

RoViSense: Active Vital-Sign Monitoring Leveraging Humanoid-Assisted mmWave Sensing

Lei Zhang
Yeshiva University
New York, NY, USA

Yucheng Xie^{*}
Yeshiva University
New York, NY, USA

Lingxi Li
Purdue University
Indianapolis, IN, USA

Honggang Wang
Yeshiva University
New York, NY, USA

Abstract—Continuous vital-sign monitoring is moving from clinical settings into homes and everyday environments, creating a need for unobtrusive sensing that does not require body-worn devices or visual imagery. Millimeter-wave (mmWave) sensing offers a contactless alternative, but most existing mmWave vital-sign systems rely on fixed or manually placed sensing positions. As a result, unfavorable viewing angles or standoff distances can substantially degrade heart-rate extraction. We develop RoViSense, an active mmWave vital-sign monitoring system in which a humanoid-mounted sensor selects and reaches a favorable sensing viewpoint before measurement. RoViSense estimates chest geometry from sparse mmWave point clouds, selects a viewpoint based on radial alignment and standoff distance, navigates the humanoid platform to that viewpoint under collision constraints, and then extracts heart rate from the stabilized mmWave phase signal. Across 8 subjects and varied sensing geometries, RoViSense estimates chest position with a mean MAE of 2.65 cm and selects viewpoints with a median position error of 3.0 cm relative to ground-truth references. Repositioning to the selected viewpoint reduces heart-rate MAE from 11.73 to 4.51 BPM, demonstrating the benefit of active viewpoint selection for contactless mmWave vital-sign monitoring.

Index Terms—mmWave Sensing, Active Viewpoint Selection, Humanoid Robot, Contactless Vital-Sign Monitoring

I. INTRODUCTION

Continuous, unobtrusive vital-sign monitoring is increasingly important as health sensing moves from clinical settings into homes and everyday environments. In these settings, people need practical ways to track physiological status without dedicated medical infrastructure. Consumer wearable devices, such as smartwatches and chest straps, require body contact and can be inconvenient for long-term or overnight monitoring [1]. Camera-based methods such as remote photoplethysmography avoid contact, but they remain sensitive to illumination and may raise privacy concerns in personal spaces such as bedrooms [2], [3].

Millimeter-wave (mmWave)-based sensing provides a contactless alternative for vital-sign monitoring. Its short wavelength makes it sensitive to cardiac micro-displacements [4]–[6]. However, the contactless nature of mmWave sensing does not eliminate its dependence on sensing geometry. Most existing works rely on fixed or manually configured sensing positions, leaving the viewing angle and sensor-to-subject distance largely determined at deployment time. Because mmWave

measurements primarily capture the radial component of chest-wall displacement, an oblique view reduces the observable cardiac motion, while an inappropriate standoff distance can weaken the reflected signal and destabilize thoracic localization. Consequently, even with the same sensing hardware and signal-processing pipeline, measurement quality can vary substantially across sensing geometries [7]. In practice, this often requires the user to adapt to the sensor rather than the sensing system adapting to the user, limiting long-term and continuous monitoring in everyday environments.

This limitation motivates a different design principle: sensing geometry should be treated as a controllable part of the vital-sign sensing pipeline rather than a fixed deployment choice. Instead of requiring the subject to remain at a favorable angle and distance, a mobile sensing platform can reposition the mmWave sensor with respect to the chest and establish a more favorable sensing configuration before measurement. Humanoid platforms are particularly suitable for this role because they are designed for human-centered indoor environments, can carry sensors at body-relevant heights, and can adjust their body pose without modifying the surrounding space [8], [9]. These properties make a humanoid platform a natural carrier for adaptive mmWave sensing in everyday indoor settings.

Realizing adaptive mmWave vital-sign sensing introduces several challenges. First, the system must estimate the subject’s chest location and facing direction from sparse mmWave observations without relying on visual sensing. Second, it must translate the estimated chest geometry into a sensing viewpoint that jointly provides favorable radial alignment and standoff distance while remaining feasible for the humanoid platform. Third, the platform must reach and stabilize at the selected viewpoint under environmental constraints before recovering the weak cardiac component from the mmWave phase signal.

To address these challenges, we develop **RoViSense**, an active mmWave vital-sign monitoring system organized as a perception–action pipeline. RoViSense first converts onboard mmWave measurements into a subject point cloud and a local obstacle map, and then estimates the chest centroid and chest-plane normal from the subject representation. Based on this geometry, it evaluates candidate sensing viewpoints according to radial alignment and standoff distance and passes the selected viewpoint to a collision-aware navigation module. After the humanoid reaches and stabilizes at the selected position, RoViSense isolates the thoracic return and estimates heart rate

^{*}Yucheng Xie is the corresponding author. Email: yucheng.xie@yu.edu.

from the mmWave phase signal. Our main contributions are summarized as follows:

- 1) We propose RoViSense, a humanoid-assisted active mmWave vital-sign sensing system that treats sensing geometry as an adaptive sensing variable rather than a fixed deployment choice.
- 2) We develop a chest-geometry-aware viewpoint adaptation method that jointly optimizes radial alignment and standoff distance while accounting for humanoid feasibility and collision constraints.
- 3) We implement and evaluate RoViSense with a humanoid-mounted mmWave sensing platform. Across 8 subjects and varied sensing geometries, repositioning reduces heart-rate MAE from 11.73 to 4.51 BPM, while the selected viewpoints achieve a median position error of 3.0 cm relative to reference viewpoints.

