Temporal-Spatial Equivalent Virtual Array Technique for Accurate Vital Sign Monitoring

Yuchen Li[#*], Jingyun Lu[#*], Shuqin Dong[#*], Changzhan Gu[#*], Junfa Mao[#*]
[#]State Key Lab of Radio Frequency Heterogeneous Integration (Shanghai Jiao Tong University), China
[*]MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University, China
changzhan@sjtu.edu.cn

***Abstract*—A novel temporal-spatial equivalent virtual array technique for accurate vital sign monitoring based on a 24GHz continuous wave (CW) radar is proposed. By leveraging the similarity between the single-input-single-output (SISO) CW radar single-channel phase signal and the single-input-multiple-output (SIMO) radar multiply-array signal, an equivalent virtual array could be synthesized, which completes the transformation from time to space. The derivation of the equivalent theory reveals that the different frequency components of human thorax surface motion can be equivalent to different direction-of-arrival (DOA) angles of waves incident on the virtual array. Subspace-based DOA estimation algorithm is employed to extract the array features and a super-resolution spectral estimation performance is achieved. A coarse-to-fine estimation method is presented for accurate heart rate and respiration rate monitoring. Experiments were carried out and the results proved that the proposed technique can accurately and stably estimate the respiration rate and heart rate of the human body with super-resolution.**

***Keywords*—CW radar, millimeter-wave, radar sensing, virtual array, vital sign monitoring.**

I. Introduction

In recent years, non-contact vital sign monitoring has attracted people's attention in many fields, such as sleep monitoring, hospital physical examination, auxiliary early warning in nursing institutions, and measurement of the health status of the elderly living alone [1]. Since vital sign signal can be obtained without physical contact, this method is more user-friendly and leverages more convenience to users.

Conventional non-contact vital sign monitoring methods mainly include camera-based solutions and millimetre-wave radar-based solutions. Camera-based solutions are limited by lighting conditions and privacy risks, which make it difficult to guarantee the accuracy and consistency of vital sign detection. Many studies have been conducted regarding the radar-based vital sign detection. Basically, vital sign detection can be formulated as a spectrum estimation problem. The most popular method at present is the discrete Fourier transform (DFT) and the time domain feature extraction method [2-4]. However, the spectral leakage and resolution limitation caused by sample data length greatly reduce the accuracy of DFT. In addition, since the amplitude of chest displacement caused by respiration is much stronger than that caused by heartbeat, the heartbeat signal can be easily overwhelmed by the third or fourth harmonic of the respiration in frequency spectrum [5].

In this paper, a novel vital sign monitoring technique with super-resolution is presented based on a signal channel CW radar. The temporal-spatial equivalent virtual array (TSEVA) principle is derived and the TSEVA-based respiration rate and heart rate monitoring method is proposed. As is shown in Fig. 1, the equivalence is divided into two levels, namely the array level and the signal feature level. The derivation of the equivalence theory and the detailed specification of the technique will be covered in the rest part of the paper.

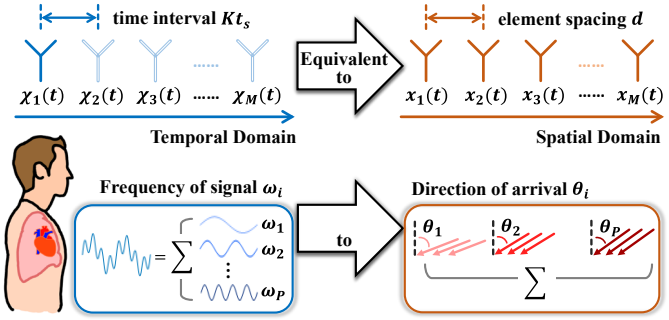

Fig. 1. Schematic diagram of the temporal-spatial equivalent virtual array technique.

