AmbiEar: mmWave Based Voice Recognition in NLoS Scenarios

Jia Zhang*[1], Yinian Zhou*[1], Rui Xi[1],
Shuai Li[1], Junchen Guo[2], Yuan He[1]

[1]Tsinghua University
[2]Alibaba DAMO Academy

The potential of mmWave sensing in voice recognition

Waveear (MobiSys'19)
VocalPrint (SenSys'20)
Wavoice (SenSys'21)
RadioMic (arXiv 2021)

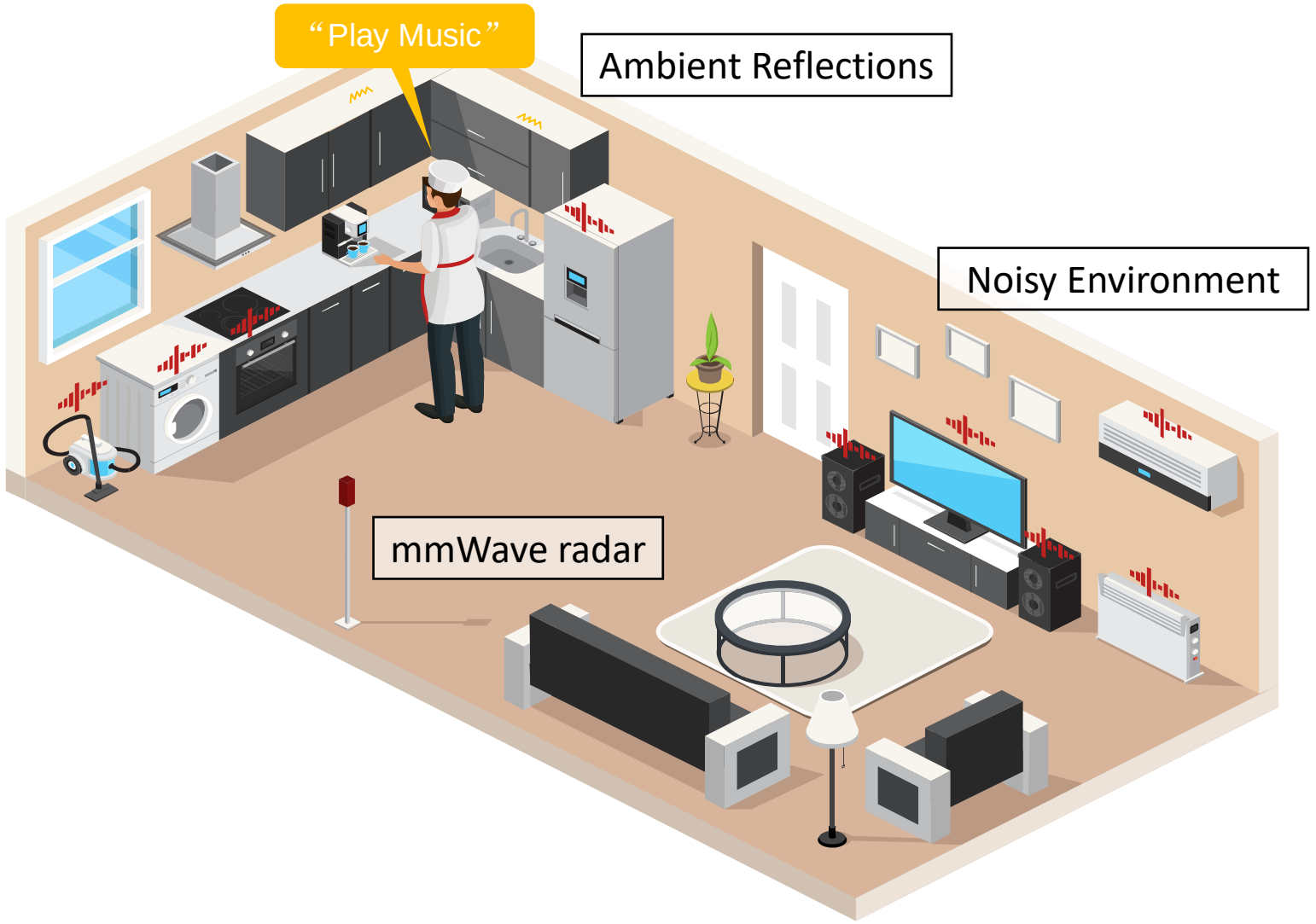
mmWave-based voice recognition in LoS scenarios

The existing direct sensing approaches can locate and track the throat in LoS scenarios.

mmWave-based voice recognition in NLoS scenarios?

The applicability of the existing direct sensing approaches is far from satisfactory in the real world.

Indirect sensing approach?