II. RELATED WORK

A. Wearable and Camera-based Vital-Sign Monitoring

Prior work has explored wearable and camera-based approaches for vital-sign monitoring. Wearable ECG and PPG devices are effective for continuous heart-rate tracking, yet they rely on sustained body contact [1], [10]. This contact can be inconvenient during long-term or overnight monitoring, especially for users with low tolerance for body-worn devices. Camera-based methods, including remote photoplethysmography, remove the need for contact [11]. Their performance, however, depends on illumination, visible skin, and unobstructed line of sight, and visual recording may raise privacy concerns in indoor monitoring settings [2], [3].

B. mmWave-based Vital-Sign Sensing

Existing mmWave vital-sign systems have shown that reflected phase can capture subtle chest-wall motion for respiration and heart-rate monitoring [5], [6], [12], [13]. Recent studies have further improved sensing performance. Chen et al. [14] enhance weak-heartbeat recovery by suppressing baseline drift and respiratory-harmonic interference, while Sadeghi and Havinga [15] improve signal-processing robustness across prescribed distances, angles, and body orientations. Despite these advances, the sensor–subject geometry remains predetermined or manually configured rather than actively adapted during sensing. As a result, changes in subject pose or location can lead to unfavorable viewing angles or standoff distances, reducing the measurable radial motion or degrading the strength and spatial stability of the mmWave measurements.

C. Active View Planning for Robotic Perception

Robotic active perception improves sensing by deliberately changing the sensor viewpoint rather than passively accepting the current observation. Existing next-best-view (NBV) methods, such as OA-NBV [16] and NeU-NBV [17], select camera viewpoints to improve target visibility under occlusion or acquire more informative observations for scene representation. These works demonstrate the benefit of coupling perception with sensor motion, but the criteria used to define a favorable viewpoint are primarily tailored to visual tasks, such

as visibility, occlusion reduction, and reconstruction quality. In mmWave vital-sign sensing, the relevant viewpoint criteria are different: the sensor geometry should favor radial alignment with chest-wall motion and an appropriate standoff distance. RoViSense therefore defines viewpoint selection using these chest-geometry-aware sensing criteria and realizes the selected viewpoint through humanoid navigation.

III. SYSTEM DESIGN AND METHODOLOGY

This section presents the design of RoViSense as an end-to-end perception–action pipeline. As illustrated in Fig. 1, RoViSense first constructs subject and obstacle representations from onboard mmWave measurements and estimates the subject’s chest geometry. It then selects a sensing viewpoint, performs collision-aware navigation, and extracts heart rate after stabilizing at the selected viewpoint.

A. mmWave-Based Subject and Obstacle Representation Generation

We first process the onboard mmWave measurements to obtain two concrete inputs for later modules: the subject point cloud $\mathcal{Q}_t$ for chest-geometry estimation and the local obstacle map $\mathcal{O}_t$ for robot collision avoidance. Specifically, the raw chirp measurements are organized by range, Doppler, and angle FFTs across the antenna array into a Range–Doppler–Angle data cube [13]. High-energy cells in this cube provide candidate mmWave detections, each associated with radial distance, Doppler velocity, azimuth, and reflection intensity [18]. The horizontal 2-TX/4-RX virtual array provides fine azimuth resolution but limited elevation discrimination, which the humanoid platform’s bipedal height modulation and torso pitch adaptation partially compensate (Sec. IV-A).

RoViSense separates these candidate mmWave detections into the subject point cloud and the local obstacle map through subject–environment separation. For each range–angle cell, a short-horizon quasi-static profile is updated from temporally consistent, near-zero-Doppler responses. The subject branch removes this quasi-static profile from the Range–Doppler–Angle cube and applies CFAR detection to the residual responses, retaining detections associated with the dominant human cluster and its body-induced motion [12], [19]. These detections form $\mathcal{Q}_t = \{(x_i, y_i, z_i, v_i, I_i)\}_{i=1}^{N_t}$, where (x_i, y_i, z_i) represents the Cartesian location, v_i denotes the Doppler velocity, and I_i signifies the reflected intensity of the i -th point. The environment branch preserves the quasi-static profile: occupied cells outside the current human cluster are projected to Cartesian coordinates and fused across frames to form $\mathcal{O}_t$ for robot collision avoidance.

