


Article

Orthogonality-Constrained CNMF-Based Noise Reduction with Reduced Degradation of Biological Sound

Naoto Murakami ¹, Shota Nakashima ^{1,*}, Katsuma Fujimoto ¹, Shoya Makihira ¹, Seiji Nishifuji ¹, Keiko Doi ², Xianghong Li ², Tsunahiko Hirano ² and Kazuto Matsunaga ²

- ¹ Division of Electrical and Electronic Engineering, Graduate School of Sciences and Technology for Innovation, Yamaguchi University, 2-16-1, Tokiwadai, Ube 755-8611, Japan; a052vgu@yamaguchi-u.ac.jp (N.M.); i070fe@yamaguchi-u.ac.jp (K.F.); b048vgv@yamaguchi-u.ac.jp (S.M.); nisifuji@yamaguchi-u.ac.jp (S.N.)
- ² Department of Respiratory Medicine and Infectious Disease, Graduate School of Medicine, Yamaguchi University, 1-1-1, Minamikogushi, Ube 755-8505, Japan; decem119@yamaguchi-u.ac.jp (K.D.); lixh@yamaguchi-u.ac.jp (X.L.); tsuna@yamaguchi-u.ac.jp (T.H.); kazmatsu@yamaguchi-u.ac.jp (K.M.)
- * Correspondence: s-naka@yamaguchi-u.ac.jp

Abstract: The number of deaths due to cardiovascular and respiratory diseases is increasing annually. Cardiovascular diseases with high mortality rates, such as strokes, are frequently caused by atrial fibrillation without subjective symptoms. Chronic obstructive pulmonary disease is another condition in which early detection is difficult owing to the slow progression of the disease. Hence, a device that enables the early diagnosis of both diseases is necessary. In our previous study, a sensor for monitoring biological sounds such as vascular and respiratory sounds was developed and a noise reduction method based on semi-supervised convolutive non-negative matrix factorization (SCNMF) was proposed for the noisy environments of users. However, SCNMF attenuated part of the biological sound in addition to the noise. Therefore, this paper proposes a novel noise reduction method that achieves less distortion by imposing orthogonality constraints on the SCNMF. The effectiveness of the proposed method was verified experimentally using the biological sounds of 21 subjects. The experimental results showed an average improvement of 1.4 dB in the signal-to-noise ratio and 2.1 dB in the signal-to-distortion ratio over the conventional method. These results demonstrate the capability of the proposed approach to measure biological sounds even in noisy environments.

Keywords: noise reduction; biological sound; vascular sound; respiratory sound; wearable device; biomedical signal processing; machine learning; non-negative matrix factorization



Citation: Murakami, N.; Nakashima, S.; Fujimoto, K.; Makihira, S.; Nishifuji, S.; Doi, K.; Li, X.; Hirano, T.; Matsunaga, K. Orthogonality-Constrained CNMF-Based Noise Reduction with Reduced Degradation of Biological Sound. *Sensors* **2021**, *21*, 7981. <https://doi.org/10.3390/s21237981>

Academic Editors: Nan Liu, Chengyu Liu and Jessilyn Dunn

Received: 12 November 2021
Accepted: 28 November 2021
Published: 29 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The number of deaths from cardiovascular diseases such as ischemic heart disease, angina pectoris, and arrhythmia is increasing annually worldwide [1]. Atrial fibrillation is a cardiovascular disease that may cause serious cerebrovascular issues. Atrial fibrillation manifests as an abnormal electrical signal in the atria and causes irregular pulsing and inability to pump sufficient blood [2]. In 2010, 33.5 million patients worldwide were estimated to suffer from this disease. In 2017, this number had increased to 37.6 million [3]. The incidence of this disease has been increasing since the 1990s [4] owing to the population aging. The initial symptoms of atrial fibrillation include abnormalities in the conduction system, such as an irregular pulse. The chronicity of these symptoms can lead to severe cerebral infarction [5,6]. The majority of affected individuals are unaware of the disease because they have no subjective symptoms [7]. Similarly, chronic obstructive pulmonary disease (COPD) is a respiratory disease responsible for an increasing number of deaths. COPD is also a slowly progressing disease and its symptoms originate from limitation of the airflow [8]. This disease is the third leading cause of death worldwide and is expected to become the first in 15 years [9,10]. The initial symptoms of COPD include respiratory abnormalities such as shortness of breath [11].

The initial symptoms of these cardiovascular and respiratory diseases are difficult to detect early by the patient because they are subtle and progress slowly. Currently, the diagnosis of COPD and atrial fibrillation is not automated and requires a visit to a clinician. To improve the efficiency of medical treatment for both diseases, it is necessary to perform repeated objective and accurate tests. This problem can be solved by developing a sensor device that can continuously measure cardiovascular and respiratory conditions. The detection of abnormal signals generated from cardiovascular and respiratory organs by the device will enable clinicians to detect and treat diseases at an early stage. In addition, the collection and analysis of measurement data can be applied as a diagnostic aid for physicians or for automated diagnosis. Several biological measurement devices have already been commercialized, such as bedside monitors and electronic stethoscopes used in hospitals [12,13]. The bedside monitor displays multiple vital signs, such as the electrocardiogram, respiratory information, and body temperature of the patient. The electronic stethoscope can record vascular and respiratory sounds even in noisy environments. However, these devices are not always available for measurement and it is difficult for people without medical knowledge to evaluate the measured data. Hence, simultaneous evaluation systems for vascular and respiratory sounds are being researched worldwide. In prior research, we developed a biological sound sensor system that could measure vascular and respiratory sounds simultaneously [14]. This system simultaneously measures these sounds from the trachea and arteries through contact with the human body, and extracts these sounds through biomedical signal processing. This development enabled patients to easily measure their own vascular and respiratory sounds. However, the biological sound sensor was negatively affected by noise from the user's surroundings.