AmbiEar: mmWave based indirect sensing for voice recognition

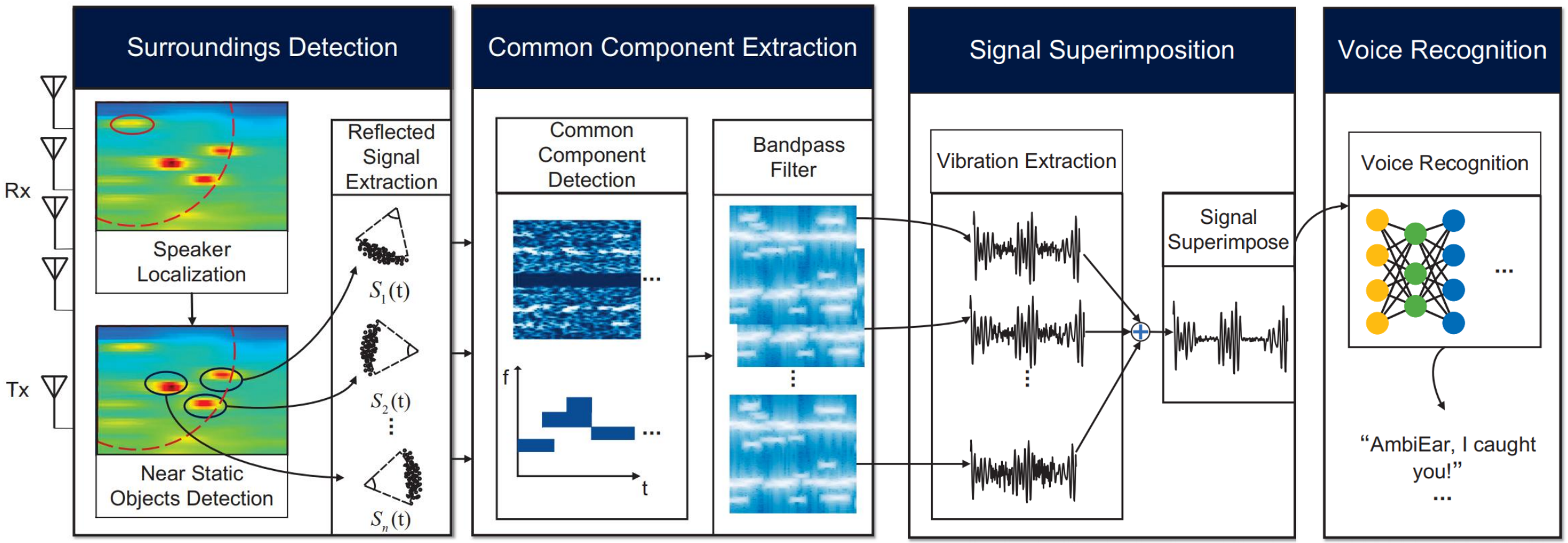
AmbiEar converts the surrounding objects into ambient “ears”.

Critical challenges from the low-quality voice-related signals

- The human location is unknown in advance and dynamic.
- The SNR of the reflected signals is low.
- The surrounding object’s vibration is intrinsically distorted.

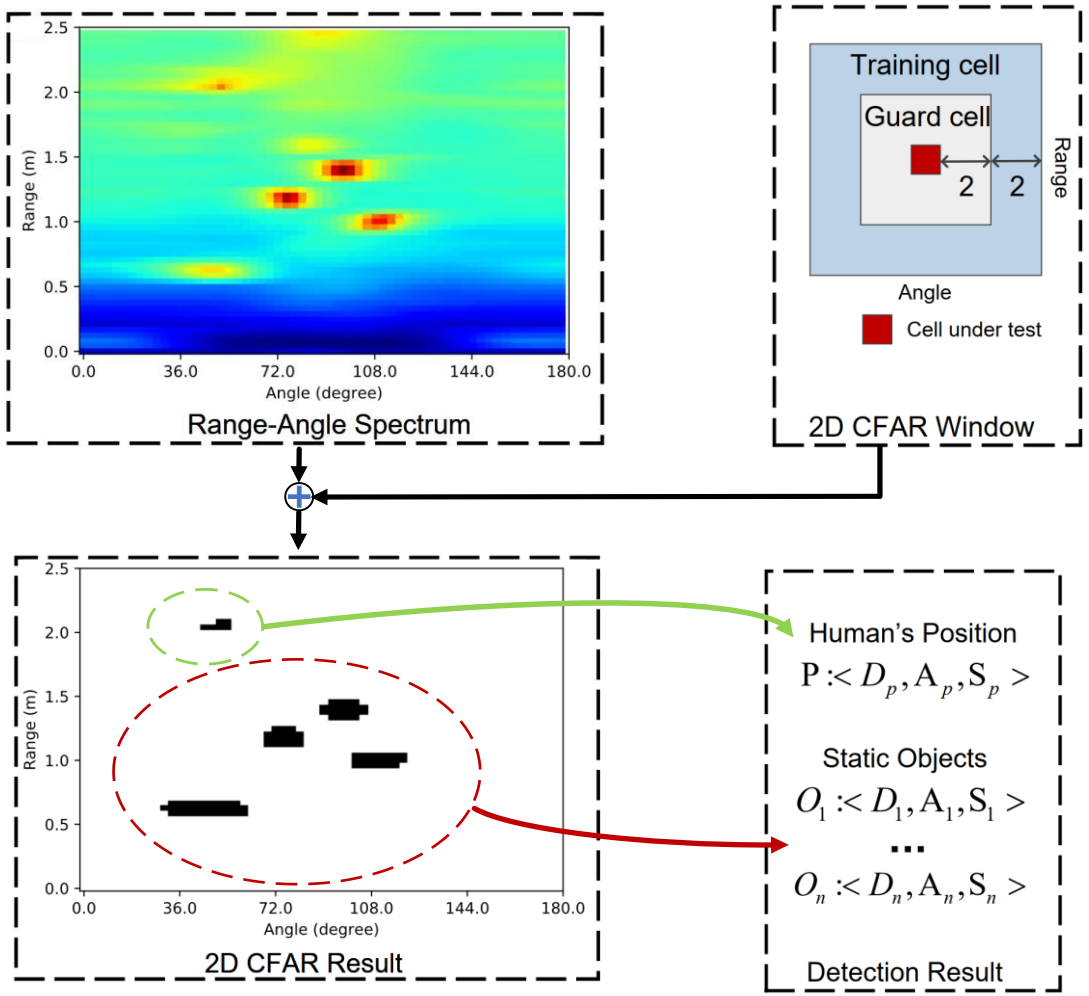
How does AmbiEar work?

