Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 34(3): 1477–1505.
DOI: 10.3934/era.2026067
Received: 14 November 2025
Revised: 21 January 2026
Accepted: 27 January 2026
Published: 12 February 2026

Research article

Robust multi-person vital-sign sensing in indoor environments using FMCW MIMO radar

Dongsheng Lai[1], **Zhiming Cai**[1,2,*], **Ganhong Tian**[1], **Jiangchao Zhang**[1] **and Long Chen**[1]

[1] School of Electronic, Electrical Engineering and Physics in Fujian University of Technology, Fuzhou 350000, China

[2] National Demonstration Center for Experimental Electronic Information and Electrical Technology Education, Fujian University of Technology, Fuzhou 350000, China

* **Correspondence:** Emai: caizm@fjut.edu.cn.

Abstract: Contactless monitoring of respiration and heart rate in shared indoor spaces—such as hospital wards, eldercare facilities, and sleep labs—demands solutions that are unobtrusive, privacy-preserving, and robust to static reflections, multipath, and inter-subject interference. This paper presents a complete millimetre-wave frequency-modulated continuous-wave (FMCW) technology - multiple-input multiple-output (MIMO) radar framework for multi-subject vital-sign sensing with three key contributions: (i) clutter-robust target discovery to stabilize detections in realistic indoor scenes; (ii) high-resolution spatial separation to suppress inter-subject leakage; and (iii) hybrid time–frequency decomposition to improve heartbeat isolation by mitigating respiration harmonics and noise. Experiments on single-, two-, and three-subject datasets achieve Mean Absolute Error/ Root Mean Squared Error (MAE/RMSE) of 2.81/4.13 Beats Per Minute (BPM) for respiration and 2.43/3.61 BPM for heart rate. Compared with filtering-, Ensemble Empirical Mode Decomposition (EEMD), Compressive Sensing - Orthogonal Matching Pursuit (CS-OMP), and mmRH-Millimetre-wave radar heart-rate (mmRH) based baselines, the proposed framework yields substantially lower heart-rate errors while maintaining reliable multi-subject operation in shared indoor environments.

Keywords: FMCW radar; vital signs monitoring; respiration rate; heart rate; multi-subject separation

1. Introduction

mmWave frequency-modulated continuous-wave (FMCW) radar enables fully non-contact monitoring of respiration and heartbeat by sensing subtle chest micro-motions, thereby offering a privacy-preserving and comfortable alternative to wearable sensors for long-term healthcare scenarios [1]. While accurate vital-sign estimation has been widely demonstrated in controlled single-subject settings,

practical indoor deployments increasingly require multi-subject monitoring under realistic layouts and non-ideal aspect angles (e.g., wall-/ceiling-mounted sensors in bedrooms or clinical wards) [2].

Extending radar vital-sign sensing from one to multiple subjects is non-trivial. In indoor environments, strong static clutter and multipath propagation produce ghost responses and angular smearing, thereby degrading detection and localization in the range–angle domain [3]. More critically, when subjects partially overlap in range, phase measurements extracted from a single range bin are easily contaminated by neighboring micro-motions, which causes spectral distortion and biased respiration rate (RR) / heart rate (HR) estimates. Oblique viewing angles further weaken chest-normal reflections and exacerbate multipath interference, thus leading to reduced Signal-to-Noise Ratio (SNR) and unstable peak tracking over long monitoring periods [4].

Existing radar-based vital-sign methods can be broadly grouped into (i) spectral-peak pipelines (band-pass filtering + Fast Fourier Transform (FFT) / Discrete Cosine Transform (DCT) with dominant-peak estimation) [5,6] and (ii) decomposition-based pipelines (e.g., Empirical Mode Decomposition (EMD)/ Variational Mode Decomposition (VMD) variants) for non-stationary separation [7–9]. Although efficient, spectral approaches remain sensitive to noise and spectral drift, while decomposition approaches often suffer from respiration-harmonic leakage into the HR band, which biases cardiac-rate estimation when the heartbeat component is weak [10,11]. Meanwhile, multi-antenna FMCW multiple-input multiple-output (MIMO) systems introduce spatial degrees of freedom for multi-subject localization and isolation [12,13], yet achieving jointly robust detection, high-resolution localization, and per-subject signal separation under range overlap and oblique angles remains challenging [14].

Accordingly, this work targets reliable *multi-person* RR and HR monitoring in realistic indoor conditions where clutter, multipath, aspect-angle variation, and partial range overlap coexist. Our goal is to unify robust multi-target detection and localization with per-subject vital-signal isolation and harmonic-robust rate estimation in an end-to-end framework. The main contributions are as follow:

1) A unified FMCW-MIMO radar framework for simultaneous multi-person respiration and heartbeat monitoring under realistic aspect angles;
2) A cascaded detection and localization pipeline integrates adaptive clutter suppression, range–angle Constant false alarm rate (CFAR), and spatial isolation to enhance multi-target discrimination in cluttered scenes; and
3) A robust vital-sign extraction module combines adaptive beamforming with harmonic-suppressed decomposition to stabilize RR/HR estimation under low-SNR and non-stationary conditions.

2. Signal model

Following the standard chirp-sequence FMCW radar model and the phase-modulation principle used in radar-based vital-sign monitoring [15,16], we briefly summarize the transmit/receive signal expressions used in this work.

To handle multiple antennas, we adopt a Transportation demand management (TDM) strategy and form a virtual array from M TX and N RX channels [17]. The complex Linear Frequency Modulation (LFM) transmit signal of the m-th TX can be expressed as

$$s_m(t) = A_m \exp\left\{ j\left[2\pi f_c t_m + \pi K t_m^2 + \phi_m(t)\right]\right\}, \tag{2.1}$$

where $t_m = t - (m-1)T_d$, f_c is the start frequency, and K is the chirp slope.

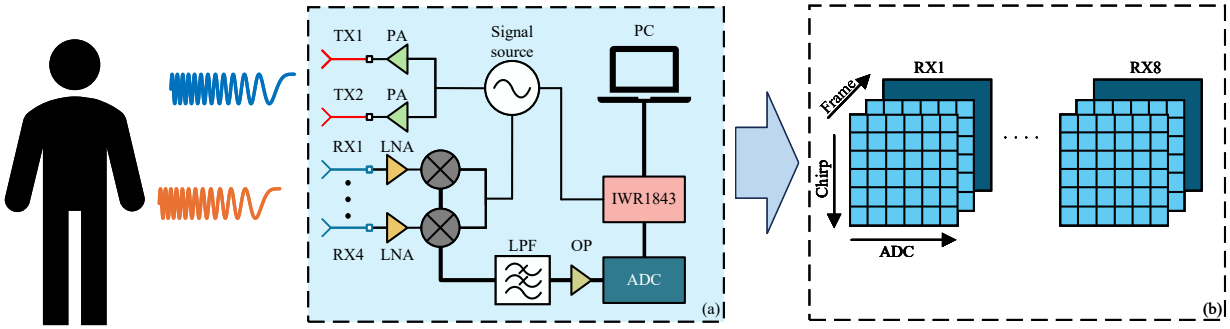


Figure 1. Signal model and processing chain of the FMCW TDM-MIMO radar.

The chest displacement induced by respiration and heartbeat is commonly modeled as a superposition of low-frequency periodic components [18]:

$$y(t) = A_r \sin(2\pi f_r t + \theta_r) + A_h \sin(2\pi f_h t + \theta_h). \tag{2.2}$$

Let d_0 be the nominal range and c the speed of light. The round-trip delay is as follows:

$$t_d = \frac{2(d_0 + y(t))}{c}. \tag{2.3}$$

In a uniform linear array, the inter-element phase term is proportional to $kd\sin\theta$ (equivalently $2\pi d\sin\theta/\lambda$). With TDM-MIMO, the m-n TX–RX pair forms one virtual element.

Thus, the received baseband signal at the (m, n) virtual channel is written as follows:

$$r_{mn}(t) = A_{mn}\exp\left\{j\Big[2\pi f_c(t_m - t_d) + \pi K(t_m - t_d)^2 - \frac{2\pi d_{mn}\sin\theta}{\lambda} + \phi_{mn}(t)\Big]\right\}. \tag{2.4}$$

After dechirping (mixing) and low-pass filtering, the intermediate frequency (IF) 112signal follows the standard FMCW sampling model and can be expressed as follows:

$$s_{IF}(t) = A_m A_{mn}\exp\left\{j\Big[2\pi f_c t_d + 2\pi K t_d t - \frac{2\pi d_{mn}\sin\theta}{\lambda} + \varphi_{mn}(t)\Big]\right\}, \tag{2.5}$$

where $\varphi_{mn}(t)$ collects residual phase terms.

Substituting (2.3) and using $f_c = c/\lambda$, while neglecting higher-order small terms, yields the widely used approximation that separates the range-dependent beat term and the displacement-induced phase term:

$$s_{IF}(t) \approx A_m A_{mn}\exp\left\{j\Big[\frac{4\pi(d_0 + y(t))}{\lambda} + \frac{4\pi K d_0}{c}t - \frac{2\pi d_{mn}\sin\theta}{\lambda}\Big]\right\}. \tag{2.6}$$

After Analog to Digital Converters (ADC) sampling ($t = nT_s$), we have the following:

$$s_{IF}[n, m] \approx A_m A_{mn}\exp\left\{j\Big[\frac{4\pi(d_0 + y[n])}{\lambda} + \frac{4\pi K d_0}{c}nT_s - \frac{2\pi d_{mn}\sin\theta}{\lambda}\Big]\right\}. \tag{2.7}$$

Hence, for a fixed d_0, the phase variation induced by chest displacement satisfies $\Delta\psi[n] = \frac{4\pi}{\lambda}y[n]$, which is the standard phase-displacement relation exploited in radar vital-sign monitoring.

3. Proposed method

3.1. Framework overview

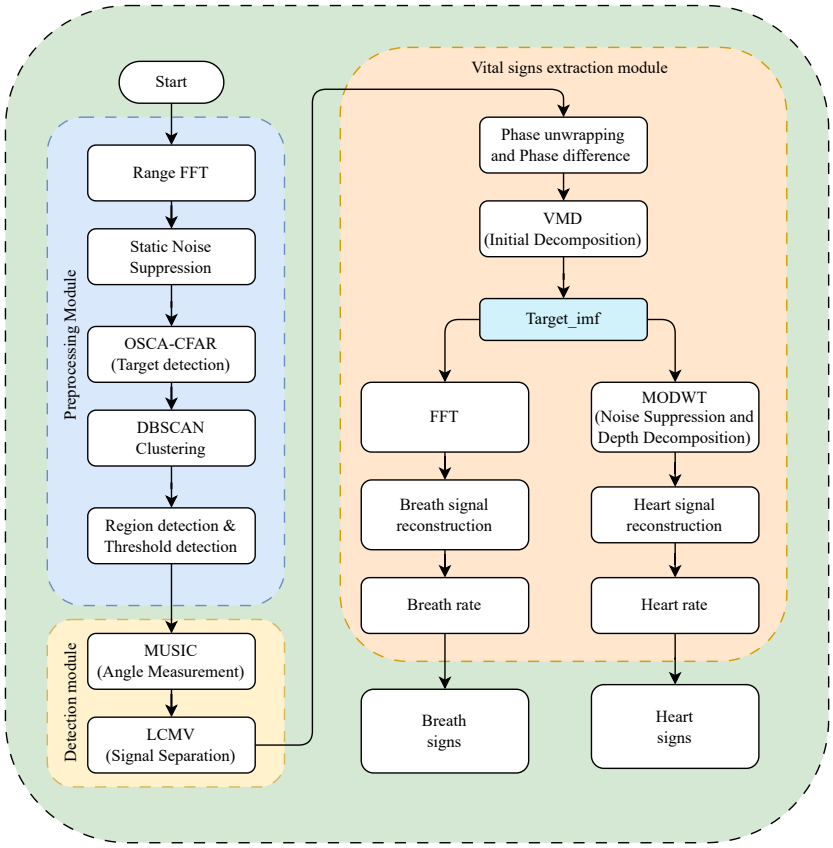


Figure 2. Algorithm flowchart.