II. Principles

A. Temporal-Spatial Equivalent Virtual Array Theory

The proposed technique is based on a single channel CW radar and the system block diagram of the radar sensor is shown in Fig. 2. The electromagnetic wave is emitted by Tx antenna, and the echo with the human chest movement information is received by Rx antenna after being reflected. A DC-coupled architecture is employed to amplify two mutually orthogonal I/Q channels IF signals. The complex form of the IF signal recombined from I/Q channels can be expressed as

$$X_{IF}(t) = A_{IF} \cdot \exp\left[j\left(\frac{4\pi}{\lambda}\left(d_0 + d(t)\right) + \Delta\theta(t)\right)\right], \quad (1)$$

where A_{IF} is the amplitude of the signal, λ is the wavelength, d_0 is the initial distance between human and radar, $d(t)$ is the displacement of the thorax surface and $\Delta\theta(t)$ is the residual phase. The displacement of the human thorax surface $d(t)$ is mainly produced by the breathing action of the lungs, the beating of the heart, and the accumulation of a series of harmonics of cardiopulmonary activities:

$$d(t) = \sum_{i=1}^{P} A_i \cdot \exp[\,j(\omega_i t + \varphi_i)], \quad (2)$$

where P is the number of all frequency components, $A_i/\omega_i/\varphi_i$ are the amplitude, angular frequency and the initial phase of

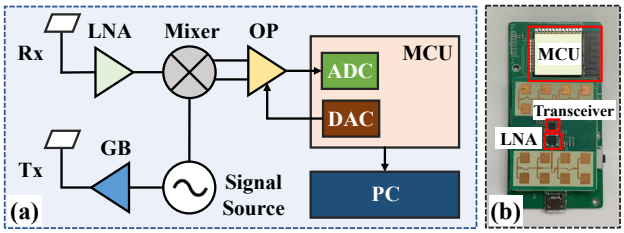

Fig. 2. Block diagram (a) and the photo (b) of the SIMO CW radar system.

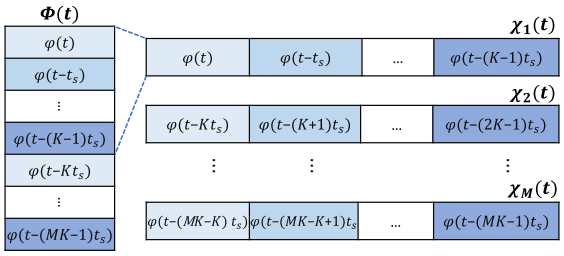

Fig. 3. Schematic diagram of the data segmental reconstruction procedure.

the *i-th* frequency component, respectively. The MDACM algorithm is employed to demodulate the phase information in (5), which can be expressed as: $\Phi(t) = d(t) + n(t)$, where $n(t)$ is the additive white Gaussian noise caused by residual phase, environmental clutter, and other types of noise introduced from the circuit and sampling process. Denote the observation time applied for vital sign signal extraction as T_{OI}, A total of N_{Total} points of data are sampled with the sampling rate f_s, where $N_{Total} = T_{OI} \times f_s$.

A data segmental reconstruction procedure is performed to transform the temporal single-channel $\Phi(t)$ signal into spatial multiple virtual array channel signals, and the schematic diagram of the procedure is shown in Figure 3. The single channel data $\Phi(t)$ is divided into M segments with data length K, and the data matrix of each segment can be regarded as the received data of each receiving element in the virtual array:

$$\chi_i(t) = \Phi(t - (i-1)Kt_s), \tag{3}$$

where $\chi_i(t)$ is the data of the *i-th* virtual array, $t_s = 1/f_s$ is sampling interval. The received data of the whole virtual array can be organized as:

