SPARSE NON-CONTACT MULTIPLE PEOPLE LOCALIZATION AND VITAL SIGNS MONITORING VIA FMCW RADAR

Yonathan Eder, Zhuoyang Liu, and Yonina C. Eldar

Faculty of Math and Computer Science, Weizmann Institute of Science, Israel
Emails: yoni.eder@weizmann.ac.il, zhuoyang.liu@weizmann.ac.il, yonina.eldar@weizmann.ac.il

ABSTRACT

Non-contact vital signs monitoring (NCVSM) of multiple people is becoming a necessity in healthcare due to increasing morbidity and manpower shortage. In meeting these requirements, frequency modulated continuous wave (FMCW) radars have shown great potential. However, current techniques present difficulties in locating and monitoring humans in noisy environments containing multiple objects. In this work, we first develop a model for NCVSM of multiple people via FMCW radar, based on a single-input-multiple-output setup. By considering the sparse nature of the modeled signals along with human-typical cardiopulmonary characteristics, we provide a joint-sparse recovery mechanism to accurately localize targets in a clutter-rich scenario where existing techniques struggle. Then, we present a robust method for NCVSM of the found individuals, with improved performance results when compared to current NCVSM techniques using several statistical metrics. Our approach offers excellent performance in a medical application where high accuracy is required.

Index Terms— Frequency modulated continuous wave radar, joint-sparse recovery, multiple people localization, non-contact vital signs monitoring.

1. INTRODUCTION

Non-contact vital signs monitoring (NCVSM) has become increasingly important in healthcare, owing to factors such as the rise in chronic health conditions, the risk of disease transmission, and the heavy burden on medical staff [1–3]. Radar technology can be ideal in these situations since it does not require users to wear, carry, or interact with any additional electronic device [4]. Initially, the family of continuous wave (CW) radars was proposed to estimate human vital signs, such as heart rate (HR) and respiration rate (RR), by detecting tiny chest wall displacements [5–7]. However, they do not provide the spatial information needed in busy multi-object environments, such as clinics and hospitals [8, 9], where multi-person monitoring is required while ignoring the other objects in space. To overcome this, high-resolution single-input multiple-output (SIMO) frequency-modulated continuous-wave (FMCW) radars can be used to spatially separate humans while simultaneously monitoring their vital signs, even if they are located at the same radial distance from the radar. Nevertheless, in noisy, cluttered environments, prevalent methods for both multiple people localization and NCVSM struggle to provide accurate estimates for future replacement of existing monitoring devices.

The task of continuous NCVSM of multiple people via radar can be divided into two signal processing components: 1. Detect the humans and estimate their spatial location 2. Monitor their vital signs (here, RR and HR). In relation to the localization part, the authors in [10] utilized the fast Fourier transform (FFT) to localize the humans by converting the radar data into the range-azimuth plane using a *range-FFT* followed by an *angle-FFT*. However, relying solely on spectral magnitudes can lead to erroneous decisions due to strong signal reflections from static objects in the radar's field of view (FOV) and poor angle resolution. Some researchers employed the multiple signal classification (MUSIC) algorithm for direction of arrival (DOA) estimation of humans in an indoor environment [11, 12]. Although the method can increase angular resolution, it is sensitive to other vibrating objects in the FOV, such as fans. As a result, a robust method with high resolution and low error rates for localizing humans in noisy environments is required.

Once the humans are correctly located, their vital signs can be extracted and monitored. The most commonly used methods to estimate the vital signs given extracted human thoracic vibration, are based on the discrete Fourier transform (DFT) spectrum, assuming that stationary resting humans have separate heartbeat and breathing frequency bands [13–18]. Despite all the well-known benefits of DFT-spectrum analysis, in NCVSM, the heartbeat signal can be easily masked by the harmonics of the respiratory signal, which reduces estimation accuracy.

In this paper, we develop an extended mathematical signal model for the problem of NCVSM of multiple people in a cluttered scenario, using SIMO FMCW radar. Based on the sparse representation of people via this model, we propose a human spatial localization using a joint-sparse recovery (JSR) mechanism [20–22]. This approach estimates both radial distance and azimuth angle and allows for computationally efficient extraction of Doppler information throughout the complete monitoring process. Then, we demonstrate high-resolution NCVSM of multiple people using the Vital Signs based Dictionary Recovery (VSDR) technique, detailed in [23], which employs a dictionary-based approach to effectively search for the vital signs over high-resolution frequency grids, corresponding to resting cardiopulmonary activity.

The performance of the proposed methodology is verified through simulations that incorporate the developed model with *in vivo* data of 3 monitored individuals from [24]. This study demonstrates accurate human localization in a multiple-object scenario as well as precise NCVSM of multiple people, when compared to state-of-the-art techniques based on [13–18], using several statistical metrics.