Building on the signal and array models established earlier, we now outline the end-to-end framework for multi-person vital-sign monitoring. As summarized in Figure 2, the pipeline follows a detect–separate–estimate paradigm: it first suppresses static clutter and robustly detects human presence in the range–time domain; then it spatially resolves multiple subjects and stabilizes phase measurements via Multiple signal classification (MUSIC)-guided adaptive beamforming; and finally extracts respiration and heartbeat using a hybrid variational mode decomposition (VMD)–wavelet scheme that suppresses respiratory harmonics. This modular organization clarifies the data flow, isolates error sources, and defines clean interfaces between stages, thus enabling robust operation in cluttered indoor scenes and straightforward ablation or substitution of individual components. The following paragraphs detail each step.

Step 1—Distance positioning and static noise suppression: The received IF signals are transformed into the range domain via FFT, followed by static clutter suppression. Then, a range–time profile is obtained, and potential targets are identified using OS-CFAR. Finally, Density-based spatial clustering

of applications with noise (DBSCAN) clustering is applied to refine the detections and determine the human position.

Step 2—Multi-user positioning and phase extraction: We employ the MUSIC algorithm to accurately estimate the positions of multiple subjects, followed by Adaptive digital beamformer (ADBF) to focus on their thoracic regions. The precision of MUSIC enables ADBF to effectively localize each subject's chest area.

Step 3—Vital signs signal extraction: First, respiratory signals are extracted using VMD. Subsequently, the Maximal overlap discrete wavelet transform (MODWT) is introduced to suppress respiratory harmonics, thereby mitigating their interference and enabling precise extraction of the cardiac signal.

3.2. Distance positioning and static noise suppression

Accurate and robust target detection constitutes a fundamental prerequisite for subsequent vital-sign monitoring. The detection process begins with range profiling, where the received IF signals are transformed into the range domain using a FFT, commonly referred to as the range-FFT. This operation can be expressed as follows:

$$X(i,m) = \sum_{n=0}^{N-1} s_{IF}[n,m], e^{-j2\pi\frac{n}{N}i}, \tag{3.1}$$

where i denotes the range-bin index, and m is the slow-time snapshot. Range bins corresponding to human targets typically exhibit higher energy than noise-only bins, thereby serving as preliminary indicators of target presence.

However, in practical indoor environments, reliable detection is hindered by strong reflections from static structures such as walls, furniture, and metallic objects. These echoes often exceed the energy of human-induced signals, thus masking or overwhelming the desired returns. Consequently, static clutter suppression becomes indispensable.

The principal difference between static and dynamic echoes lies in their temporal characteristics: reflections from static scatterers remain nearly constant across slow time, whereas human chest echoes periodically vary due to respiration and heartbeat motion. This property enables clutter mitigation through temporal averaging. Specifically, a background estimate is obtained by averaging IF signals across multiple slow-time snapshots, thus yielding a reference dominated by static components. Since dynamic echoes fluctuate, they tend to cancel out during averaging. Subtracting this reference from each instantaneous measurement suppresses static clutter while retaining dynamic information:

$$\bar{X}(i) = \frac{1}{M}\sum_{m=1}^{M} X(i,m), \tag{3.2}$$

$$\hat{X}(i,m) = X(i,m) - \bar{X}(i), \tag{3.3}$$

where M denotes the number of slow-time samples.

Following clutter suppression, detection is performed in the range spectrum. Because propagation attenuation leads to distance-dependent signal strength, which is stronger for nearby targets and weaker for distant ones, a fixed threshold is unsuitable. It would either miss weak yet valid targets or generate excessive false alarms. To address this, a CFAR detector is employed, which dynamically adjusts the

detection threshold according to the local noise and interference level, thereby maintaining a constant false-alarm probability across varying ranges and environments.

For multi-subject scenarios, OS-CFAR scheme is adopted to mitigate masking effects and improve the detection reliability. In this method, the Cell under test (CUT) is compared with an adaptive threshold estimated from neighboring reference cells. Guard cells are inserted around the CUT to prevent target spillover contamination. The interference power is estimated as the minimum average from the reference windows and scaled by a threshold factor to form the detection threshold T. A detection is declared when the CUT power Y exceeds T. Compared with conventional CFAR schemes, OS-CFAR enhances robustness in dense multi-target conditions by avoiding overestimation of the interference power.

Nevertheless, due to the volumetric nature of the human body, a single subject may generate multiple reflections (e.g., from the head, chest, and limbs), thus resulting in several consecutive range bins being flagged as targets. Without post-processing, these fragmented detections would lead to redundant candidate bins and unstable downstream processing. To address this issue, we perform density-based clustering along the range dimension to form *range gates* that represent contiguous target regions.

In this study, we adopt DBSCAN as a *range-gating* module. DBSCAN groups adjacent CFAR-positive range bins into high-density clusters according to a distance threshold (in range-bin units) and a minimum cluster size, while isolated bins are treated as outliers. Then, each cluster is summarized by a representative range bin (e.g., the energy-weighted centroid or the peak-energy bin) and/or a short interval of bins. The resulting set of range gates $\mathcal{R} = \{\mathcal{R}_j\}$ is used to (i) localize candidate subject ranges, (ii) reduce computation by restricting subsequent direction of arrival (DOA) estimation to $\mathcal{R}$, and (iii) stabilize covariance estimation via range-neighborhood aggregation. Importantly, in this work, DBSCAN is not used to determine the angular model order of MUSIC.

3.3. Multi-person positioning and phase extraction

Based on the array signal model in Section II, the IF samples from all virtual channels are stacked into a data matrix and processed with subspace methods. To spatially separate multiple respiration and heartbeat sources, we first apply the MUSIC algorithm for high-resolution DOA estimation [19]. In the controlled recordings of this work, the MUSIC model order is set to the known subject count ($K = K_{\mathrm{gt}}$), while DBSCAN is only used for range gating. Concretely, we form the sample covariance of the received data and perform eigen-decomposition to partition the signal and noise subspaces. By projecting array steering vectors onto the noise subspace, we obtain a pseudospectrum whose peaks indicate the target DOA.

Then, we employ ADBF to enhance and extract the signals at the identified directions. Using the estimated DOA, we design linearly constrained weights that ensure unit response toward the target while minimizing the output power. This spatial filtering suppresses interference and noise, thereby improving SNR and the separability of the respiratory and cardiac components. In combination, MUSIC and ADBF enable multi-person positioning and yield phase signals suitable for subsequent vital-sign analysis.

In our setup, a TDM-MIMO FMCW radar with a Uniform linear array (ULA) captures echoes from multiple subjects. Sequential activation of the transmitters synthesizes a virtual array with an extended aperture, thus improving the angular resolution and facilitating multi-target localization.

We consider an M-sensor array that receives $D(< M)$ far-field, narrowband sources impinging from directions $\{\theta_i\}_{i=1}^{D}$ at the carrier frequency f_c (with $\omega_0 = 2\pi f_c$ and $\lambda = c/f_c$). After quadrature

demodulation and anti-alias filtering, the complex baseband signal at sensor k is as follows:

$$x_k(t) = \sum_{i=1}^{D} g_{ki}\, s_i(t - \tau_{ki}) + n_k(t), \tag{3.4}$$

where $s_i(t)$ denotes the i-th source, $g_{ki} \in \mathbb{C}$ aggregates propagation and element responses, τ_{ki} is the inter-sensor delay determined by the array geometry, and $n_k(t)$ is zero-mean noise. Under the standard narrowband criterion $B_s \max_{k,i} \tau_{ki} \ll 1$ (with B_s the baseband bandwidth),

$$s_i(t - \tau_{ki}) \approx s_i(t)\, e^{-j\omega_0 \tau_{ki}}, \tag{3.5}$$

which yields

$$x_k(t) = \sum_{i=1}^{D} g_{ki}\, e^{-j\omega_0 \tau_{ki}}\, s_i(t) + n_k(t). \tag{3.6}$$

Stacking $x_k(t)$ over all sensors gives the following customary array model:

$$\mathbf{X}(t) = \mathbf{A}\,\mathbf{S}(t) + \mathbf{N}(t), \tag{3.7}$$

where $\mathbf{X}(t) = [x_1(t), \ldots, x_M(t)]^{\mathsf{T}}$, $\mathbf{S}(t) = [s_1(t), \ldots, s_D(t)]^{\mathsf{T}}$, $\mathbf{N}(t) = [n_1(t), \ldots, n_M(t)]^{\mathsf{T}}$, and $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_D)]$.

The steering matrix is as follows:

$$\mathbf{A} = \big[\mathbf{a}(\theta_1),\ \ldots,\ \mathbf{a}(\theta_D)\big] \in \mathbb{C}^{M \times D}. \tag{3.8}$$

Assuming $g_{ki} = 1$, the ULA steering vector with inter-element spacing d is as follows:

$$\mathbf{a}(\theta) = \big[\,1,\ \mathrm{e}^{-j\frac{2\pi}{\lambda} d \sin\theta},\ \ldots,\ \mathrm{e}^{-j\frac{2\pi}{\lambda}(M-1) d \sin\theta}\,\big]^{\mathsf{T}}. \tag{3.9}$$

The array covariance matrix is as follows:

$$\mathbf{R} = \mathbb{E}\big[\mathbf{X}(t)\mathbf{X}^{\mathsf{H}}(t)\big] = \mathbf{A}\,\mathbf{R}_S\,\mathbf{A}^{\mathsf{H}} + \sigma^2 \mathbf{I}, \tag{3.10}$$

with $\mathbf{R}_S = \mathbb{E}[\mathbf{S}(t)\mathbf{S}^{\mathsf{H}}(t)]$. The eigen-decomposition yields the following:

$$\mathbf{R} = \mathbf{U}_S \mathbf{\Sigma}_S \mathbf{U}_S^{\mathsf{H}} + \mathbf{U}_N \mathbf{\Sigma}_N \mathbf{U}_N^{\mathsf{H}}, \tag{3.11}$$

where $\mathbf{U}_S$ and $\mathbf{U}_N$ span the signal and noise subspaces, respectively, and $\mathbf{a}^{\mathsf{H}}(\theta)\mathbf{U}_N \approx \mathbf{0}$ at the true DOAs.

To mitigate finite-snapshot and model-mismatch effects, we apply diagonal loading as follows:

$$\hat{\mathbf{R}}_\delta = \hat{\mathbf{R}} + \delta\, \frac{\mathrm{tr}(\hat{\mathbf{R}})}{M}\, \mathbf{I}, \quad \delta = 10^{-6}, \tag{3.12}$$

thereby using $\hat{\mathbf{R}}_\delta$ in place of $\hat{\mathbf{R}}$ in subsequent subspace decomposition and beamformer design.