Over the years, researchers have proposed methods to reduce noise in measured sounds. Examples of noise reduction methods include the application of Wiener filters, spectral subtraction, and empirical mode decomposition (EMD) [15–19]. These reduction effects are related to the number of microphones and the computational complexity. From the viewpoint of miniaturization in wearable devices, a noise reduction method that functions with only a single microphone is desirable. A noise reduction method for monaural sources was formulated based on non-negative matrix factorization (NMF). It has been reported that the noise reduction performance of NMF is superior to those of methods using EMD and other methods described above [20]. NMF decomposes a non-negative matrix into two matrices; it has been widely applied in various fields such as image processing, text analysis, and speech processing [21–23]. In our previous study, we proposed a noise reduction method based on semi-supervised convolutive non-negative matrix factorization (SCNMF), which is an extension of NMF [24]. Vascular and respiratory sounds have a time dependence on the frequency; SCNMF was developed for the analysis of these sounds. Nevertheless, this method may misclassify some parts of biological sounds as noise and consequently distort biological sounds. In this paper, we propose a novel NMF extension model, orthogonality-constrained convolutive NMF (OCNMF), which imposes a similarity constraint between noise and biological sounds on the SCNMF. The OCNMF-based noise reduction method can prevent the distortion of biological sounds due to misclassification. The proposed noise reduction can contribute greatly to health care innovations in biomedical signal processing, such as heartbeat analysis and blood pressure prediction [25,26]. The effectiveness of OCNMF was verified experimentally by using biological sounds acquired from 21 subjects.

This paper is organized as follows. Section 2 explains the vascular and respiratory sounds and the structure and principles of biological sound sensors. Section 3 discusses the related research and explains the proposed OCNMF. Section 4 describes the experimental setup and verification process. Finally, Section 5 concludes the paper.

2. Biological Sound Sensor

Biological sounds such as vascular, respiratory, and swallowing sounds are generated in the body through physiological functions. Biological sounds are one of the most impor-

tant sources of information about a person's health status. Hence, physicians perform initial diagnosis through auscultation. We designed a sensor to measure vascular and respiratory sounds for diagnosing cardiac and pulmonary diseases. In the following subsections, we explain vascular and respiratory sounds and the structure and principle of the biological sound sensor used in this study.

2.1. Vascular Sound

Vascular sound is generated by the opening and closing of the heart valves due to beating. This sound is divided into four sounds: S1, S2, S3, and S4. S1 is generated by the closing the mitral and tricuspid valves and S2 is generated by the closing of the aortic and pulmonary valves. S3 and S4 have low-frequency components and are identified by their very low amplitude. Therefore, it is difficult to hear S3 and S4 [27]. Vascular sounds have a main frequency range of 75–200 Hz [28].

2.2. Respiratory Sound

Respiratory sounds are generated by the flow of air through the airway during exhalation and inhalation [29]. This sound has a frequency range of 200–2000 Hz for both exhalation and inhalation. The high-frequency components are attenuated by the propagation medium in the body, such as bone and soft tissue. For example, the respiratory sound generated by the chest has a frequency range of approximately 200–600 Hz [30].

2.3. Biological Sound Sensor

The biological sound sensor shown in Figure 1 is a wearable sensor that can measure vascular and respiratory sounds [14]. The sensor used in this study consists of two parts: a holding unit and a sensor unit. Figure 2 shows the photograph and cross-sectional illustration of the sensor unit. The holding unit is made of an elastoplastic material. The unit is designed to keep the sensor unit in contact with the human body. The sensor unit consists of an electret condenser microphone (ECM), a contact part, and a case. The ECM EM-258 (Primo Co., Ltd., Tokyo, Japan) was designed with its diaphragm exposed to accurately measure biological sounds. Figure 3 shows the ECM before and after exposure. The ECM is housed in a light-curing resin case, and its exposed diaphragm is covered with a polyurethane elastomer with a hardness of 15 HITOHADA gel (Exseal Co., Ltd., Gifu, Japan).

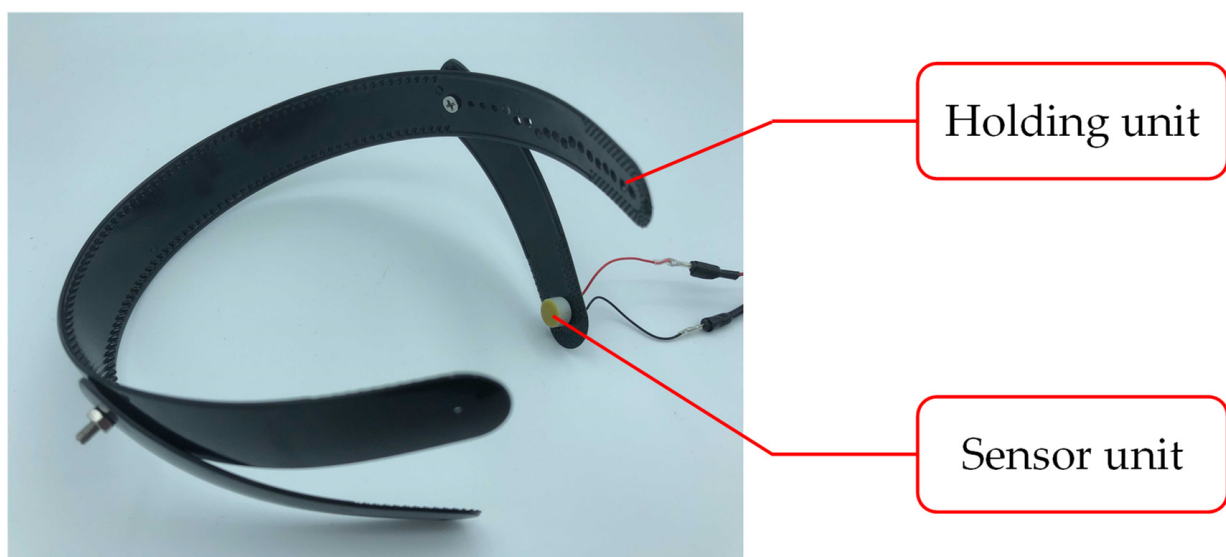


Figure 1. Biological sound sensor.