System design



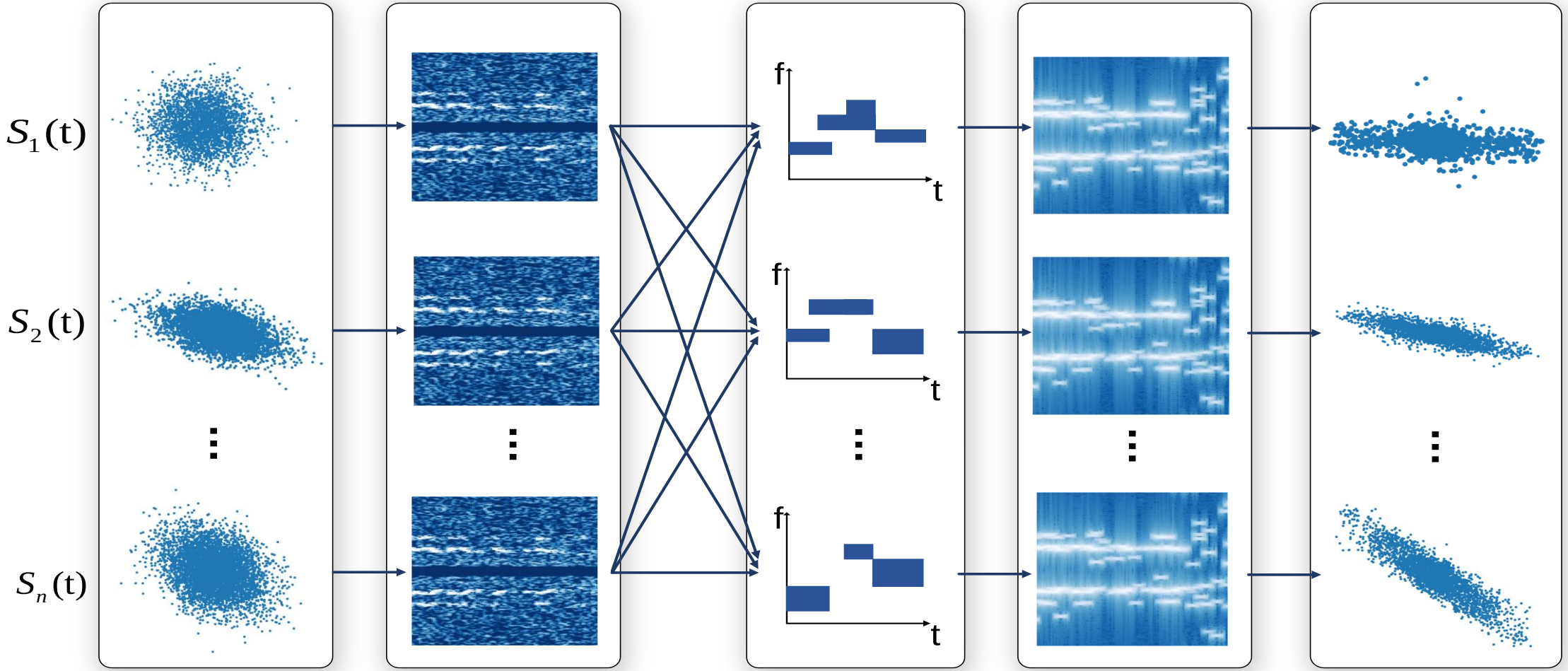
Surrounding detection

Challenge 1: The human location is unknown in advance and dynamic.