$$\begin{aligned} \boldsymbol{\chi}(t) &= \begin{bmatrix} \chi_1(t) \\ \chi_2(t) \\ \dots \\ \chi_M(t) \end{bmatrix} = A(\omega)S(t) + N(t) \\ &= \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{j\omega_1 Kt_s} & e^{j\omega_2 Kt_s} & \dots & e^{j\omega_P Kt_s} \\ \dots & \dots & \dots & \dots \\ e^{j(M-1)\omega_1 Kt_s} & e^{j(M-1)\omega_2 Kt_s} & \dots & e^{j(M-1)\omega_P Kt_s} \end{bmatrix} \\ &\times \begin{bmatrix} A_1 e^{j(\omega_1 t+\varphi_1)} \\ A_2 e^{j(\omega_2 t+\varphi_2)} \\ \dots \\ A_P e^{j(\omega_P t+\varphi_P)} \end{bmatrix} + \begin{bmatrix} n(t) \\ n(t-Kt_s) \\ \dots \\ n(t-(M-1)Kt_s) \end{bmatrix} \end{aligned} \tag{4}$$

where $A(\omega) = [a(\omega_1), a(\omega_2), \dots, a(\omega_P)]$ stands for the array manifold vector, $a(\omega_i) = \left[1, e^{j\omega_i Kt_s}, \dots, e^{j(M-1)\omega_i Kt_s}\right]^T$ is the steering vector of ω_i, $S(t) = [s_1(t), s_2(t), \dots, s_P(t)]^T$ is the signal vector, and $s_i(t) = A_i e^{j(\omega_i t+\varphi_i)}$ is the signal of the *i-th*

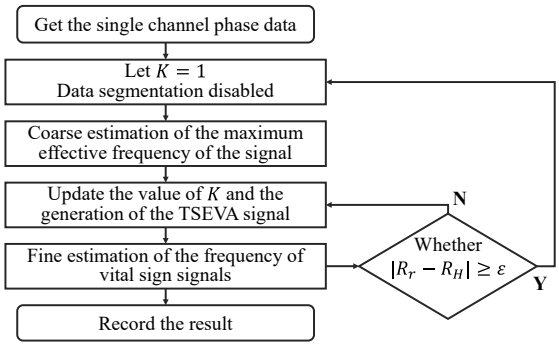

Fig. 4. Flow chart of the TSEVA-based vital sign monitoring method.

frequency component.

After the above processing, a temporal-spatial equivalent virtual array signal is generated. Considering a uniform linear array (ULA) of M elements with a spacing of d, assuming that there are P signals incident on the array with different DOA angles, the overall received signal $\boldsymbol{X}(t)$ can be expressed as:

$$\boldsymbol{X}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T = A'(\theta)S'(t) + N'(t) \tag{5}$$

where $x_i(t)$ is the received signal of the *i-th* element, $A'(\theta) = [a'(\theta_1), a'(\theta_2), \dots, a'(\theta_P)]$ is the array manifold vector and $a'(\theta_i) = \left[1, e^{j2\pi d\sin\theta_i/\lambda}, \dots, e^{j(M-1)2\pi d\sin\theta_i/\lambda}\right]^T$ is the steering vector of DOA angle θ_i, $S'(t) = [s_1'(t), s_2'(t), \dots, s_P'(t)]^T$ is the signal source matrix and $N'(t)$ stands for the white noise.

It can be concluded that the received signal of the M element ULA is very similar to the signal of the temporal-spatial equivalent virtual array through the comparison between (4) and (5). Let $A(\omega)$ be equal to $A'(\theta)$, the equivalent condition can be obtained as $a(\omega_i) = a'(\theta_i)$:

$$e^{j\omega_i Kt_s} = e^{j2\pi d\sin\theta_i/\lambda} \tag{6}$$

The above equation can be further simplified when the element spacing $d = \lambda/2$. Then, the relationship between the motion frequency f_i and the incoming wave DOA angle θ_i is:

$$\theta_i = \sin\left(2K\frac{f_i}{f_s}\right)^{-1}. \tag{7}$$

It should be noted that the $\sin(\cdot)$ function is more sensitive when angle is close to 0, so a large value of K is more appropriate for a more accurate estimation f_i. However, in order to avoid ambiguity when solving angles in (6), there is a upper bound for K, which is $\omega_i Kt_s \le \pi$. In conclusion, the number of equivalent linear array elements $M = T_{OI}/Kt_s$, the angle of equivalent incident signal $\theta_i = \sin(2Kf_i/f_s)^{-1}$, and the optimal number of each channel's snapshots $K = \lfloor f_s/2f_{Max} \rfloor$ where f_{Max} is the maximum frequency in $d(t)$.