B. mmWave-Based Chest Geometry Estimation

From the point cloud $\mathcal{Q}_t$, RoViSense estimates the chest centroid $\mathbf{c}$, the chest-plane normal $\hat{\mathbf{n}}_{\text{chest}}$, and the aspect angle γ between the mmWave boresight and this normal. Here, $\hat{\mathbf{n}}_{\text{chest}}$ approximates the torso facing direction, while γ quantifies the angular offset relevant to radial chest-wall sensing. To estimate these quantities, RoViSense aggregates consecutive point-cloud frames to reduce the ambiguity caused

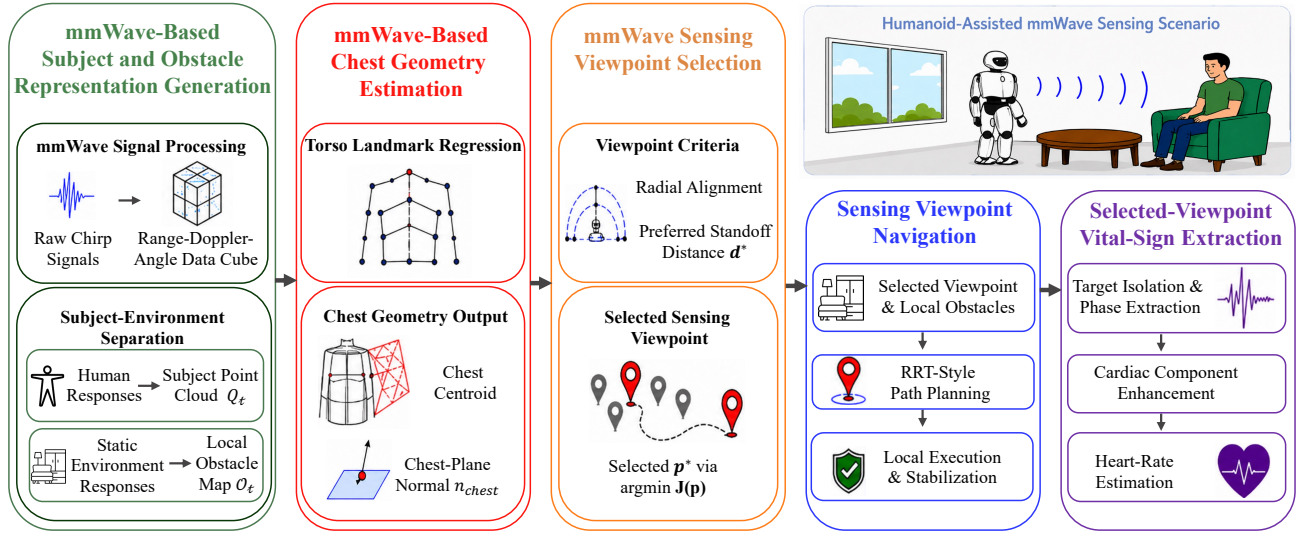


Fig. 1. End-to-end architecture of RoViSense. The system converts mmWave measurements into subject and obstacle representations, estimates chest geometry from the subject representation, selects a sensing viewpoint using radial-alignment and standoff-distance criteria, and navigates the humanoid platform to the selected viewpoint. After stabilization, RoViSense estimates heart rate from the mmWave phase signal through selected-viewpoint vital-sign extraction.

by sparse and intermittent mmWave returns. The aggregated point features are then fed to a regression network that recovers torso landmarks [18], [20].

From the reconstructed body landmarks, RoViSense uses the left shoulder $\mathbf{J}_{LS}$, right shoulder $\mathbf{J}_{RS}$, and mid-spine $\mathbf{J}_{MS}$ to define the chest geometry. The chest centroid $\mathbf{c} = \frac{1}{3}(\mathbf{J}_{LS} + \mathbf{J}_{RS} + \mathbf{J}_{MS})$ serves as the target location for viewpoint selection, while the chest-plane normal is computed as

$$\hat{\mathbf{n}}_{\text{chest}} = \frac{(\mathbf{J}_{RS} - \mathbf{J}_{MS}) \times (\mathbf{J}_{LS} - \mathbf{J}_{MS})}{\|(\mathbf{J}_{RS} - \mathbf{J}_{MS}) \times (\mathbf{J}_{LS} - \mathbf{J}_{MS})\|}. \quad (1)$$

With the mmWave boresight defined as $\vec{u}_{\text{radar}} = [0, 1, 0]^T$, the raw angular mismatch is $\gamma_{\text{raw}} = \arccos(\hat{\mathbf{n}}_{\text{chest}} \cdot \vec{u}_{\text{radar}})$. Because front and back views are both treated as aligned with the chest-plane normal axis, we use the folded aspect angle

$$\gamma = \min(\gamma_{\text{raw}}, \pi - \gamma_{\text{raw}}), \quad (2)$$

where $\gamma \in [0^\circ, 90^\circ]$ and smaller γ indicates better radial alignment for chest-wall motion. We empirically use 15° as the threshold: when $\gamma > 15^\circ$, the current view is treated as geometrically unfavorable and the system triggers autonomous repositioning [5], [15].

C. mmWave Sensing Viewpoint Selection

In this stage, we define a sensing viewpoint as the sensor position relative to the subject's chest. We select this viewpoint from candidate sensor positions $\mathbf{p} \in \mathbb{R}^3$ by modeling two controllable sensing conditions: appropriate standoff distance and radial alignment with the chest motion.