The MUSIC pseudospectrum is as follows:

$$P_{\mathrm{MUSIC}}(\theta) = \frac{1}{\mathbf{a}^{\mathsf{H}}(\theta)\, \mathbf{U}_N \mathbf{U}_N^{\mathsf{H}}\, \mathbf{a}(\theta)}. \tag{3.13}$$

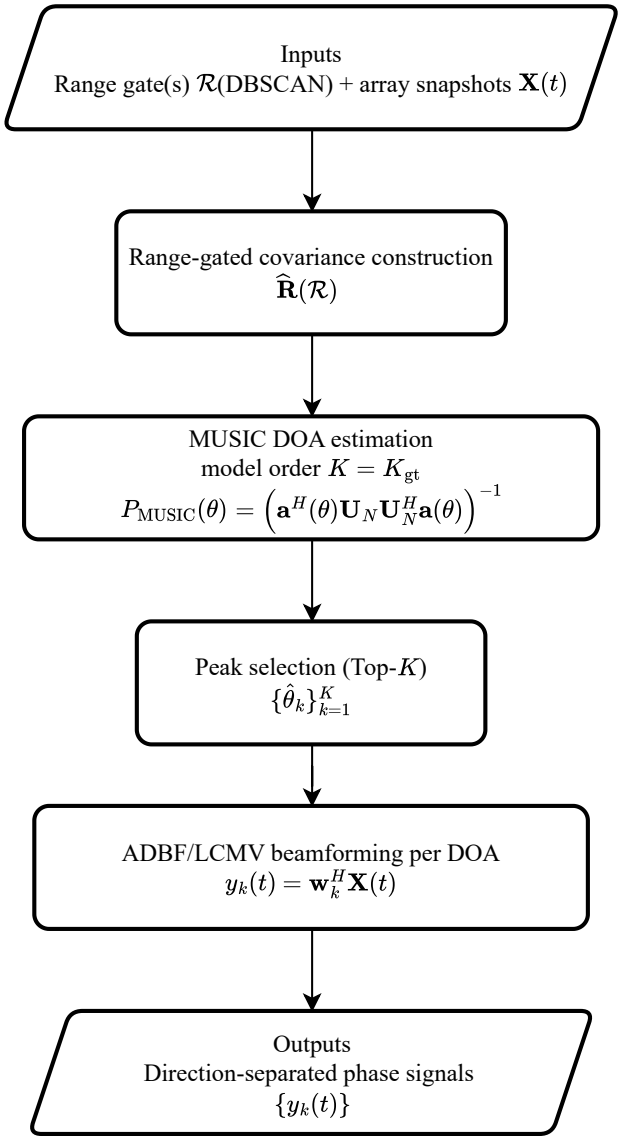


Figure 3. MUSIC–ADBF pipeline with DBSCAN range gating and fixed model order ($K = K_{\mathrm{gt}}$).

MUSIC requires the model order (i.e., the number of sources) to form the noise subspace.In our framework, the DBSCAN range-cluster count K_r does not determine the number of angular sources. Figure 3 summarizes the chapter-level processing chain in this section, thus highlighting that DBSCAN outputs range gate(s) $\mathcal{R}$ for covariance construction, while the MUSIC model order is set as $K = K_{\mathrm{gt}}$ in controlled recordings. Then, the resulting DOA estimates are used to design linearly constrained minimum variance (LCMV) beamformers for direction-separated phase extraction.

In the controlled experiments reported in this paper, the number of subjects in each recording is known (single-/two-/three-subject configurations). Therefore, we directly set the MUSIC model order directly as follows:

$$K = K_{\mathrm{gt}}, \tag{3.14}$$

where $K_{\mathrm{gt}} \in \{1, 2, 3\}$ denotes the known subject count for the current recording. Accordingly, the noise subspace is formed by taking the eigenvectors associated with the smallest $M - K$ eigenvalues.

While MUSIC yields accurate DOA estimates, it does not, by itself, produce directly isolated time series suitable for vital-sign analyses. Therefore, we synthesize an ADBF of the LCMV type for each detected DOA $\hat{\theta}_k$ to extract chest micro-motions as a one-dimensional signal [20].

Let $y_k(t) = \mathbf{w}_k^{\mathsf{H}}\mathbf{X}(t)$ denote the beamformer output toward the k-th target, where $\mathbf{w}_k \in \mathbb{C}^M$ is the weight vector, and $\mathbf{X}(t) \in \mathbb{C}^M$ stacks the sensor snapshots. The LCMV design solves the following:

$$\min_{\mathbf{w}_k} \ \mathbf{w}_k^{\mathsf{H}}\,\hat{\mathbf{R}}\,\mathbf{w}_k \quad \text{s.t.} \quad \mathbf{C}_k^{\mathsf{H}}\mathbf{w}_k = \mathbf{f}_k, \tag{3.15}$$

where $\hat{\mathbf{R}}$ is the sample covariance, and $\mathbf{C}_k$ collects linear constraints with desired responses $\mathbf{f}_k$. In the simplest "single-look" setting,

$$\mathbf{C}_k = \mathbf{a}(\hat{\theta}_k), \qquad \mathbf{f}_k = 1,$$

which enforces unit gain toward the estimated DOA. To suppress dominant interferers, $\mathbf{C}_k$ can be augmented with steering vectors of undesired directions, e.g.,

$$\mathbf{C}_k = [\mathbf{a}(\hat{\theta}_k),\ \mathbf{a}(\hat{\theta}_{i_1}), \ldots, \mathbf{a}(\hat{\theta}_{i_q})], \quad \mathbf{f}_k = [1,\ 0,\ \ldots,\ 0]^{\mathsf{T}}.$$

The closed-form solution of (3.15) is as follows:

$$\mathbf{w}_k^{\star} = \hat{\mathbf{R}}^{-1}\mathbf{C}_k(\mathbf{C}_k^{\mathsf{H}}\hat{\mathbf{R}}^{-1}\mathbf{C}_k)^{-1}\mathbf{f}_k. \tag{3.16}$$

Optionally, we estimate

$$\hat{\mathbf{R}} = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{X}(t_\ell)\mathbf{X}^{\mathsf{H}}(t_\ell), \tag{3.17}$$

with range-neighborhood averaging and forward–backward averaging to improve conditioning.

To mitigate finite-snapshot and model-mismatch effects, we apply the same diagonal loading in (3.12) and replace $\hat{\mathbf{R}}$ with $\hat{\mathbf{R}}_\delta$ in (3.16).

For extended thoracic targets or slight DOA errors, sector constraints can improve robustness by enforcing a near-flat response across a small angular interval $[\hat{\theta}_k - \Delta, \hat{\theta}_k + \Delta]$, e.g.,

$$\mathbf{C}_k = [\mathbf{a}(\hat{\theta}_k - \Delta),\ \mathbf{a}(\hat{\theta}_k),\ \mathbf{a}(\hat{\theta}_k + \Delta)], \quad \mathbf{f}_k = \mathbf{1}.$$

Applying $\mathbf{w}_k^{\star}$ to the array data yields the beamformed time series $y_k(t)$, from which respiration and heartbeat components are subsequently extracted. For simultaneous extraction across all K detected subjects, we stack the beams as follows:

$$\mathbf{y}(t) = \mathbf{W}^{\mathsf{H}}\mathbf{X}(t), \qquad \mathbf{W} = [\mathbf{w}_1^{\star}, \ldots, \mathbf{w}_K^{\star}], \tag{3.18}$$

thereby providing one clean waveform per subject for downstream vital-sign processing.

In cluttered indoor environments, we adopt three practical refinements to stabilize the DOA estimation. First, rather than relying on a single range bin, we build the sample covariance from a small neighborhood around each candidate bin and apply distance-dependent weights to fuse adjacent bins; this range-neighborhood fusion reduces the estimator variance and improves the robustness of the signal/noise subspace separation under finite snapshots. Second, we apply Two-dimensional non-maximum suppression (2D-NMS) on the computed angle–range pseudospectrum to suppress redundant or fragmented local maxima and retain well-separated peaks. We emphasize that 2D-NMS selects among existing peaks in the pseudospectrum; it cannot create peaks that are absent due to an underestimated model order. Third, we re-score the surviving peaks with a composite criterion that combines the peak SNR relative to the local noise floor, an eigenvalue-ratio indicator of the subspace separability, local energy compactness, and a penalty for detections near angular boundaries.

Specifically, for each candidate range bin r_0, we fuse covariance estimates within $\mathcal{N}(r_0) = \{r_0 - R_f, \ldots, r_0 + R_f\}$, where R_f denotes the neighborhood half-width in *range-bin units*. The fused covariance is as follows:

$$\widehat{\mathbf{R}}(r_0) = \sum_{\Delta r=-R_f}^{R_f} \widetilde{w}(\Delta r)\, \widehat{\mathbf{R}}(r_0 + \Delta r), \tag{3.19}$$

where $\widehat{\mathbf{R}}(r) = \frac{1}{C}\sum_{c=1}^{C} \mathbf{X}_{f,c,r}\mathbf{X}_{f,c,r}^{\mathsf{H}}$. We adopt a Hann taper over $L = 2R_f + 1$ bins as follows:

$$w_n = \tfrac{1}{2}\left(1 - \cos\left(\tfrac{2\pi(n-1)}{L-1}\right)\right), \quad n = 1, \ldots, L, \tag{3.20}$$

and normalize it as $\widetilde{w}_n = w_n / \sum_{i=1}^{L} w_i$, with index mapping $n = \Delta r + R_f + 1$. Unless otherwise stated, we set $R_f = 2$ (five range bins).

The remaining refinements are implemented as summarized in Algorithm 1 and described in the subsequent paragraphs.

After 2D-NMS, each remaining candidate peak at (θ, r) is assigned a composite score. Let $P(\theta, r)$ denote the MUSIC pseudospectrum and $S_{\mathrm{dB}}(\theta, r) = 10\log_{10}(P(\theta, r) + \epsilon)$. For each range bin r, we estimate a per-range noise floor using a low percentile across angles as follows:

$$N_{\mathrm{dB}}(r) = \mathrm{prctile}_{\theta}(S_{\mathrm{dB}}(\theta, r), 20), \qquad S_{\mathrm{snr}}(\theta, r) = S_{\mathrm{dB}}(\theta, r) - N_{\mathrm{dB}}(r). \tag{3.21}$$

1) Peak SNR. We use $S_{\mathrm{snr}}(\theta, r)$ as the peak SNR score at (θ, r).

2) Eigenvalue-ratio indicator. We compute an eigenvalue-ratio measure from the snapshot covariance at the candidate range bin:

$$S_{\mathrm{er}}(r) = \log\Big(1 + \frac{\lambda_1(r)}{\lambda_2(r) + \epsilon}\Big), \tag{3.22}$$

where $\lambda_1(r) \geq \lambda_2(r) \geq \cdots$ are eigenvalues of the corresponding covariance matrix.

3) Local blob score. We quantify the local energy concentration around (θ, r) via

$$S_{\mathrm{blob}}(\theta, r) = \sum_{(\theta', r') \in \Omega(\theta, r)} \max(0,\, S_{\mathrm{snr}}(\theta', r')), \tag{3.23}$$

where $\Omega(\theta, r)$ is a small neighborhood in the angle–range grid .

4) Optional edge penalty. A boundary penalty $S_{\mathrm{edge}}(\theta)$ can be included to discourage detections close to the angular search limits.

The final composite score is as follow:

$$\text{Score}(\theta, r) = w_{\text{peak}}\, S_{\text{snr}}(\theta, r) + w_{\text{eig}}\, S_{\text{er}}(r) + w_{\text{blob}}\, S_{\text{blob}}(\theta, r) - w_{\text{edge}}\, S_{\text{edge}}(\theta). \tag{3.24}$$

Unless otherwise stated, we set $(w_{\text{peak}}, w_{\text{eig}}, w_{\text{blob}}, w_{\text{edge}}) = (1.0, 0.8, 0.8, 0)$. The weights were tuned on a validation subset and then fixed for all experiments.

Collectively, these refinements produce cleaner and more stable DOA estimates, thus furnishing reliable inputs to the subsequent ADBF beamforming for vital-sign extraction. The end-to-end processing chain is summarized in Algorithm 1; unless otherwise noted, all experiments in the remainder of this paper adopt this robustified MUSIC–ADBF pipeline.

3.4. Vital signs signal extraction

In this section, the raw signal is first decomposed into a set of modes using VMD. According to the center frequency of each mode, the components corresponding to respiration and heartbeat are reconstructed, respectively. To further suppress the interference of respiratory harmonics on the heartbeat signal, the heartbeat-related modes are processed with MODWT for filtering and reconstruction, which results in a more stable heartbeat signal. This approach is particularly effective in multi-person scenarios, enabling reliable separation and extraction of respiration and heartbeat signals in complex environments.