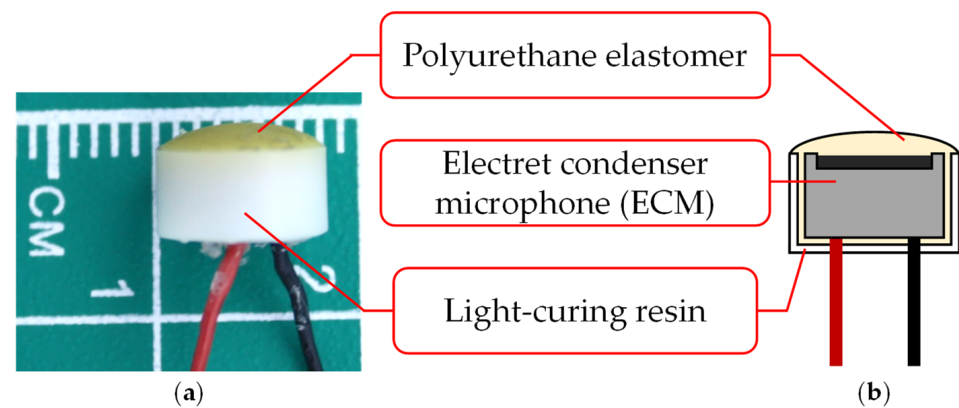


Figure 2. Sensor unit: (a) photograph; (b) cross section.

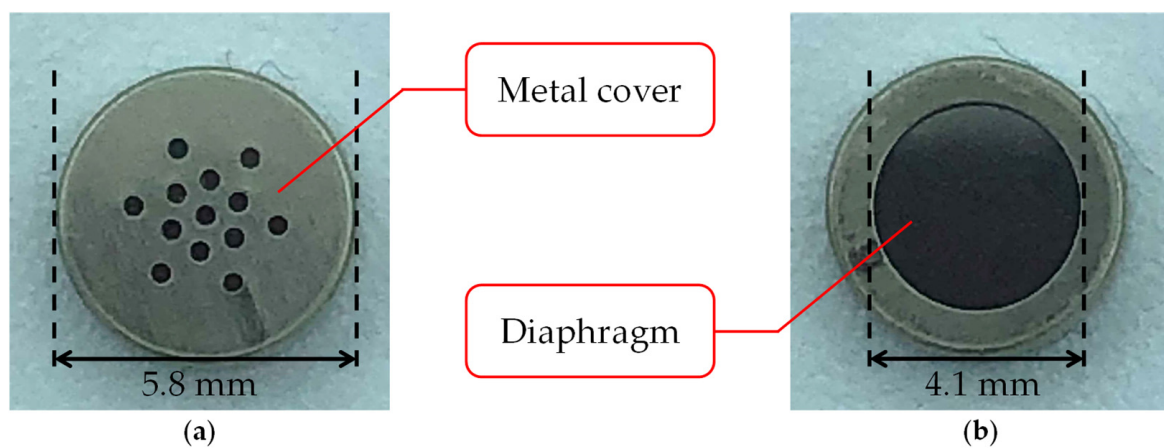


Figure 3. Electret condenser microphone (ECM): (a) before exposure; (b) after exposure.

The ECM is a condenser microphone with an electret. The condenser microphone consists of a capacitor with two electrodes and a diaphragm. The distance between the two electrodes is lengthened or shortened by vibrations of the input sound wave. This mechanism converts the sound vibrations into voltage changes, which enables the condenser microphone to receive sound as an electrical signal. The condenser microphones require a high voltage, known as a phantom source, for charging. In contrast, the ECM can be used at low voltages because it has an electret element that can be charged semi-permanently [31]. The ECM is suitable for wearable devices because of its small size and non-requirement of additional power supply. Signal processing is necessary to compensate for the lower sound quality of the ECM when compared with condenser microphones.

The polyurethane elastomer is used to accurately measure biological sounds with the ECM. The biological sounds are attenuated at the interface between the skin and the sensor because of the natural law that sound waves propagating in a medium with different acoustic properties will be reflected. Their acoustic impedances are matched by passing through a polyurethane elastomer, which has acoustic properties similar to those of the skin.

We decided to measure the biological sound around the mastoid process of the subjects. The mastoid process is a cone-shaped bony prominence located in the posterior inferior portion of the temporal bone [32]. This position enables simultaneous measurement of vascular and respiratory sounds because of its proximity to blood vessels and the trachea. Furthermore, it alleviates the discomfort felt by the user when wearing the sensor. Figure 4 shows a photograph of a subject wearing the biological sound sensor.

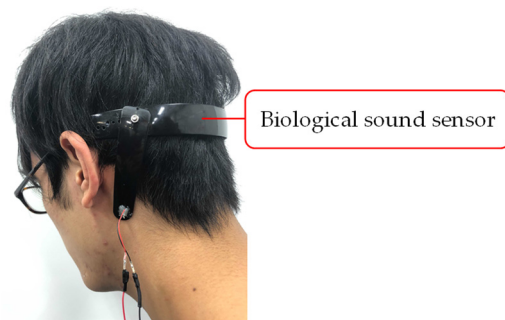


Figure 4. Photograph of a subject wearing the biological sound sensor.

3. Noise Reduction Methods

3.1. Related Research

In the field of audio signal processing, noise reduction methods have been designed with a focus on time and frequency properties. In the following subsections, we introduce non-negative matrix factorization (NMF), a monaural noise reduction method that has been studied extensively in recent years, and its extended models.

3.1.1. NMF and Semi-Supervised NMF

NMF is a method for decomposing non-negative data into additive components [21–23]. Based on previous studies, additivity is assumed valid in an amplitude spectrogram, which is an absolute value of the audio signal processed by short-time Fourier transform (STFT). Given a non-negative matrix $Y \in \mathbb{R}_+^{F \times T}$, NMF approximates it by the product $Y \cong HU$ of a basis matrix $H \in \mathbb{R}_+^{F \times R}$ and an activation matrix $U \in \mathbb{R}_+^{R \times T}$, where the input Y is the amplitude spectrogram of the audio signal, F is the number of frequency bins, T is the number of frames, and R is the number of bases of the NMF. H contains the spectral patterns found in the original spectrogram and U contains the time variation of the amplitude of each spectral pattern. The computational complexity and decomposition accuracy of NMF depend on an objective function that represents the norm between the target matrix and the output.

