

Heart Rate Estimation of Car Driver Using Radar Sensors and Blind Source Separation

Keito Murata*, Daichi Kitamura*, Ryo Saito†, and Daichi Ueki†

* National Institute of Technology, Kagawa College, Kagawa, Japan

† Murata Manufacturing Co., Ltd., Kyoto, Japan

Abstract—Monitoring a driver’s health condition is essential for advanced driver assistance systems. To detect the driver’s health condition, we propose a new system using a contactless biometric radar sensor to monitor the driver’s heart rate. The radar sensor measures a microdisplacement of the driver’s body surface, and heartbeat signals can be obtained. However, the observed signals are contaminated by noise signals such as body movement and car vibration. The proposed system utilizes blind source separation algorithms and extracts the driver’s heartbeat signal from the noisy observation. Furthermore, the estimated heartbeat signal is input to a heart rate estimation algorithm. In experiments, we built a vibration measurement system to simulate a moving car, and an actual noisy heartbeat signal from a subject sitting on a driver’s seat was measured. The results show that the proposed system can accurately estimate the subject’s heart rate by removing the noise signals.

I. INTRODUCTION

It is fatal for a driver to have a seizure or lose consciousness while driving a car. Thus, monitoring the driver’s condition in a car is a crucial issue. In this paper, to address this problem, we propose a new system for measuring and estimating the driver’s heart rate while driving a car. The proposed system uses a contactless biometric radar sensor [1] (hereafter referred to as a *radar sensor*), as shown in Fig. 1 (a). This radar sensor can measure a microdisplacement of the driver’s body surface, and signals originating from the heartbeat can be obtained.

To simulate a moving car condition, we built a vibration measurement system depicted in Fig. 1 (b). The radar sensor is embedded in the back of the driver’s seat so that the microdisplacement of the driver’s back is measured. Our objective is to estimate the heart rate from the microdisplacement of the body surface. However, the observed signal is contaminated by enormous noise, mainly a car vibration. A vibration plate simulates this noise signal in our vibration measurement system, where the plate vibrates in the vertical direction. In addition, the driver’s breathing also causes a significant movement of the body surface and appears in the observed signal.

This paper aims to improve the accuracy of heart rate estimation using blind source separation (BSS) [2], [3]. BSS is applied to the observed signals measured by the radar sensor in the vibration measurement system and extracts the displacement of the body surface derived from the driver’s heartbeat. In particular, we apply independent vector analysis (IVA) [4], [5], independent low-rank matrix analysis (ILRMA) [6], [7], and ILRMA based on the complex Student’s *t* distribution (*t*-ILRMA) [8], [9]. These algorithms have been well investigated

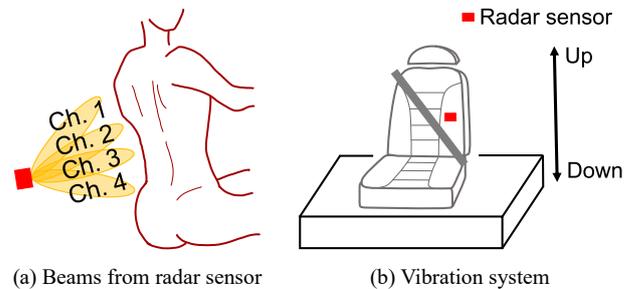


Fig. 1. Vibration measurement system with the radar sensor embedded in a back of seat.

in the field of audio BSS. We compare the precision of heart rate estimation with these BSS algorithms.

II. OVERVIEWS OF VIBRATION MEASUREMENT SYSTEM AND OBSERVED SIGNAL

A. Conditions for Vibration Measurement System

The vibration measurement system built for an experiment is shown in Fig. 1 (b). In this system, the radar sensor is embedded in the back of the seat, and the microdisplacement of the subject’s back surface can be measured. We measured for 420 s in total, and the vibration plate was operated only in 60–360 s. The input signal to the vibration plate was a sinusoidal wave with a 10 mm amplitude and 1.2 Hz frequency towards the vertical direction, which simulates car movements of idling or driving. The radar sensor emits millimeter waves with four different directivities, as shown in Fig. 1 (a), and receives the reflected waves. Thus, the microdisplacements of four points can be simultaneously measured as a four-channel signal. The sampling frequency of this radar sensor is 40 Hz. In addition, a contact-type electrocardiograph (ECG) sensor (hereafter referred to as an *ECG sensor*) is attached to the subject’s chest and is operated on during the measurement. A reference heart rate can be obtained by analyzing the ECG sensor signal, which is used in the evaluation.

B. Observed Spectrograms of Radar and ECG Sensors

The spectrograms of the observed signals obtained by the radar sensor (ch. 1) and the ECG sensor are shown in Fig. 2. We can confirm that the observed signal of the radar sensor contains components of not only heartbeat but also the vibration plate noise, breathing motion, and background noise. On the other hand, the ECG sensor captures the components of

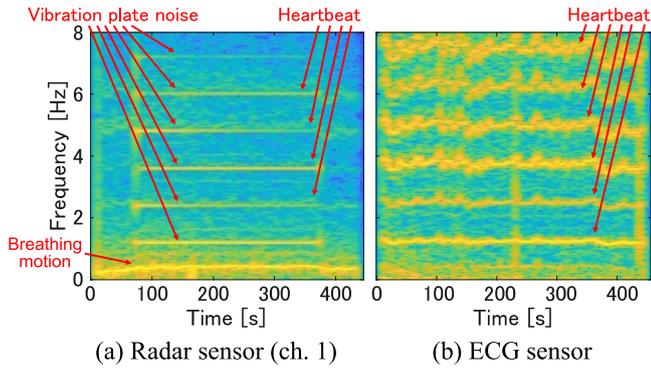


Fig. 2. Spectrograms of the observed signals obtained by the radar sensor (left) and the ECG sensor (right).

the heartbeats with a high signal-to-noise ratio. This is because the ECG sensor is a contact-type device and is attached to the chest.

To extract only the heartbeat components from the radar sensor signal, we apply a BSS technique and a heart rate estimation algorithm. The accuracy of the estimated heart rates is evaluated by comparing the reference heart rates calculated from the ECG sensor signal.