The rest of the paper is organized as follows. In Section 2, we propose a model for NCVSM of multiple people via SIMO FMCW radar. Based on this model, we present sparsity-based multiple people localization and NCVSM in Section 3. In Section 4, we assess the quality of the suggested algorithms and compare them with existing techniques. Section 5 concludes this work.

2. SIMO FMCW MODEL FOR NCVSM OF MULTIPLE PEOPLE

A typical linear FMCW radar transmits a saw-tooth waveform at each given time frame, called chirp [19], whose frequency linearly increases over time. The reflected echo signals are mixed with versions of the transmitted ones to obtain analog base-band signals, called beat signals [16, 18]. For each frame, the beat signals are sequentially sampled by the ADC, resulting in discrete signals of length N. We consider a SIMO Uniform Linear Array (ULA) with a single transmitter and K receivers spaced by $r_k \triangleq (k-1)\lambda/2$, $k = 1, ..., K$ where λ denotes the chirp's maximal wavelength. To that end, we extend the single-input-single-output (SISO) signal model suggested by Eder *et al.* [23] for L given frames of N beat samples using $K > 1$ receivers, to the following 3-D discrete beat signal based on PM objects in the radar's FOV

$$y[n,k,l] = \sum_{p=1}^{P} \sum_{m=1}^{M} x_{m,p} e^{j\left(2\pi f_m n T_f + \psi_{m,p}[l] + \phi_p[k]\right)} + w[n,k,l]$$

(1)

where $n = 1, ..., N$, $k = 1, ..., K$, $l = 1, ..., L$ and $\{w[n,k,l]\}$ is a 3-D sequence of zero mean i.i.d. complex Gaussian noise with variance σ^2. Here, T_f denotes the ADC sampling interval and $x_{m,p}$ denotes the received amplitude of the $\{m,p\}$'th object based on its radar cross section (RCS).

Each frequency f_m is distinct and proportional to a different radial distance from the radar d_m:

$$f_m \triangleq \frac{2S}{c} d_m, \quad m = 1, \ldots, M \le N,$$

(2)

where c denotes the speed of light and $S \triangleq B/T_c$ corresponds to the rate of the frequency sweep of each chirp, with B and T_c respectively being the chirp's total bandwidth and duration. The *slow-time* varying phase $\psi_{m,p}[l]$ of each component is given by

$$\psi_{m,p}[l] \triangleq \frac{4\pi}{\lambda}(d_m + v_{m,p}[l]), \quad l = 1, \ldots, L$$

(3)

where the function $v_{m,p}[l]$ refers to small vibrations caused by human thoracic displacements or other vibrating objects, such as fans, and is therefore generally modeled as follows

$$v_{m,p}[l] \triangleq \sum_{q=1}^{Q} a_{m,p,q} \cos(2\pi g_{m,p,q} l T_s), \quad l = 1, \ldots, L,$$

(4)

with the pairs $\{a_{m,p,q}, g_{m,p,q}\}_{q=1}^{Q}$ denoting the corresponding amplitudes and frequencies and $T_s \triangleq 1/f_s$ is the frame duration, also known as the *slow-time* sampling interval. The generic vibration model in (4) allows for adequate representation of both static and vibrating objects, using appropriate values for $\{a_{m,p,q}\}_{q=1}^{Q}$ and $\{g_{m,p,q}\}_{q=1}^{Q}$. Particularly, for the case of Z people located in the radar's FOV, we assume that the frequency set $\{g_{m,p,q}\}_{q=1}^{Q}$ includes their HR and RR, denoted by $f_h^{(z)}$ and $f_r^{(z)}$, respectively, for $z = 1, ..., Z$. Finally, we consider in this work a total of P possible azimuth angles $\{\theta_p\}_{p=1}^{P}$, reflected in the following phase shifts

$$\phi_p[k] = \frac{2\pi}{\lambda} r_k \sin\theta_p, \quad k = 1, \ldots, K \le P.$$

(5)

Define the *slow-time* varying complex amplitude

$$\tilde{x}_{m,p}[l] \triangleq x_{m,p} e^{j\psi_{m,p}[l]}, \quad l = 1, \ldots, L.$$

(6)

For each frame l we can assemble the samples of $y[n,k,l]$ (1) into the matrix $\mathbb{Y}_l \in \mathbb{C}^{N \times K}$, which satisfies

$$\mathbb{Y}_l = \mathbf{A}\mathbb{X}_l^T \mathbf{B}^T + \mathbb{W}_l, \quad l = 1, \ldots, L$$

(7)

where $\mathbf{A} \in \mathbb{C}^{N \times M}$, $M \le N$, $\mathbb{X}_l \in \mathbb{C}^{P \times M}$, $\mathbf{B} \in \mathbb{C}^{K \times P}$, $K \le P$ and $\mathbb{W}_l \in \mathbb{C}^{N \times K}$, whose entries are respectively given by $\mathbf{A}(n,m) \triangleq e^{j2\pi f_m n T_f}$, $\mathbb{X}_l(p,m) \triangleq \tilde{x}_{m,p}[l]$ (6), $\mathbf{B}(k,p) \triangleq e^{j\phi_p[k]}$ and $\mathbb{W}_l(n,k) \triangleq w[n,k,l]$. One can verify that for $K = 1$, the model in (7) coincides with the SISO model in [23, Eq. (14)].