The objective functions of the NMF have been proposed by many researchers. In this study, we focus on β -divergence, which was devised in previous studies [33,34]. The β -divergence is expressed in Equation (1) using the variables β , x , and y . In Equation (1), $D_\beta(y|x)$ is the pseudodistance of y to x . This is the Itakura–Saito (IS) distance when $\beta = 0$, the Kullback–Leibler (KL) distance when $\beta = 1$, and the Euclidean distance when $\beta = 2$.

$$D_\beta(y|x) = \frac{y^\beta}{\beta(\beta-1)} + \frac{x^\beta}{\beta} - \frac{yx^{\beta-1}}{\beta-1} \quad (1)$$

The objective function of the NMF with β -divergence as the norm is expressed in Equation (2).

$$E_\beta(\theta) = D_\beta(Y|HU), \quad (2)$$

where θ is the parameter to be optimized. NMF estimates H and U such that $E_\beta(\theta)$ is minimized. It is not possible to solve this optimization problem analytically. However, the solution can be obtained indirectly by iteratively minimizing the auxiliary function, which is the upper bound of the objective function [35]. The multiplicative update rules for the objective function with β -divergence are expressed as follows:

$$H \leftarrow H \odot \left(\frac{(HU)^{\beta-2} Y U^T}{(HU)^{\beta-1} U^T} \right)^{\varphi(\beta)}, \quad (3)$$

$$\mathbf{U} \leftarrow \mathbf{U} \odot \left(\frac{\mathbf{H}^T (\mathbf{H}\mathbf{U})^{\beta-2} \mathbf{Y}}{\mathbf{H}^T (\mathbf{H}\mathbf{U})^{\beta-1}} \right)^{\varphi(\beta)}, \quad (4)$$

where the operator \odot represents the adamantine product, and the operator $(\cdot)^T$ represents the transpose operation. $\varphi(\beta)$ is defined using the following equation:

$$\varphi(\beta) = \begin{cases} 1/(2-\beta) & (\beta < 1) \\ 1 & (1 \leq \beta \leq 2) \\ 1/(\beta-1) & (2 < \beta) \end{cases}. \quad (5)$$

It is difficult to interpret the meaning of the spectral patterns represented by each basis because NMF is an unsupervised learning process. In addition, a single basis may contain extensive signal information. Thus, the extraction of the target signal using NMF is very problematic.

Target signal extraction methods have been devised using the spectral patterns of the signal as supervised information [36–38]. Semi-supervised NMF (SNMF) is an extension of NMF to supervised clustering. SNMF consists of two phases: pre-training and noise analysis. In pre-training, the amplitude spectrogram $\mathbf{Y}_{target} \in \mathbb{R}_+^{F \times T}$ of the target signal sound without noise is decomposed to the product of $\mathbf{H} \in \mathbb{R}_+^{F \times R}$ and $\mathbf{Q} \in \mathbb{R}_+^{R \times T}$ through NMF. The basis matrix \mathbf{H} is incorporated into the NMF as supervised information for the target signal. In noise analysis, \mathbf{Q} is not necessary because it represents the activity level of the pre-trained sound. SNMF decomposes the amplitude spectrogram $\mathbf{Y} \in \mathbb{R}_+^{F \times T}$ of the unknown mixed signal, as shown in the following equation:

$$\mathbf{Y} \cong \mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G}, \quad (6)$$

where $\mathbf{F} \in \mathbb{R}_+^{F \times J}$ is the basis matrix that contains information other than the signal to be extracted in the noisy input signal, $\mathbf{G} \in \mathbb{R}_+^{J \times T}$ is the activation matrix corresponding to \mathbf{F} , and J is the number of bases in the basis matrix \mathbf{F} . \mathbf{H} is the basis matrix of the extraction target signal obtained through pre-training. Note that \mathbf{H} is not updated during the decomposition of Equation (6) because \mathbf{H} contains the spectral pattern of the extraction target signal. The extraction target signal can be separated from the other signals by decomposing the input sound into additive components, as described above. The objective function $I_\beta(y|x)$ of SNMF with β -divergence is defined as follows:

$$I_\beta(\theta) = D_\beta(\mathbf{Y} | \mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G}). \quad (7)$$

The multiplicative update rules that minimize the objective function of SNMF using the auxiliary function as well as Equations (3) and (4) can be derived as follows:

$$\mathbf{F} \leftarrow \mathbf{F} \odot \left(\frac{(\mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G})^{\beta-2} \mathbf{Y}\mathbf{G}^T}{(\mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G})^{\beta-1} \mathbf{G}^T} \right)^{\varphi(\beta)}, \quad (8)$$

$$\mathbf{G} \leftarrow \mathbf{G} \odot \left(\frac{\mathbf{F}^T (\mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G})^{\beta-2} \mathbf{Y}}{\mathbf{F}^T (\mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G})^{\beta-1}} \right)^{\varphi(\beta)}, \quad (9)$$

$$\mathbf{U} \leftarrow \mathbf{U} \odot \left(\frac{\mathbf{H}^T (\mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G})^{\beta-2} \mathbf{Y}}{\mathbf{H}^T (\mathbf{H}\mathbf{U} + \mathbf{F}\mathbf{G})^{\beta-1}} \right)^{\varphi(\beta)}. \quad (10)$$

SNMF is utilized as a noise reduction method by reconstructing the sound from the information $\mathbf{H}\mathbf{U}$ of the target signal in the mixed signal obtained in this manner.