III. BSS AND HEART RATE ESTIMATION ALGORITHM

This section presents BSS formulation in the time-frequency domain and the algorithms of IVA, ILRMA, and t -ILRMA. Also, an algorithm for estimating heart rate is described.

A. Formulation of BSS

As described in Sects. II-A and II-B, we assume that the observed radar sensor signal consists of a four-channel time-domain signal and is contaminated by some noise sources. Let $\tilde{\mathbf{s}}$ be the N source signals that include heartbeat and the other $N - 1$ noise signals:

$$\tilde{\mathbf{s}}[l] = [\tilde{s}_1[l], \tilde{s}_2[l], \dots, \tilde{s}_n[l], \dots, \tilde{s}_N[l]]^T \in \mathbb{R}^N, \quad (1)$$

where $l = 1, 2, \dots, L$ and $n = 1, 2, \dots, N$ are the indices of the discrete-time samples and the source signals, respectively, and \cdot^T denotes the transpose. These source signals are mixed and observed as an M channel signal $\tilde{\mathbf{x}}$:

$$\tilde{\mathbf{x}}[l] = [\tilde{x}_1[l], \tilde{x}_2[l], \dots, \tilde{x}_m[l], \dots, \tilde{x}_M[l]]^T \in \mathbb{R}^M, \quad (2)$$

where $m = 1, 2, \dots, M$ is the index of the channels. Estimated (separated) signals are also defined as

$$\tilde{\mathbf{y}}[l] = [\tilde{y}_1[l], \tilde{y}_2[l], \dots, \tilde{y}_n[l], \dots, \tilde{y}_N[l]]^T \in \mathbb{R}^N. \quad (3)$$

In this paper, we assume $N = M = 4$. The assumption of $N = M$ is often called the determined condition [3], which is necessary for the BSS techniques used in this paper.

To model the convolutive mixture in the time domain, we transform the signals (1)–(3) into the time-frequency (spectrogram) domain via the short-time Fourier transform (STFT).

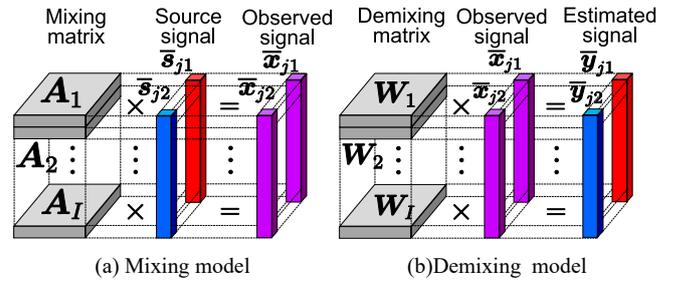


Fig. 3. Mixing and demixing models ($M = N = 2$)

The source, observed, and estimated signals in the time-frequency domain are defined as

$$\mathbf{s}_{ij} = [s_{ij1}, s_{ij2}, \dots, s_{ijn}, \dots, s_{ijN}]^T \in \mathbb{C}^N, \quad (4)$$

$$\mathbf{x}_{ij} = [x_{ij1}, x_{ij2}, \dots, x_{ijm}, \dots, x_{ijM}]^T \in \mathbb{C}^M, \quad (5)$$

$$\mathbf{y}_{ij} = [y_{ij1}, y_{ij2}, \dots, y_{ijn}, \dots, y_{ijN}]^T \in \mathbb{C}^N, \quad (6)$$

respectively, where $i = 1, 2, \dots, I$ and $j = 1, 2, \dots, J$ are the indices of frequency bins and time frames. Note that \mathbf{s}_{ij} and \mathbf{y}_{ij} are “multisource” signals, and \mathbf{x}_{ij} is a “multichannel” signal in each time-frequency slot. We also define the time-frequency matrices (spectrograms) of (4)–(6) as $\mathbf{S}_n \in \mathbb{C}^{I \times J}$, $\mathbf{X}_m \in \mathbb{C}^{I \times J}$, and $\mathbf{Y}_n \in \mathbb{C}^{I \times J}$, respectively.

In time-frequency domain BSS, the mixing system of the sources is modeled by a time-invariant instantaneous mixture in each frequency bin. This model can be represented as

$$\mathbf{x}_{ij} = \mathbf{A}_i \mathbf{s}_{ij}, \quad (7)$$

where $\mathbf{A}_i \in \mathbb{C}^{M \times N}$ is a frequency-wise mixing matrix. Then, BSS estimates the inverse system of (7):

$$\mathbf{y}_{ij} = \mathbf{W}_i \mathbf{x}_{ij}, \quad (8)$$

where $\mathbf{W}_i = [\mathbf{w}_{i1} \ \mathbf{w}_{i2} \ \dots \ \mathbf{w}_{iN}]^H \in \mathbb{C}^{N \times M}$ is a frequency-wise demixing matrix, which ideally coincides with $\mathbf{W}_i = \mathbf{A}_i^{-1}$ under the determined condition, and \cdot^H denotes the Hermitian transpose.

The mixing and demixing systems (7) and (8) are depicted in Fig. 3, where $\bar{\mathbf{s}}_{jn}$, $\bar{\mathbf{x}}_{jn}$, and $\bar{\mathbf{y}}_{jn}$ in this figure are the vectors that consist of all the frequency components as

$$\bar{\mathbf{s}}_{jn} = [s_{1jn}, s_{2jn}, \dots, s_{ijn}, \dots, s_{Ijn}]^T \in \mathbb{C}^I, \quad (9)$$

$$\bar{\mathbf{x}}_{jm} = [x_{1jm}, x_{2jm}, \dots, x_{ijm}, \dots, x_{Ijm}]^T \in \mathbb{C}^I, \quad (10)$$

$$\bar{\mathbf{y}}_{jn} = [y_{1jn}, y_{2jn}, \dots, y_{ijn}, \dots, y_{Ijn}]^T \in \mathbb{C}^I. \quad (11)$$

The algorithms of time-frequency domain BSS aim to estimate the frequency-wise demixing matrix \mathbf{W}_i from the multichannel observed signal $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_M)$ without any assumptions about the mixing system \mathbf{A}_i , e.g., positions of the sensors or sources.

