

# Natural Voice Interface for the Next Generation of Smart Spaces

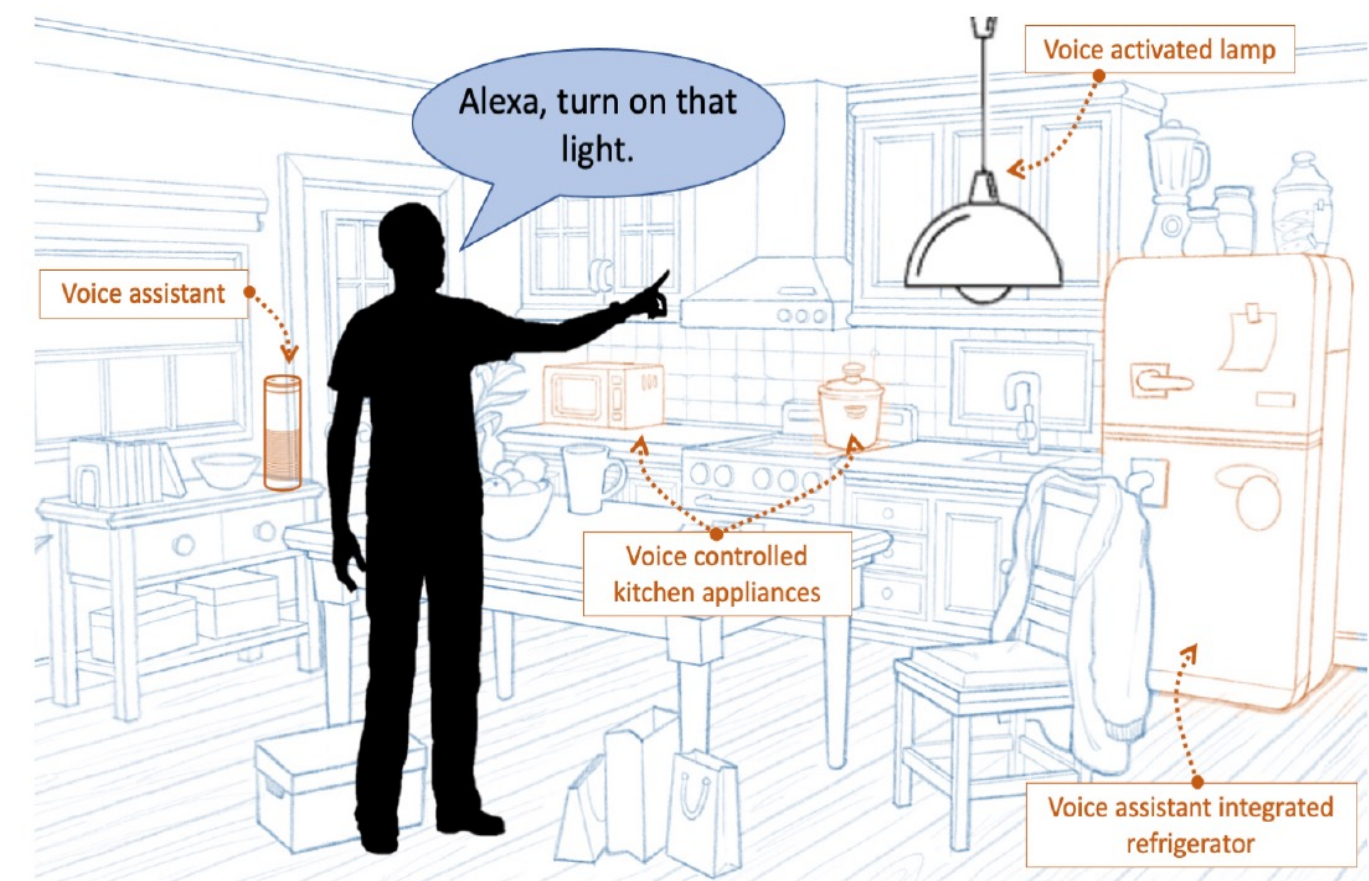


Yang Bai\*, Nakul Garg\*, Harshvardhan Takawale\*, Anupam Das $\diamond$ , Nirupam Roy\*  
 North Carolina State University $\diamond$ , University of Maryland College Park\*