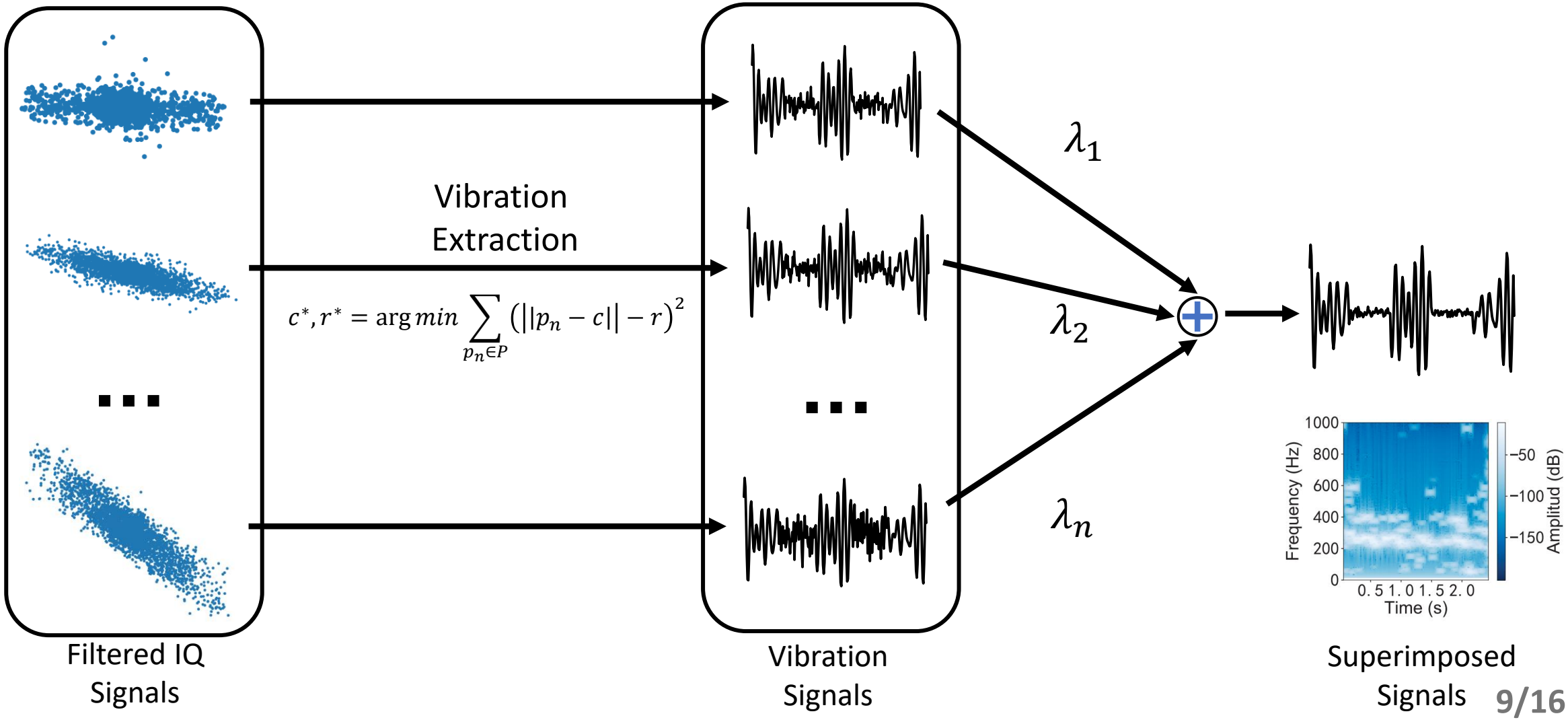
Common component extraction

Challenge 2: The SNR of the reflected signals is low.



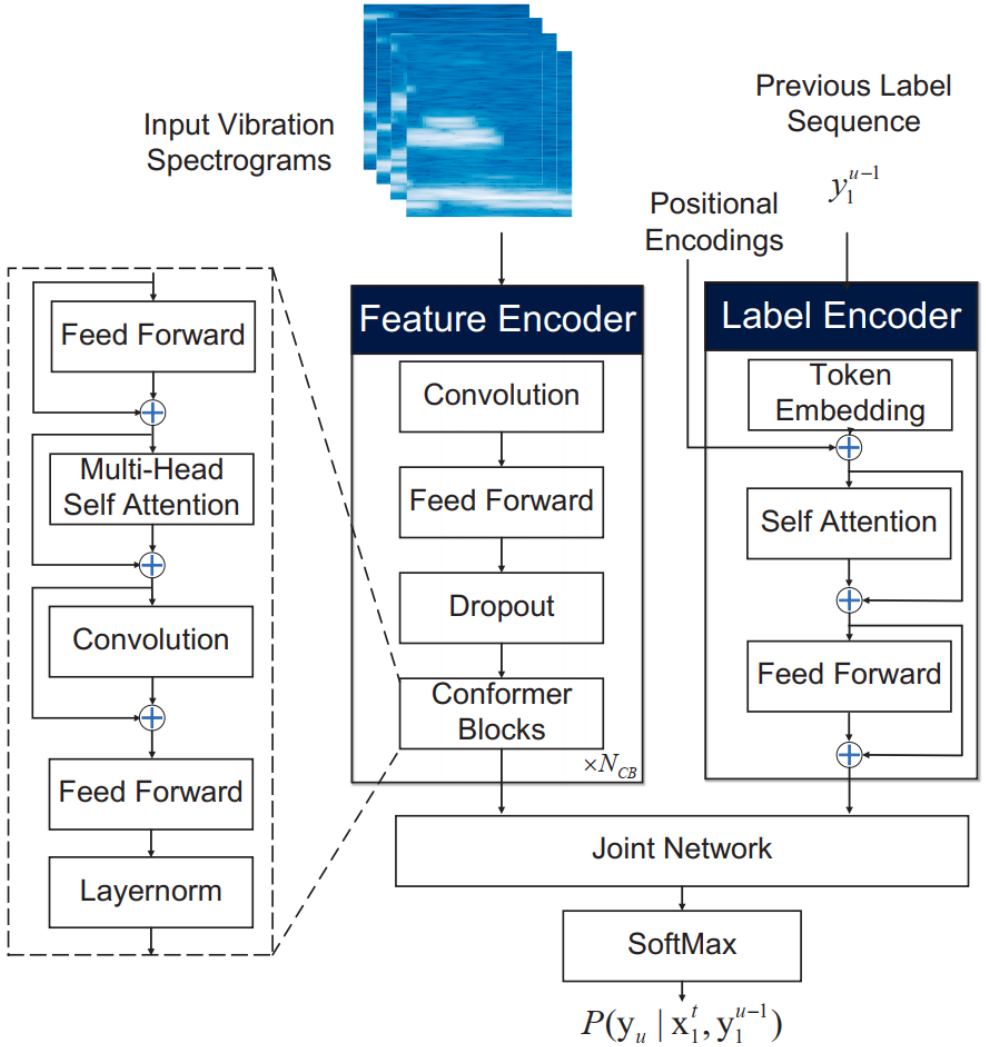
Signal superimposition

Challenge 2: The SNR of the reflected signals is low.