Spatial accessibility is required because each candidate position must correspond to a kinematically feasible, collision-free robot pose. We therefore treat the local obstacle map $\mathcal{O}_t$ as a standard collision-avoidance constraint: candidate positions whose corresponding robot footprint intersects occupied cells in $\mathcal{O}_t$ are pruned from the search, ensuring that every retained viewpoint is physically reachable by the humanoid platform.

The standoff distance controls the strength and spatial stability of the observation: long range weakens reflected energy and coarsens spatial resolution, whereas overly short range overfills the sensing field and destabilizes thoracic localization [5], [6]. We thus penalize deviation from a preferred distance $d^* = 0.9\text{m}$ (empirically selected from the low-error range of our workspace sweep) as $(d(\mathbf{p}) - d^*)^2$, where $d(\mathbf{p}) = \|\mathbf{p} - \mathbf{c}\|$.

Radial alignment governs how much chest-wall motion is measurable: because the sensor observes only the radial projection of displacement, a more oblique view reduces the component available for vital-sign extraction [6], [7]. We therefore use $\gamma(\mathbf{p})^2$ as the angular penalty, with the folded aspect angle $\gamma(\mathbf{p})$ computed between the viewing direction $(\mathbf{c} - \mathbf{p})/\|\mathbf{c} - \mathbf{p}\|$ and the chest-plane normal $\hat{\mathbf{n}}_{\text{chest}}$.

We combine the two criteria into the cost function

$$J(\mathbf{p}) = w_\gamma \gamma(\mathbf{p})^2 + w_d (d(\mathbf{p}) - d^*)^2, \quad (3)$$

where w_γ and w_d are balancing weights. We evaluate $J(\mathbf{p})$ over a discretized candidate set $\mathcal{P}$ constructed by sampling humanoid-reachable, kinematically feasible sensing positions around the estimated chest centroid. The sampling varies the bearing around the subject and the standoff distance within the admissible sensing range, while excluding poses that are unreachable or collision-prone. The selected viewpoint is $\mathbf{p}^* = \arg \min_{\mathbf{p} \in \mathcal{P}} J(\mathbf{p})$. We then pass $\mathbf{p}^*$ to the navigation module as the sensing goal.

D. Sensing Viewpoint Navigation

After viewpoint selection, we use the selected sensing position $\mathbf{p}^*$ as the navigation goal for the humanoid robot. Each repositioning episode starts from the latest stationary sensing update, which provides $\mathbf{p}^*$ from Sec. III-C and the local obstacle representation $\mathcal{O}_t$ from Sec. III-A. The navigation module then uses these inputs to move the robot to the selected viewpoint [21], [22].

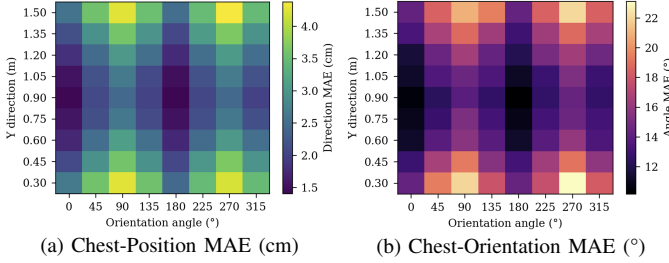


Fig. 2. Chest-geometry estimation accuracy over the distance–yaw workspace. Rows indicate radial distance ($Y = 0.30\text{--}1.50\text{ m}$), and columns indicate subject yaw angle ($0^\circ\text{--}315^\circ$). The lowest-error region appears near $Y = 0.90\text{ m}$ and $\theta = 180^\circ$, where both the chest-centroid position and chest-plane normal are estimated most reliably.

We implement navigation as a two-level planning stack over $\mathcal{O}_t$. The global layer uses an RRT-style sampling-based planner to search the collision-free workspace and generate a waypoint path toward $\mathbf{p}^*$. The execution layer tracks this path with the robot motion controller while enforcing local collision constraints from $\mathcal{O}_t$. A new stationary sensing update can refresh the chest geometry and obstacle representation, from which the next sensing viewpoint is selected [23], [24]. In our evaluation, this navigation stack is instantiated in the MuJoCo digital twin described in Sec. IV-A.

Once the robot reaches $\mathbf{p}^*$, it stabilizes its body pose and sensor orientation before vital-sign measurement begins. The output of this stage is therefore the stabilized sensing configuration at $\mathbf{p}^*$, which is passed to the stationary vital-sign extraction pipeline [25].

E. Selected-Viewpoint Vital-Sign Extraction

After reaching the selected viewpoint $\mathbf{p}^*$, RoViSense enters stationary measurement. The improved radial alignment and standoff geometry enhance observability of chest-wall motion, yet the mmWave phase remains dominated by respiration while the cardiac component stays weak. We therefore recover heart rate by estimating chest-wall displacement and then extracting its cardiac component.