To facilitate the analysis, in the following we convert the model in (7) to a single matrix representation for L given frames. To that end, we first assume that the *fast-time* frequencies $\{f_m\}_{m=1}^{M}$ (2) lie on the Nyquist grid, i.e.,

$$f_m = \frac{f_{\text{ADC}}}{N} i_m, \quad i_m = 0, \ldots, M-1,$$

(8)

where $f_{\text{ADC}} \triangleq 1/T_f$ is determined by the ADC component. Then, by the structure of $\mathbf{A}$ in (7) given (8), we can construct $\widehat{\mathbb{Y}}_l = \mathbf{B}\mathbb{X}_l + \widehat{\mathbb{W}}_l$, $l = 1, ..., L$ where $\widehat{\mathbb{Y}}_l \triangleq \frac{1}{N}\left(\mathbf{A}^H \mathbb{Y}_l\right)^T \in \mathbb{C}^{K \times M}$ and $\widehat{\mathbb{W}}_l \triangleq \frac{1}{N}\left(\mathbf{A}^H \mathbb{W}_l\right)^T \in \mathbb{C}^{K \times M}$. Here, we used the fact that $\mathbf{A}^H \mathbf{A} = N\mathbf{I}_M$, where $\mathbf{I}_M$ denotes a size-M identity matrix, since by (8), $\mathbf{A}(n,m) = e^{j2\pi \frac{im}{N} n}$. We note that since $\mathbb{W}_l$ consists of i.i.d. Gaussian random variables, the statistical properties of the model in (7) are preserved. Next, we concatenate all M columns of $\widehat{\mathbb{Y}}_l$, to obtain a vectorized representation $\widehat{\mathbf{y}}_l \in \mathbb{C}^{KM \times 1}$

$$\widehat{\mathbf{y}}_l = \mathbf{C}\tilde{\mathbf{x}}_l + \widehat{\mathbf{w}}_l, \quad l = 1, \ldots, L,$$

(9)

where $\mathbf{C} \in \mathbb{C}^{KM \times PM}$ denotes a block matrix containing M blocks of $\mathbf{B}$ on its main-diagonal, $\widehat{\mathbf{w}}_l \in \mathbb{C}^{KM \times 1}$ is the transformed noise vector and based on (6) the entries of $\tilde{\mathbf{x}}_l \in \mathbb{C}^{PM \times 1}$ are given by

$$\tilde{\mathbf{x}}_l(i_{m,p}) = x_{i_{m,p}} e^{j\psi_{i_{m,p}}[l]}, \quad l = 1, \ldots, L,$$

(10)

with the row index $i_{m,p} = (m-1)P + p$.

In order to perform continuous NCVSM, the radar should operate and generate data frames throughout the entire monitoring duration. To this end, at each predefined time interval T_{int}, we form the sequence $\{\widehat{\mathbf{y}}_l\}_{l=1}^{L}$ (9) by collecting the last L frames up to that time. The number of frames to be processed, L, is determined by a predefined time window T_{win} according to $L = T_{\text{win}} f_s$, where the units of T_{win} and f_s are $[s]$ and $[1/s]$, respectively. Finally, we reformulate the observations in (9) for L given frames, as follows

$$\mathbf{Y} = \mathbf{C}\tilde{\mathbf{X}} + \mathbf{W},$$

(11)

where $\mathbf{Y} \triangleq \left[\widehat{\mathbf{y}}_1, ..., \widehat{\mathbf{y}}_L\right] \in \mathbb{C}^{KM \times L}$, $\tilde{\mathbf{X}} \triangleq [\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_L] \in \mathbb{C}^{PM \times L}$ and $\mathbf{W} \triangleq [\mathbf{w}_1, ..., \mathbf{w}_L] \in \mathbb{C}^{KM \times L}$ is the noise matrix. We assume here that only $Z \ll PM$ stationary humans are being monitored. Using the above settings, this assumption induces a row-wise sparsity in $\tilde{\mathbf{X}}$, meaning that the vectors $\{\tilde{\mathbf{x}}_l\}_{l=1}^{L}$ share a joint support.

Based on the model in (11), the first goal is to recover the row-coordinates of $\tilde{\mathbf{X}}$ associated with humans in the radar's FOV, denoted by the support $\mathcal{S}$, by which the spatial location of each human can be estimated. The second goal, using the recovered $\mathcal{S}$, is to continuously evaluate the RR and HR of each detected human throughout the complete monitoring duration. Mathematically, every T_{int} we seek to estimate the corresponding $\{f_r^{(z)}, f_h^{(z)}\}_{z=1}^{Z}$ (4).

