## 8. Experiments

**Data Process:** The ground truth is generated using a RealSense depth camera. Specifically, the depth camera captures the 3D point cloud of the surrounding environment. This point cloud is then transformed into the radar coordinate system using pre-calibrated intrinsic and extrinsic matrices. To create ground truth for wall estimation, we project all 3D points onto the radar's horizontal (XZ) plane by collapsed along Z. This projection provides a reliable reference for evaluating our wall detection algorithm. For object detection, we first use the depth camera to identify the positions of objects in the scene. Then, using geometric cues from the depth data and simple measurement rules, we estimate the actual dimensions of each object. Based on the estimated position and size, we construct 2D bounding boxes in the radar coordinate system, which serve as ground truth for evaluating object detection performance.

## 9. Detail for Section - 4.2. Multipath Inversion

### 9.1. Ghost Target Formation and Identification

To make the ghost and human target identification process transparent, we provide a clear step-by-step pseudocode description in Algorithm 1.

### 9.2. Reflector Point Estimation

#### 9.2.1. First-Bounce Ghost

The first-bounce ghost point, denoted as $G_1'$, is geometrically related to the source $S'$ as follows:

$$2|sg_1'| = |sh| + |hc_1| + |sc_1|, \tag{11}$$

where $|sg_1'|$ represents the distance from the radar to the first-bounce ghost point, $|sh|$ is the distance from the radar to the human being, $|hc_1|$ is the distance from the human being to the reflector point, and $|sc_1|$ is the distance from the radar to the reflector point.

Applying the cosine law to the triangle formed by $S$, $C_1$, and $H$, we obtain:

$$|sh|^2 + |sc_1|^2 - 2|sh||sc_1|\cos(\theta_2^s - \theta_1^s) = |hc_1|^2. \tag{12}$$

where $\theta_2^s - \theta_1^s$ is the angle between $\vec{sh}$ and $\vec{sg_1}$ From the above relationships, we derive $|sc_1|$ as:

$$|sc_1| = \frac{2|sg_1'|^2 - 2|sg_1'||sh|}{2|sg_1'| - |sh|\cos(\theta_2^s - \theta_1^s) - |sh|} \tag{13}$$

The coordinates of the reflector point $C_1$ can then be determined as:

$$c_1 = s + \frac{\vec{sg'}}{|sg'|} \cdot |sc_1| \tag{14}$$

where $\frac{\vec{sg'}}{|sg'|}$ is the unit vector in the direction of $sg'$.

**Algorithm 1:** Ghost Target Formation and Identification

**Input:** Clusters $\mathcal{C} = \{(r_i, \theta_i, m_i)\}$ obtained from CFAR detection, where $r_i$ denotes the range, $\theta_i$ the angle, and $m_i$ the reflection magnitude. Threshold ratio $\tau = 0.4$; Range tolerance $\delta_r = 0.15$m; Angle tolerance $\delta_\theta = 15°$.

**Output:** Human $H$, first-order ghosts $G_1, G_1'$, second-order ghosts $G_2, G_2'$.

**Step 1: Identify Human $H$;**

$M_{\max} \leftarrow \max_i m_i$

$\mathcal{C}_{\text{valid}} \leftarrow \{c \in \mathcal{C} \mid m_c > \tau M_{\max}\}$

$H \leftarrow \arg\min_{c \in \mathcal{C}_{\text{valid}}} r_c$

**Step 2: Identify $G_1$ (same direction, slightly larger range);**

$\mathcal{G}_1 \leftarrow \{c \in \mathcal{C} \mid |\theta_c - \theta_H| < \delta_\theta,\ 0 < r_c - r_H < \delta_r\}$

$G_1 \leftarrow \arg\min_{c \in \mathcal{G}_1} r_c$ if $\mathcal{G}_1 \neq \emptyset$, else $\emptyset$

**Step 3: Identify $G_1'$ (same range, different direction);**

$\mathcal{G}_{1'} \leftarrow \{c \in \mathcal{C} \mid |r_c - r_{G_1}| < \delta_r,\ |\theta_c - \theta_{G_1}| \geq \delta_\theta\}$

$G_1' \leftarrow \arg\max_{c \in \mathcal{G}_{1'}} m_c$ if $\mathcal{G}_{1'} \neq \emptyset$, else $\varnothing$

**Step 4 (Optional): Identify $G_2$ (same direction as $H$, farther than $G_1$);**

$\mathcal{G}_2 \leftarrow \{c \in \mathcal{C} \mid |\theta_c - \theta_H| < \delta_\theta,\ r_c > r_{G_1}\}$

$G_2 \leftarrow \arg\min_{c \in \mathcal{G}_2} r_c$ if $\mathcal{G}_2 \neq \emptyset$, else $\varnothing$

**Step 5: Identify $G_2'$ (aligned with $\vec{G_1'}$, farther range);**

$\mathcal{G}_{2'} \leftarrow \{c \in \mathcal{C} \mid r_c > r_{G_1'},\ |\theta_c - \theta_{G_1'}| < \delta_\theta\}$

$G_2' \leftarrow \arg\min_{c \in \mathcal{G}_{2'}} r_c$ if $\mathcal{G}_{2'} \neq \emptyset$, else $\varnothing$

