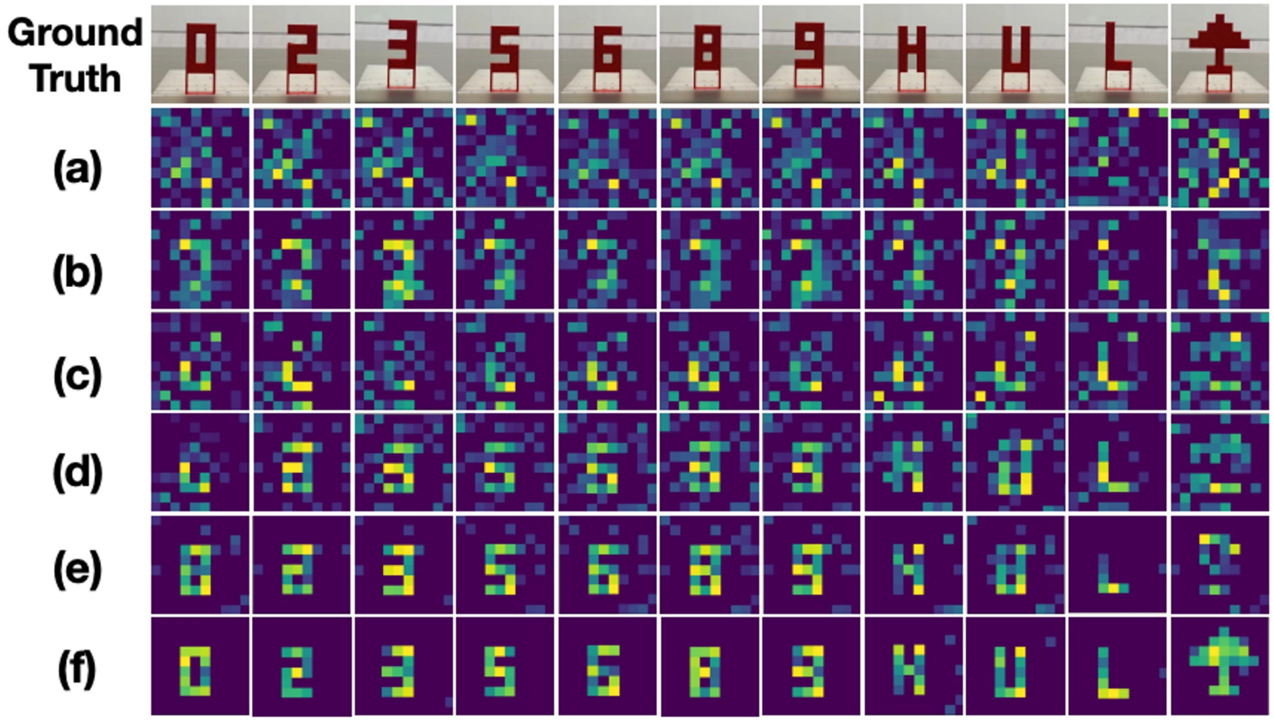
Overall Performance

Table 1: Summary of schemes for comparison.

	speaker		microphone	
	beamforming	**metasurface**	**beamforming**	**metasurface**
(a)	no	no	no	no
(b)	joint opt.	no	no	no
(c)	joint opt.	random	no	no
(d)	joint opt.	joint opt.	no	no
(e)	joint opt.	joint opt.	no	joint opt.
(f)	joint opt.	joint opt.	joint opt.	joint opt.