3. SPARSITY-BASED MULTIPLE PEOPLE LOCALIZATION AND NCVSM

Our sparsity-based solution is divided into two main stages after a preliminary processing. In the first, the support $\mathcal{S}$ is recovered, by which the spatial location of each human is estimated. In the second, we show simultaneous NCVSM of the detected humans using $\mathcal{S}$ and the VSDR method [23], for the discussed scenario.

3.1. Preliminary Processing

At each T_{int}, we assemble the matrix $\mathbf{Y}$ (11) from the given radar data samples in (1) by following relations (7)-(11). To do this, we should fisrt set the dictionary matrices $\mathbf{A}$ and $\mathbf{B}$ (7). For the matrix $\mathbf{A}$, the frequencies $\{f_m\}_{m=1}^M$ satisfy (8). As for $\mathbf{B}$, the phase shift $\phi_p[k]$ (5) is set according to the ULA defined by $\{r_k\}_{k=1}^K$ (Sec. 2) and angle grid covering FOV of 180 degrees

$$\theta_p = -90 + i_p \Delta_\theta, \quad i_p = 0, \dots, P-1 \quad, P = \frac{180}{\Delta_\theta}, \quad (12)$$

where Δ_θ denotes the spacing of the angle grid. Thus, by (5) and (12), $\mathbf{B}(k,p) = e^{j\pi(k-1)\sin\left(-90+i_p\Delta_\theta\right)}$.

3.2. Multiple People Spatial localization

In the first monitoring iteration, we use the assembled matrix $\mathbf{Y}$ (11) to estimate the spatial locations of all stationary humans by recovering $\tilde{\mathbf{X}}$ and its support $\mathcal{S}$. This allows us to efficiently extract the relevant Doppler samples, from which we monitor each individual's vital signs, for the remainder of the monitoring process.

To aid in the separation of humans from static or vibrating clutter, before recovering $\tilde{\mathbf{X}}$, we perform spectral filtering of $\mathbf{Y}$ in the *slow-time* axis based on prior knowledge of human-typical pulse and breathing rates. To that end, we denote by $B^{(R)}$ and $B^{(H)}$ the frequency bands of normal respiration and heartbeat, respectively. The filtered signal is then given by

$$\bar{\mathbf{Y}} = \frac{1}{L}\left(\mathbf{F}_L^H\left(\mathbf{\Pi} \odot \mathbf{F}_L\mathbf{Y}^T\right)\right)^T, \quad (13)$$

where $\mathbf{F}_L$ is a full L-size DFT matrix, $\mathbf{\Pi}$ denotes an ideal window corresponding to the vital frequencies in $B^{(H)} \cup B^{(R)}$ and $\odot$ denotes the element-wise product.

Since we assume that $\tilde{\mathbf{X}}$ is a row-sparse matrix, we now recover it from $\bar{\mathbf{Y}}$ using $\mathbf{C}$ and a JSR technique [20–22], formulated by the following optimization problem

$$\min_{\tilde{\mathbf{X}} \in \mathbb{C}^{PM \times L}} \left\|\bar{\mathbf{Y}} - \mathbf{C}\tilde{\mathbf{X}}\right\|_F^2 + \gamma\left\|\tilde{\mathbf{X}}\right\|_{2,1}. \quad (14)$$

Here, to promote the row sparsity of $\tilde{\mathbf{X}}$, inspired by [20], we use the regularization parameter $\gamma \geq 0$ and the mixed $l_{2,1}$ norm defined by $\|\mathbf{X}\|_{2,1} \triangleq \sum_i \|\mathbf{x}^i\|_2$, with $\mathbf{x}^i$ denoting the i'th row of a matrix $\mathbf{X}$. Similarly to [20], we solve (14) using the fast iterative soft-thresholding algorithm (FISTA) [25, 26].

Now, the support $\mathcal{S}$ can be obtained by selecting the most dominant rows of the recovered $\tilde{\mathbf{X}}$ according to the average power [16] of each row. Note by (10) and since $\tilde{\mathbf{X}} \triangleq [\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_L]$ that each row index of $\tilde{\mathbf{X}}$ satisfies $i_{m,p} = (m-1)P + p$. Hence, by using the modulo operation on the indexes $i_{m,p} \in \mathcal{S}$, we can estimate both the distance d_m and angle θ_p of each monitored individual w.r.t. the radar, through (2) and the grids in (8) and (12). We note that since the humans are stationary, the coordinates of $\mathcal{S}$ are fixed throughout the monitoring, implying that we can recover $\mathcal{S}$ only once, and use it for all subsequent iterations.