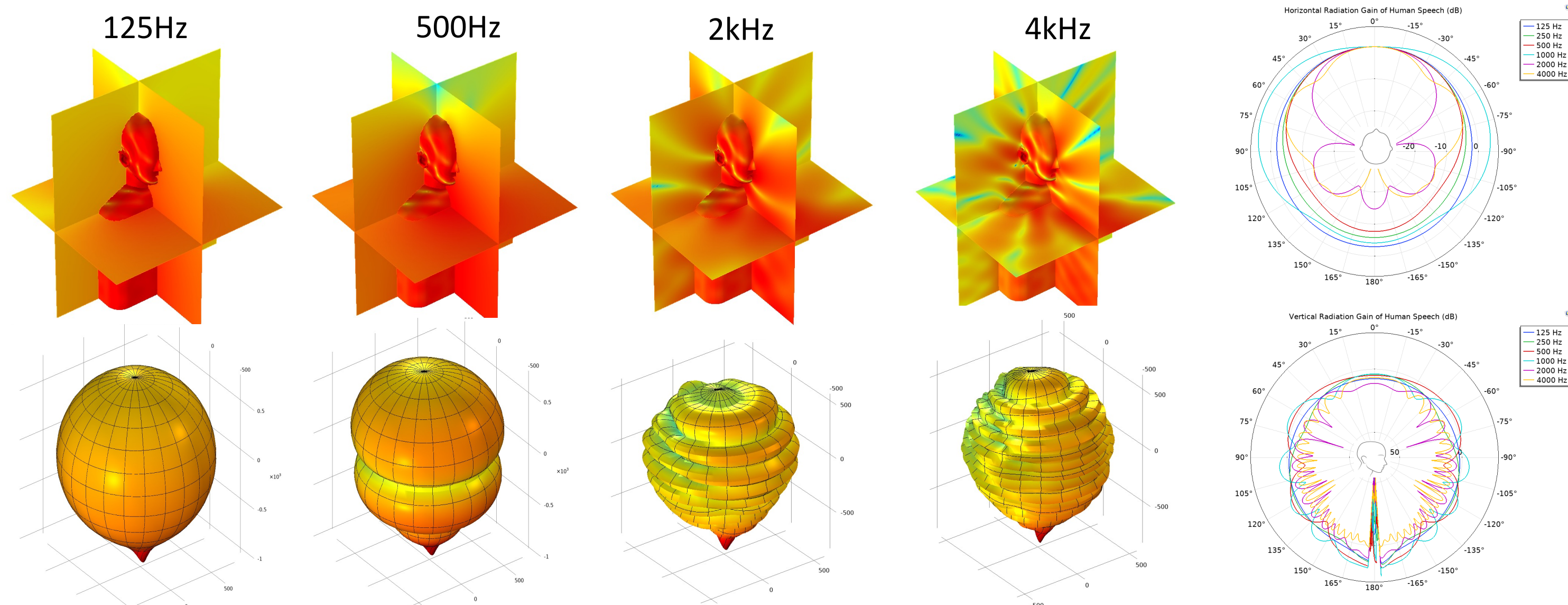


## 1. Vision

*Our vision: Combine directionality pattern of voice signal with multipath of the room to estimate if the speaker is facing towards a device.*