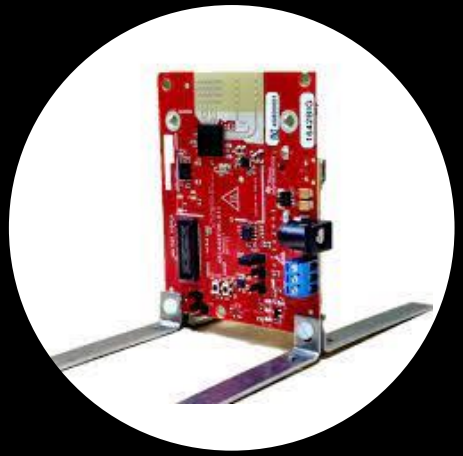


Voice Recognition

Challenge 3: The surrounding object's vibration is intrinsically distorted.

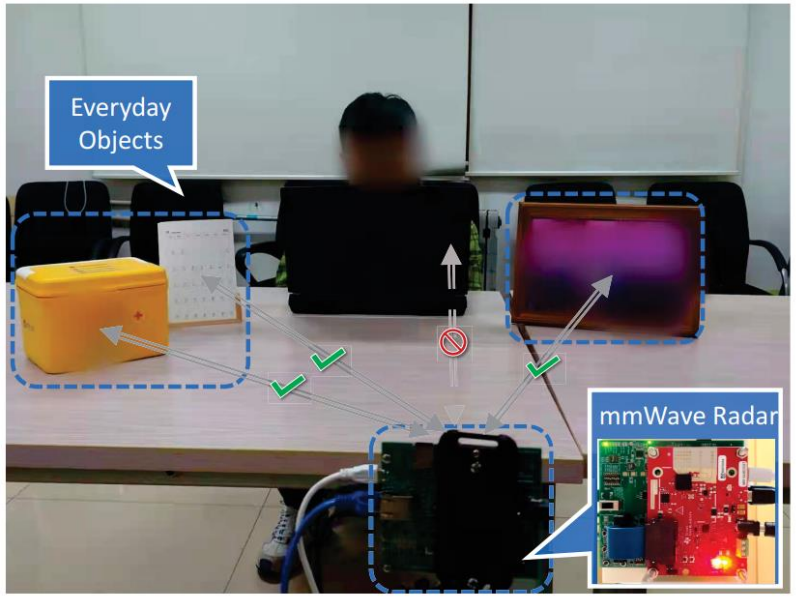


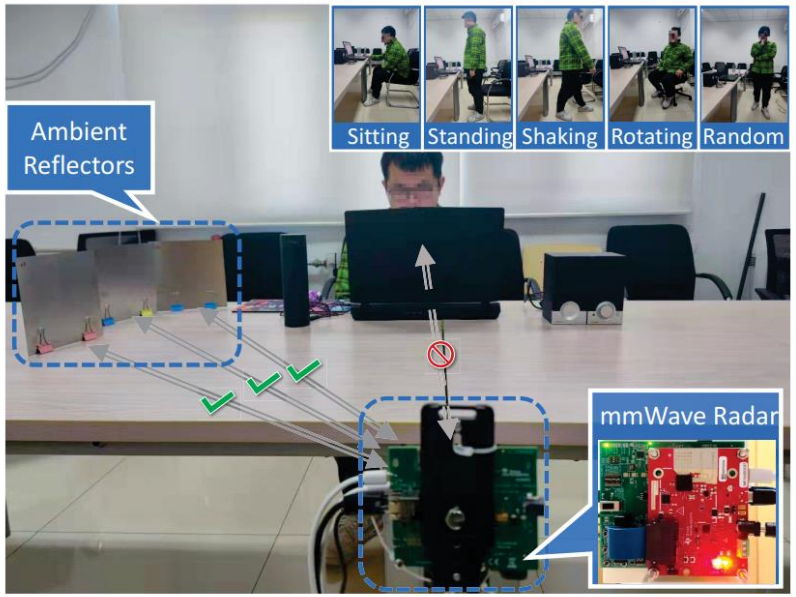
Evaluation



Implementation

- Implementation on a COTS mmWave radar TI IWR1642.
- A public voice data set TSRC is used to generate the training set (3000s).
- Six volunteers are instructed to generate the testing set (1000s).





