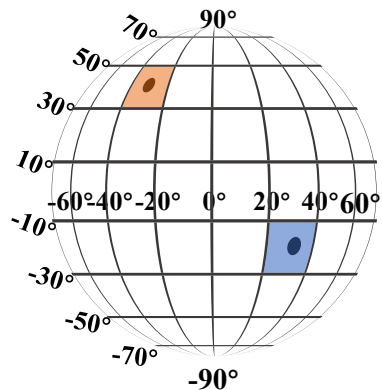
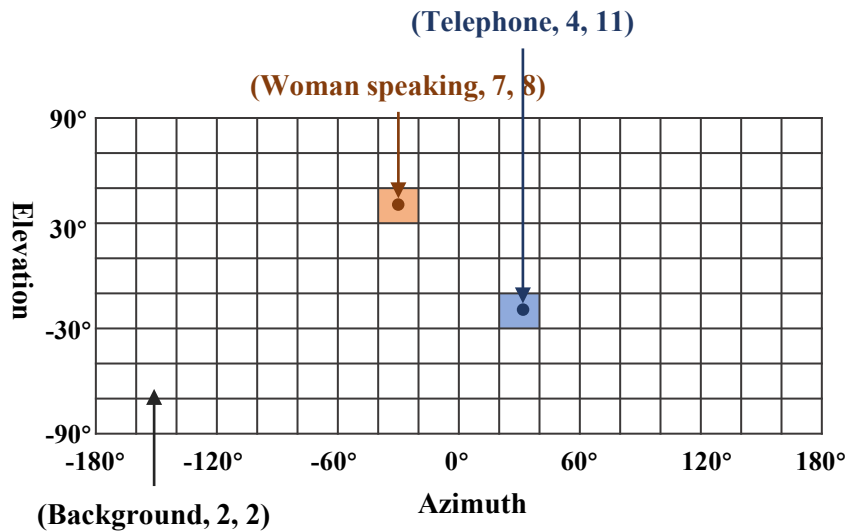


Spatial Mapping



3D to 2D plane
at frame t

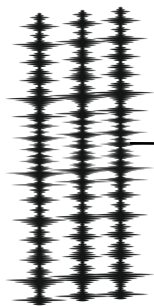


$$y \in R^{T' \times (I * J) \times M}$$

Localization
regression loss

$$\hat{y} \in R^{T' \times (I * J) \times M}$$

FOA format
acoustic signal



Network Structure