Since the radar sensor has four-channel directivities ($M = 4$), the BSS algorithm outputs four estimated signals $(\mathbf{Y}_1, \dots, \mathbf{Y}_4)$ that ideally correspond to one heartbeat signal and three other noise sources. However, the order of the estimated sources depends on the initial values of BSS parameters

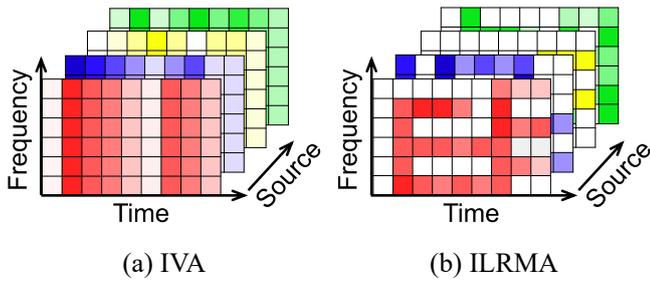


Fig. 4. Source models assumed in IVA and ILRMA.

and an optimization algorithm. This paper treats the method for selecting the estimated heartbeat signal from the four estimated signals as future work. In all the experiments, we manually select the estimated heartbeat signal by comparing BSS outputs with the reference heartbeat signal obtained by the ECG sensor.

B. IVA

The first BSS algorithm applied in this paper is called IVA [4], [5]. To estimate the demixing matrix \mathbf{W}_i , IVA assumes both “statistical independence between sources” and “co-occurrence of all the frequency components in each source” [4]. The co-occurrence assumption means that all the components in $\bar{\mathbf{s}}_{jn}$ or $\bar{\mathbf{y}}_{jn}$ are assumed to have large powers at the same time j , as depicted in Fig. 4 (a). This sourcewise time-frequency model is valid in many sources to some extent. For example, clear harmonic structures can be confirmed in the heartbeat and vibration plate components in Fig. 2, resulting in the co-occurrence of fundamental and harmonic frequencies.

In IVA, the estimation of \mathbf{W}_i is formulated as the following minimization problem [4]:

$$\underset{\mathbf{W}}{\text{Minimize}} \quad -2J \sum_i \log |\det \mathbf{W}_i| + \sum_{j,n} \|\bar{\mathbf{y}}_{jn}\|_2, \quad (12)$$

where $\mathbf{W} = \{\mathbf{W}_i\}_{i=1}^I$ is the set of optimization parameters and $\|\cdot\|_2$ is the L_2 norm, namely, $\|\bar{\mathbf{y}}_{jn}\|_2 = \sqrt{\sum_i |y_{ijn}|^2}$. Although the closed-form solution of (12) has not been found, an efficient iterative optimization algorithm was proposed, which is called auxiliary-function-based IVA (AuxIVA) [5]. In AuxIVA, the following update rule for \mathbf{W} is iteratively calculated:

$$\mathbf{G}_{in} = \frac{1}{J} \sum_j \frac{1}{\sqrt{\sum_i |\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2}} \mathbf{x}_{ij} \mathbf{x}_{ij}^H, \quad (13)$$

$$\mathbf{w}_{in} \leftarrow (\mathbf{W}_i \mathbf{G}_{in})^{-1} \mathbf{e}_n, \quad (14)$$

$$\mathbf{w}_{in} \leftarrow \mathbf{w}_{in} (\mathbf{w}_{in}^H \mathbf{G}_{in} \mathbf{w}_{in})^{-\frac{1}{2}}, \quad (15)$$

where $\mathbf{e}_n \in \{0, 1\}^N$ denotes the unit vector with the n th element equal to unity. The update rule (13)–(15) is called iterative projection (IP) and guarantees a monotonic decrease (or non-increasing) of the cost function (12). Thus, we can obtain the estimated demixing matrix \mathbf{W}_i by iterating (13)–(15) until the value of (12) is converged.

C. ILRMA

The second BSS algorithm, ILRMA [6], [7], assumes the independence and “low-rank time-frequency structure of each source” instead of the co-occurrence assumption in IVA. This assumption is depicted in Fig. 4 (b). The power spectrogram of source signals, $|\mathbf{S}_n|^2$ or $|\mathbf{Y}_n|^2$, is assumed to be well approximated by a low-rank matrix, where the absolute and dotted exponent for matrices represent entry-wise absolute and exponent, respectively. Namely, the power spectrogram of the same source includes many repetitions of similar spectral patterns. For this low-rank modeling, ILRMA uses nonnegative matrix factorization (NMF) [10].

Fig. 5 illustrates the BSS principle of ILRMA. The demixing matrix \mathbf{W}_i is updated so that the estimated power spectrograms $|\mathbf{Y}_n|^2$ have a low-rank time-frequency structure and are modeled by NMF. During the low-rank modeling, NMF enhances the components of one source and reduces the other source components, which facilitates the separation in the optimization of \mathbf{W}_i .

As shown in Fig. 2, the components of both the heartbeat and the vibration plate have repetitions of harmonic spectral patterns along with the time frames. Thus, the low-rank source model in ILRMA (Fig. 4 (b)) has the potential to precisely represent the time-frequency structures of each source, and more accurate BSS can be achieved compared with IVA in many cases.