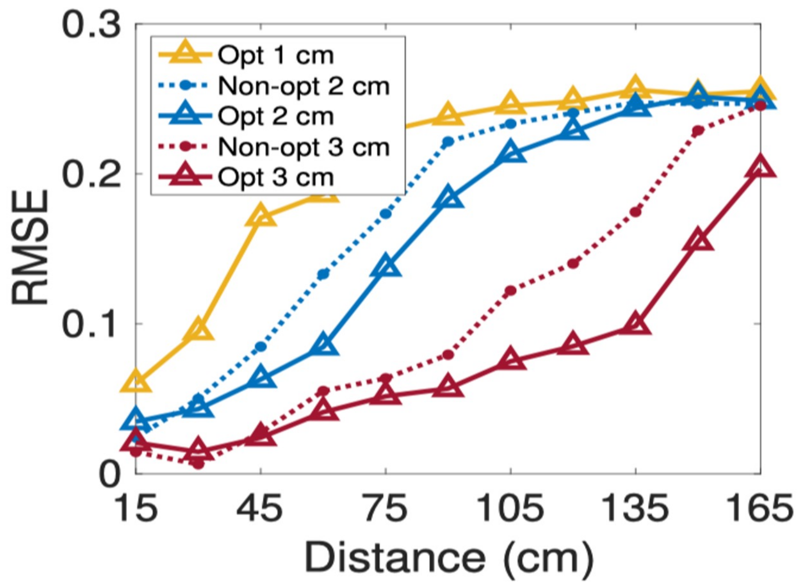
Performance of various schemes



2D Imaging examples

Achieving 83.1% error reduction compared to no metasurface and optimization

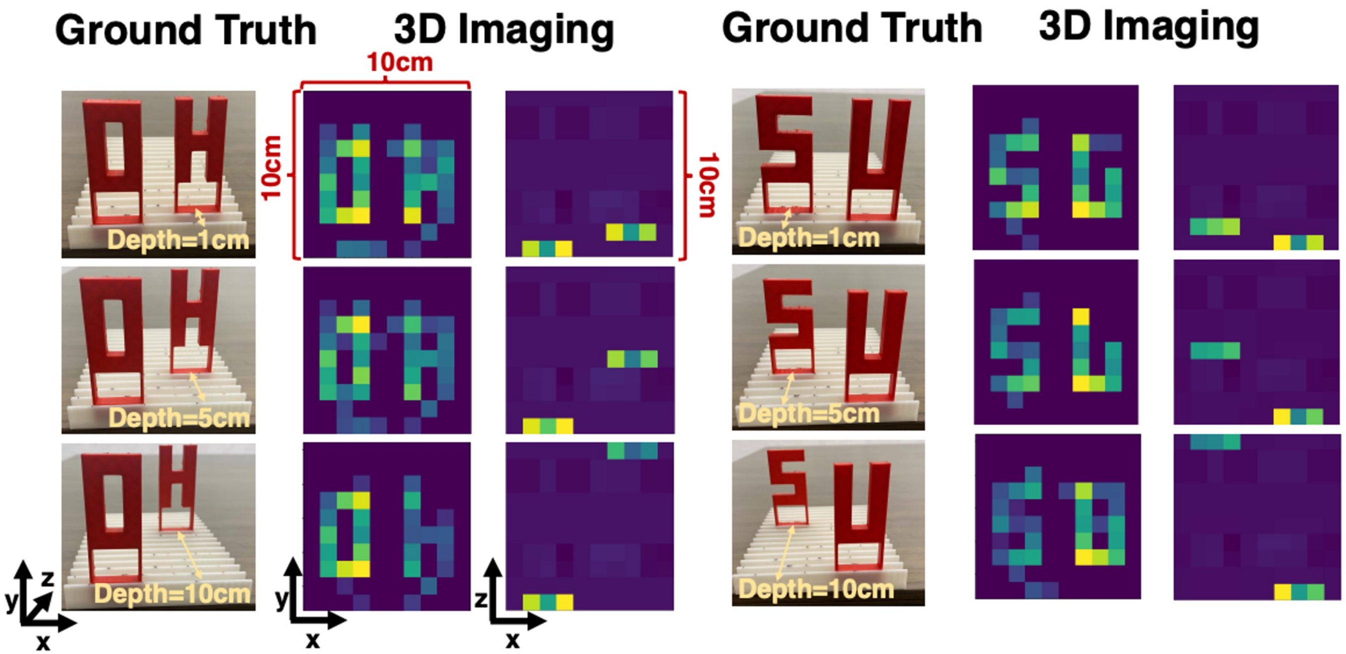
Imaging Performance in Varying distance



Varying distance

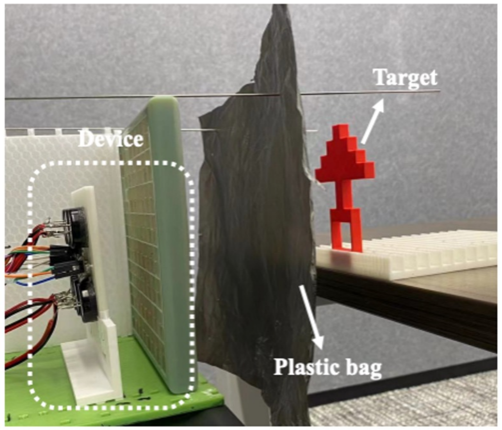
Enhancing the imaging distance to 135cm with <0.1 RMSE

3D Imaging



3D imaging achieves <0.11 RMSE.