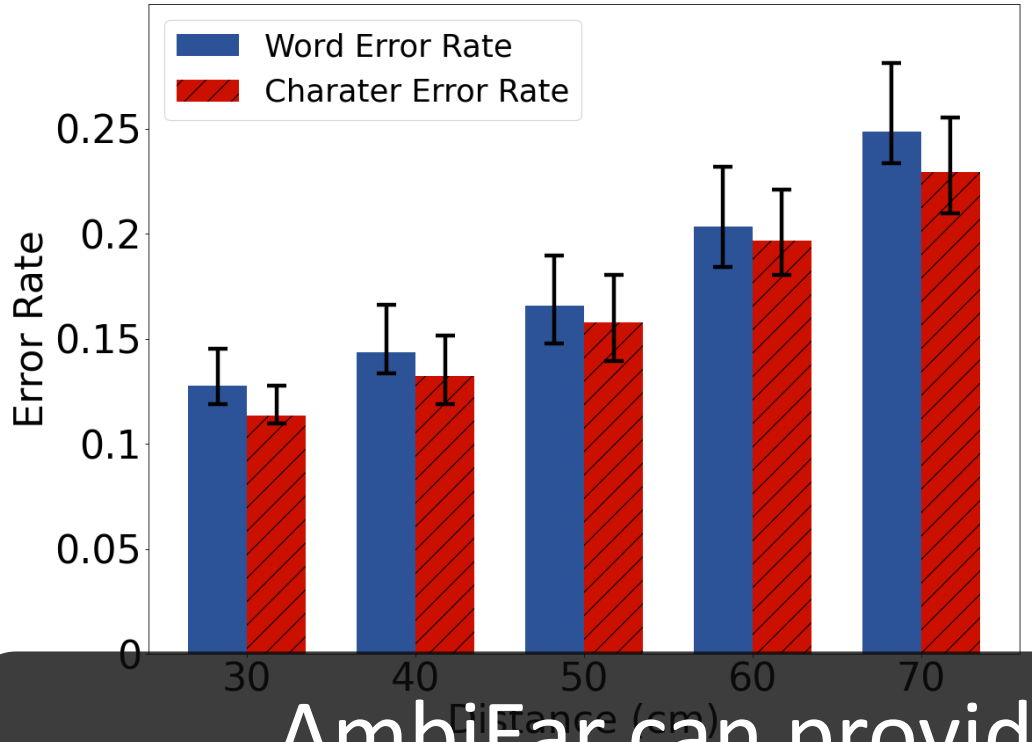
The experiment scenario

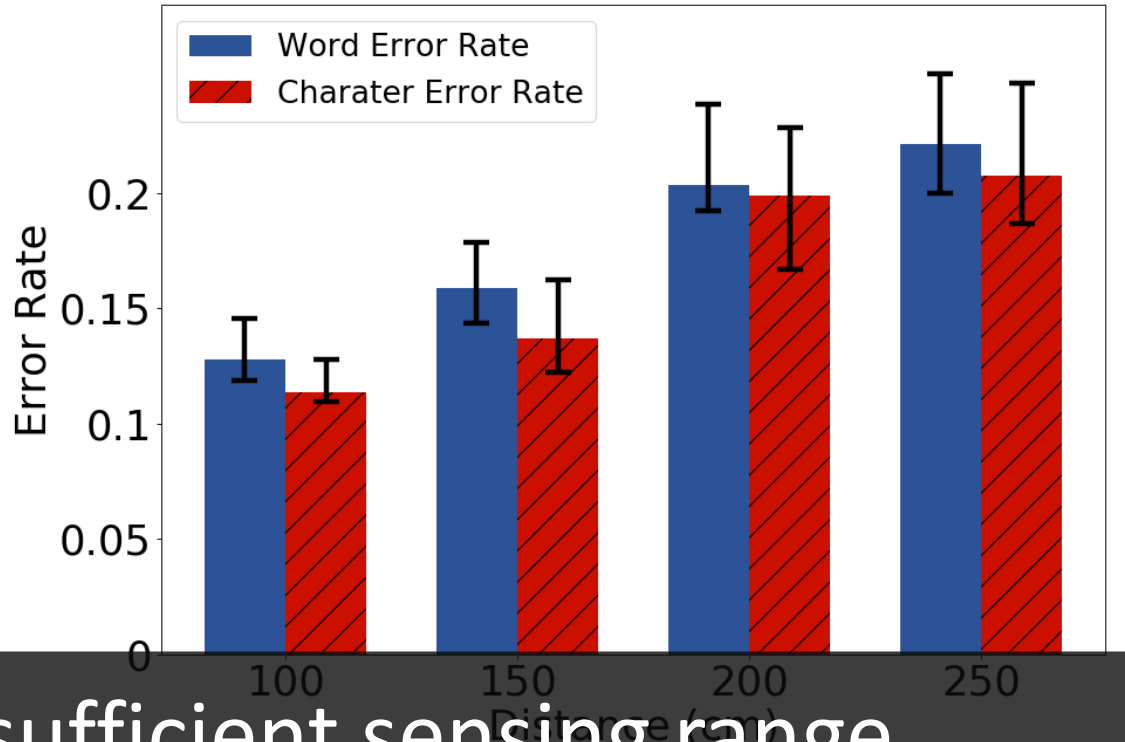
Overall performance

Scene	Method	WER		
		LoS	NLoS	Agg.
Meeting Room	AmbiEar	15.01%	16.19%	15.60%
	WaveEar	5.44%	95.92%	50.68%
Dormitory	AmbiEar	15.19%	16.58%	15.88%
	WaveEar	5.44%	96.37%	50.91%