3.3. Sparsity Based NCVSM of Multiple People

Since $\tilde{\mathbf{X}}$ is a row-sparse matrix based on $\mathcal{S}$, the model in (11) can be written as

$$\mathbf{Y} = \mathbf{C}_\mathcal{S}\tilde{\mathbf{X}}_\mathcal{S} + \mathbf{W}, \quad (15)$$

with $\mathbf{C}_\mathcal{S} \in \mathbb{C}^{KM \times Z}$ and $\tilde{\mathbf{X}}_\mathcal{S} \in \mathbb{C}^{Z \times L}$ respectively being the atoms of $\mathbf{C}$ and the rows of $\tilde{\mathbf{X}}$ corresponding to $\mathcal{S}$, where $Z = |\mathcal{S}|$. By

holding the support, for each T_{int} we can directly estimate $\tilde{\mathbf{X}}_\mathcal{S}$ from $\mathbf{Y}$, using the Least-Squares (LS) solution [27] given by

$$\widehat{\mathbf{X}}_\mathcal{S} = \left(\mathbf{C}_\mathcal{S}^H\mathbf{C}_\mathcal{S}\right)^{-1}\mathbf{C}_\mathcal{S}^H\mathbf{Y}, \quad (16)$$

considering that $Z \ll KM$.

Using (10) and the definition of $\tilde{\mathbf{X}}$ below (11), we have that $\tilde{\mathbf{X}}_\mathcal{S}(i_{m,p}, l) = x_{i_{m,p}}e^{j\psi_{i_{m,p}}[l]}$, $i_{m,p} \in \mathcal{S}$. That is, $\widehat{\mathbf{X}}_\mathcal{S}$ estimates the *slow-time* varying phasor terms associated with humans in the radar's FOV that contain their vital information in the phase terms $\psi_{i_{m,p}}[l]$, $i_{m,p} \in \mathcal{S}$ (3). Hence, similarly to [23], we use the four quadrant arctangent function on each element of $\widehat{\mathbf{X}}_\mathcal{S}$ followed by unwrapping based on [16], to yield the vibration matrix $\mathbf{V} \in \mathbb{R}^{L \times Z}$. The matrix $\mathbf{V}$ can be viewed as a chain of vectors corresponding to the thoracic vibration pattern of each detected human, i.e., $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_Z]$ where each $\mathbf{v}_z \in \mathbb{R}^L$ contains a scaled approximation of the samples $\{v_{i_{m,p}}[l]\}_{l=1}^L$ (4) for $i_{m,p} \in \mathcal{S}$.

In the final stage of each iteration, both the RR and HR of each individual, $\{f_r^{(z)}, f_h^{(z)}\}_{z=1}^Z$, are estimated simultaneously given $\mathbf{V}$, $B^{(R)}$ and $B^{(H)}$, and recorded for continuous NCVSM. Here, we use the VSDR method, elaborated in [23], which exploits appropriate dictionaries to search for the desired rates over high-resolution grids corresponding to human cardiopulmonary activity.

4. NUMERICAL RESULTS

In this section, the performance of the proposed approach is evaluated and compared to existing techniques, through a simulation based on the model in (1) that involves real Electrocardiography (ECG) and impedance data from [24], which is divided into two parts. The first investigates spatial localization of multiple people in a cluttered environment and the second examines NCVSM given the extracted thoracic vibration of each detected human, versus SNR.

To that end, we generated 5 different objects (of which $Z = 3$ humans) in the radar's FOV using the proposed data model in (1). Each object is characterized by a different set of $x_{m,p}$ (1), d_m (2), θ_p (5) and $\{v_{m,p}[l]\}_{l=1}^L$ (4), as detailed in Table 1. Note that at a distance of $d_m = 0.98$ [m] from the radar there are 2 humans and a vibrating fan (with $a_{\text{fan}} = 10^{-2}$ and $f_{\text{fan}} = 20$ [Hz]), whose position differs only in their azimuth angle. To relate to real thoracic vibrations, we used the 10-minute long 100 [Hz] impedance cardiography signals of subjects $4-6$ from [24]'s resting scenario (in which participants were told to breath calm and avoid large movements), for Table 1's vibrations $\{v_{m,p}[l]\}_{l=1}^L$, with proper adjustments.

Since a variety of respiratory parameters can be extracted from the impedance signal, including RR, [28,29], the raw signal serves as a reference for comparing the RR estimates. As to the HR reference, we used the gold-standard 2000 [Hz] lead-2 ECG signal from [24], and down-sampled it to 100 [Hz] to correspond to $T_s = 10$ [ms]. The rest of the parameters for assembling the model in (1) were set as follows: $\lambda = 3.9$ [mm], $T_c = 57$ [μs], $f_{\text{ADC}} = 4$ [MHz], $S = 70$ [MHz/μs], $M = N/2 = 100$ and $K = 4$. Furthermore, to examine the impact of environmental noise, we used an SNR term that controls the variance of $\{w[n,k,l]\}$ (1) via SNR $\triangleq 1/\sigma^2$. Finally, the frequency bands of respiration and heartbeat were set to $B^{(R)} = [0.1 \ \ 0.4]$ [Hz] and $B^{(H)} = [0.78 \ \ 1.67]$ [Hz], respectively, corresponding to a normal resting state.