1) Variational Mode Decomposition [21]: We decompose the phase signal $f(m)$ into K narrowband modes with distinct center frequencies $\{\omega_k\}_{k=1}^{K}$, thus enforcing perfect reconstruction as follows:

$$\begin{aligned} &\min_{\{u_k\},\{\omega_k\}} \sum_{k=1}^{K} \left\| \partial_m \Big[\Big(\delta(m) + \tfrac{j}{\pi m}\Big) * u_k(m) \Big]\, e^{-j\omega_k m} \right\|_2^2 \\ &\text{s.t.} \quad \sum_{k=1}^{K} u_k(m) = f(m). \end{aligned} \tag{3.25}$$

Introducing a quadratic penalty term and a Lagrange multiplier $\lambda(m)$, the augmented Lagrangian is written as:

$$\begin{aligned} \mathcal{L}(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_{k=1}^{K} &\left\| \partial_m \Big[\Big(\delta(m) + \tfrac{j}{\pi m}\Big) * u_k(m) \Big]\, e^{-j\omega_k m} \right\|_2^2 \\ &+ \left\langle \lambda(m),\, f(m) - \sum_{k=1}^{K} u_k(m) \right\rangle + \frac{\tau}{2} \left\| f(m) - \sum_{k=1}^{K} u_k(m) \right\|_2^2, \end{aligned} \tag{3.26}$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2 inner product (and $\Re\{\cdot\}$ is implied for complex-valued signals).

Using Alternating Direction Method of Multipliers(ADMM) in the frequency domain (one-sided spectrum $\omega \geq 0$), the updates are as follows:

$$\hat{u}_k^{l+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^{l+1}(\omega) - \sum_{i>k} \hat{u}_i^{l}(\omega) + \frac{1}{2}\hat{\lambda}^{l}(\omega)}{1 + 2\alpha\Big(\omega - \omega_k^{l}\Big)^2}, \tag{3.27}$$

$$\omega_k^{l+1} = \frac{\displaystyle\int_0^{\infty} \omega \left|\hat{u}_k^{l+1}(\omega)\right|^2 d\omega}{\displaystyle\int_0^{\infty} \left|\hat{u}_k^{l+1}(\omega)\right|^2 d\omega}, \tag{3.28}$$

Algorithm 1: Robust Multi-person Phase Extraction (MUSIC–ADBF/LCMV)

Input: Range profiles $\mathbf{X} \in \mathbb{C}^{F\times C\times R\times M}$;range gates $\mathcal{R}$ (from DBSCAN); model order K (set as $K = K_{\mathrm{gt}}$ and bounded by $K_{\max}$); fusion radius R_f

Output: Range–angle spectrum $S(\theta, r, f)$; beamformed signals $\{y_{f,r,k}(t)\}$

for $f \leftarrow 1$ **to** F **do**
  foreach $r_0 \in \mathcal{R}$ **do**
    $\widehat{\mathbf{R}}_x \leftarrow \mathbf{0}_{M\times M}$;
    $L \leftarrow 2R_f + 1$;
    compute normalized Hann weights $\{\widetilde{w}_n\}_{n=1}^{L}$ with $\sum_n \widetilde{w}_n = 1$;
    for $r \leftarrow r_0 - R_f$ **to** $r_0 + R_f$ **do**
      $r \leftarrow \mathrm{clip}(r, 1, R)$;
      $\Delta r \leftarrow r - r_0$; $n \leftarrow \Delta r + R_f + 1$;
      $\widehat{\mathbf{R}}_x \leftarrow \widehat{\mathbf{R}}_x + \widetilde{w}_n \cdot \frac{1}{C}\sum_{c=1}^{C} \mathbf{X}_{f,c,r}\mathbf{X}_{f,c,r}^{\mathsf{H}}$;
    $\mathbf{J} \leftarrow \mathrm{fliplr}(\mathbf{I}_M)$;
    $\mathbf{R}_{\mathrm{fb}} \leftarrow \frac{1}{2}(\widehat{\mathbf{R}}_x + \mathbf{J}\widehat{\mathbf{R}}_x^{*}\mathbf{J})$;
    $\widetilde{\mathbf{R}} \leftarrow \mathbf{R}_{\mathrm{fb}} + \delta \cdot \max(\frac{\mathrm{tr}(\mathbf{R}_{\mathrm{fb}})}{M}, 1)\mathbf{I}$; $\delta = 10^{-6}$;
    $[\mathbf{U}, \mathbf{\Sigma}] \leftarrow \mathrm{EVD}(\widetilde{\mathbf{R}})$ with eigenvalues in descending order;
    $\mathbf{U}_N \leftarrow$ eigenvectors associated with the smallest $M - K$ eigenvalues;
    foreach $\theta \in \Theta$ **do**
      $S(\theta, r_0, f) \leftarrow [\mathbf{a}^{\mathsf{H}}(\theta)\mathbf{U}_N\mathbf{U}_N^{\mathsf{H}}\mathbf{a}(\theta)]^{-1}$;
  $C_f \leftarrow \mathrm{CollectCandidates}(S(\cdot, \cdot, f), \mathcal{R}, R_{\mathrm{scan}})$;
  $C_f \leftarrow \mathrm{NMS2D}(C_f)$;
  foreach $(\hat{\theta}, r) \in C_f$ **do**
    compute $S_{\mathrm{snr}}(\hat{\theta}, r)$, $S_{\mathrm{er}}(r)$, $S_{\mathrm{blob}}(\hat{\theta}, r)$ and optional $S_{\mathrm{edge}}(\hat{\theta})$;
    compute $\mathrm{Score}(\hat{\theta}, r)$ using Eq (3.24);
  $\{(\hat{\theta}_k, r_k)\}_{k=1}^{K} \leftarrow \mathrm{TopK}(C_f, \mathrm{Score})$;
  for $k \leftarrow 1$ **to** K **do**
    $\mathbf{c} \leftarrow \mathbf{a}(\hat{\theta}_k)$; $\mathbf{f} \leftarrow 1$;
    $\mathbf{w}_k \leftarrow \widetilde{\mathbf{R}}^{-1}\mathbf{c}(\mathbf{c}^{\mathsf{H}}\widetilde{\mathbf{R}}^{-1}\mathbf{c})^{-1}\mathbf{f}$;
    $y_{f,r_k,k}(t) \leftarrow \mathbf{w}_k^{\mathsf{H}}\mathbf{X}_{f,:,r_k}(t)$;
$S \leftarrow \mathrm{AggregateAcrossFrames}(S)$;
return $S(\theta, r, f)$, $\{y_{f,r,k}(t)\}$;

$$\hat{\lambda}^{l+1}(\omega) = \hat{\lambda}^{l}(\omega) + \tau\left[\hat{f}(\omega) - \sum_{k=1}^{K} \hat{u}_k^{l+1}(\omega)\right], \tag{3.29}$$

where l is the iteration index, and $\tau > 0$ is the dual ascent step size.

The iterations stop when

$$\frac{\sum_{k=1}^{K} \left\|\hat{u}_k^{l+1} - \hat{u}_k^{l}\right\|_2^2}{\sum_{k=1}^{K} \left\|\hat{u}_k^{l}\right\|_2^2} < \varepsilon, \tag{3.30}$$

with tolerance $\varepsilon > 0$. Here, $\hat{f}(\omega)$, $\hat{u}_k(\omega)$, and $\hat{\lambda}(\omega)$ denote the Fourier transforms of $f(m)$, $u_k(m)$, and $\lambda(m)$, respectively.

Parameter notes K (number of modes) and α (bandwidth penalty) chiefly control decomposition granularity and mode narrowness; an overly large K leads to redundant modes, while an overly small K under-decomposes. A larger α enforces narrower bands (risking over-smoothing), whereas a smaller α improves detail but is more noise-sensitive. The dual step τ affects the convergence speed/stability (e.g., $0 < \tau \leq 2$), and ε sets the stopping accuracy.

2) Maximal Overlap Discrete Wavelet Transfor: In practical scenarios, signals are typically non-stationary. While the Wavelet Transfor (WT) has been demonstrated to be effective in capturing the features of non-stationary signals, its performance can be limited by insufficient joint time–frequency resolution at extreme frequency bands. Consequently, MODWT is employed for signal feature extraction [22,23]. MODWT supports the analysis of data sets with arbitrary sample sizes and is translation-invariant with respect to time shifts. In addition, MODWT provides improved asymptotic wavelet variance estimation compared with the decimated DWT.

Assume the discrete signal is $\{X_\iota\}_{\iota=0}^{N-1}$. The DWT decomposes the signal into wavelet coefficients at the dyadic scale 2^j and the translation index k. Using the notation $\varphi_X(j,k)$ for the wavelet coefficient, a conventional expression for the (decimated) DWT coefficients is as follows:

$$\varphi_X(j,k) \;=\; \sum_{\iota=0}^{N-1} X_\iota \, \psi_{j,k}(\iota), \tag{3.31}$$

where the wavelet basis is

$$\psi_{j,k}(\iota) \;=\; 2^{-j/2}\, \psi\!\left(\frac{\iota - 2^j k}{2^j}\right), \tag{3.32}$$

and $\psi(\cdot)$ denotes the mother wavelet.

Accordingly, for MODWT, we denote the level-c wavelet coefficients by $\varphi_{c,t}$ and the scaling (approximation) coefficients by $\eta_{c,t}$, where t is the time index. Because MODWT omits downsampling, coefficients are computed by circular convolution with rescaled filters. The level-c MODWT coefficients may be expressed as follows:

$$\varphi_{c,t} = \sum_{\ell=0}^{L_c-1} \tilde{h}_{c,\ell}\, X_{(t-\ell) \bmod N}, \tag{3.33}$$

$$\eta_{c,t} = \sum_{\ell=0}^{L_c-1} \tilde{g}_{c,\ell}\, X_{(t-\ell) \bmod N}, \tag{3.34}$$

where $\{\tilde{h}_{c,\ell}\}$ and $\{\tilde{g}_{c,\ell}\}$ are the level-c MODWT wavelet and scaling filter coefficients of length L_c, respectively, and $(\cdot) \bmod N$ enforces periodic indexing.

The MODWT filters are related to the corresponding (decimated) WT filters $\{h_{c,\ell}\}$ and $\{g_{c,\ell}\}$ via the following normalization:

$$\tilde{h}_{c,\ell} = \frac{h_{c,\ell}}{2^{c/2}}, \qquad \tilde{g}_{c,\ell} = \frac{g_{c,\ell}}{2^{c/2}}. \tag{3.35}$$

Here, c denotes the decomposition level, and ℓ is the filter-tap index.

The coefficients associated with the scaling function (the η's) are referred to as the approximation coefficients a, which capture the low-frequency content; the coefficients associated with the wavelet function (the φ's) are referred to as detail coefficients d, which capture high-frequency content. For decomposition up to level C, the MODWT output at each time t can be represented by the following collection:

$$\{\, d_1[t], d_2[t], \ldots, d_C[t], a_C[t] \,\},$$

where we have used the common notation $d_C[t] \equiv \varphi_{c,t}$ and $a_C[t] \equiv \eta_{C,t}$.

The original signal is reconstructed by summing the approximation and detail contributions as follows:

$$X[t] = a_C[t] + \sum_{c=1}^{C} d_c[t]. \tag{3.36}$$

In applications, the choice of mother wavelet, filter length L_c, and decomposition depth C should be determined according to the signal's spectral properties and the desired time–frequency resolution.