SNMF is based on the regular NMF, which assumes that the spectral pattern of the audio signal is time-invariant. This assumption has been reported to deteriorate the

extraction accuracy for signals with time-varying amplitude spectra, such as voices [39]. In the next subsection, we describe the convolutive NMF for time-varying signal analysis.

3.1.2. Convolutive NMF and Semi-Supervised Convolutive NMF

Convolutive NMF (CNMF) is an extended NMF capable of analyzing audio signals with a time-varying amplitude spectrum [40]. Convolution is an operation used to estimate the relationship between the neighboring spectra on the time axis. CNMF decomposes the amplitude spectrogram \mathbf{Y} of the audio signal as follows:

$$\mathbf{Y} \cong \sum_{\tau=0}^{K-1} \mathbf{H}_{\tau} \overset{\rightarrow}{\mathbf{U}}, \quad (11)$$

where $\mathbf{H}_{\tau} \in \mathbb{R}_+^{F \times R}$ represents the basis matrix and K represents the number of time frames to be convolved. Note that the operator $\overset{\rightarrow}{\mathbf{U}}$ shifts the matrix column to the right by τ and sets the elements of the shifted column to zero from outside the matrix. In contrast, the operator $\overset{\leftarrow}{\mathbf{U}}$ shifts the matrix column to the left by τ and sets the elements of the shifted column to zero from outside the matrix. As shown in Equation (11), CNMF decomposes the amplitude spectrogram of the audio signal into a shared activation matrix and a set of time-frame-shifted basis matrices. From this calculation, CNMF can estimate the time variation and relationship between the neighboring spectra.

Semi-supervised CNMF (SCNMF) is an extended model of SNMF that utilizes CNMF for pre-training and noise analysis. Figure 5 shows a conceptual diagram of SNMF and SCNMF. It is noteworthy that the spectral pattern of respiratory sounds changes from the beginning to the end of the respiratory cycle. We have already reported that SCNMF has a high noise reduction effect in respiratory sounds with a time-varying frequency distribution [24].

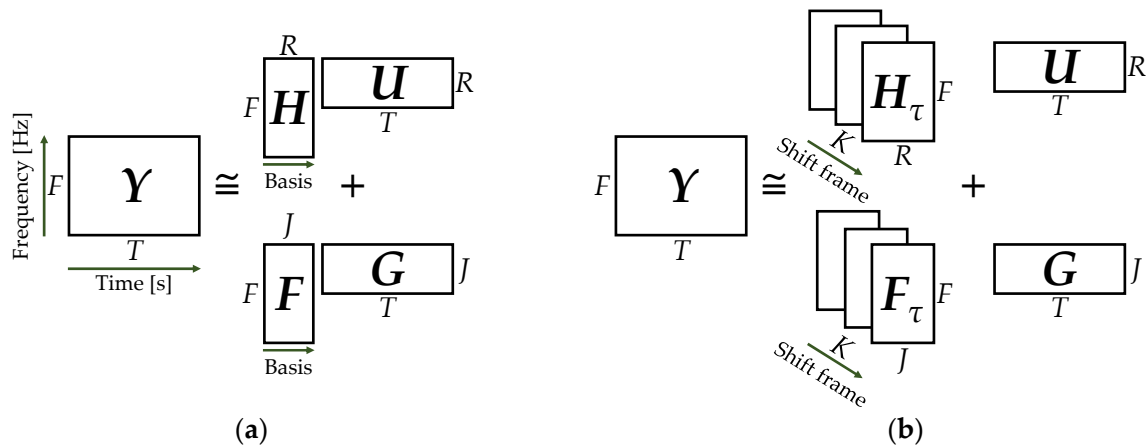


Figure 5. Conceptual diagram of semi-supervised non-negative matrix factorization (SNMF) and semi-supervised convolutive NMF (SCNMF): (a) SNMF; (b) SCNMF.

The objective function $L_{\beta}(\theta)$ of SCNMF and its multiplicative update rules derived as followed in Equations (8)–(10) are given below. Note that the estimated result of the SCNMF, \mathbf{Z} , is defined in Equation (12).

$$\mathbf{Z} = \sum_{\tau=0}^{K-1} \mathbf{H}_{\tau} \overset{\rightarrow}{\mathbf{U}} + \sum_{\tau=0}^{K-1} \mathbf{F}_{\tau} \overset{\rightarrow}{\mathbf{G}} \quad (12)$$

$$L_{\beta}(\theta) = D_{\beta}(\mathbf{Y}|\mathbf{Z}), \quad (13)$$

$$F_{\tau} \leftarrow F_{\tau} \odot \left(\frac{\mathbf{Z}^{\beta-2} \mathbf{Y}^{\tau \rightarrow T} \mathbf{U}}{\mathbf{Z}^{\beta-1} \tau \mathbf{G}} \right)^{\varphi(\beta)}, \quad (14)$$

$$\mathbf{G} \leftarrow \mathbf{G} \odot \left(\frac{\mathbf{F}_{\tau}^T \mathbf{Z}^{\leftarrow \tau} \beta^{-2} \mathbf{Y}}{\mathbf{F}_{\tau}^T \mathbf{Z}^{\leftarrow \tau} \beta^{-1}} \right)^{\varphi(\beta)}, \quad (15)$$

$$\mathbf{U} \leftarrow \mathbf{U} \odot \left(\frac{\mathbf{H}_{\tau}^T \mathbf{Z}^{\leftarrow \tau} \beta^{-2} \mathbf{Y}}{\mathbf{H}_{\tau}^T \mathbf{Z}^{\leftarrow \tau} \beta^{-1}} \right)^{\varphi(\beta)}. \quad (16)$$

Through these multiplicative update rules, information other than the target signal in the input spectrogram is contained in the basis matrix F . However, NMF and SCNMF have the risk of incorrectly storing the target signal in F , because NMF originally has no uniqueness and has high dependency on the initial value in the solution. In the next subsection, we describe the details of an extended model of the SNMF, called the orthogonality-constrained NMF, which solves this problem.