Beyond the camera



Occlusion setup

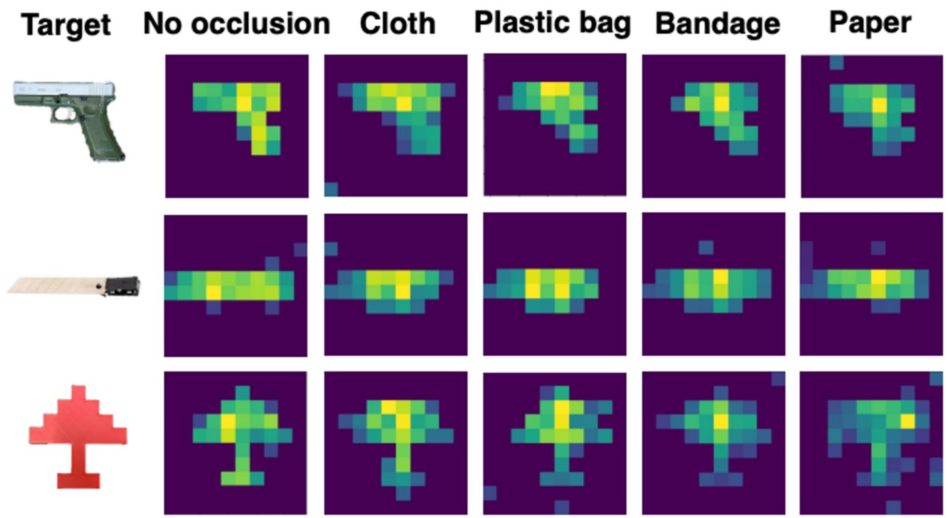
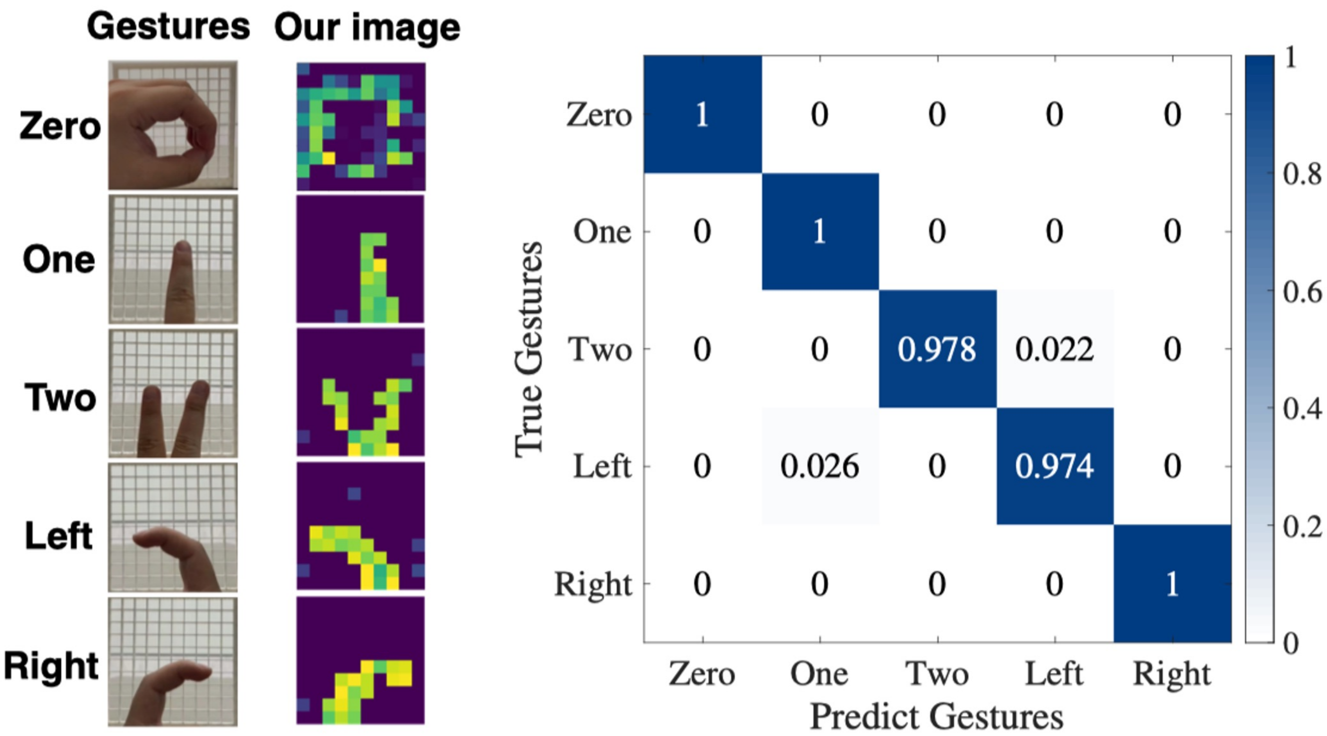


Image results

Achieving 0.07~0.12 RMSE under occlusion.