B. TSEVA-based Vital Sign Monitoring

Since the array generated by TSEVA is a uniform linear array, several super-resolution spectrum estimation algorithms based on array features can be adopted, such as the estimating signal parameters via rotational invariance techniques signal parameters via rotational invariance techniques (ESPRIT). TLS-ESPRIT algorithm is adopted in this paper in order to achieve a high-resolution spectrum estimation result. Divide

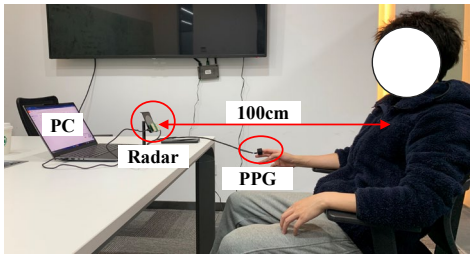

Fig. 5. Photograph of the experimental setup.

the virtual array into two $M-1$ elements sub-arrays $\boldsymbol{\chi_1}$ and $\boldsymbol{\chi_2}$, where $\boldsymbol{\chi_1} = [\chi_1(t), \chi_2(t), \ldots, \chi_{M-1}(t)]^T$ and $\boldsymbol{\chi_2} = [\chi_2(t), \chi_3(t), \ldots, \chi_M(t)]^T$. There is a diagonal matrix ψ that reflects the phase difference between the two subarrays' manifold vector, and the elements on the diagonal of ψ represent the frequencies of the individual motion components on the thorax surface: $\psi = \mathrm{diag}[e^{j\omega_1 K t_s}, e^{j\omega_2 K t_s}, \ldots, e^{j\omega_P K t_s}]$. By solving the eigenvalues of the covariance matrix of the two sub-arrays, the value of the diagonal array ψ can be finally obtained, and thus the frequency components contained in the surface displacement of the thorax can be estimated.

The flow chart of TSEVA-based vital sign monitoring method is shown in Fig. 4. The first step is to obtain the phase information of a single channel. Next, set $K = 1$ to obtain the largest equivalent array aperture. Due to the small number of snapshots, MUSIC algorithm is used for rough spectrum estimation to obtain the frequency interval where the heartbeat signal is located. Generally, the heart rate of a normal person is within the interval [0.1Hz, 0.3Hz]. Therefore, the maximum value of the spectrum peak in this interval is considered as a rough estimate of the heart rate, and then K is updated and the TSEVA signal is generated. In the next step, TLS-ESPRIT algorithm is employed to find the frequency corresponding to the largest eigenvalue in the heart rate interval as the heart rate estimation value R_h and that in the respiration rate interval as the respiration rate estimation value R_r. Also, R_h will be used to compare with the historical value of heart rate R_H. When the difference exceeds ε (e.g., 0.1Hz), it is considered that the heart rate has changed greatly, and the coarse-fine heart rate estimation procedure needs to be restarted.

III. Experiments and Results

A custom-designed interferometric radar operating at 24GHz is employed in the experimental part. The photograph of the radar sensor is shown in Fig. 2(b) and the experimental setup is shown in Fig. 5. The experiments are conducted in office environment. The people sit in a chair and maintain a stable posture, and the distance between the chest and the radar is about 1m. The photoplethysmography (PPG) device worn on the finger is sampled synchronously with the radar signal and it is served as the ground truth.