3.1.3. Orthogonality-Constrained NMF

A method has been devised to improve the separation accuracy of the target signal by imposing constraints on the multiplicative update rules of the SNMF. Orthogonality-constrained NMF (ONMF) imposes a constraint that maximizes the cosine distance between the pre-training basis matrix and the other basis matrices [41]. By orthogonalizing the basis vectors of the pre-training basis matrix H with those of the other basis matrices F , they become uncorrelated. ONMF imposes the following constraints on the product of F and H .

$$\min_{\mathbf{F}} (\mathbf{H}\mathbf{F})^2 \quad (17)$$

With the imposition of this constraint, the objective function converges by normalizing each column of the basis matrix. The update rule for each element $F_{f,j}$ of the basis matrix F other than the target signal can be obtained by solving the optimization problem of the auxiliary function, which is the upper bound of the objective function. The update rule is as follows:

$$F_{f,j} \leftarrow F_{f,j} \left(\frac{\sum_t \left(\sum_r H_{f,r} U_{r,t} + \sum_j F_{f,j} G_{j,t} \right)^{\beta-2} Y_{f,t} G_{j,t}}{\sum_t \left(\sum_r H_{f,r} U_{r,t} + \sum_j F_{f,j} G_{j,t} \right)^{\beta-1} G_{j,t} + \mu F_{f,j} \sum_r H_{f,r}^2} \right)^{\varphi(\beta)}, \quad (18)$$

where $H_{f,r}$ denotes each element of the basis matrix H , $U_{r,t}$ belongs to the activation matrix U , $G_{j,t}$ belongs to the activation matrix G , and $Y_{f,t}$ is the input amplitude spectrogram Y . The parameter μ is the weight variable of the orthogonality constraint. When the parameter $\mu = 0$, the ONMF process is the same as that of the regular SNMF. In the next subsection, a novel noise reduction method based on these NMF models is proposed to prevent distortion of the respiratory sound signal information.

3.2. Proposed Method

We propose an orthogonality-constrained convolutive NMF (OCNMF) framework for noise reduction, as shown in Figure 6. First, the biological sound measured by the sensor is preprocessed using a band-pass filter (BPF) and harmonic percussion sound separation (HPSS), as described below, to separate the sound into vascular and respiratory sounds. The purpose of this preprocessing is to independently learn vascular and respiratory sounds, which have different characteristics, as described above. We used BPF and HPSS, which are fast processing methods, because the focus of this study is to improve the efficiency of healthcare. Second, the separated respiratory sounds are converted into an am-

plitude spectrogram by using STFT for OCNMF analysis. OCNMF consists of pre-training and noise analysis phases as well as SNMF. In pre-training, OCNMF learns a dataset, $Y_{train} \in \mathbb{R}_+^{F \times T_{train}}$, which consists only of signal information because it requires no information of unknown noise. In the noise analysis, OCNMF decomposes the preprocessed as well as the pre-training input amplitude spectrogram Y . The mask M shown in Equation (19) is generated from the matrices obtained through OCNMF and post-processed with a Wiener filter multiplied by Y to reconstruct the amplitude spectrogram of the denoising result.

$$M = \frac{\left(\sum_{\tau=0}^{K-1} H_{\tau} \vec{U} \right)^2}{\left(\sum_{\tau=0}^{K-1} H_{\tau} \vec{U} \right)^2 + \left(\sum_{\tau=0}^{K-1} F_{\tau} \vec{G} \right)^2} \quad (19)$$

Finally, the product of the denoised amplitude spectrogram and the phase of the input signal Y_p are processed using inverse STFT (ISTFT) to obtain the denoised signal. In the following subsections, we explain the preprocessing and details of the proposed OCNMF.

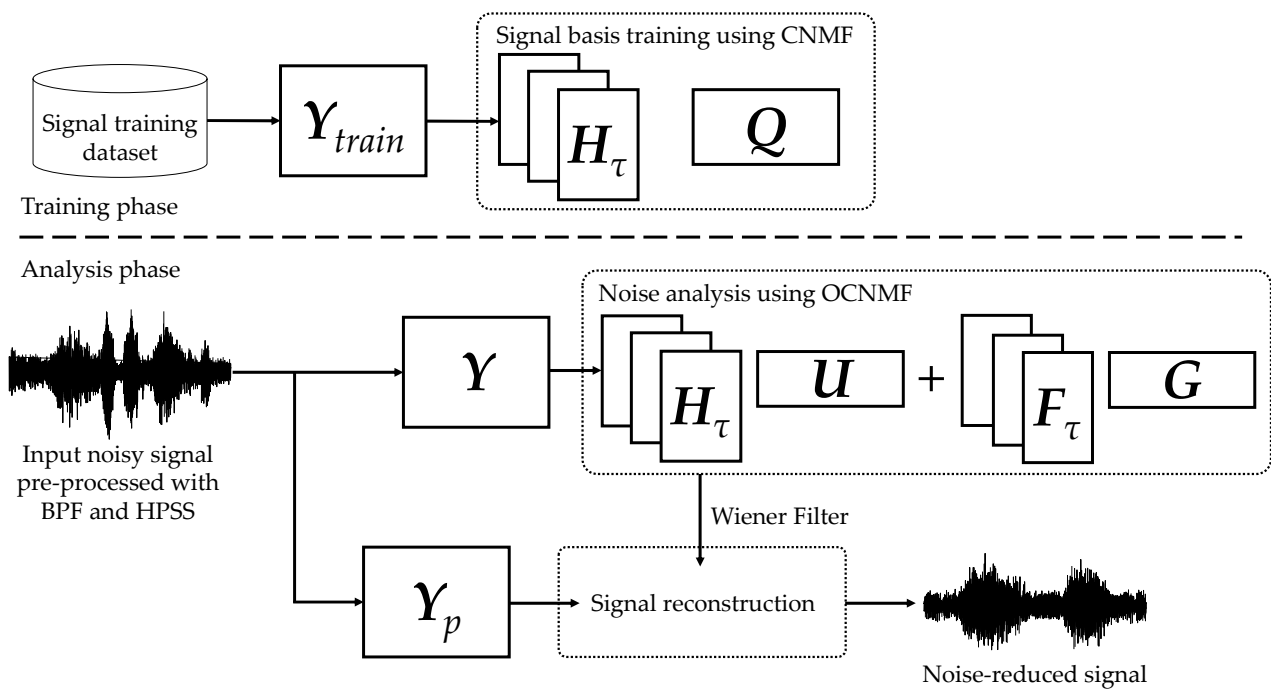


Figure 6. Proposed noise reduction framework based on orthogonality-constrained CNMF.