Using the chest centroid $\mathbf{c}$ from the chest-geometry module (Sec. III-B), we revisit the complex data cube at $\mathbf{p}^*$ and define a thoracic range–angle region Ω_c around the projected chest. We select the strongest-power bin, $(\hat{r}, \hat{\theta}) = \arg \max_{(r, \theta) \in \Omega_c} P(r, \theta)$, extract its complex signal $x_{\hat{r}, \hat{\theta}}(t)$ across frames, and convert its phase $\phi(t) = \angle x_{\hat{r}, \hat{\theta}}(t)$ to displacement $s(t) = \frac{\lambda}{4\pi} \phi(t)$ [6], [13].

Since $s(t)$ is dominated by low-frequency respiration while the cardiac micro-displacement is weaker and higher-frequency, we apply the second-derivative operator $\ddot{s}(t)$ to obtain an acceleration-enhanced signal [26]. With $S_c(f)$ its spectrum, the Harmonic Product Spectrum $H(f) = \prod_{m=1}^{K_h} |S_c(mf)|$ [6] yields heart rate $\widehat{\text{HR}}$ from its dominant peak, while $\ddot{s}(t)$ provides the recovered cardiac waveform.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Platform, Data Collection, and Metrics

Hardware Platform. The system integrates a Texas Instruments AWR1642 FMCW radar, which we mount rigidly on the torso of a Unitree G1 humanoid robot. The radar

operates over 77–81 GHz with a sweep bandwidth of 4.0 GHz and a frame rate of 20 fps. Its 2-TX/4-RX antenna configuration forms 8 virtual channels through TDM-MIMO. Bipedal height modulation and torso pitch adaptation of the humanoid platform partially compensate for the limited native elevation discrimination of the AWR1642. For ground truth, an Apple Watch Ultra provides an electrocardiogram (ECG)-derived reference heart rate, time-synchronized via a common trigger. An NVIDIA Jetson Orin edge AI platform handles all real-time computation onboard.

Simulation Environment. Prior to deployment, we constructed a high-fidelity MuJoCo [24] simulation with the Unitree G1 URDF and a 1:1 digital twin of the physical laboratory room (table, chair, and static obstacles). We use the simulation exclusively to benchmark the path-planning module (Sec. IV-D); all vital-sign and orientation experiments are conducted with the physical robot. The simulated workspace replicates the $4 \times 5\text{ m}$ laboratory room, and we define three obstacle-complexity levels by the number of static obstacles placed in the scene: Low (3–4 obstacles), Medium (6–8 obstacles), and High (10–12 obstacles). For each complexity level we run 40 randomized navigation trials with pseudo-randomly repositioned obstacles, for a total of 120 navigation trials. Consequently, the quantitative sim-to-real transfer of navigation performance remains an acknowledged limitation.

Participants and Data Collection. Eight healthy adults (4M/4F, ages 22–35, BMI 19.5–27.2 kg/m²) participated with informed consent. Their resting heart rate ranged from 68 to 82 BPM. Each participant was evaluated under sitting-upright, standing, and lateral-decubitus postures in the $4 \times 5\text{ m}$ laboratory room, with obstacles pseudo-randomly repositioned between trials. For heart-rate evaluation, each recording window spans 30 s at 20 Hz, corresponding to 600 radar frames per window. To avoid identity leakage, we adopt a subject-level split for cross-subject validation, training on six participants and testing on the two held-out participants.

Evaluation Metrics. We evaluate three axes matching the architectural layers. *Chest-geometry accuracy* against the ground truth uses **Chest-Position MAE** (cm) for the chest centroid and **Chest-Normal Orientation MAE** ($^\circ$) for the chest-plane normal [18]. *Navigation* uses **Success Rate** $\mathcal{S}_r$ (collision-free trials reaching $\mathbf{p}^*$), and **Navigation Time** $\mathcal{T}_{\text{nav}}$ [23]. *HR fidelity* against the ECG reference uses **MAE**, **RMSE**, **MedAE**, and the percentage within $\pm 5\text{ BPM}$ [26].

B. Performance of Chest Geometry Estimation

We first evaluate whether the chest-geometry estimates used for viewpoint selection remain reliable across different sensing distances and subject yaw angles. For each participant, we conduct a full distance–yaw sweep. The subject is positioned at nine distances along the sensor boresight, ranging from 0.30 to 1.50 m in 0.15 m increments. At each distance, the subject is rotated through eight yaw angles relative to the sensor ($0^\circ, 45^\circ, \dots, 315^\circ$). This design yields $9 \times 8 = 72$ distance–yaw configurations per participant. For each configuration, we compare the mmWave-derived chest position and

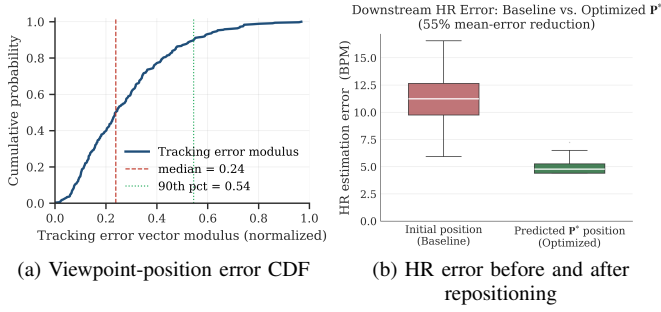


Fig. 3. Selected-viewpoint accuracy and heart-rate improvement. (a) Normalized position error between the selected viewpoint $\hat{\mathbf{p}}^*$ and the reference viewpoint $\mathbf{p}_{\text{ref}}^*$. (b) Heart-rate error at the initial viewpoint and after repositioning to the selected viewpoint.

chest-normal orientation against their corresponding ground-truth labels.