In ILRMA, the optimization of parameters is formulated as

$$\underset{\mathbf{W}, \mathbf{T}, \mathbf{V}}{\text{Minimize}} \quad -2J \sum_i \log |\det \mathbf{W}_i| + \sum_n \mathcal{D}(|\mathbf{Y}_n|^2 | \mathbf{T}_n \mathbf{V}_n), \quad (16)$$

where $\mathbf{T} = \{\mathbf{T}_n\}_{n=1}^N$ and $\mathbf{V} = \{\mathbf{V}_n\}_{n=1}^N$ are the sets of NMF variables $\mathbf{T}_n \in \mathbb{R}_{\geq 0}^{I \times K}$ and $\mathbf{V}_n \in \mathbb{R}_{\geq 0}^{K \times J}$, respectively, which are often called basis and activation matrices in the context of NMF [6], [7]. They model the power spectrogram of the estimated signals as $|\mathbf{Y}_n|^2 \approx \mathbf{T}_n \mathbf{V}_n$ (see Fig. 5). Furthermore, $\mathcal{D}(\cdot)$ is the sum of the entrywise Itakura–Saito divergence [10] between the input matrices, and $K \in \mathbb{N}$ is the number of NMF basis vectors. The parameters in ILRMA, \mathbf{W} , \mathbf{T} , and \mathbf{V} , can be optimized by iterating the following update rules:

$$t_{ikn} \leftarrow t_{ikn} \sqrt{\frac{\sum_j |\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2 v_{kjn} (\sum_{k'} t_{ik'n} v_{k'jn})^{-2}}{\sum_j v_{kjn} (\sum_{k'} t_{ik'n} v_{k'jn})^{-1}}}, \quad (17)$$

$$v_{kjn} \leftarrow v_{kjn} \sqrt{\frac{\sum_i |\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2 t_{ikn} (\sum_{k'} t_{ik'n} v_{k'jn})^{-2}}{\sum_i t_{ikn} (\sum_{k'} t_{ik'n} v_{k'jn})^{-1}}}, \quad (18)$$

$$\mathbf{U}_{in} = \frac{1}{J} \sum_j \frac{1}{\sum_k t_{ikn} v_{kjn}} \mathbf{x}_{ij} \mathbf{x}_{ij}^H, \quad (19)$$

$$\mathbf{w}_{in} \leftarrow (\mathbf{W}_i \mathbf{U}_{in})^{-1} \mathbf{e}_n, \quad (20)$$

$$\mathbf{w}_{in} \leftarrow \mathbf{w}_{in} (\mathbf{w}_{in}^H \mathbf{U}_{in} \mathbf{w}_{in})^{-\frac{1}{2}}, \quad (21)$$

where t_{ikn} and v_{kjn} are the elements of \mathbf{T}_n and \mathbf{V}_n , respectively, and $k = 1, 2, \dots, K$ is the index of the basis vectors.

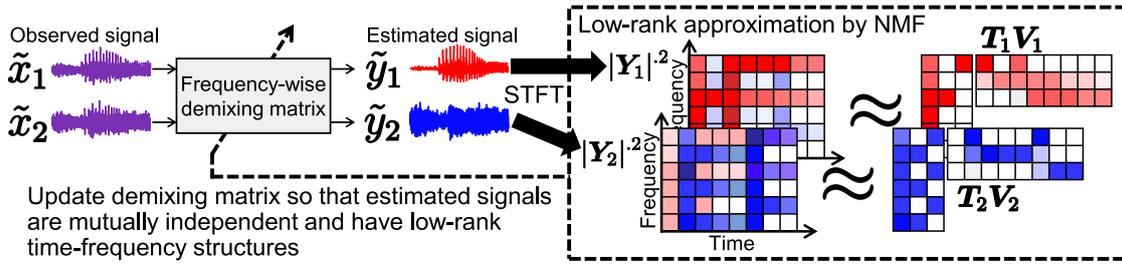


Fig. 5. BSS principle of ILRMA.

Similar to IP in AuxIVA, this iterative optimization algorithm also guarantees a monotonic decrease of the cost function (16).

D. t -ILRMA

The third BSS is an extension of ILRMA, which is called t -ILRMA [8], [9]. ILRMA uses NMF based on the Itakura-Saito divergence (ISNMF). In ISNMF, the time-frequency components of source signals are assumed to follow the zero-mean isotropic complex Gaussian distribution [10], namely,

$$y_{ijn} \sim \mathcal{N}_{\mathbb{C}}(y_{ijn}; 0, \sigma_{ijn}^2), \quad (22)$$

$$\mathcal{N}_{\mathbb{C}}(c; \mu, \sigma^2) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|c - \mu|^2}{\sigma^2}\right), \quad (23)$$

$$\sigma_{ijn}^2 = \sum_k t_{ikn} v_{kjn}, \quad (24)$$

where $\mu \in \mathbb{C}$ and $\sigma^2 > 0$ are the mean and variance of the isotropic complex Gaussian distribution $\mathcal{N}_{\mathbb{C}}(c; \mu, \sigma^2)$ and $c \in \mathbb{C}$ is a random variable. ISNMF in ILRMA decomposes the power spectrogram (time-frequency variance σ_{ijn}^2) based on the generative model (22).

The generative model (22) can be extended by using the isotropic complex Student's t distribution, which is called t -NMF [11], [12], as

$$y_{ijn} \sim \mathcal{T}_{\mathbb{C}}(y_{ijn}; 0, \sigma_{ijn}^2, \nu), \quad (25)$$

$$\mathcal{T}_{\mathbb{C}}(c; \mu, \sigma^2, \nu) = \frac{1}{\pi\sigma^2} \left(1 + \frac{2}{\nu} \frac{|c - \mu|^2}{\sigma^2}\right)^{-\frac{2+\nu}{2}}, \quad (26)$$

$$\sigma_{ijn}^p = \sum_k t_{ikn} v_{kjn}, \quad (27)$$

where $\nu > 0$ is a degree-of-freedom parameter. Note that (24) is also generalized to (27) by introducing the domain parameter $p > 0$, i.e., $|\mathbf{Y}_n|^p \approx \mathbf{T}_n \mathbf{V}_n$. If we set $\nu \rightarrow \infty$ and $p = 2$, (25) converges to (22). Thus, t -NMF is interpreted as a generalization of ISNMF.

It has been reported that t -NMF performs better than ISNMF in specific NMF-based tasks, such as sparse noise reduction [12] and multichannel audio BSS [13]. t -NMF tends to ignore components that do not contribute to low-rank modeling, such as sparse outliers, and focuses on capturing truly low-rank structures in the mixture. This property is enhanced when we set ν to a small value. t -ILRMA utilizes this nature by replacing ISNMF in ILRMA with t -NMF [8], [9]. Therefore, if the sources in Fig. 2 (left) have a truly low-rank structure (but

are contaminated by background noise), t -ILRMA provides better separation performance than standard ILRMA.