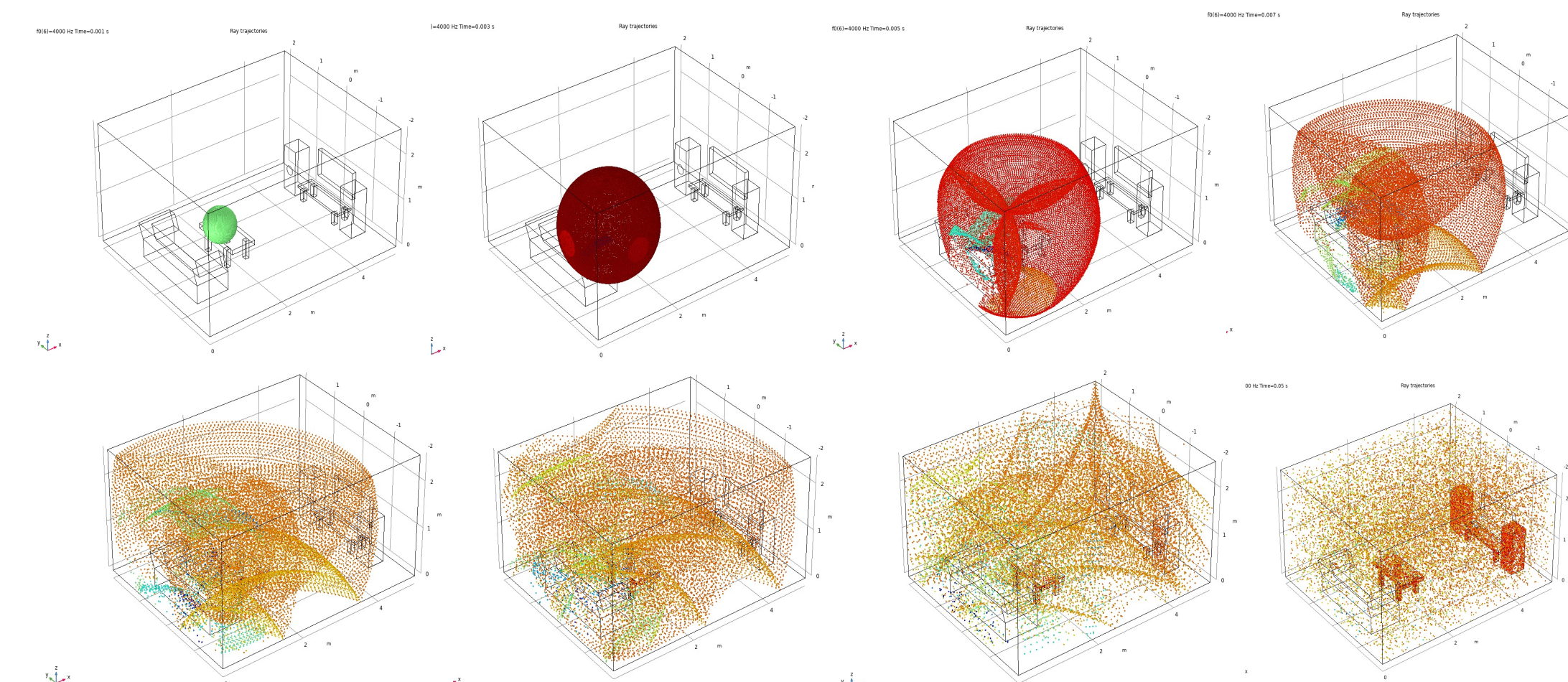


## 2. Polar pattern of voice signal



*Human voice shows directional diversity, especially in high frequencies.*

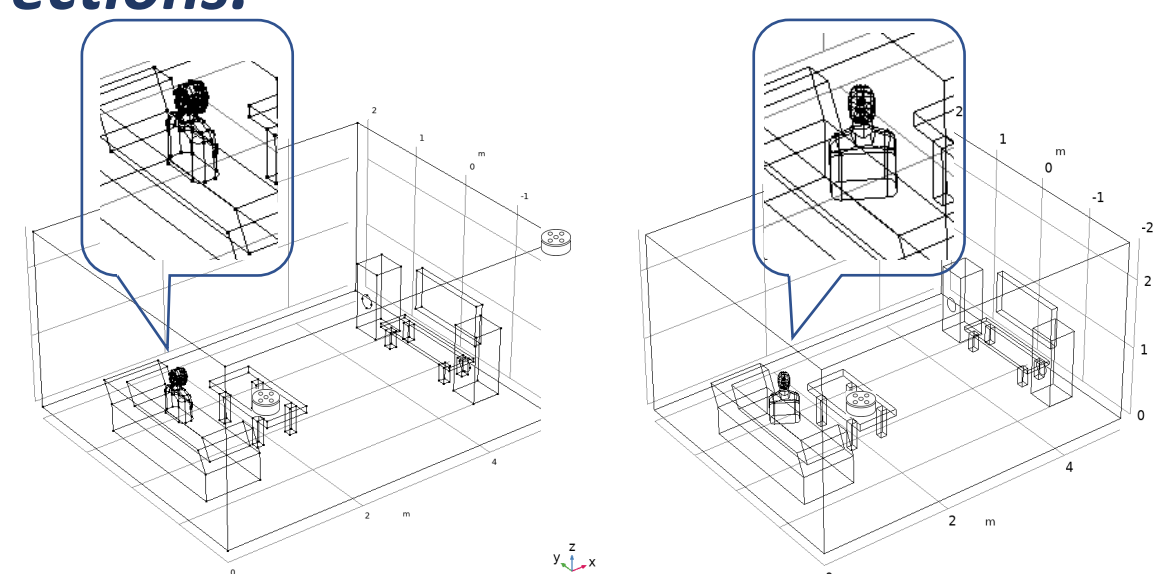