Across these configurations (Fig. 2), the mmWave-based chest-geometry module maintains centimeter-level position accuracy and stable chest-normal estimation, with chest-position MAE ranging from 1.41 cm to 4.38 cm (mean 2.65 cm) and chest-normal orientation MAE ranging from 10.2° to 23.1° (mean 15.1°). The smallest errors are observed at $Y=0.90$ m and $\theta = 180^\circ$. When the results are averaged over front/back views ($0^\circ/180^\circ$) and side views ($90^\circ/270^\circ$), side views show larger errors: chest-position MAE increases from 1.92 cm to 3.30 cm, and chest-normal orientation MAE increases from 12.2° to 17.6° . These results show that the perception module provides reliable chest-centroid and chest-normal inputs for viewpoint selection, and they give empirical support for using $d^* = 0.9$ m in Eq. (3) and favoring folded aspect angles near $\gamma \approx 0^\circ$ as defined in Eq. (2).

C. Performance of Viewpoint Selection and Heart-Rate Estimation

We evaluate whether the sensing viewpoint selected by RoViSense provides a measurable benefit for heart-rate extraction. The experiment includes 8 participants, 3 postures, and a distance–orientation grid that covers different starting distances and subject yaw angles, with 3 recording windows (30 s at 20 Hz) per condition. Specifically, the evaluation protocol first requires each subject to directly face the radar boresight at an initial orientation of 0° . Measurements are then captured at radial distances from 0.30 to 1.50 m in 0.15 m increments, and at each distance the subject is systematically rotated through eight distinct yaw angles relative to the sensor ($0^\circ, 45^\circ, \dots, 315^\circ$), establishing a total grid of $9 \times 8 = 72$ evaluation configurations [12], [15]. For each condition, we perform paired measurements at the initial viewpoint $\mathbf{p}_{\text{init}}$ and at the selected viewpoint $\hat{\mathbf{p}}^*$ computed from the mmWave-derived chest geometry. Both measurements are processed with the same vital-sign extraction pipeline.

We first compare the selected viewpoint $\hat{\mathbf{p}}^*$ with a reference viewpoint $\mathbf{p}_{\text{ref}}^*$ computed from the ground-truth chest geometry. Across the tested cases, the viewpoint-position error is 2.6 ± 1.4 cm, with a median of 3.0 cm and a 90th percentile of 6.3 cm, as shown in Fig. 3(a). This indicates that the mmWave-derived chest geometry keeps the selected sensing position

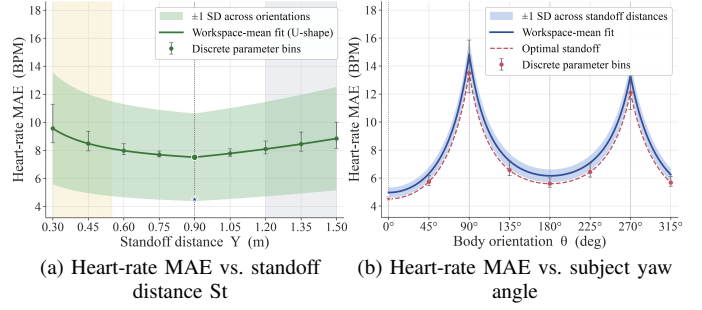


Fig. 4. Heart-rate error as a function of sensing geometry. Error is lowest near the preferred standoff distance $Y = 0.90$ m and increases at shorter or longer distances. Across yaw angles, error is lower near front/back-facing views and higher near side-facing views.

close to the reference solution. Compared with the ECG reference, the heart-rate MAE decreases from 11.73 to 4.51 BPM after moving from $\mathbf{p}_{\text{init}}$ to $\hat{\mathbf{p}}^*$, which corresponds to a $2.6\times$ reduction. In addition, 80 % of the selected-viewpoint windows fall within ± 5 BPM of the ECG reference. The corresponding RMSE and error-distribution statistics are reported in Fig. 3(b).

Fig. 4 examines the two geometric factors in Eq. (3) separately, with markers showing the MAE measured at each discrete parameter bin and the curve showing the mean trend through them. In Fig. 4(a), the orientation-averaged MAE follows a shallow U-shaped trend with standoff distance, reaching its minimum near the preferred distance $d^* = 0.9$ m. In contrast, Fig. 4(b) shows a much stronger dependence on body orientation: the distance-averaged MAE is lowest for front- and back-facing configurations and rises sharply near side-facing orientations, where the radial component of chest-wall motion is reduced. The ± 1 SD band across orientations in Fig. 4(a) stays wide at every distance while the band across distances in Fig. 4(b) stays narrow at every orientation, indicating that radial alignment has a stronger influence on heart-rate accuracy over the evaluated workspace.