The parameter optimization in t -ILRMA is formulated as

$$\begin{aligned} \text{Minimize}_{\mathbf{W}, \mathbf{T}, \mathbf{V}} \quad & -2J \sum_i \log |\det \mathbf{W}_i| \\ & + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \log \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2}{z_{ijn}}\right) \right. \\ & \left. + \frac{2}{p} \log \sum_k t_{ikn} v_{kjn} \right], \end{aligned} \quad (28)$$

Similar to ILRMA, the minimization problem (28) can be solved by iterating the following update rules:

$$t_{ikn} \leftarrow t_{ikn} \left(\frac{\sum_j \frac{|\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2}{b_{ijn} \sum_{k'} t_{ik'n} v_{k'jn}} v_{kjn}}{\sum_j \frac{1}{\sum_{k'} t_{ik'n} v_{k'jn}} v_{kjn}} \right)^{\frac{p}{p+2}}, \quad (29)$$

$$v_{kjn} \leftarrow v_{kjn} \left(\frac{\sum_i \frac{|\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2}{b_{ijn} \sum_{k'} t_{ik'n} v_{k'jn}} t_{ikn}}{\sum_i \frac{1}{\sum_{k'} t_{ik'n} v_{k'jn}} t_{ikn}} \right)^{\frac{p}{p+2}}, \quad (30)$$

$$\mathbf{U}_{in} = \frac{1}{J} \left(1 + \frac{2}{\nu}\right) \sum_j \frac{1}{z_{ijn} (\sum_k t_{ikn} v_{kjn})^{\frac{2}{p}}} \mathbf{x}_{ij} \mathbf{x}_{ij}^H, \quad (31)$$

$$\mathbf{w}_{in} \leftarrow (\mathbf{W}_i \mathbf{U}_{in})^{-1} \mathbf{e}_n, \quad (32)$$

$$\mathbf{w}_{in} \leftarrow \frac{\mathbf{w}_{in}}{\sqrt{\mathbf{w}_{in}^H \mathbf{U}_{in} \mathbf{w}_{in}}}, \quad (33)$$

where

$$b_{ijn} = \frac{\nu}{\nu + 2} \left(\sum_{k'} t_{ik'n} v_{k'jn} \right)^{\frac{2}{p}} + \frac{2}{\nu + 2} |\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2, \quad (34)$$

$$z_{ijn} = 1 + \frac{2}{\nu} \frac{|\mathbf{w}_{in}^H \mathbf{x}_{ij}|^2}{(\sum_k t_{ikn} v_{kjn})^{\frac{2}{p}}}. \quad (35)$$

The monotonic decrease of the cost function (28) is guaranteed.

E. Algorithm of Heart Rate Estimation

After applying BSS to the observed radar sensor signal, we obtain the spectrogram of the estimated heartbeat signal. Let $\tilde{y}_{\text{HB}}[l]$ be the estimated heartbeat signal in the time domain, calculated via an inverse STFT.

A popular method to calculate heart rates from the heartbeat signal is to detect a QRS peak (a spiky waveform that

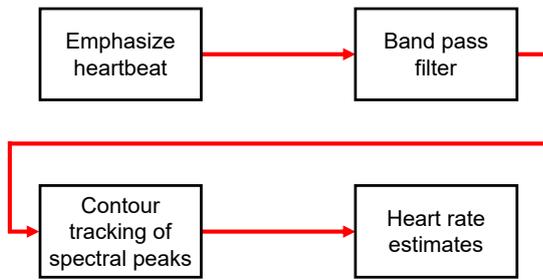


Fig. 6. Process flow of the heart rate estimation algorithm.

appears in the heartbeat signal) using, e.g., the Pan–Tompkins algorithm [14]. Since an interbeat interval corresponds to a fundamental period of the heartbeat, the heart rate can be estimated. However, this approach is designed for ECG sensor signals that clearly include the QRS peak. Since the heartbeat signal estimated by BSS, $\tilde{y}_{\text{HB}}[l]$, does not include clear QRS peaks, in this paper, we apply a heart rate estimation algorithm summarized in Fig. 6. In this algorithm, a band-pass filter enhances the harmonic structure in $\tilde{y}_{\text{HB}}[l]$ while the low-frequency components are removed. The signal is then converted to a spectrogram using STFT, and the contours of the spectral peaks are estimated. Finally, the heart rate for each time frame is obtained from the estimated contours.

IV. EXPERIMENTS

A. Preprocess for Eliminating Breathing Components

As shown in Fig. 2 (left), the components of the breathing motion tend to be strongly measured by the radar sensor. These components often degrade the accuracy of heart rate estimation.

To solve this problem, we reduced the low-frequency energy using a high-pass filter before applying the BSS algorithm. The amplitude and phase responses of the high-pass filter are shown in Fig. 7. The filter was designed as 170th-order finite impulse response, and its cutoff frequency was set to 1.5 Hz. After applying this filter to all channels, $\tilde{x}_1[l], \dots, \tilde{x}_4[l]$, we input them to BSS. Although the fundamental frequency of the heartbeat may also be reduced by the high-pass filter, the algorithm described in Sect. III-E can predict accurate heart rates due to the processes shown in Fig. 6.

B. Experimental Conditions of BSS algorithms

To confirm the best BSS algorithm for radar sensor signals, we applied three BSS algorithms, IVA, ILRMA, and t -ILRMA, to reduce the noise components. The spectrograms of the high-pass-filtered signals were calculated via STFT, where a 1.6-s-long hamming window with 0.1-s-long shifting was used¹.

The number of iterations of each BSS algorithm was set to 100. The initial values of the demixing matrix W_i were set to the identity matrix for all frequencies. Also, the initial values

¹This STFT setting is different from that showing Figs. 2 and 8–11. The spectrograms of Figs. 2 and 8–11 are calculated by STFT with a 25.6-s-long hamming window with 0.8-s-long shifting for the visibility of components in the signal.