3.2X

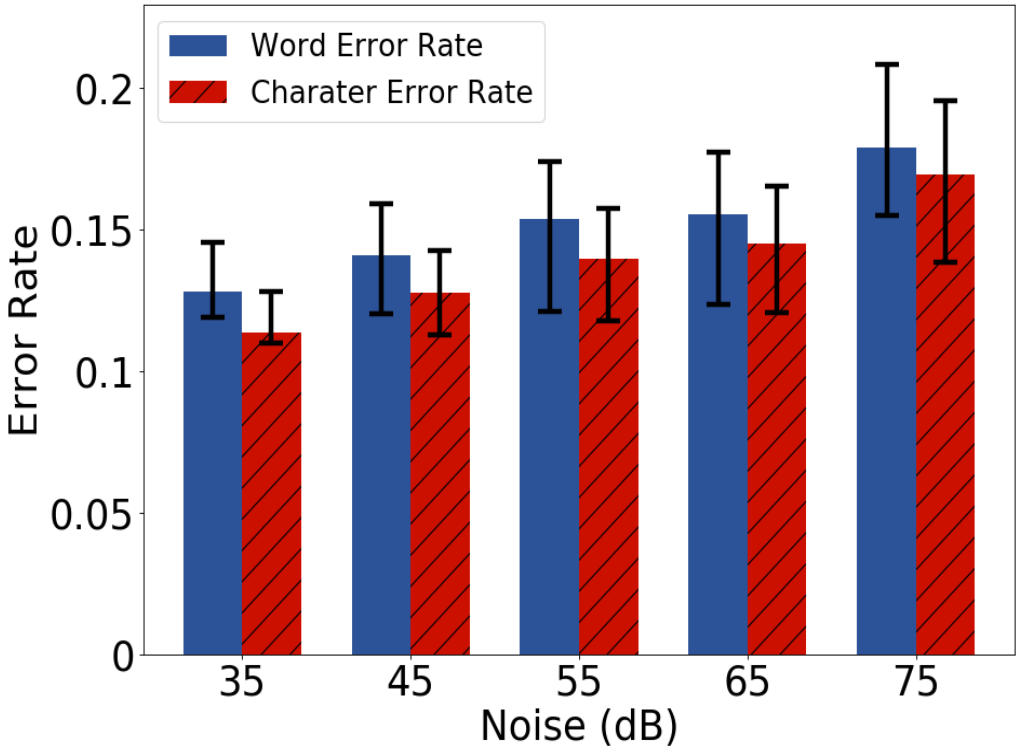
Accuracy v.s. distance

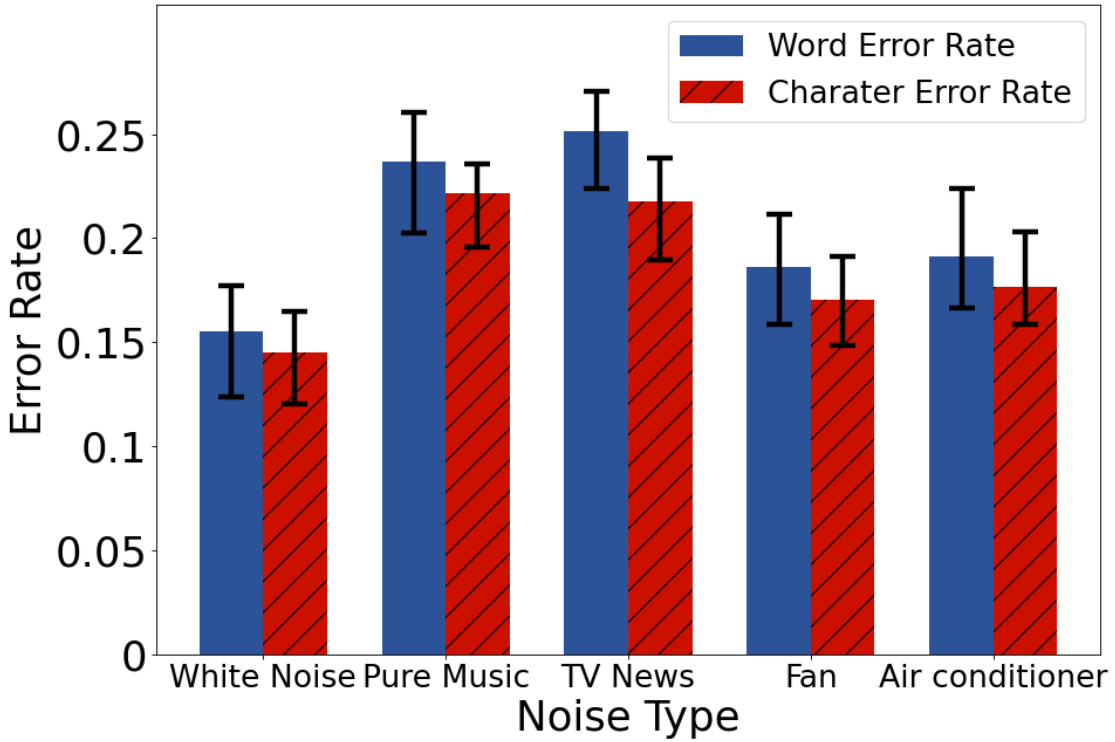




AmbiEar can provide sufficient sensing range in everyday scenarios.