Using all the above specifications, we simulate a localization and 10-minute NCVSM of 3 people simultaneously, with RR and HR estimates computed every $T_{\text{int}} = 0.05$ [s], using L data frames from the last $T_{\text{win}} = 30$ [s], starting at T_{win}.

Object type	$x_{m,p}$	d_m [m]	θ_p [°]	$\{v_{m,p}[l]\}_{l=1}^L$
Static clutter	1	0.5	0	0
Human #1	0.3	0.98	-30	Based on [24]
Vibrating fan	0.3	0.98	0	$a_{\text{fan}}\cos\left(2\pi f_{\text{fan}} l T_s\right)$
Human #2	0.3	0.98	$+30$	Based on [24]
Human #3	0.2	1.5	$+15$	Based on [24]

Table 1: Setup of multiple objects scenario.

4.1. Multiple People Localization

Here, we compared the proposed JSR localization method to those based on the *Angle-FFT* [10] and *MUSIC-DOA* [11,12] approaches, for localizing the considered humans.

The parameters of the proposed JSR were set as follows. The frequencies of $\mathbf{\Pi}$ in (13) were drawn from the length-L Nyquist grid determined by f_s, the parameters for solving (14) using FISTA [25] were set to $\gamma = 5$, the Lipschitz constant to $4.5e6$ and 1000 iterations were considered. Finally, the angle grid spacing was set to $\Delta_\theta = 1$. For a fair comparison to the other techniques, all are based on the same angular grid.

Fig. 1 shows the spatial estimates by the examined techniques for SNR $= 0$ [dB]. One can observe that only the proposed JSR method indicates the correct locations (both distance and angle) of the humans compared to the intensity-based *Angle-FFT* method [10] and the *MUSIC-DOA* approach [11,12]. Notice that the *Angle-FFT* method suffers from gross errors especially for $d_m = 0.98$ [m] where there are multiple equidistant objects, since the theoretical angular resolution is $\approx 30°$ for $K = 4$ receivers. The *MUSIC-DOA* approach seeks for highly oscillating objects and thus for $d_m = 0.98$ it incorrectly selects the vibrating fan over the humans. In contrast to the compared localization techniques, the proposed JSR exploits both characteristics of human-typical vital frequencies and prior assumption on the sparse structure of $\tilde{\mathbf{X}}$.
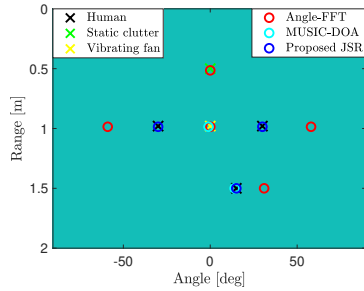

Fig. 1: Multiple people localization by the setup of Table 1, for SNR $= 0$ [dB]. Only the proposed JSR method properly detects the humans in the given scenario.

4.2. NCVSM vs. SNR

Given a successful human localization from the previous study, we examined the performance of VSDR [23] for NCVSM, and compared it to other state-of-the-art techniques based on [13–18].

To compare fairly (regardless of localization performance), we examined only the last step of our algorithm, which estimates human vital signs given the extracted matrix $\mathbf{V}$. In addition, all methods used the same frequency bands $B^{(R)}$ and $B^{(H)}$, and the same stability enhancement procedure by replacing the RR and HR estimates with the average of the last 0.7 seconds' estimates.

The VSDR approach [23] was first compared to the method detailed in [13] for estimating RR and HR given the phase of an FMCW signal, called here Phase-Reg. Moreover, VSDR was compared to a FFT-based peak selection in each frequency band, with zero-padding

(FFT w/ ZP) [18] and without (FFT w/o ZP) [13–17]. The padding of FFT w/ ZP was set to fit a 60-second time window corresponding to frequency resolution of 1 [bpm]. The RR and HR of the reference data was estimated via the DFT spectrum, similarly to [13–16], with padding to fit a 60-second time window for increased resolution, presuming they are noise-free.

Here, performance is evaluated using the following metrics: 1. Success Rate (SR) - 2 bpm, defined here as the percentage of times the estimate differed from the reference output by less than 2 [bpm]. 2. Pearson Correlation Coefficient (PCC). 3. Mean-Absolute Error (MAE), and 4. Root-Mean-Square Error (RMSE). Since we investigate NCVSM of $Z = 3$ people for different SNR cases, the performance score produced for each metric and SNR is regarded as the average across the participants' scores. We note that the simulation settings brings to 11400 non-contact HR/RR estimates for each participant to compare to the contact HR/RR reference estimates.