D. Performance of Selected-Viewpoint Navigation

To evaluate whether the selected sensing viewpoints can be reached by the humanoid platform, we use the viewpoints generated in Sec. IV-C as navigation goals in the MuJoCo-based simulation environment described in Sec. IV-A. We evaluate the navigation stack from Sec. III-D, which combines global path generation with local collision-constrained execution, and compare it with five representative baselines: grid-based search (A), sampling-based planning (RRT), reactive obstacle avoidance (DWA and Score-Weighted APF), and learning-based navigation (standalone SAC). The comparison is conducted under Low, Medium, and High obstacle complexity.

The proposed planner sustains a Success Rate S_r near 95 % at Low/Medium complexity and approximately 90 % at High complexity (Fig. 5), whereas classical and reactive baselines degrade sharply with clutter (APF and DWA fail at High complexity). In the high-clutter setting the proposed planner attains the shortest Navigation Time $\mathcal{T}_{\text{nav}}$ —a 24 % reduction over the strongest baseline (SAC) (Fig. 5). These results support the feasibility of reaching the selected sensing viewpoints under the evaluated environment layouts.

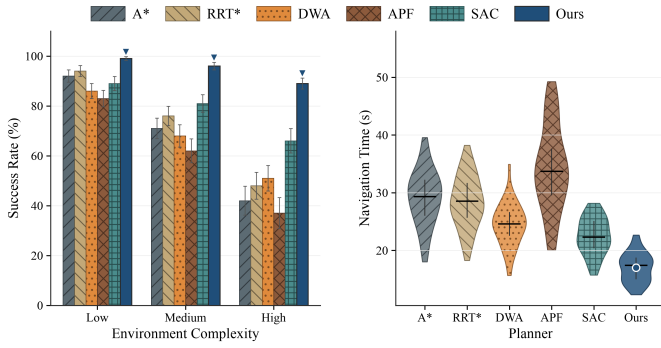


Fig. 5. Navigation benchmark under low, medium, and high obstacle complexity. Left: Navigation success rate $\mathcal{S}_r$ across obstacle-complexity levels. Right: Navigation time $\mathcal{T}_{nav}$ in the high-clutter setting.

V. CONCLUSION

We develop RoViSense, an active mmWave vital-sign monitoring system that adapts sensing geometry using a humanoid-mounted sensor. RoViSense estimates chest geometry from sparse mmWave observations and selects sensing viewpoints that jointly account for radial alignment and standoff distance. Experiments with 8 subjects show a mean chest-position MAE of 2.65 cm and a median selected-viewpoint position error of 3.0 cm, while repositioning reduces heart-rate MAE from 11.73 to 4.51 BPM. A MuJoCo-based navigation benchmark further evaluates the feasibility of reaching the selected viewpoints under varying obstacle conditions. More broadly, these results demonstrate the value of treating sensing geometry as an adaptive sensing variable rather than a fixed deployment choice for contactless mmWave vital-sign monitoring. Future work will extend the navigation evaluation to physical deployments and quantify sim-to-real transfer on the humanoid platform.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant No. 2428595 and by the Enhancing Faculty Research in STEM and Health Grant from the Katz School of Science and Health, Yeshiva University.

REFERENCES

- [1] N. Alugubelli, H. Abuissa, and A. Roka, "Wearable devices for remote monitoring of heart rate and heart rate variability: What we know and what is coming," *Sensors*, vol. 22, no. 22, p. 8903, 2022.
- [2] K. Harlow, H. Jang, T. D. Barfoot, A. Kim, and C. Heckman, "A new wave in robotics: Survey on recent mmwave radar applications in robotics," *arXiv preprint arXiv:2305.01135v4*, 2024.
- [3] H. B. D. H. Acharya B, Saakyan W, "The reliability of remote photoplethysmography under low illumination and elevated heart rates," *npj Digital Medicine*, vol. 8, p. 478, 2025.
- [4] H.-W. Huang, J. Chen, P. R. Chai, C. Ehmke, P. Rupp, F. Z. Dadabhoi, A. Feng, C. Li, A. J. Thomas, M. da Silva, E. W. Boyer, and G. Traverso, "Mobile robotic platform for contactless vital sign monitoring," *Cyborg and Bionic Systems*, vol. 2022, p. 9780497, 2022, mIT, Harvard Medical School, Boston Dynamics.
- [5] X. Zhao, D. Zhang, H. Yin, D. Mao, and H. Li, "Remote vital sign monitoring based millimeter-wave sensing," *IEEE Access*, vol. 12, pp. 142 760–142 771, 2024.
- [6] B. R. Upadhyay, A. B. Baral, and M. Torlak, "Vital sign detection via angular and range measurements with mmwave MIMO radars: Algorithms and trials," *IEEE Access*, vol. 10, pp. 106 017–106 032, 2022.