4. Experiments and results analysis

4.1. Experimental setup and measurement configurations

4.1.1. Radar hardware and acquisition

A FMCW millimetre-wave radar platform based on the Texas Instruments IWR1843BOOST evaluation module was employed for vital-sign sensing. The IWR1843BOOST was interfaced to a DCA1000EVM data-acquisition module, thus enabling raw IF capture and PC-based recording. The radar operated in a MIMO configuration with $N_{\mathrm{TX}} = 3$ transmitters and $N_{\mathrm{RX}} = 4$ receivers, thus forming $N_{\mathrm{v}} = N_{\mathrm{T}} N_{\mathrm{RX}} = 12$ virtual channels for spatial processing.

The digitized IF samples were streamed through the DCA1000EVM via Ethernet to a host computer for offline processing. Figure 4 illustrates the measurement setup. Table 1 summarizes the radar configuration used in all experiments.

4.1.2. Environment, participants, and protocols

All measurements were conducted in a typical indoor environment containing furniture and surrounding walls, thus introducing realistic multipath and clutter. The radar was mounted at a height of 1 m. The participants were seated facing the radar at predefined positions under line-of-sight conditions. The

Algorithm 2: Vital-Sign Extraction via VMD–MODWT

Input: Beamformed signals $\{y_{f,r,k}(t)\}$; sampling rate f_s; respiration band $B_b = [0.05, 0.7]$ Hz; heartbeat band $B_h = [0.8, 2.5]$ Hz

Hyperparameters: VMD parameters $(K, \alpha, \tau, \mathrm{DC}, \mathrm{init}, \mathrm{tol})$; MODWT wavelet w and level J; Welch parameters (L_w, N_{FFT}); detrend order p; Tukey taper β

Output: Respiration signals $\{\texttt{breathSig}_{f,r,k}(t)\}$, heartbeat signals $\{\texttt{heartSig}_{f,r,k}(t)\}$, respiration rates $\{\mathrm{breathRate}_{f,r,k}\}$, heartbeat rates $\{\mathrm{heartRate}_{f,r,k}\}$

foreach (f, r, k) **do**

- $y(t) \leftarrow y_{f,r,k}(t)$;
- $y \leftarrow \mathrm{HPF}(y,\ 0.03\ \mathrm{Hz})$;
- $y \leftarrow \mathrm{Detrend}(y;\ \mathrm{order} = p)$;
- $y \leftarrow \mathrm{Taper}(y;\ \mathrm{Tukey},\ \beta)$;
- $y \leftarrow \mathrm{ZScore}(y)$;
- $\{u_m(t)\}_{m=1}^{K},\ \{f_m\}_{m=1}^{K} \leftarrow \mathrm{VMD}(y;\ K, \alpha, \tau, \mathrm{DC}, \mathrm{init}, \mathrm{tol})$;
- $\mathcal{K}_b \leftarrow \{\, m \mid f_m \in B_b \,\}$;
- $\texttt{breathSig}(t) \leftarrow \sum_{m \in \mathcal{K}_b} u_m(t)$;
- $\mathcal{K}_h \leftarrow \{\, m \mid f_m \in B_h \,\}$;
- $\texttt{heartRaw}(t) \leftarrow \sum_{m \in \mathcal{K}_h} u_m(t)$;
- $W \leftarrow \mathrm{MODWT}(\texttt{heartRaw};\ w, J)$;
- $\mathcal{L}_h \leftarrow \mathrm{LevelsFromBand}(B_h, f_s, J)$;
- $\texttt{heartSig}(t) \leftarrow \mathrm{InverseMODWT}(W;\ \mathrm{keep} = \mathcal{L}_h)$;
- $P_b(f) \leftarrow \mathrm{PSD}(\texttt{breathSig}, f_s, L_w, N_{\mathrm{FFT}})$;
- $f_b \leftarrow \arg\max_{f \in B_b} P_b(f)$;
- $\mathrm{breathRate} \leftarrow 60\, f_b$;
- $P_h(f) \leftarrow \mathrm{PSD}(\texttt{heartSig}, f_s, L_w, N_{\mathrm{FFT}})$;
- $f_h \leftarrow \arg\max_{f \in B_h} P_h(f)$;
- $\mathrm{heartRate} \leftarrow 60\, f_h$;
- $\texttt{breathSig}_{f,r,k}(t) \leftarrow \texttt{breathSig}(t)$;
- $\texttt{heartSig}_{f,r,k}(t) \leftarrow \texttt{heartSig}(t)$;
- $\mathrm{breathRate}_{f,r,k} \leftarrow \mathrm{breathRate}$;
- $\mathrm{heartRate}_{f,r,k} \leftarrow \mathrm{heartRate}$;

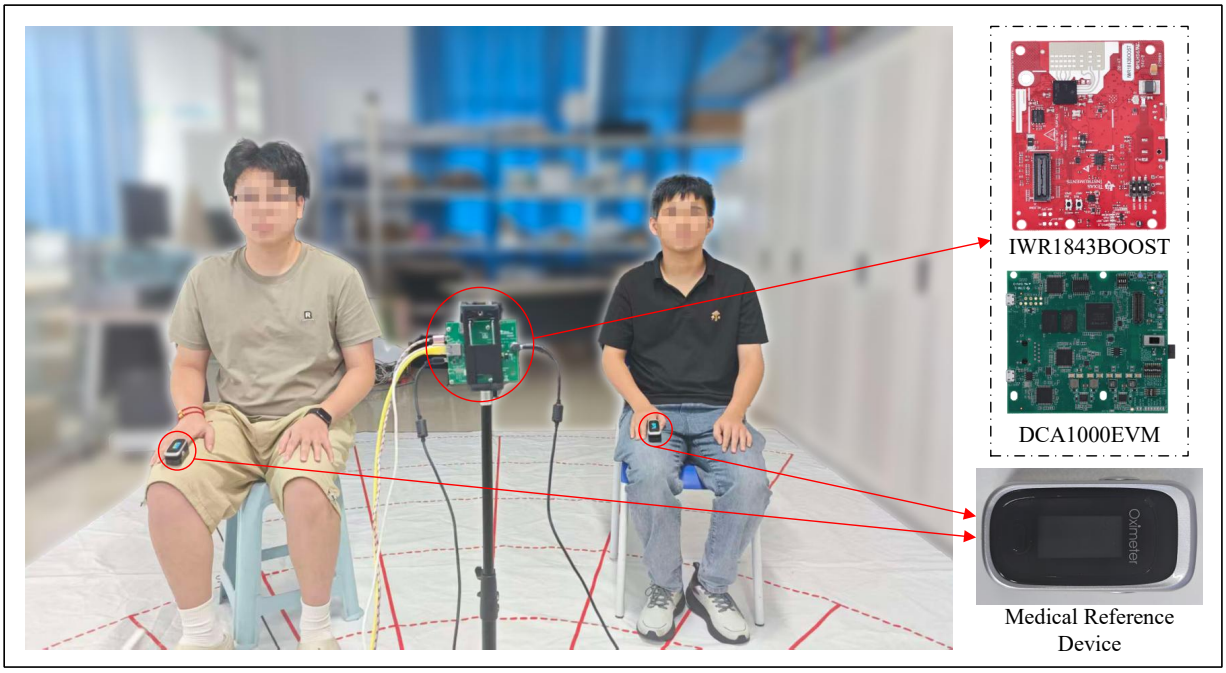


Figure 4. Experimental setup of the FMCW radar measurement system.

radar-to-subject distance ranged from 0.5 m to 1.2 m, with azimuth angles spanning from −30° to 30° across scenarios.

The healthy volunteers were recruited for this study, comprising nine individuals spanning an age range of 18–50 years. The cohort had a mean age of 24 years; the oldest participant was 45 years old and the youngest was 21 years old. The sample included five males and four females.

Each measurement lasted 60 s. To emulate realistic monitoring conditions, small voluntary motions were allowed, while the participants were instructed to remain seated and avoid large body movements.

4.1.3. Scene configurations and dataset size

Measurements were conducted under a set of predefined spatial layouts to evaluate robustness to different angular separations and range placements. As illustrated in Figure 5, the subjects were positioned at azimuth angles selected from $\{-30°, 0°, +30°\}$ with radar-to-subject distances spanning 0.5–1.2 m. Specifically, the single-subject cases place one subject at either 0° or +30° (at 0.5–1.0 m). The two-subject cases place two subjects at ±30° with unequal ranges (0.5 m and 1.2 m) and equal ranges (both at 1.0 m), respectively. The three-subject cases place three subjects at +30°, 0°, and −30° with equal ranges (all at 1.0 m) and with unequal ranges (0.7 m, 1.0 m, and 1.2 m), respectively.

In total, we recorded $N_{\text{rec}} = 7$ trials (single-subject: 3 trials, two-subject: 2 trials, three-subject: 2 trials). Since multi-subject trials yield multiple subject-wise estimates, these trials correspond to $N = 13$ target instances for metric computations.

4.1.4. Ground truth acquisition

For the quantitative evaluation, ground-truth HR and RR were acquired using a finger-clip pulse oximeter (M170, Shenzhen Fitfaith Technology Co., Ltd., China). The radar and the reference device

Table 1. Radar configuration and acquisition parameters.

Parameter	Value
Carrier frequency f_c	77 GHz
Slope S	70.006 MHz/μs
ADC sampling rate f_s	4 MSps
Samples per chirp N_s	128
Chirps per frame N_c	200
Frame rate f_{frame}	20 Hz
TX/RX ($N_{\text{TX}}/N_{\text{RX}}$)	3/4
Virtual channels N_{v}	12
Bandwidth B	2.24 GHz
Angle search grid (MUSIC)	$-40°$ to $40°$, step $1°$

were simultaneously started, and the HR/RR readings were averaged over the same 60 s interval used for the radar estimates.

4.1.5. Evaluation metrics

To ensure fair and reproducible comparisons across methods, the proposed algorithm was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Root Mean Squared Error (RMSE). Here, N denotes the number of evaluation instances. In single-subject trials, one recording yields one instance, whereas in multi-subject trials, each subject-wise estimate is treated as one instance. Therefore, our dataset contains $N_{\text{rec}} = 7$ recordings and $N = 13$ target instances in total. The MAE is defined as follows:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|f_i^{\text{est}} - f_i^{\text{gt}}\right|, \tag{4.1}$$

where f_i^{est} and f_i^{gt} denote the estimated and ground-truth heart (or respiration) rate for the i-th recording, respectively.

The RMSE is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_i^{\text{est}} - f_i^{\text{gt}}\right)^2}. \tag{4.2}$$

The MRE is expressed as follows:

$$\text{MRE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|f_i^{\text{est}} - f_i^{\text{gt}}\right|}{f_i^{\text{gt}}} \times 100\%. \tag{4.3}$$

All results are reported in beats per minute (BPM). Frequency values in hertz were converted to BPM using BPM = $60 \times$ Hz. Additionally, the detection accuracy was reported as the proportion of estimates within a tolerance threshold (e.g., ± 5 BPM). Values are presented as mean $\pm$ standard deviation across N recordings.