The top 3×5 block of Fig. 2 shows NCVSM of the detected humans compared to the references by VSDR and the other examined techniques, for SNR $= 0$ [dB]. It can be seen how both the HR and RR estimates by the VSDR approach show great resemblance to those of the reference, compared to the others in which the noisy setup impairs their assessments. The bottom 2×4 block of Fig. 2 shows the SR - 2 bpm, PCC, MAE and RMSE for both HR and RR estimation by all examined methods, as a function of the SNR. One sees that VSDR outperforms the other compared methods in all 4 metrics, for every SNR value.
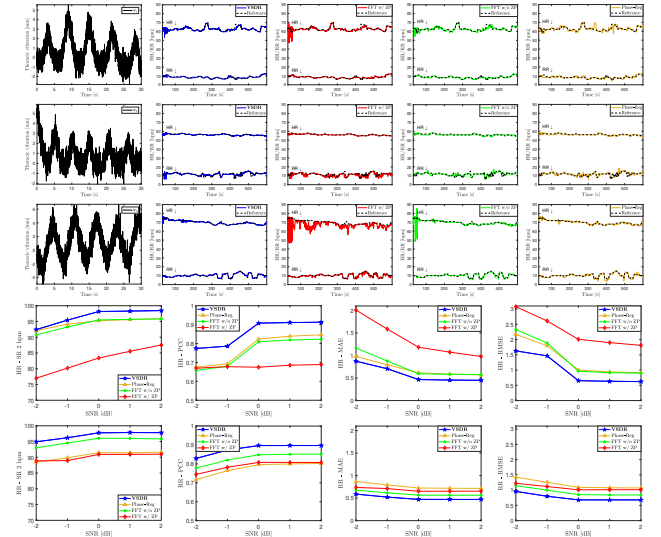

Fig. 2: **Top** 3×5 **block:** Comparison of NCVSM techniques given $\mathbf{V}$, w.r.t. ground truth. (Rows) NCVSM of Table 1's humans #1-#3, for SNR $= 0$ [dB]. (Columns) $\mathbf{v}_1$-$\mathbf{v}_3$ for some T_{int}, VSDR, FFT w/ ZP, FFT w/o ZP, Phase-Reg. **Bottom** 2×4 **block:** mean NCVSM performance vs. SNR. (Rows) HR, RR. (Columns) SR - 2 bpm, PCC, MAE, RMSE.

5. CONCLUSION

In this paper, we proposed a model for NCVSM of multiple people based on SIMO FMCW radar. We presented a JSR approach using this model that can accurately localize targets in a clutter-rich scenario involving equidistant people, where known localization methods perform poorly. Furthermore, the VSDR method was used to estimate the RR and HR of the detected individuals, yielding improved performance results when compared to contemporary NCVSM techniques using 4 evaluation metrics.

6. REFERENCES

[1] S. J. Brownsell, G. Williams, D. A. Bradley, R.Bragg, P. Catlin, and J. Carlier, "Future systems for remote health care," *Journal of Telemedicine and Telecare,* vol. 5, no. 3, pp. 141–152, 1999.

[2] O. Boric-Lubeke and V. M. Lubecke, "Wireless house calls: using communications technology for health care and monitoring," *IEEE Microwave Magazine,* vol. 3, no. 3, pp. 43–48, 2002.

[3] V. Nangalia, D. Prytherch, and G. Smith, "Health technology assessment review: Remote monitoring of vital signs-current status and future challenges," *Critical Care,* vol. 14, no. 5, pp. 1–8, 2010.

[4] F. Fioranelli, J. Le Kernec, and S. Shah, "Radar for health care: Recognizing human activities and monitoring vital signs," *IEEE Potentials,* vol. 38, no. 4, pp. 16–23, 2019.

[5] Y. Xiao, J. Lin, O. Boric-Lubecke, and V.M. Lubecke, "A Ka-band low power Doppler radar system for remote detection of cardiopulmonary motion," *In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference,* pp. 7151–7154, Jan, 2006.

[6] C. Gu, Y. He, and J. Zhu, "Noncontact vital sensing with a miniaturized 2.4 G[Hz] circularly polarized Doppler radar," *IEEE Sensors Letters,* vol. 3, no. 7, pp. 1–4, 2019.

[7] H. Zhao, H, Hong, D. Miao, Y. Li, H. Zhang, Y. Zhang, C. Li, and X. Zhu, "A noncontact breathing disorder recognition system using 2.4-GHz digital-IF Doppler radar," *IEEE journal of biomedical and health informatics,* vol. 23, no. 1, pp. 208–217, 2018.

[8] J. J. Hillhouse, and C. M. Adler, "Investigating stress effect patterns in hospital staff nurses: results of a cluster analysis," *Social Science & Medicine,* vol. 45, no. 12, pp. 1781–1788, 1997.