Gestures recognition



(a) Examples. (b) Confusion matrix.

High gesture recognition accuracy!

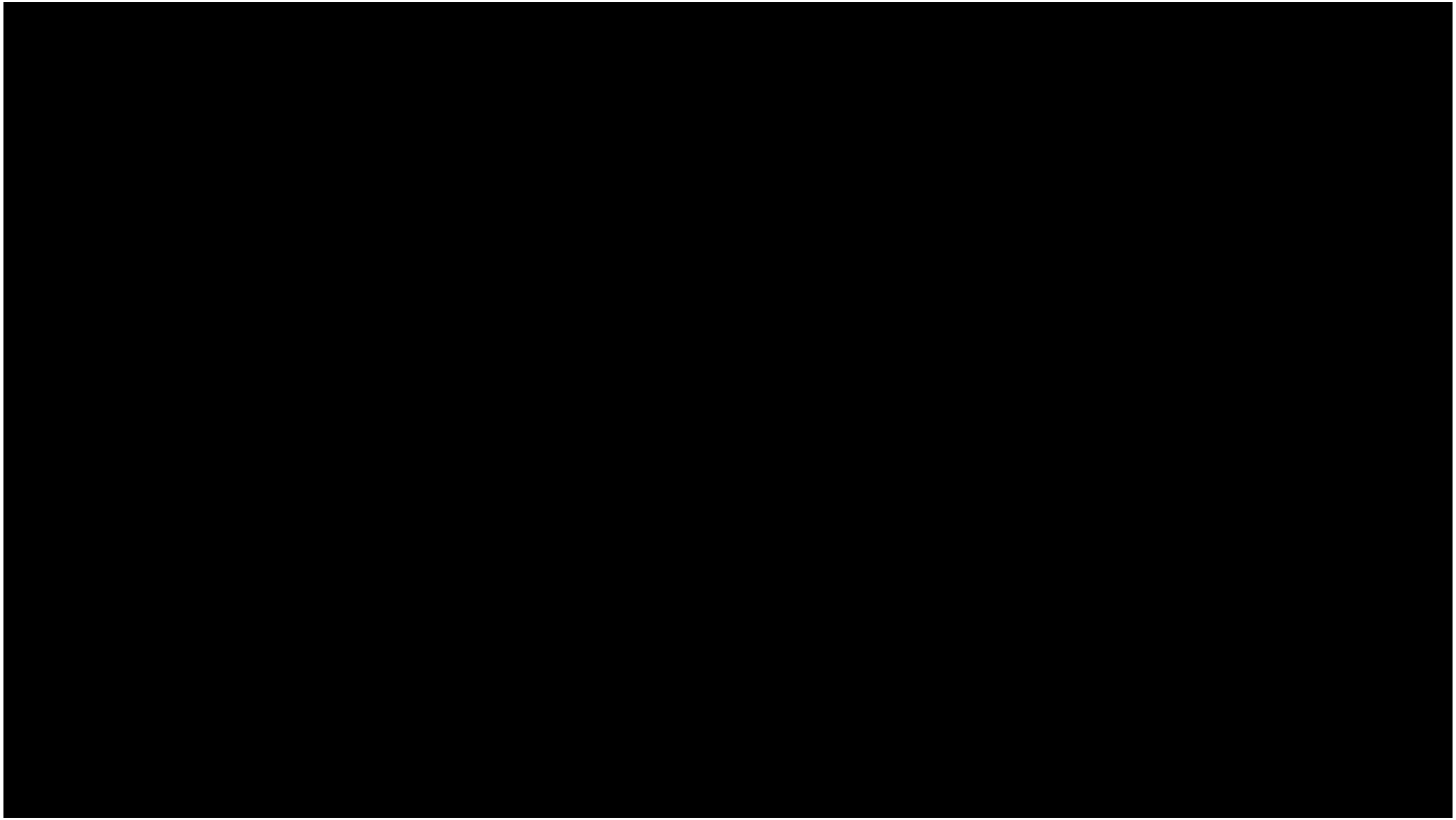
Computational Cost

Table 2: Reconstruction algorithm overhead using a GeForce RTX 3090 on the server.

	LR	ADMM	UNet	ADMMNet	Ours
Time (ms)	4.4	175.1	1.7	9.3	9.7
Model size (MB)	\	\	116.2	1.8	3.6

Low computation cost (<10ms) !

Demo Video



Contributions

- **Imaging performance is limited by the rank defect and propose to use metasurface and neural priors for imaging.**

- **A joint optimization framework for maximizing the contribution of each designable component.**

- **A prototype and conduct extensive evaluation to demonstrate the effectiveness of the imaging capabilities.**