## 3. Ray tracing in room



*Environmental reflections induce multipath to voice signal.*

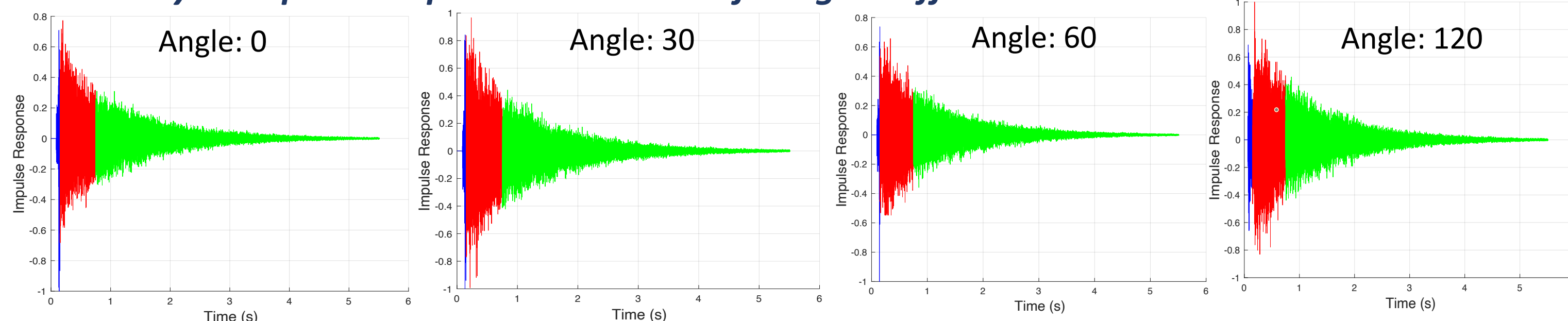
## 4. Simulation setup

*The speaker is facing towards different directions.*