[9] E. R. Greenglass, R. J. Burke, and L. Fiksenbaum, "Workload and burnout in nurses," *Journal of community & applied social psychology,* vol. 11, no. 3, pp. 211–215, 2001.

[10] A. Ahmad, J. Roh, D. Wang, and A. Dubey, "Vital signs monitoring of multiple people using a FMCW millimeter-wave sensor," *In 2018 IEEE Radar Conference (RadarConf18),* pp. 1450—1455, 2018.

[11] D. Sasakawa, N. Honma, T. Nakayama and S. Iizuka, "Fast Living-Body Localization Algorithm for MIMO Radar in Multipath Environment," in IEEE Transactions on Antennas and Propagation, vol. 66, no. 12, pp. 7273-7281, Dec. 2018, doi: 10.1109/TAP.2018.2870405.

[12] K. Konno, M. Nango, N. Honma, K. Nishimori, N. Takemura and T. Mitsui, "Experimental evaluation of estimating living-body direction using array antenna for multipath environment," *IEEE Antennas and Wireless Propagation Letters,* vol. 13, pp. 718–721, 2014.

[13] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R.C. Miller, "Smart homes that monitor breathing and heart rate," *In Proceedings of the 33rd annual ACM conference on human factors in computing systems,* pp. 837–846, 2015.

[14] M. Mercuri, I. Lorato, Y.H. Liu, F. Wieringa, C. Van Hoof, and T. Torfs, "Vital-sign monitoring and spatial tracking of multiple people using a Non-contact radar-based sensor," *Nature Electronics,* vol. 2, no. 6, pp. 252–262, 2019.

[15] G. Sacco, E. Piuzzi, E. Pittella, and S. Pisa, "An FMCW radar for localization and vital signs measurement for different chest orientations," *Sensors,* vol. 20, no. 12, pp. 3489, 2020.

[16] M. Alizadeh, G. Shaker, J. De Almeida, P. Morita, and S. Safavi-Naeini, "Remote monitoring of human vital signs using mm-wave FMCW radar," *IEEE Access,* vol. 7, pp. 54958–54968, 2019.

[17] E. Antolinos, F. García-Rial, C. Hernández, D. Montesano, J. I. Godino-Llorente, and J. Grajal , "Cardiopulmonary Activity Monitoring Using Millimeter Wave Radars," *Remote Sensing,* vol. 12, no. 14, pp. 2265, 2020.

[18] E. Turppa, J. Kortelainen, O. Antropov, and T. Kiuru, "Vital sign monitoring using FMCW radar in various sleeping scenarios," *Sensors,* vol. 20, no. 22, pp. 6505, 2020.

[19] IWR1642 Single-Chip 76- to 81-GHz mmWave Sensor (Rev. B). Available: https://www.ti.com/document-viewer/IWR1642/datasheet.

[20] U. Rossman, R. Tenne, O. Solomon, I. Kaplan-Ashiri, T. Dadosh, Y. C. Eldar, and, D. Oron, "Rapid quantum image scanning microscopy by joint sparse reconstruction," *Optica,* vol. 6, no. 10, pp. 1290–1296, 2019.

[21] Y. C. Eldar, *Sampling theory: Beyond bandlimited systems,* U.K., Cambridge Univ. Press, 2015.

[22] Y. C. Eldar, and G. Kutyniok *Compressed sensing: theory and applications,* U.K., Cambridge Univ. Press, 2012.

[23] Y. Eder, and Y. C. Eldar, "Sparsity Based Non-Contact Vital Signs Monitoring of Multiple People Via FMCW Radar," *arXiv preprint arXiv:2205.05152,* 2022.

[24] S. Schellenberger et al., "A dataset of clinically recorded radar vital signs with synchronised reference sensor signals," *Scientific data,* vol. 7, no. 1, pp. 1–11, 2020.

[25] A. Beck, and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences,* vol. 2, no. 1, pp. 183–202, 2009.

[26] D. P. Palomar, and Y. C. Eldar, *Convex optimization in signal processing and communications,* Cambridge Univ. Press, 2010.

[27] D. Ruppert, and M. P. Wand, "Multivariate locally weighted least squares regression," *The annals of statistics,* pp. 1346–1370, 1994.

[28] J. M. Ernst, D. A. Litvack, D. L. Lozano, J. T. Cacioppo, and G. G. Berntson, "Impedance pneumography: Noise as signal in impedance cardiography," *Psychophysiology,* vol. 36, no. 3, pp. 333–338, 1999.

[29] A. Sherwood, M. T. Allen, J. Fahrenberg, R. M. Kelsey, W. R. Lovallo, and L. J. Van Doornen, "Methodological guidelines for impedance cardiography," *Psychophysiology,* vol. 27, no. 1, pp. 1–23, 1990.