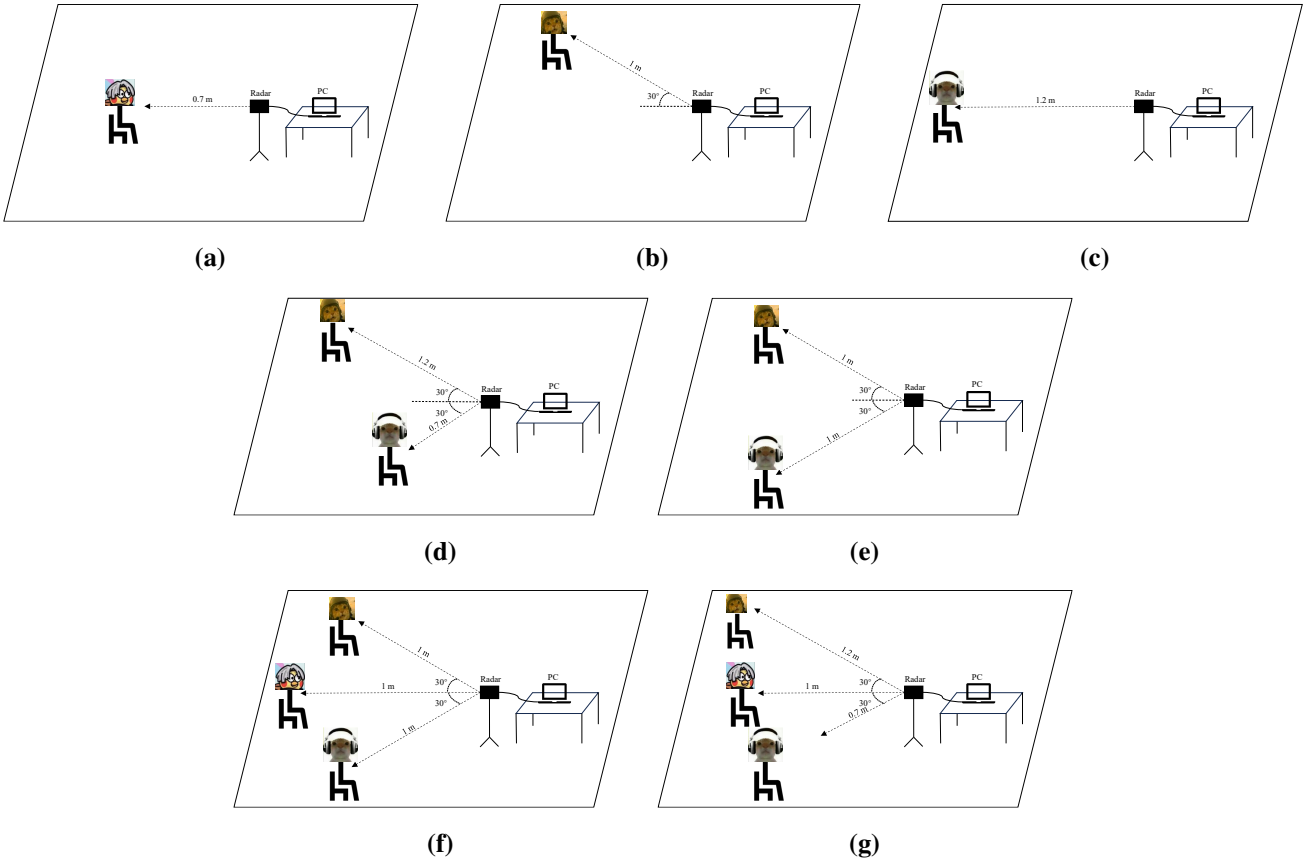


Figure 5. Schematic diagrams of subject placement in the considered scenarios. Subjects were positioned at azimuth angles in $\{-30°, 0°, +30°\}$ with radar-to-subject distances in the range 0.5–1.2 m.

4.1.6. Parameter settings

The key parameters used in the algorithm are summarized in Table 2. These parameters were carefully chosen through extensive experimentation to ensure robust performances across different experimental scenarios. Below, we describe the rationale behind each parameter's selection.

The VMD parameters K = 9 and α = 4000 were selected because they effectively separate the respiratory and heartbeat signals in the frequency domain. These values were chosen based on extensive experimentation to ensure that the decomposition accurately isolates the vital-sign components from noise and other interference. The choice of K = 9 ensures that the VMD algorithm generates nine modes, which sufficiently captures the main features of the respiratory and heartbeat signals, while α = 4000 balances the decomposition's stability and its ability to separate the components.

The OS-CFAR parameters were selected as follows: Training Window Size = 16, Guard Window Size = 4, and Pfa = 5×10^{-2}, thus ensuring robust target detection with minimal false alarms. These values were chosen based on experimentation to balance the target sensitivity and the interference rejection. Additionally, N = 16 and Guard Total = 4 provide sufficient protection against clutter, thus ensuring reliable detection performances.

For the DBSCAN clustering algorithm, the parameters Epsilon (ϵ) = 2.0, MinPts = 5, and MinClusterSize = 150 were chosen. These values ensure the accurate clustering of vital-sign signals while

Table 2. Algorithm parameter settings.

Parameter	Value	Description
OS-CFAR		
Training Window Size	16	Number of bins for training CFAR detector.
Guard Window Size	4	Number of bins for protection against interference.
Pfa (False Alarm Probability)	5×10^{-2}	Target false alarm probability.
Order of Statistical Exponent	2	Exponent for CFAR threshold.
DBSCAN Clustering		
Epsilon (ϵ)	2.0	Maximum distance between two range bins to be considered neighbors.
MinPts	5	Minimum number of neighbors for a core point in DBSCAN.
MinClusterSize	150	Clusters with fewer points are discarded as outliers.
Range Neighborhood Fusion		
Neighborhood Radius (R_f)	2	Range around each candidate bin for covariance fusion.
Weight Function	Hann Window	Weighting function for fused range bins.
2D-NMS and Peak Re-scoring		
Angle Tolerance ($nmsAngleTolDeg$)	5°	Angular tolerance for NMS.
Range Bin Tolerance ($nmsBinTol$)	3	Range tolerance for NMS.
Static Clutter Suppression		
Thresholding Method	Median Filtering	Used for static clutter suppression.
Diagonal Loading Parameter (δ)	1×10^{-6}	Used for covariance matrix conditioning.
VMD		
Number of modes (K)	9	Number of decomposed modes.
Bandwidth penalty (α)	4000	Controls mode bandwidth (larger α yields narrower-band modes).
Noise tolerance (τ)	0	Lagrange multiplier update step (set to 0 for noise-slack-free decomposition).
DC component	0	Whether to force a DC mode (0: disabled).
Initialization	1	Center-frequency initialization (1: uniform initialization).
Stopping criterion (tol)	1×10^{-7}	Convergence tolerance.

discarding noise and small, irrelevant clusters, thus improving the signal extraction reliability.

For the Range Neighborhood Fusion, the Neighborhood Radius (R_f) was set to 2, and a Hann Window weight function was used. These choices were made to ensure an effective spatial fusion of the candidate range bins while preserving the signal integrity. The radius was optimized to ensure that closely located bins meaningfully contribute to the fused signal.

The Overlap Threshold Filtering (OTF) was used with a threshold value of 3, which was experimentally determined to balance the sensitivity to valid signals and the rejection of unnecessary data. This parameter helps to ensure that overlapping components from different subjects are adequately separated, thus enhancing the overall signal accuracy.

Finally, for static clutter suppression, the Median Filtering method was applied, with a Diagonal Loading Parameter (δ) set to 1×10^{-6}. This method was chosen for its ability to effectively remove static clutter without distorting the vital-sign signals, and the diagonal loading parameter ensures stability in the covariance matrix, thus improving the accuracy of the clutter suppression process.

These parameters were selected to optimize the detection and separation of vital-sign signals while maintaining robustness against noise and interference.

4.2. Experimental results

Single-subject experiments were conducted to validate the effectiveness of the proposed signal-processing framework. As shown in Figure 6, the radar signals were first decomposed using VMD to extract intrinsic oscillatory components. The second VMD mode exhibited a dominant spectral peak that closely aligned with the reference respiratory frequency, thus enabling the accurate estimation

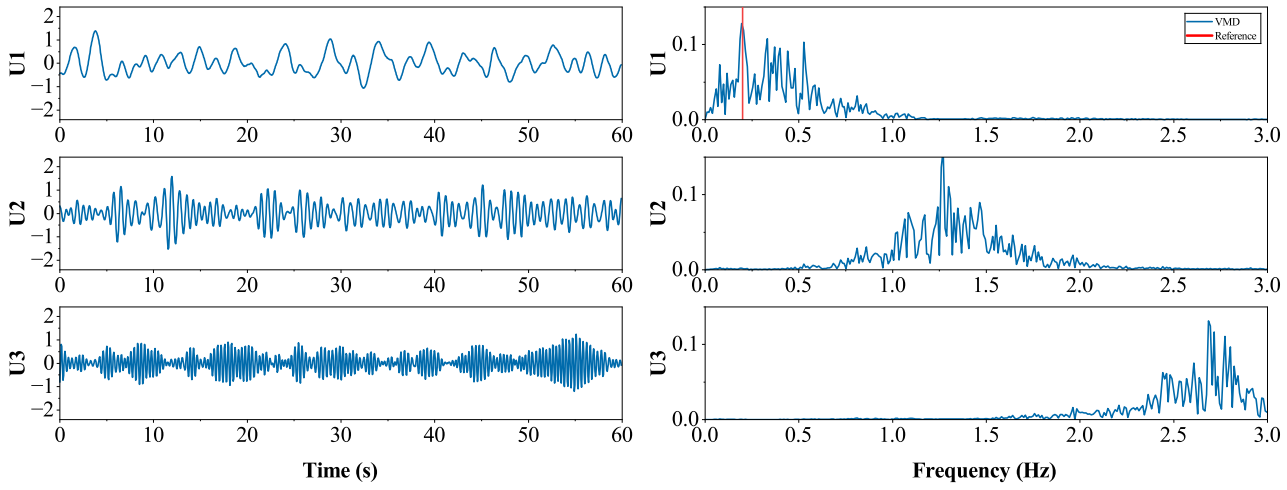


Figure 6. Experimental results for a single subject after VMD processing. (a) Time-domain waveform of the extracted respiratory component. (b) Corresponding spectral distribution showing a dominant peak at the respiration frequency.

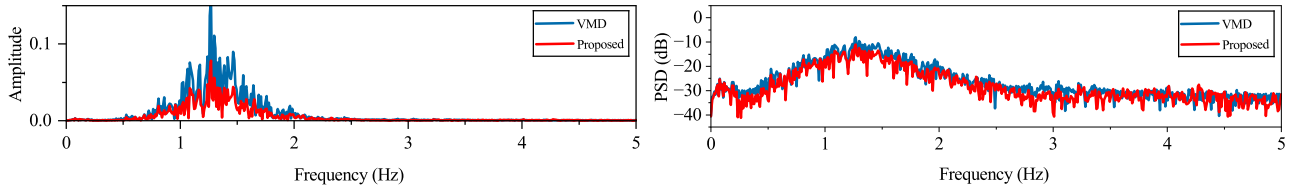


Figure 7. Enhanced post-processing results after applying the MODWT algorithm to the VMD-derived signal. (a) Reconstructed heartbeat waveform after MODWT-based denoising. (b) Power spectral density (PSD) illustrating effective suppression of residual noise components while retaining the dominant cardiac peak.

of RR.

To further enhance the detection of cardiac micro-motions, the MODWT was subsequently applied to the VMD-derived signal. This post-processing stage effectively suppressed the residual wideband noise while preserving the spectral energy of the target heartbeat component. As illustrated in Figure 7, the application of MODWT yielded a smoother temporal waveform and a more distinct cardiac peak in the PSD, thus indicating an improved signal clarity and noise resilience. In representative cases, the dominant frequency around 1.25 Hz closely corresponded to the ground-truth HR. Quantitative single-subject results are summarized in Table 3.

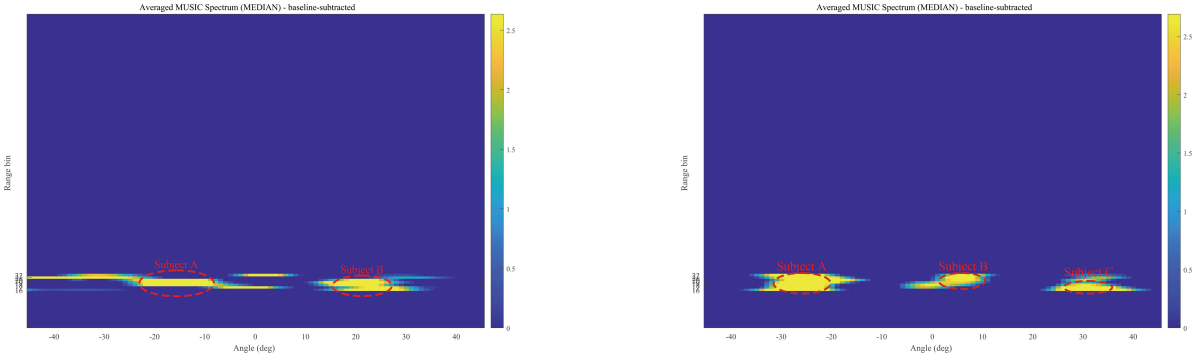
To further validate the scalability and generalization beyond single-subject cases, additional experiments were conducted under multi-subject scenarios.