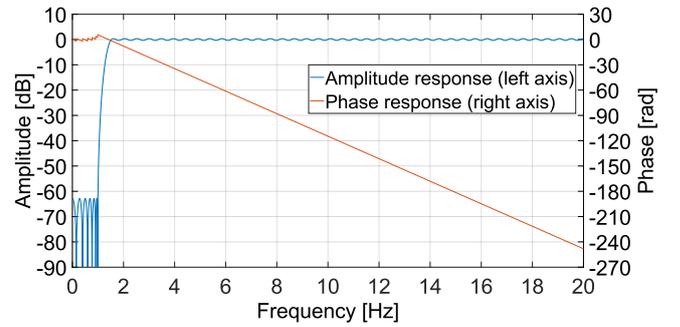


Fig. 7. Amplitude response (left axis) and phase response (right axis) of the high-pass filter used in the preprocessing of observed signals.

of the NMF matrices T_n and V_n in ILRMA and t -ILRMA were set to uniformly random values in the range $(0, 1)$, and the number of NMF bases was set to $K = 3$. For t -ILRMA, we set $p = 1$ (amplitude domain) and $\nu = 5$.

C. Experimental Results

Fig. 8 shows the spectrograms of the observed radar sensor signals, where the high-pass filter is applied. As we discussed in Sect. II-B, the heartbeat components are contaminated by noise. The spectrograms of the signals estimated by IVA, ILRMA, and t -ILRMA are shown in Figs. 9, 10, and 11, respectively. In all cases, the third estimated signal corresponds to the heartbeat signal. From these results, we can confirm the following points:

- The vibration plate and background noise components tend to split into several estimated signals (the first, second, and fourth signals in Figs. 9, 10, and 11).
- All BSS algorithms can extract only the heartbeat components from the observed signal, although their accuracy is different.
- In the IVA results (Fig. 9), some residual components of the vibration plate can be confirmed at around 3.6 Hz. This estimation error is mitigated in the results of ILRMA and t -ILRMA (Figs. 10 and 11).
- In the ILRMA results (Fig. 10), some components of the heartbeat at 200–400 s are split into the second signal, which is not confirmed in the result of t -ILRMA (Fig. 11). This difference will affect to the estimation accuracy of heart rates.

From these results, we expect that t -ILRMA will provide the best performance for the heart rate estimation.

Fig. 12 shows the results of the heart rate estimation algorithm, where the estimated heartbeat signals obtained by each of IVA, ILRMA, and t -ILRMA (the third signals in Figs. 9, 10, and 11) were input, respectively. In addition, estimated heart rates using the ECG sensor signal are shown as a reference for estimation. The IVA-based result (Fig. 12 (a)) includes some errors, particularly at 60–160 s. The results of ILRMA and t -ILRMA show a comparable accuracy of estimation, but the result of t -ILRMA is slightly better. These results show the

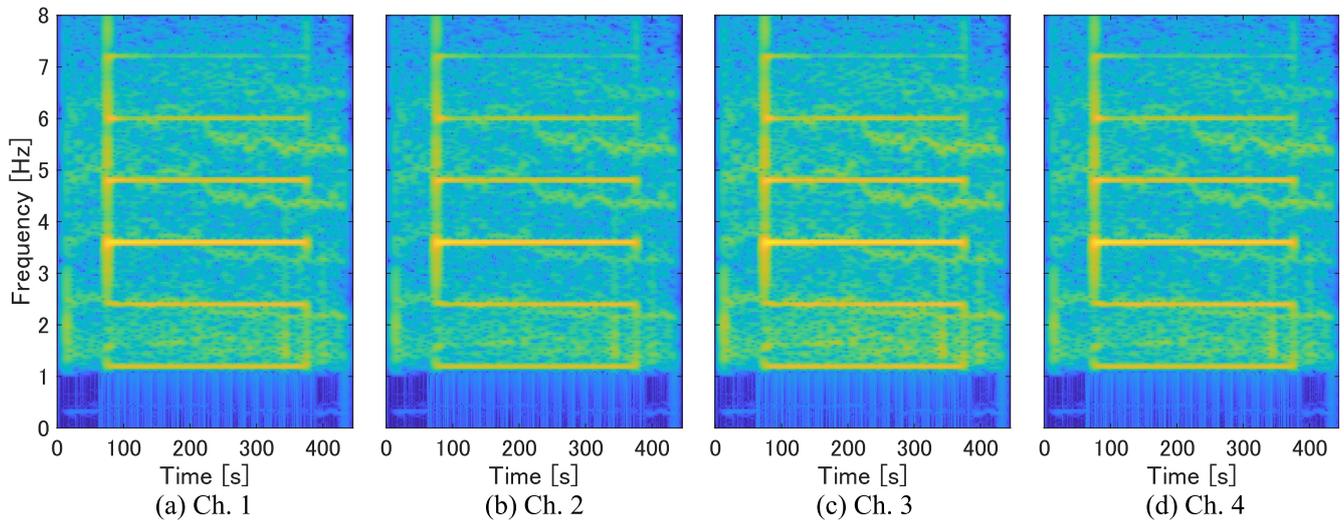


Fig. 8. Four-channel spectrograms of the observed radar sensor signal after applying the high-pass filter.