3.2.1. Preprocessing with BPF and HPSS

To pre-determine the number of bases other than the target signal, the SNMF should narrow down the types of sounds in the preprocessing. Frequency filters designed to match the frequency distribution of vascular and respiratory sounds can extract the respective sounds [14]. The specifications of the BPF designed in this study are listed in Table 1. Our system is intended for practical use in telemedicine and is expected to demonstrate real-time performance. Hence, we designed an infinite impulse response (IIR) filter with low computational complexity. A Butterworth filter was selected because the signal distortion caused by the passband ripple is more detrimental to OSCNMF than the slower out-of-band attenuation. The order was set to 12 to sufficiently attenuate the out-of-band frequency components. Nevertheless, this BPF process is not sufficient to completely extract only the respiratory sounds. The reasons for this are the wide range of secondary frequency components of vascular sounds and the pressure of these sounds, which is approximately 20 dB higher than that of the respiratory sounds.

Table 1. Specification of frequency filters for vascular sound signal (VSS) and respiratory sound signal (RSS).

Target	Frequency [Hz]	Response	Filter	Order
VSS	75–200	Infinite impulse response	Butterworth	12
RSS	200–2000			

In this paragraph, we describe the characteristics of vascular and respiratory sounds and explain the process of removing residual vascular noise from respiratory sounds. As mentioned in the previous paragraph, vascular sounds are short in the time axis and wide in the frequency axis. In contrast, respiratory sounds are longer in the time axis and narrower in the frequency axis. HPSS can be used to separate such sources with temporal and frequency differences [42]. Vascular sounds have percussive characteristics and respiratory sounds have harmonic characteristics. HPSS first applies a one-dimensional median filter with a predefined filter width to each column and row of the amplitude spectrogram of the mixed signal. While setting this filter width in the HPSS, it is necessary to consider the time width of the harmonic and percussive sounds to be extracted. The following conditional equation is imposed on the filter width m_t in the time axis direction for harmonic extraction, which depends on the time width p_t of the percussive sound and time width h_t of the harmonic sound.

$$2p_t < m_t < h_t \quad (20)$$

In this study, the filter width was set to a time frame equivalent to 0.61 s based on the time length of the respiratory and vascular sounds. A mask was calculated from the respiratory sound enhanced by the median filter, as in Equation (19), and multiplied by the original mixture signal to obtain the separation result.

The HPSS processes any signal with a predefined filter width. Hence, it has the risk of causing musical noise by over-reducing when unexpected noise characteristics are mixed in [43]. Our proposed OCNMF is expected to recover from the degradation caused by non-adaptive HPSS because it reduces the noise of the input based on the supervised signal.

3.2.2. Noise Reduction with OCNMF

As mentioned in the previous subsection, CNMF is effective in analyzing signals with time-varying frequency patterns, such as biological sounds. Based on Equation (11), the multiplicative update rules for the basis matrix \mathbf{H}_τ and activation matrix \mathbf{Q} of the target signal information are given as follows:

$$\mathbf{H}_\tau \leftarrow \mathbf{H}_\tau \odot \left(\frac{\mathbf{Z}^{\beta-2} \mathbf{Y}_{train} \overset{\rightarrow T}{\mathbf{Q}}}{\mathbf{Z}^{\beta-1} \overset{\rightarrow T}{\mathbf{Q}}} \right)^{\varphi(\beta)}, \quad (21)$$

$$\mathbf{Q} \leftarrow \mathbf{Q} \odot \left(\frac{\overset{\leftarrow \tau}{\mathbf{H}_\tau^T} \mathbf{Z}^{\beta-2} \overset{\leftarrow \tau}{\mathbf{Y}_{train}}}{\overset{\leftarrow \tau}{\mathbf{H}_\tau^T} \mathbf{Z}^{\beta-1}} \right)^{\varphi(\beta)}, \quad (22)$$

where \mathbf{Z} is the estimated result of CNMF. The procedure for pre-training using CNMF is outlined in Algorithm 1. i_{CNMF} represents the number of update rule iterations required for solution convergence. The number of iterations, i_{CNMF} , depends on the size of the input matrix and the number of shift frames K . A previous study reported that it converged at approximately 100 [39]. The β -divergence should be determined based on the NMF model characteristics and the target sound generation process. NMF with the IS, KL, and Euclidean distances as norms is equivalent to the maximum likelihood estimation assuming exponential, Poisson, and Gaussian distributions for the generation process [44].

Algorithm 1. Signal basis training using CNMF

Input: Spectrogram of signal dataset Y_{train} , priori basis number R , shift length of CNMF K , type of β -divergence β

number of iterations in CNMF i_{CNMF}

Output: The basis matrix of signal H_τ

- 1: Initialize H_τ and Q with random non-negative values
- 2: Normalize columns of H_τ
- 3: **for** $i = 1, \dots, i_{CNMF}$ **do**
- 4: **for** $\tau = 0, \dots, K - 1$ **do**
- 5: Compute Z
- 6: Update H_τ and Q using Equations (21) and (22)
- 7: Normalize columns of H_τ
- 8: **end for**
- 9: **end for**