4.3. Multi-subject experiments and results

Figure 8 illustrates representative angle–range heatmaps obtained in the two- and three-subject experiments. Each map shows the spatial-spectrum energy averaged across the observation interval. The

Table 3. Single-subject results.

Group	Dist. (m)	Angle (°)	Meas. RR	Meas. HR	GT RR	GT HR
First	1.0	0	20.0	75.60	12	75
Second	1.0	30	18.2	66.80	15	70
Third	0.5	0	13.2	65.70	13	65



(a) Two-subject angle–range spatial spectrum.

(b) Three-subject angle–range spatial spectrum.

Figure 8. Angle–range heatmaps depicting the averaged spatial-spectrum energy over the observation window. The horizontal and vertical axes represent azimuth angle (°) and range (m), respectively, while color denotes a normalized spectrum magnitude (dB).

proposed pipeline produces clearly localized and well-separated spectral peaks with minimal angular overlap, which indicates an accurate association between detected targets and their true positions. In particular, in the equal-range three-subject case (Scenario S3), DBSCAN produces a single range gate due to range-only grouping, yet MUSIC is executed with $K = K_{\text{gt}} = 3$, which results in three resolvable angular peaks within the same range gate. These results support the low RR and HR estimation errors summarized in Table 6, thus confirming the framework's robustness in multi-target detection and discrimination.

Experiments involving two and three participants were conducted to evaluate the system's ability to simultaneously separate multiple subjects and extract their individual vital signs. During each trial, the participants were seated at predefined distances and azimuths relative to the radar, as detailed in Tables 4 and 5. Each subject was equipped with commercial respiration and heart-rate sensors to provide synchronized ground-truth measurements for validation.

Table 4. Two-subject experimental results.

Subject	Dist. (m)	Angle (°)	Meas. RR	Meas. HR	GT RR	GT HR
First	1.2	30	15.8	62.7	14	67
	0.5	-30	19.9	77.3	18	76
Second	1.0	30	18.2	73.8	18	70
	1.0	-30	24.0	70.9	23	71

Across the multi-subject recordings, the proposed framework achieved an average MAE of 2.52 BPM

Table 5. Three-subject experimental results.

Subject	Dist. (m)	Angle (°)	Meas. RR	Meas. HR	GT RR	GT HR
First	1.0	30	17.00	62.10	16	63
	1.0	0	15.50	77.30	15	76
	1.0	-30	13.80	74.50	14	75
Second	0.7	30	11.72	63.90	14	62
	1.0	0	29.80	85.50	22	83
	1.2	-30	20.50	64.45	12	75

Table 6. Multi-subject performance summary.

Scenario	Metric	Subject 1	Subject 2	Subject 3
S1 (2P, +30°)	MAE_RR (BPM)	1.8	0.2	—
	MAE_HR (BPM)	4.3	3.8	—
S2 (2P, −30°)	MAE_RR (BPM)	1.9	1.0	—
	MAE_HR (BPM)	1.3	0.1	—
S3 (3P, equal range)	MAE_RR (BPM)	1.0	0.5	0.2
	MAE_HR (BPM)	0.9	1.3	0.5
S4 (3P, unequal range)	MAE_RR (BPM)	2.28	7.8	8.5
	MAE_HR (BPM)	1.9	2.5	10.55

for RR and 2.72 BPM for HR. In addition, 90% of HR estimates fell within a ±5 BPM tolerance, thus indicating stable multi-target separation and vital-sign extraction under limited angular spacing.

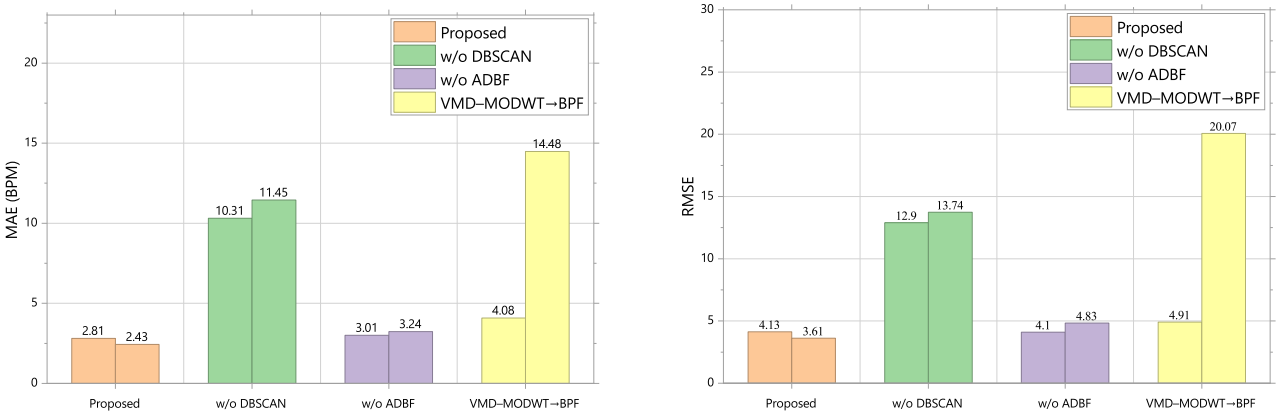
4.4. Ablation study

To quantify the contribution of each processing stage, we performed an ablation study by removing or replacing one module at a time while keeping all remaining settings unchanged. To decouple the module effects from target-count estimation errors, the number of subjects P is provided as a controlled input in all variants. Performance is evaluated using RR/HR MAE and RMSE over all target instances (7 recordings, 13 instances in total). In addition, we report the valid output rate for HR, which is defined as the proportion of instances for which a dominant peak can be reliably identified in the HR band.

As illustrated in Figure 9, DBSCAN plays a critical role in stabilizing the subsequent spatial and physiological estimation. Without DBSCAN-based range gating, the CFAR detections contain a considerable number of outliers and are more likely to lock onto strong static reflectors (e.g., walls or furniture) rather than the thoracic range bin. This leads to a severe performance degradation, with HR MAE increasing from 2.43 to 11.45 BPM and RR MAE increasing from 2.81 to 10.31 BPM (with the corresponding RMSE rising accordingly).

Replacing the ADBF with a non-adaptive FFT beamforming baseline results in a moderate but consistent accuracy loss. Although the coarse angular localization remains acceptable, the reduced interference suppression causes noticeable degradation in vital-sign extraction, thereby increasing HR MAE to 3.24 BPM (RR MAE 3.01 BPM), as shown in Figure 9.

Finally, substituting the proposed VMD–MODWT stage with a conventional band-pass filter reveals that simple filtering is insufficient for a reliable heartbeat estimation in the presence of respiratory



(a) MAE comparison across ablation variants.

(b) RMSE comparison across ablation variants.

Figure 9. Ablation results of the proposed pipeline. the left corresponds to respiration and the right corresponds to heartbeat.

harmonics and residual motion components. While the respiratory component can still be recovered reasonably well (RR MAE 4.08 BPM), the heartbeat estimation substantially deteriorates (HR MAE 14.48 BPM) and the HR validity drops to 92.3%, thus indicating that the MODWT-based secondary decomposition is important to enhance weak cardiac micro-motion and to prevent it from being masked by clutter and harmonics.

4.5. Computational complexity analysis

Let F be the number of frames, N_c be the chirps per frame, N_r be the range-FFT size, and N_v be the number of virtual channels. Let P denote the number of detected targets entering the back-end, R_g be the gated range bins per target, and N_a be the MUSIC scan angles. The front-end Moving target indication and range FFT cost is $O(FN_cN_vN_r \log N_r)$, while the OS-CFAR over range bins is near-linear $O(FN_r)$. For MUSIC–ADBF, the main costs include covariance accumulation $O(PR_gFN_cN_v^2)$, eigen-decomposition $O(PR_gN_v^3)$, and spectrum scanning $O(PR_gN_aN_v^2)$. For VMD, with K modes, I iterations, and slow-time length $L \approx F$, the complexity is approximately $O(PIKL \log L)$. Overall, one trial is dominated by the spatial separation and decomposition stages: $O(FN_cN_vN_r \log N_r)+O\Big(PR_g(FN_cN_v^2 + N_v^3 + N_aN_v^2)\Big)+O(PIKL \log L)$. In our setup, $N_v = 12$ and $N_a = 81$ (-40° to 40° with 1° step), with $F = 1200$, $N_c = 200$, and $N_r = 128$.

4.6. Performance evaluation and comparative analysis

Although the proposed pipeline outputs physiologically plausible RR/HR estimates, small systematic deviations can remain due to spectral peak bias (windowing/leakage), residual clutter, and occasional harmonic locking (particularly in the RR band). In practice, these effects can be well approximated by a first-order measurement model as follows:

$$y_{\text{gt}} = a\, y_{\text{meas}} + b + \varepsilon, \tag{4.4}$$

where y_{meas} denotes the raw radar estimate, y_{gt} is the reference-sensor rate, and (a, b) capture the dominant multiplicative and additive biases. Therefore, we adopt an affine calibration (linear regression)

Table 7. Error metrics before and after linear calibration.

Signal	MAE (BPM)	RMSE (BPM)	MRE (%)
RR (before calib.)	2.81	4.13	19.4
RR (after calib.)	1.75	2.42	11.9
HR (before calib.)	2.43	3.61	3.4
HR (after calib.)	2.05	3.01	3.0

as a parsimonious, interpretable correction that reduces systematic bias while avoiding overfitting to a specific dataset.

Ten representative datasets spanning single-, two-, and three-subject trials were used to estimate (a, b) via least squares: $\widehat{\mathrm{RR}} = 0.5051\,\mathrm{RR}_{\mathrm{meas}} + 6.6139$, $\widehat{\mathrm{HR}} = 0.7216\,\mathrm{HR}_{\mathrm{meas}} + 20.2866$. To assess the robustness, we performed a leave-one-dataset-out evaluation: calibration coefficients were fitted on 9 datasets and applied to the held-out dataset. The improvements were consistent across folds, thus indicating that the affine mapping captures a stable systematic bias rather than memorizing a specific trial. In addition, applying coefficients learnt from single-subject datasets to multi-subject datasets (and vice versa) yielded comparable gains, thus suggesting that the correction is not restricted to a particular subject count or geometry.

Table 7 reports the aggregate performance before and after calibration. The calibrated results are only used when reporting our final accuracy against ground truth; for cross-study comparisons in Table 8, we report the raw (uncalibrated) outputs to avoid confounding differences in calibration protocols across papers.

Calibration improves the estimation accuracy and consistency with the reference sensors. For RR, the MAE and RMSE decrease from 2.81 BPM and 4.13 BPM to 1.75 BPM and 2.42 BPM, respectively, and the MRE decreases from 19.4% to 11.9%. For HR, the MAE amd RMSE reduce from 2.43 BPM and 3.61 BPM to 2.05 BPM and 3.01 BPM, respectively, with the MRE decreasing from 3.4% to 3.0%. These results confirm that affine calibration effectively mitigates systematic deviations and improves the overall agreement with the ground-truth vital-sign measurements.