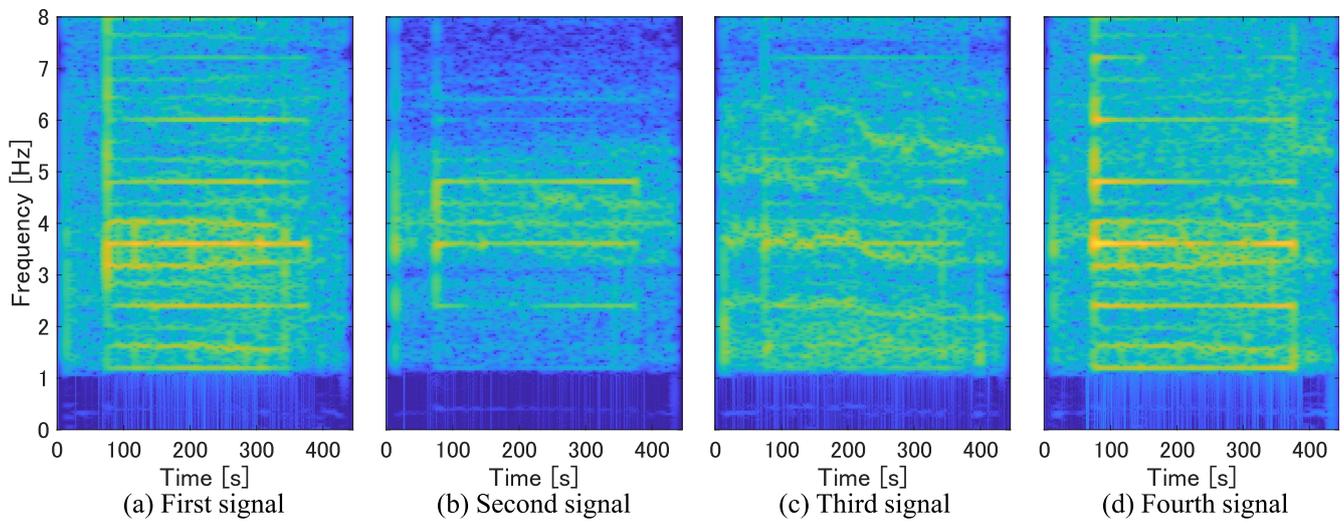


Fig. 9. Spectrograms of the estimated signals obtained by IVA.

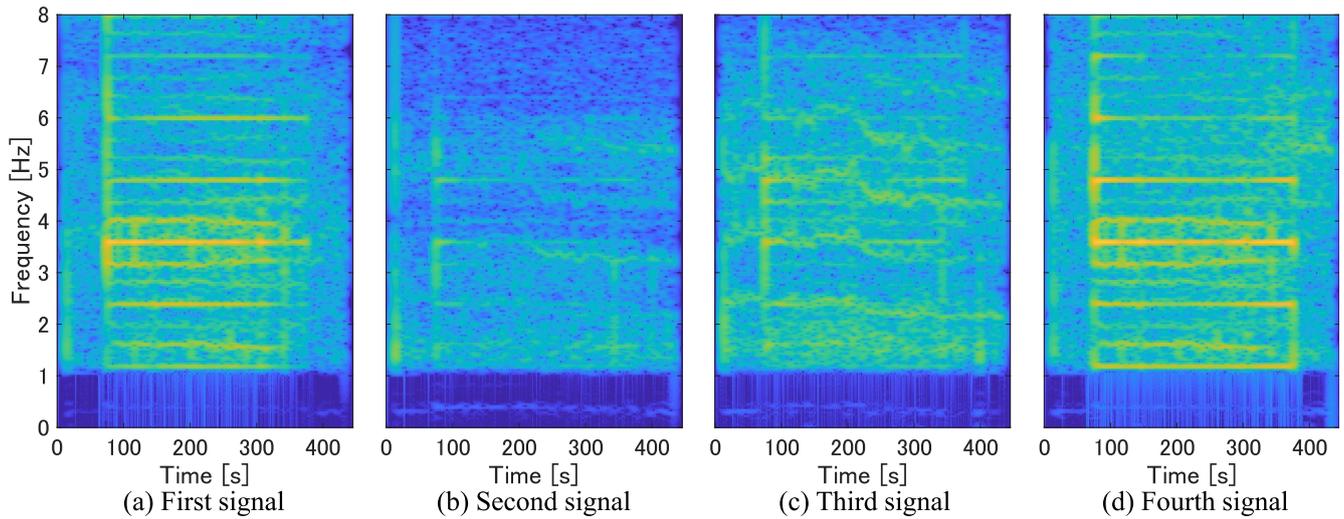


Fig. 10. Spectrograms of the estimated signals obtained by ILRMA.

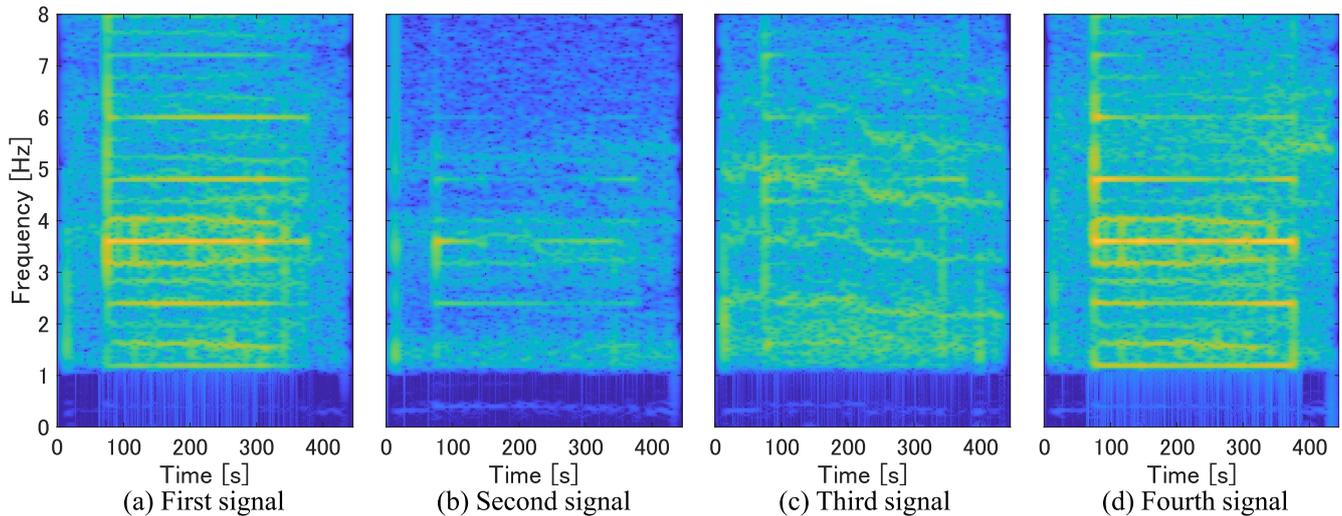


Fig. 11. Spectrograms of the estimated signals obtained by t -ILRMA.