- [7] J. Choi, S. Hor, S. Yang, and A. Arbabian, "MVDoppler-Pose: Multi-modal multi-view mmwave sensing for long-distance self-occluded human walking pose estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.
- [8] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang, "Expressive whole-body control for humanoid robots," *arXiv preprint arXiv:2402.16796*, 2024, uC San Diego, CMU, UIUC, MIT.
- [9] J. S. J. Y. Y. G. Weiji Xie, Chenjia Bai, "Humanoid whole-body locomotion on narrow terrain via dynamic balance and reinforcement learning," *arXiv preprint arXiv:2502.17219*, 2025, unitree H1-2 deployment.
- [10] H. D. Nguyen, T.-T. Pham, N. Le, and V. Nguyen, "TolerantECG: A foundation model for imperfect electrocardiogram," in *Proceedings of the 33rd ACM International Conference on Multimedia*. ACM, 2025, pp. 1–9.
- [11] P. Pirzada, A. Wilde, G. H. Doherty, and D. Harris-Birtill, "Remote photoplethysmography for heart rate and blood oxygen measurement: A review," *IEEE Sensors Journal*, 2024, university of St Andrews.
- [12] M. Zhou, Y. Liu, S. Wu, C. Wang, Z. Chen, and H. Li, "A novel scheme of high-precision heart rate detection with a mm-wave fmcw radar," *IEEE Access*, vol. 11, pp. 85 118–85 136, 2023.
- [13] S. Iyer, L. Zhao, M. P. Mohan, J. Jimeno, M. Y. Siyal, A. Alphones, and M. F. Karim, "mm-wave radar-based vital signs monitoring and arrhythmia detection using machine learning," *Sensors*, vol. 22, no. 9, p. 3106, 2022.
- [14] Y. Chen, J. Yuan, and J. Tang, "A high precision vital signs detection method based on millimeter wave radar," *Scientific Reports*, vol. 14, p. 25535, 2024.
- [15] E. Sadeghi and P. Havinga, "Breaking barriers in health monitoring: Multi-scenario vital sign detection using mm-wave mimo fmcw radar," *arXiv preprint arXiv:2508.20864*, 2025.
- [16] B. Hu, C. Chang, J. Ge, M. Namgung, X. Lin, A. Krieger, and T. Mohsenin, "OA-NBV: Occlusion-aware next-best-view planning for human-centered active perception on mobile robots," *arXiv preprint arXiv:2603.11072*, 2025, Johns Hopkins University, University of South Florida.
- [17] L. Jin *et al.*, "NeU-NBV: Next best view planning using uncertainty estimation in image-based neural rendering," *arXiv preprint arXiv:2303.01284*, 2023, university of Bonn.
- [18] H. Xue, Y. Ju, C. Miao, Y. Wang, S. Wang, A. Zhang, and L. Su, "mmesh: Towards 3d real-time dynamic human mesh construction using millimeter-wave," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '21)*. ACM, 2021, pp. 269–282.
- [19] B. Zhu, Z. He, W. Xiong, G. Ding, J. Liu, T. Huang, W. Chen, and W. Xiang, "Probradarm3f: mmwave radar-based human skeletal pose estimation with probability map guided multi-format feature fusion," *IEEE Internet of Things Journal*, 2024.
- [20] S. An and U. Y. Ogras, "Mars: Mmwave-based assistive rehabilitation system for smart healthcare," *ACM Trans. Embed. Comput. Syst.*, vol. 20, no. 5s, Sep. 2021. [Online]. Available: <https://doi.org/10.1145/3477003>
- [21] N. Pico, E. Montero, M. Vanegas, J. M. E. Ayon, E. Auh, J. Shin, M. Doh, S.-H. Park, and H. Moon, "Integrating radar-based obstacle detection with deep reinforcement learning for robust autonomous navigation," *Applied Sciences*, vol. 15, no. 1, p. 295, 2024.
- [22] P. Shi, X. Yang, C. Wu, and H. Cheng, "Research on an obstacle avoidance system for unmanned vessels based on millimeter-wave radar," *Journal of Marine Science and Engineering*, vol. 14, no. 3, p. 306, 2026.
- [23] M. Lam, J. Herrera, S. S. Afzal, K. Zhou, and F. Adib, "Minav: Autonomous drone navigation indoors using millimeter-waves," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 9, no. 3, p. 97, 2025.
- [24] E. Todorov, T. Erez, and Y. Tassa, "MuJoCo: A physics engine for model-based control," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 5026–5033.
- [25] K. Pfeiffer, Y. Jia, M. Yin, A. K. Veldanda, Y. Hu, A. Trivedi, J. Zhang, S. Garg, E. Erkip, S. Rangan *et al.*, "Path planning under uncertainty to localize mmwave sources," *arXiv preprint arXiv:2303.03739v2*, 2023.
- [26] L. Zhao, R. Lyu, H. Lei, Q. Lin, A. Zhou, H. Ma, J. Wang, X. Meng, C. Shao, Y. Tang *et al.*, "AirECG: Contactless electrocardiogram for cardiac disease monitoring via mmwave sensing and cross-domain diffusion model," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 3, pp. 1–27, 2024.