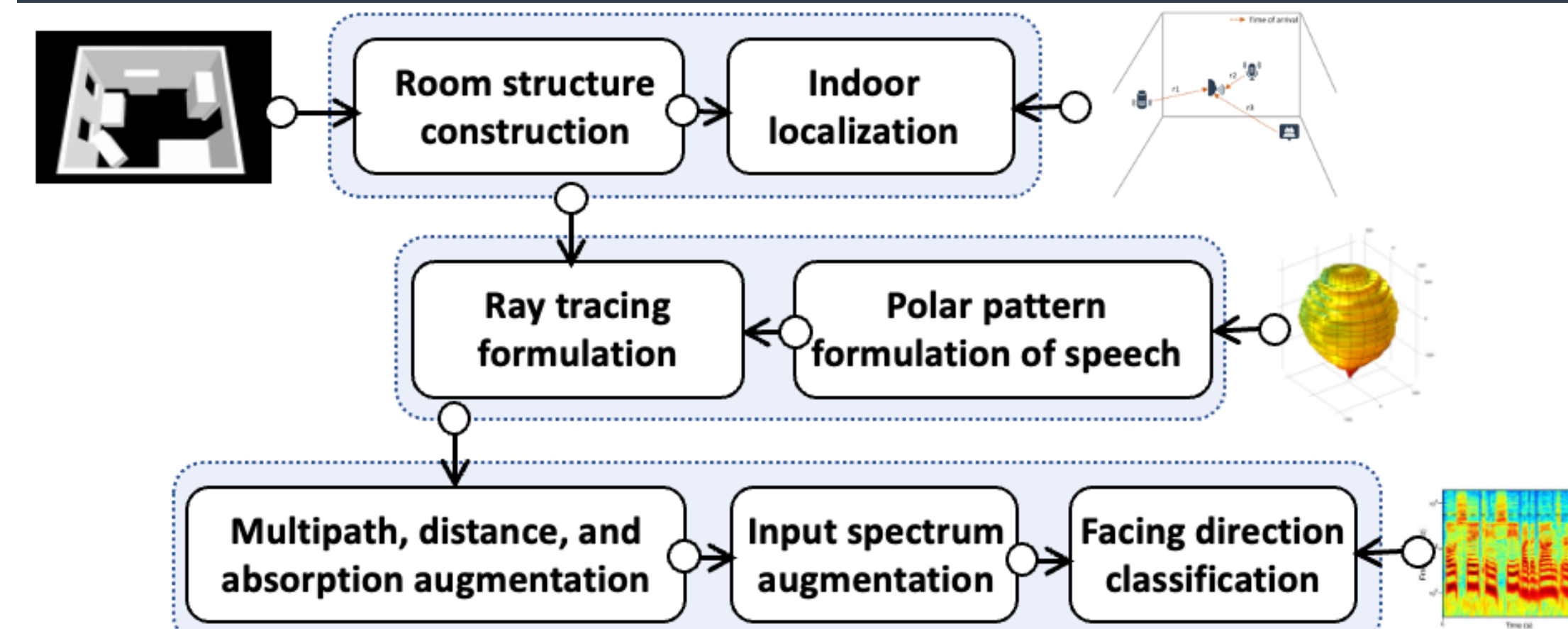


## 5. Feasibility study

*Diversity in impulse response exists when facing to different directions.*

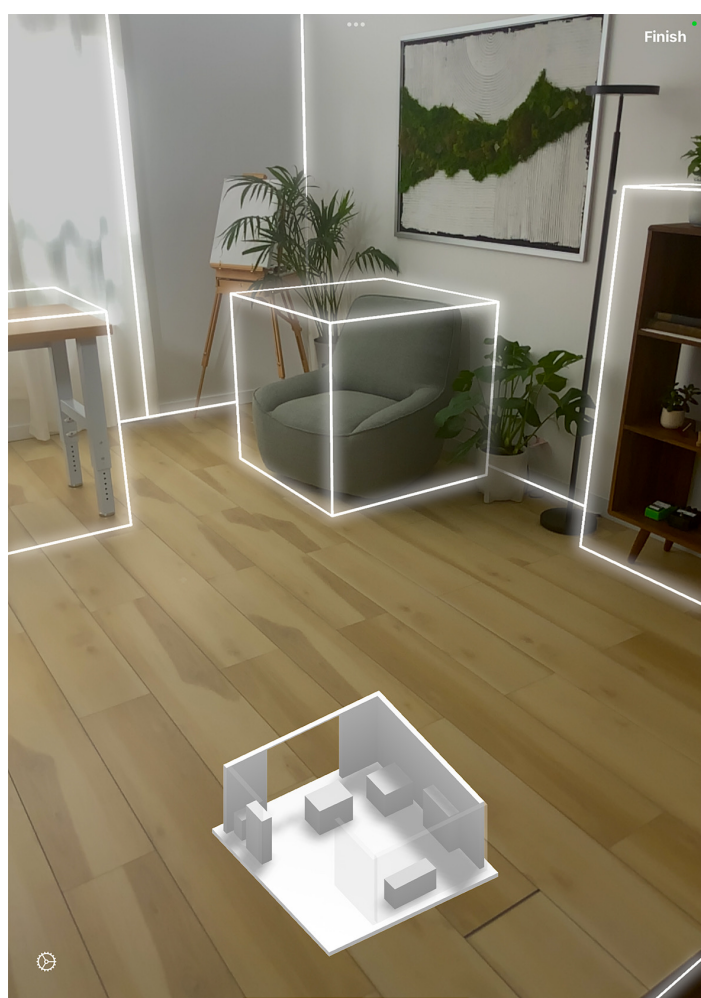


## 6. System design

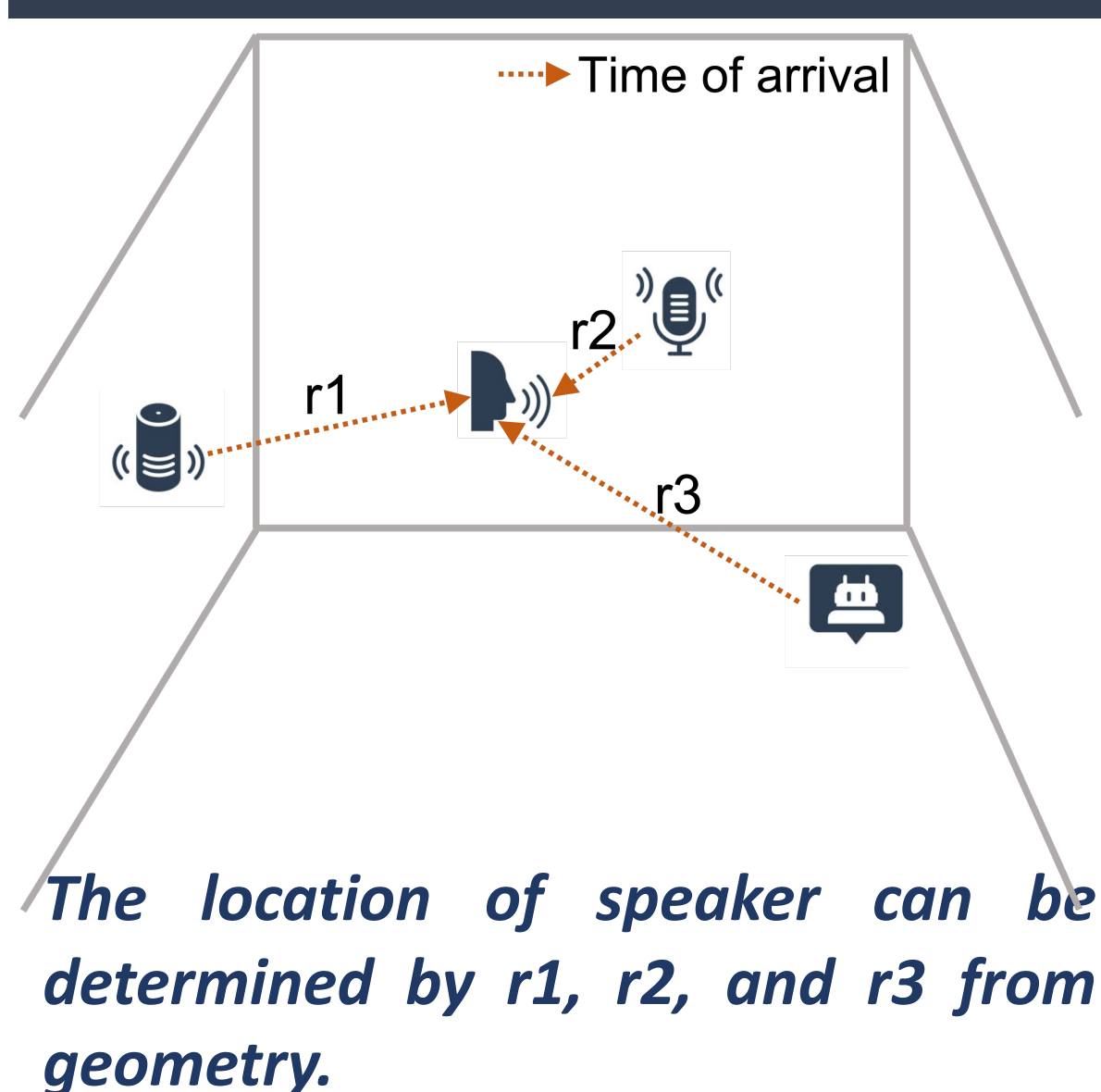