Following the calibration analysis, we provide an indicative cross-study comparison with representative baseline methods reported in prior work. Since these baselines were evaluated under different hardware setups, datasets, and protocols, the comparison is not a controlled head-to-head evaluation; the numbers are quoted from the original papers to offer context. For consistency with typical reporting in the literature, Table 8 reports the raw (uncalibrated) outputs of our method.

As summarized in Table 7, affine calibration consistently reduces both RR and HR errors (notably RR RMSE from 4.13 to 2.42 BPM and HR RMSE from 3.61 to 3.01 BPM), confirming improved agreement with the reference sensors.

Finally, Figure 9 further verifies the necessity of the key modules (DBSCAN range gating, MUSIC–ADBF separation, and VMD–MODWT decomposition), where removing any of them leads to a clear degradation in RR/HR accuracy and/or HR validity.

5. Conclusions

This paper developed an integrated FMCW MIMO radar framework for non-contact vital-sign monitoring in single- and multi-subject indoor scenarios. Across seven recordings (N = 13 subject-wise

Table 8. Indicative cross-study comparison with representative methods reported in the literature.

Vital-sign type	Filtering [24]	EEMD [7]	CS-OMP [25]	mmRH [26]	Proposed (ours)
Respiration					
MAE (BPM)	3.61	2.96	2.22	1.73	2.81
RMSE (BPM)	3.71	3.06	2.29	1.83	4.13
MRE (%)	18.8	15.6	11.7	9.0	19.4
Heartbeat					
MAE (BPM)	8.83	6.43	4.58	4.02	2.43
RMSE (BPM)	9.17	6.73	4.80	4.29	3.61
MRE (%)	11.3	8.2	5.8	5.1	3.4

instances, 60 s per estimate), the proposed method achieved RR MAE/RMSE of 2.81/4.13 BPM and HR MAE/RMSE of 2.43/3.61 BPM, and improved to 1.75/2.42 BPM (RR) and 2.05/3.01 BPM (HR) after affine calibration. These results highlight a practical design implication: robust multi-person vital-sign sensing should prioritize spatial separation (MUSIC-guided LCMV beamforming) before spectral decomposition (VMD–MODWT), with lightweight calibration used to correct the residual systematic bias.

6. Limitations

The proposed framework integrates multiple processing stages, which inevitably increases the computational cost. In particular, subspace decomposition in MUSIC and the subsequent decomposition-based filtering (VMD–MODWT) dominate the runtime, thus making the overall pipeline heavier than several representative baselines. Moreover, our current experiments are based on an offline PC implementation, so real-time performance and embedded feasibility have not yet been validated on a dedicated hardware platform. In addition, the VMD component is not always optimally tuned: although it generally improves respiration/heartbeat separation, some cases still exhibit residual spectral leakage or suboptimal mode allocation, thus indicating that further optimization is possible. Finally, K is fixed according to the known number of subjects in each controlled trial; an automatic model-order selection will be investigated in future work.

7. Future work

Future work will focus on reducing the computational cost and enabling real-time, long-duration monitoring by migrating the proposed pipeline from an offline PC implementation to an embedded hardware platform. We will profile the end-to-end runtime and replace the most expensive blocks with lower-complexity alternatives suitable for resource-constrained deployment, while maintaining the estimation accuracy. In addition, we will explore learning-aided localization to accelerate the DOA/positioning front-end, such as using neural networks to approximate or refine the covariance estimation in subspace-based DOA, or employing CNN-based models to directly recognize human locations from radar features. These directions aim to improve processing throughput and robustness for practical continuous vital-sign sensing.

LIST OF ABBREVIATIONS

2D-NMS	Two-dimensional non-maximum suppression
ADC	Analog-to-digital converter
ADBF	Adaptive digital beamformer
ADMM	Alternating Direction Method of Multipliers
CFAR	Constant false alarm rate
CUT	Cell under test
DBSCAN	Density-based spatial clustering of applications with noise
DCT	Discrete cosine transform
DOA	Direction-of-arrival
DWT	Discrete wavelet transform
EMD	Empirical mode decomposition
EEMD	Ensemble empirical mode decomposition
FFT	Fast Fourier transform
FMCW	Frequency-modulated continuous wave
HR	Heart rate
IF	Intermediate frequency
IMF	Intrinsic mode function
LCMV	Linearly constrained minimum variance
MFCW	Multiple-frequency continuous wave
MIMO	Multiple-input multiple-output
mmRH	Millimetre-wave radar heart-rate (estimation)
mmWave	Millimetre-wave
MODWT	Maximal overlap discrete wavelet transform
MRE	Mean relative error
MUSIC	Multiple signal classification
MTI	Moving target indication (static-clutter suppression)
MAE	Mean absolute error
OS-CFAR	Ordered-statistics CFAR
PSD	Power spectral density
RF	Radio frequency
RR	Respiration rate
RX	Receiver
RR/HR	Respiration rate / heart rate
RMSE	Root mean square error
SNR	Signal-to-noise ratio
TDM	Time-division multiplexing
TX	Transmitter
ULA	Uniform linear array
WT	Wavelet transform

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by the Research on the Application of Millimeter Wave Radar in Human Body Detection under Grant No. M820125057, and the Start-up Research Projects of Fujian University of Technology under Grant No. GY-Z21064.

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. H. Reggad, X. Jiang, X. Wu, R. Amirtharajah, D. Matthews, X. Liu, A single-chip single-antenna radar for remote vital sign monitoring, *IEEE Trans. Microw. Theory Tech.*, **71** (2023), 4519–4532. https://doi.org/10.1109/TMTT.2023.3267554
2. J. Liu, Y. Li, C. Li, C. Gu, J. F. Mao, Accurate measurement of human vital signs with linear fmcw radars under proximity stationary clutters, *IEEE Trans. Biomed. Circuits Syst.*, **15** (2021), 1393–1404. https://doi.org/10.1109/TBCAS.2021.3123830
3. M. Mercuri, G. Sacco, R. Hornung, P. Zhang, H. J. Visser, M. Hijdra, et al., 2-d localization, angular separation and vital signs monitoring using a siso fmcw radar for smart long-term health monitoring environments, *IEEE Internet Things J.*, **8** (2021), 11065–11077. https://doi.org/10.1109/JIOT.2021.3051580
4. J. Benny, N. N. Moudhgalya, M. Khan, H. K. Meena, M. Wajid, A. Srivastava, Scalable multisubject vital sign monitoring with mmWave FMCW radar and FPGA prototyping, *IEEE Sensors J.*, **25** (2025), 3571–3583. https://doi.org/10.1109/JSEN.2024.3507951
5. J. Tu, J. Lin, Fast acquisition of heart rate in noncontact vital sign radar measurement using time-window-variation technique, *IEEE Trans. Instrum. Meas.*, **65** (2016), 112–122. https://doi.org/10.1109/TIM.2015.2479103
6. M. Li, J. Lin, Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8-ghz cw doppler radar, *IEEE Trans. Microw. Theory Tech.*, **66** (2018), 568–576. https://doi.org/10.1109/TMTT.2017.2730182
7. L. Sun, S. Huang, Y. Li, C. Gu, H. Pan, H. Hong, et al., Remote measurement of human vital signs based on joint-range adaptive eemd, *IEEE Access*, **8** (2020), 68514–68524. https://doi.org/10.1109/ACCESS.2020.2985286
8. Z. L. Xia, X. Wang, H. B. Wei, Y. Xu, Detection of vital signs based on variational mode decomposition using fmcw radar, in *2021 International Conference on Microwave and Millimeter Wave Technology (ICMMT)*, 2021, 1–3. https://doi.org/10.1109/ICMMT52847.2021.9617967

9. L. Qu, C. Liu, T. Yang, Y. Sun, Vital sign detection of fmcw radar based on improved adaptive parameter variational mode decomposition, *IEEE Sensors J.*, **23** (2023), 25048–25060. https://doi.org/10.1109/JSEN.2023.3312513
10. Y. Xiong, Z. Peng, C. Gu, S. Li, D. Wang, W. Zhang, et al., Differential enhancement method for robust and accurate heart rate monitoring via microwave vital sign sensing, *IEEE Trans. Instrum. Meas.*, **69** (2020), 7108–7118. https://doi.org/10.1109/TIM.2020.2978347
11. P. Zheng, C. Zheng, X. Li, H. Chen, A. Wang, Y. Luo, Second harmonic weighted reconstruction for non-contact monitoring heart rate, *IEEE Sensors J.*, **22** (2022), 5815–5823. https://doi.org/10.1109/JSEN.2022.3148003
12. F. Weishaupt, I. Walterscheid, O. Biallawons, J. Klare, Vital sign localization and measurement using an lfmcw mimo radar, in *2018 19th International Radar Symposium (IRS)*, 2018, 1–8. https://doi.org/10.23919/IRS.2018.8448229
13. M. J. López, C. Palacios Arias, J. Romeu, L. Jofre-Roca, In-cabin mimo radar system for human dysfunctional breathing detection, *IEEE Sensors J.*, **22** (2022), 23906–23914. https://doi.org/10.1109/JSEN.2022.3221052
14. J. K. Park, J. H. Park, D. K. Kang, J. H. Jeong, K. T. Kim MPSK-MIMO FMCW Radar-Based Indoor Multipath Recognition, *IEEE Sensors J.*, **24** (2024), 27824–27835. https://doi.org/10.1109/JSEN.2024.3430082
15. G. Paterniani, D. Sgreccia, A. Davoli, G. Guerzoni, P. Di Viesti, A. C. Valenti, et al., Radar-based monitoring of vital signs: A tutorial overview, *Proc. IEEE*, **111** (2023), 277–317. https://doi.org/10.1109/JPROC.2023.3244362
16. J. Bechter, F. Roos, C. Waldschmidt, Compensation of motion-induced phase errors in tdm mimo radars, *IEEE Microw. Wireless Compon. Lett.*, **27** (2017), 1164–1166.
17. S. Rao, Mimo radar (rev. a), Texas Instruments Application Report, SWRA554A, 2018.
18. A. B. Obadi, P. J. Soh, O. Aldayel, M. H. Al-Doori, M. Mercuri, D. Schreurs, A survey on vital signs detection using radar techniques and processing with fpga implementation, *IEEE Circuits Syst. Mag.*, **21** (2021), 41–74. https://doi.org/10.1109/MCAS.2020.3027445
19. A. Ahmad, J. C. Roh, D. Wang, A. Dubey, Vital signs monitoring of multiple people using a fmcw millimeter-wave sensor, in *2018 IEEE Radar Conference (RadarConf18)*, (2018), 1450–1455. https://doi.org/10.1109/RADAR.2018.8378778
20. B. D. Van Veen, K. M. Buckley, Beamforming: A versatile approach to spatial filtering, *IEEE ASSP Mag.*, **5** (1988), 4–24. https://doi.org/10.1109/53.665
21. K. Dragomiretskiy, D. Zosso, Variational mode decomposition, *IEEE Trans. Signal Process.*, **62** (2014), 531–544. https://doi.org/10.1109/TSP.2013.2288675
22. D. B. Percival, A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, 2000.
23. MathWorks, modwt: Maximal overlap discrete wavelet transform, MathWorks Documentation.
24. U. Pangerc, F. Jager, Robust detection of heart beats in multimodal records using slope- and peak-sensitive band-pass filters, *Physiol. Meas.*, **36** (2015), 1645. https://doi.org/10.1088/0967-3334/36/8/1645

25. Y. Wang, W. Wang, M. Zhou, A. Ren, Z. Tian, Remote monitoring of human vital signs based on 77-ghz mm-wave fmcw radar, *Sensors*, **20** (2020), 2999. https://doi.org/10.3390/s20102999

26. L. Liu, J. Zhang, Y. Qu, S. Zhang, W. Xiao, mmRH: Noncontact vital sign detection with an FMCW mm-wave radar, *IEEE Sensors J.*, **23** (2023), 8856–8866. https://doi.org/10.1109/JSEN.2023.3250500

AIMS Press