DFT and the original MUSIC algorithm was performed to make a comparison with the proposed method. Since the ESPRIT algorithm directly obtains the frequency values and cannot generate a continuous spectrum, a 2000 times Monte-Carlo simulation is carried out to obtain the distribution of the TSEVA-based results and illustrate its spectrum composition. Because the PPG signal is a relatively more stable signal with

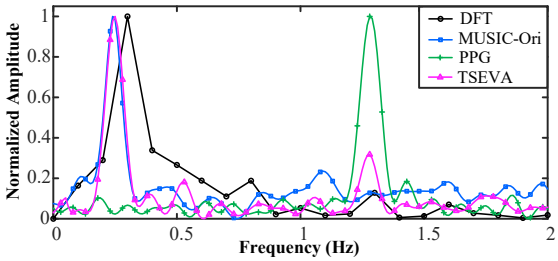

Fig. 6. Comparison between the results of DFT, original MUSIC algorithm, the proposed TSEVA-based method and the ground truth PPG signal.

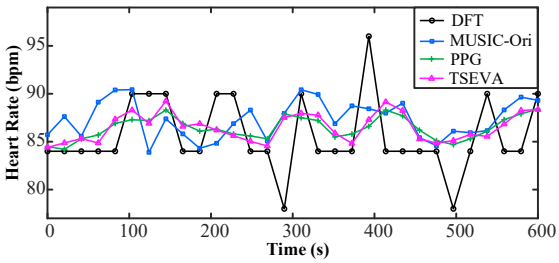

Fig. 7. Heart rate estimation results over a 10-minute data.

high SNR, super-resolution algorithm is performed on it and its result is regarded as the ground truth. The comparison of the results is shown in Fig. 6, and $T_{OI} = 10$s and $f_s = 100$Hz. Since the frequency resolution of DFT is only 0.1Hz, it cannot obtain accurate respiration rate and heart rate. The original MUSIC algorithm achieves good performance on respiration rate estimation, but it cannot locate the exact frequency peak of the heartbeat. The result obtained by the proposed TSEVA-based method is 0.253Hz for respiration and 1.261Hz for heartbeat, which is the same as the result of PPG.

The heart rate estimation results by different methods for a period of 10 minutes data is shown in Fig. 7. Because of the resolution limitation, the estimation result of DFT method is the worst with a RMSE=5.72. The results obtained by the original MUSIC algorithm achieve good accuracy in some time periods, but there are large fluctuations and the overall RMSE=2.31. The proposed TSEVA-based method performed best, with high accuracy and high stability over the entire time period, and RMSE=0.875.

IV. Conclusion

In this paper, a novel temporal-spatial equivalent virtual array technique is proposed for accurate vital sign monitoring. A virtual array is established, and the different frequency components of human thorax surface motion are equivalent to different incoming angles of waves incident on the virtual array. A coarse-to-fine estimation method is presented for respiration rate and heart rate monitoring. Experiments were carried out and the results show that the proposed technique can detect human vital sign signal with super-resolution and maintain high accuracy.

Acknowledgement

This work was supported in part by the National Key R&D Program of China under Grant 2020YFB1807304, and Natural Science Foundation of China under Grant 62171277 and Grant 62188102.

REFERENCES

[1] C. Li, *et al.*, "A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 2046–2060, May 2013.

[2] W. Xia, Y. Li, and S. Dong, "Radar-Based High-Accuracy Cardiac Activity Sensing," IEEE *Trans. Instrum. Meas.*, vol. 70, pp. 1-13, 2021.

[3] C. Zhu *et al.*, "Doppler cardiogram detected by a V-band Doppler radar sensor," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 1, pp. 521–531, Jan. 2022.

[4] E. Cardillo, C. Li and A. Caddemi, "Vital Sign Detection and Radar Self-Motion Cancellation Through Clutter Identification," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 3, pp. 1932-1942, March 2021.

[5] S. Dong, *et al.*, "Contactless Measurement of Human Systolic Time Intervals Based on Doppler Cardiograms in Clinical Environment," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 6, pp. 796-799, June 2022.