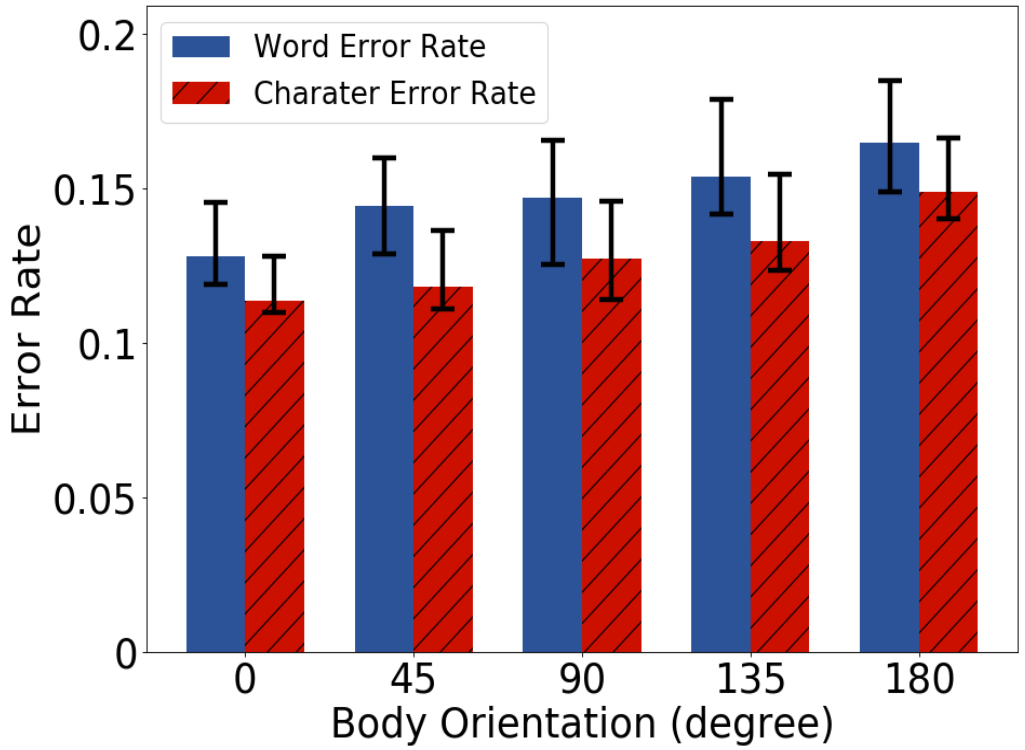
Accuracy v.s. noise

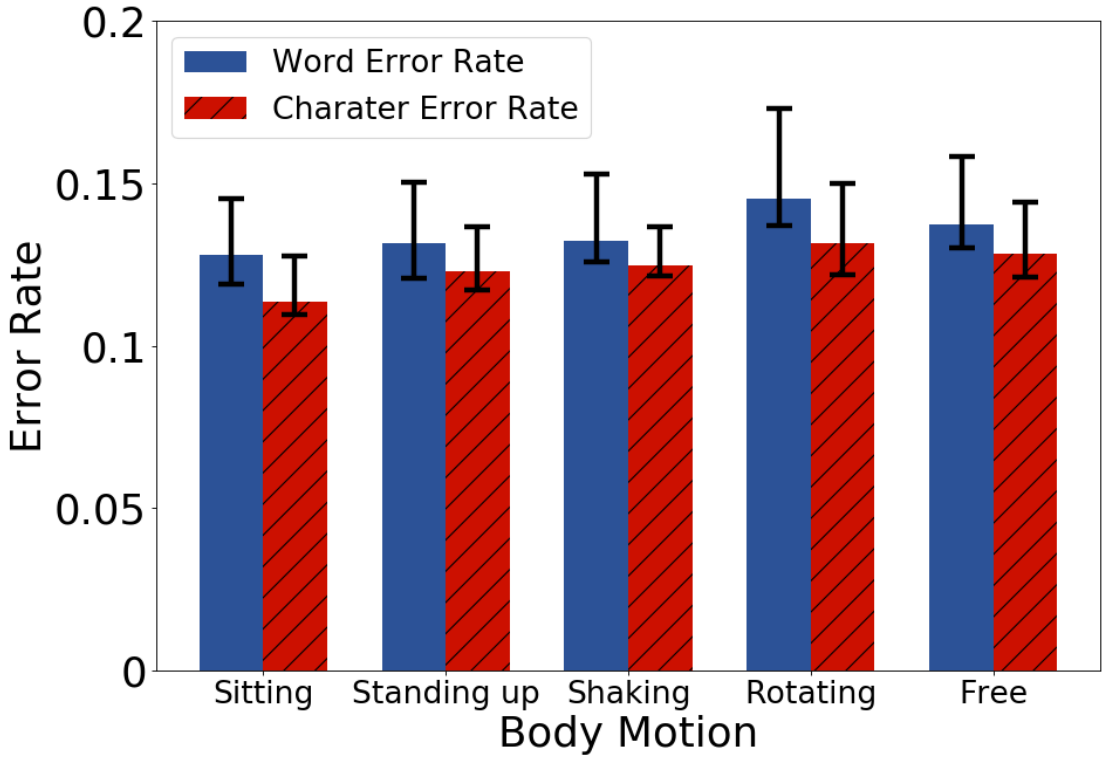




AmbiEar can handle various environment noise well.

Accuracy v.s. movement





AmbiEar is indeed applicable in dynamic scenarios.

AmbiEar

- First-of-its-kind approach for mmWave based voice recognition in NLoS scenarios.
- Provides a tailored design to utilize the low-SNR and semantically incomplete vibration signals for voice recognition.
- Implemented on the commercial device TI IWR1642 board and evaluated through experiments conducted under various settings.
- Has the potential to be applied to smart home, industrial control, smart cities, etc.

Jia Zhang
j-zhang19@mails.Tsinghua.edu.cn