**return** $H, G_1, G_1', G_2, G_2'$

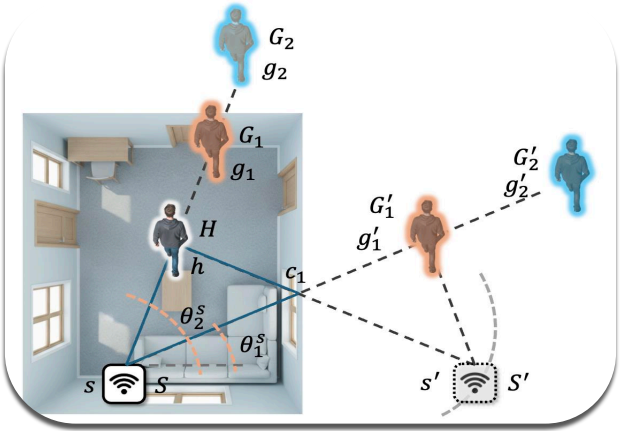


Figure 10. **Multipath Inversion.** Geometric relationships between ghost targets and reflectors.

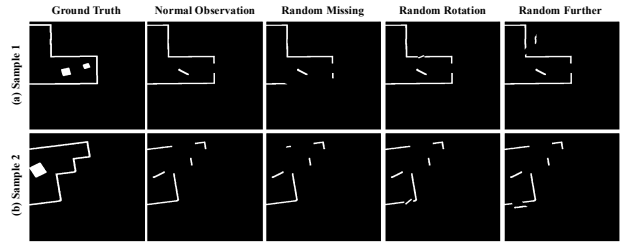


Figure 11. **Illustration of Simulator and Data Augmentation.** The first column represents the ground truth in the simulator, while the second column shows partial observations resulting from occlusion. The third column depicts randomly missing regions, the fourth column illustrates random rotations, and the fifth column demonstrates random scaling.

#### 9.2.2. Second-Bounce Ghost

The second-bounce ghost point, denoted as $G_2'$, follows a geometric relationship similar to the first-bounce case. The total path length from the radar to $G_2'$ can be expressed as:

$$|sg_2'| = |hc_1| + |sc_1|, \tag{15}$$

where $|sg_2'|$ represents the total distance from the radar to the second-bounce ghost point, $|hc_1|$ is the distance from the human to the reflector, and $|sc_1|$ is the distance from the radar to the reflector.

Using the cosine law and substituting the known relationships, we derive:

$$|sc_1| = \frac{|sg_2'|^2 - |sh|^2}{2|sg_1'| - 2|sh|\cos(\theta_2^s - \theta_1^s)}, \tag{16}$$

where $\theta_1^s$ and $\theta_2^s$ denote the respective angles between the reflection points and the radar.

Subsequently, the remaining steps outlined in the first-bounce ghost section can be applied to determine the location of the mirror radar.

### 9.3. Post Processing Reflector Points

After estimating reflector points from ghost targets, we obtain a sparse and noisy set of 2D samples that approximate the underlying wall and object boundaries. To extract coherent geometric structures, we refine these points in two stages.

First, we apply a Gaussian Mixture Model (GMM) to group the reflector points into spatially coherent clusters. The GMM models the distribution of reflector points as a mixture of Gaussian components, which enables soft probabilistic assignments and naturally captures elongated point distributions associated with physical surfaces. This clustering step separates points belonging to different walls or objects while remaining robust to uneven sampling and local noise.

Next, for each cluster, we perform RANSAC line fitting to recover the dominant structural direction and reject outliers introduced by multipath noise or incomplete reflections. RANSAC iteratively samples minimal point sets, fits line hypotheses, and selects the model with the largest inlier set under a point-to-line distance threshold. This procedure produces clean and geometrically consistent line segments for each cluster.

The resulting set of RANSAC-refined line structures forms a robust initial layout estimate, which is subsequently fed into our diffusion-based refinement module to obtain complete wall boundaries and object footprints.

## 10. Discussion

RISE demonstrates that a single static mmWave radar can reliably recover indoor layout and object information by leveraging multipath reflections and generative modeling. While our system has several limitations, we emphasize that it remains highly practical and broadly useful in many real-world settings.

A key limitation is that RISE relies on human motion to stimulate diverse multipath paths. This motion provides the geometric variation necessary for uncovering occluded regions, and completely static environments remain challenging. However, this requirement aligns well with many deployment scenarios: homes, clinics, elder-care facilities, offices, and rehabilitation environments naturally contain frequent human movement. In such settings, RISE can operate fully passively and unobtrusively, without requiring additional devices, active user participation, or intrusive sensors. Even short or routine movements, walking through a hallway or moving to sit in a chair, can generate sufficient multipath diversity for reliable reconstruction.

Second, blind regions inherent to fixed sensor placement cannot be eliminated entirely, although our bi-angular enhancement significantly reduces them. These blind spots typically correspond to uncommon reflection geometries and often do not affect the global layout.

Third, the current system outputs 2D top-down geometry rather than full 3D meshes or semantic object models. While this limits fine-grained reconstruction, top-down geometry is already sufficient for many high-impact applications such as indoor navigation, elder monitoring, health assessment, fall detection, and safety analysis—domains where privacy constraints prohibit cameras.

Despite these limitations, RISE remains an effective, privacy-preserving perception system capable of recovering meaningful indoor structure from a single static radar. It offers a lightweight, low-cost alternative to camera- or LiDAR-based approaches, especially in privacy-sensitive environments. These strengths, combined with its ability to operate passively using natural human motion, underscore its practical value and establish a foundation for future advances in radar-based indoor understanding.