## 7. RoomPlan

*RoomPlan uses LiDAR to construct room structure.*

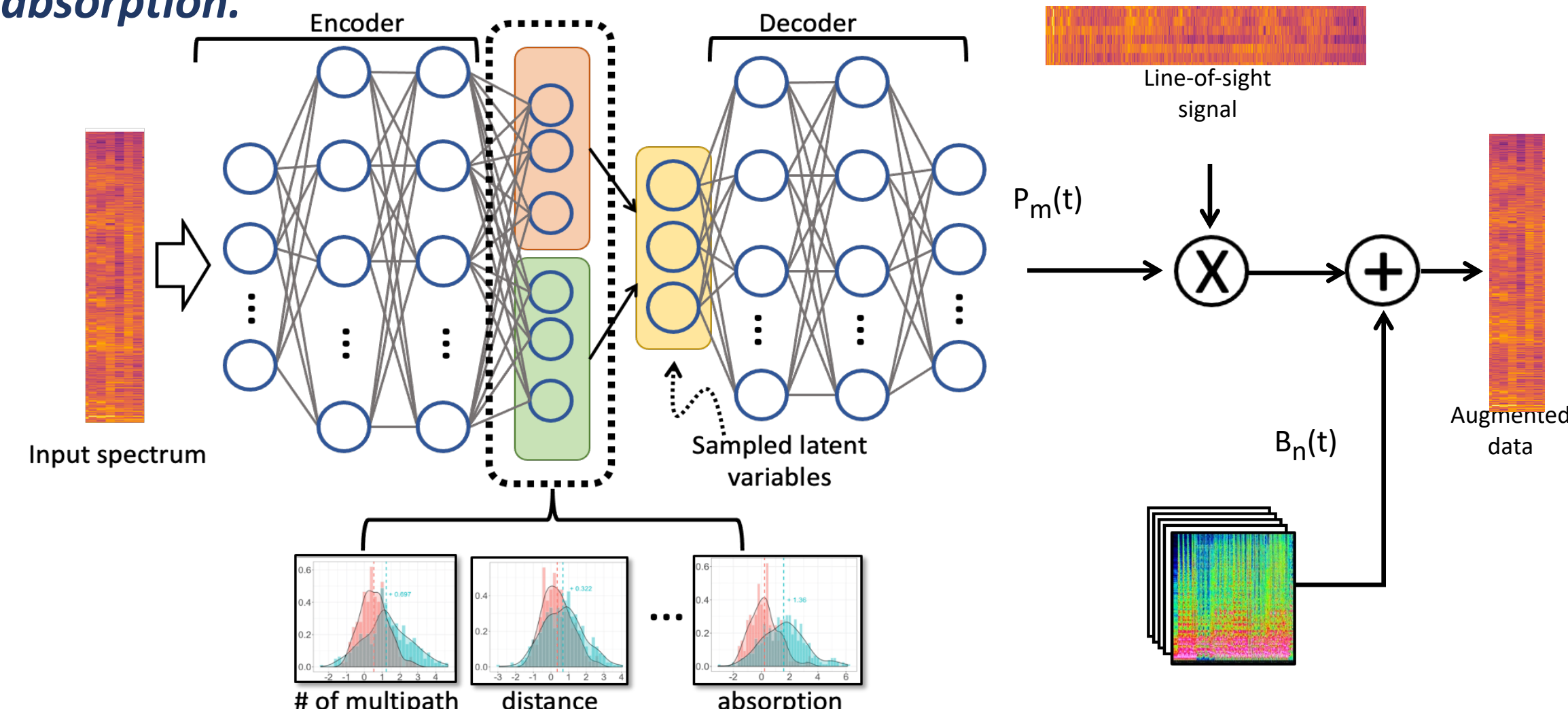


## 8. Indoor localization



## 9. Input spectrum augmentation

*Augment input spectrum using the constructed multipath, distance, and absorption.*



## 10. Facing direction classification

*Facing direction classification using the RNN-LSTM model trained with augmented data.*