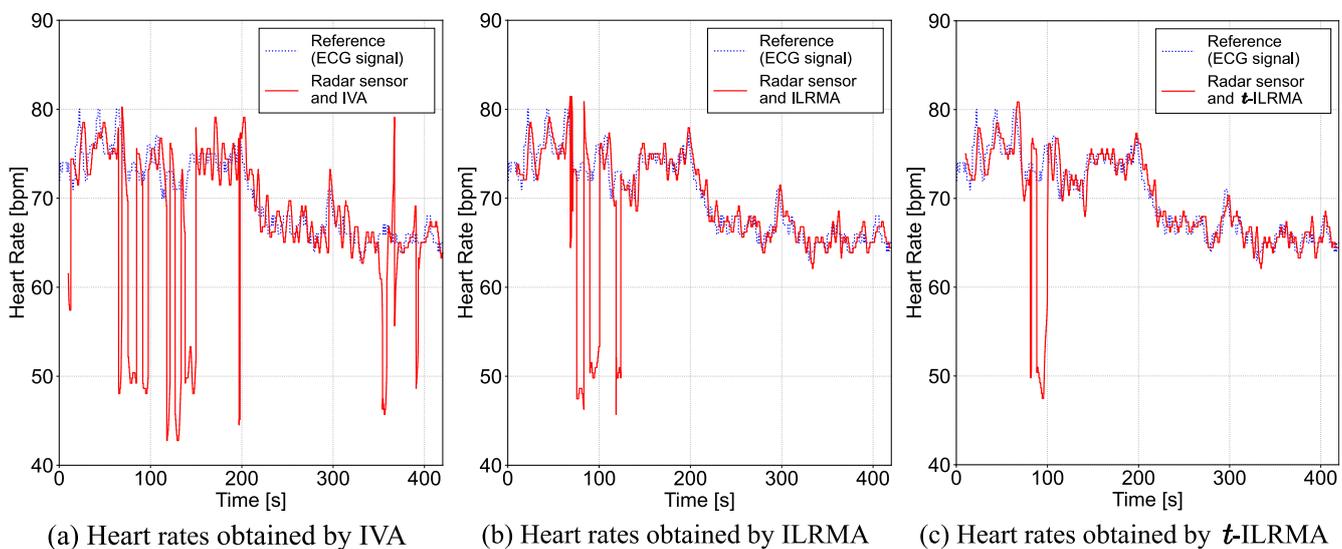


Fig. 12. Estimated (red) and reference (blue) heart rates obtained by BSS algorithms.

validity of the proposed unification of the radar sensor and the BSS algorithms to estimate the driver’s heart rate.

V. CONCLUSION

This paper proposed a new system to monitor the driver’s heart rate. The unification of radar sensors and BSS algorithms was proposed. We conducted a car-simulated experiment using the vibration plate, and the actually observed signals were analyzed using the proposed system. The experimental result confirmed that ILRMA and t -ILRMA could improve the accuracy of estimating the driver’s heart rate by reducing the noise components.

In future work, we will evaluate heart rate estimation accuracy using a quantitative criterion and various conditions of observed signals. In addition, actual car-driving noise is measured and used to simulate more realistic observed signals.

REFERENCES

- [1] M. Alizadeh, G. Shaker, J. C. M. D. Almeida, P. P. Morita, and S. S.-Naeini, “Remote monitoring of human vital signs using mm-wave FMCW radar,” *IEEE Access*, vol. 7, pp. 54958–54968, 2019.
- [2] P. Comon, “Independent component analysis, A new concept?,” *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994.
- [3] H. Sawada, N. Ono, H. Kameoka, D. Kitamura, and H. Saruwatari, “A review of blind source separation methods: Two converging routes to ILRMA originating from ICA and NMF,” *APSIPA Trans. Signal and Information Processing*, vol. 8, no. e12, pp. 1–14, 2019.
- [4] T. Kim, H. T. Attias, S.-Y. Lee, and T.-W. Lee, “Blind source separation exploiting higher-order frequency dependencies,” *IEEE Trans. Audio, Speech, and Language Processing*, vol. 15, no. 1, pp. 70–79, 2007.
- [5] N. Ono, “Stable and fast update rules for independent vector analysis based on auxiliary function technique,” *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 189–192, 2011.
- [6] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, “Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization,” *IEEE/ACM Trans. Audio, Speech, and Language Processing*, vol. 24, no. 9, pp. 1626–1641, 2016.
- [7] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, “Determined blind source separation with independent low-rank matrix

- analysis,” *Audio Source Separation*, S. Makino, Ed., pp. 125–155. Springer, Cham, 2018.
- [8] S. Mogami, D. Kitamura, Y. Mitsui, N. Takamune, H. Saruwatari, and N. Ono, “Independent low-rank matrix analysis based on complex Student’s t -distribution for blind audio source separation,” *Proc. IEEE International Workshop on Machine Learning for Signal Processing*, 2017.
- [9] D. Kitamura, S. Mogami, Y. Mitsui, N. Takamune, H. Saruwatari, N. Ono, Y. Takahashi, and K. Kondo, “Generalized independent low-rank matrix analysis using heavy-tailed distributions for blind source separation,” *EURASIP J. Advances in Signal Processing*, vol. 2018, no. 28, pp. 1–25, 2018.
- [10] C. Févotte, N. Bertin, J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis,” *Neural Computation*, vol. 21, no. 3, pp. 793–830, 2009.
- [11] K. Yoshii, K. Itoyama, and M. Goto, “Student’s t nonnegative matrix factorization and positive semidefinite tensor factorization for single-channel audio source separation,” *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 51–55, 2016.
- [12] D. Kitamura, “Nonnegative matrix factorization based on complex generative model,” *Acoustical Science and Technology*, vol. 40, no. 3, pp. 155–161, 2019.
- [13] K. Kitamura, Y. Bando, K. Itoyama, and K. Yoshii, “Student’s t multichannel nonnegative matrix factorization for blind source separation,” *Proc. IEEE International Workshop on Acoustic Signal Enhancement*, 2016.
- [14] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” *IEEE Trans. Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.