OCNMF imposes a constraint similar to Equation (17) on the product of the basis matrix H_τ of the target signal information and the basis matrix F_τ of the noise. Based on the constraints, the update rule for each element $F_{\tau,f,j}$ of F_τ in OCNMF is given as follows:

$$F_{\tau,f,j} \leftarrow F_{\tau,f,j} \left(\frac{\sum_t Z_{f,t} \beta^{-2} Y_{f,t} G_{j,t}}{\sum_t Z_{f,t} \beta^{-1} G_{j,t} + \mu F_{\tau,f,j} \sum_r H_{\tau,f,r}^2} \right)^{\varphi(\beta)}, \quad (23)$$

where Z is the estimated result of the OCNMF denoted by Equation (12). Note that the convergence of OCNMF is conditional on the normalization of the basis vectors, as is the case with ONMF. Noise analysis using OCNMF and noise reduction through masking are performed as shown in Algorithm 2. Figure 7 shows the amplitude spectrograms of the four types of respiratory sounds (Figure 7a: noiseless, Figure 7b: noise added, Figure 7c: processed using SCNMF, Figure 7d: processed using the proposed OCNMF method). Based on preliminary experiments investigating the time length of changes in the frequency pattern of the respiratory sound, the number of shift frames K is set to be equivalent to 0.3 s. To scale the two terms in the denominator of the right side of Equation (23), the weight parameter μ of the orthogonality constraint was set to 1.0×10^6 . As shown in Figure 7c,d, our proposed OCNMF has less respiratory sound distortion than the SCNMF.

Algorithm 2. Noise analysis and reduction using OCNMF

Input: Spectrogram of an input signal Y , priori basis number R , an undesired basis number J , type of β -divergence β ,

number of iterations in OCNMF i_{OCNMF} , shift length of CNMF K , signal basis matrix H_τ ,

weight parameter μ , phase matrix of the input signal Y_p

Output: Noise reduced signal Y_{cl}

- 1: Initialize F_τ , U , and G with random non-negative values
- 2: Normalize columns of F_τ
- 3: **for** $i = 1, \dots, i_{OCNMF}$ **do**
- 4: **for** $\tau = 0, \dots, R - 1$ **do**
- 5: Compute Z
- 6: Update F_τ , U , and G using Equations (15), (16) and (23)
- 7: Normalize columns of F_τ
- 8: **end for**
- 9: **end for**
- 10: Compute M using Equation (19) and $Y_{cl} = Y \odot M$
- 11: **ISTFT** ($Y_{cl} \odot Y_p$)

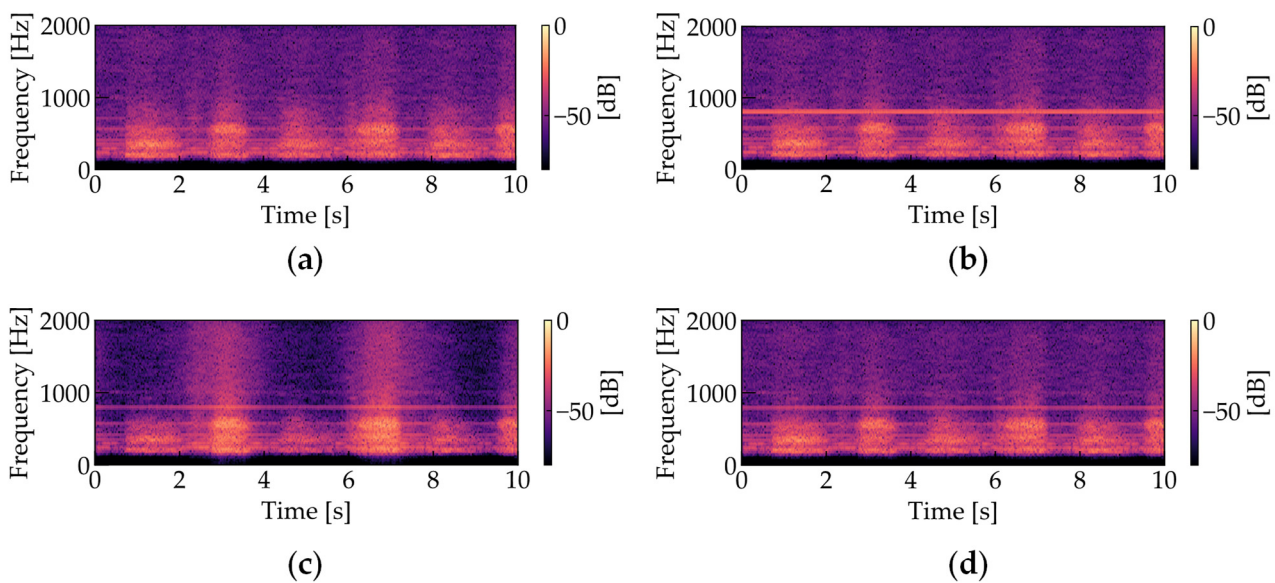


Figure 7. Amplitude spectrograms of respiratory sounds (The parameters were set as follows: sampling frequency = 44.1 kHz, window length in STFT = 1024 points, window length in STFT = 512 points): (a) noiseless; (b) noise-added; (c) processed with SCNMF ($R = 30, J = 15, K = 10, i_{CNMF} = 200, i_{SCNMF} = 200, \beta = 2$); (d) processed with the proposed OCNMF ($R = 30, J = 15, K = 10, i_{CNMF} = 200, i_{OCNMF} = 200, \beta = 2, \mu = 1.0 \times 10^6$).

4. Experimental Verification

4.1. Setup

An experiment was conducted to verify the effectiveness of the noise reduction based on our proposed OCNMF. Sounds that were measured in noiseless environments were used as the pre-training data, and sounds to which noise was added on the computer were used as the test data. The method of separating the pre-training data and test data is explained in the next subsection. As shown in Figure 3, the biological sound sensor was attached to the area around the mastoid process of a subject in a seated resting state. The sensor was connected directly to the PC with a phone connector for power supply and recording control. Biological sounds were measured for 10 s in 21 subjects who consented to participate in this study. The health status of the subjects is shown in Table 2. The experimental design was approved by the Institutional Review Board of Yamaguchi University (Approval number: H2021-031). The respiratory rate was arbitrarily set by the subjects. The sampling frequency and bit depth of the measurement were set to 44.1 kHz and 16 bits, respectively, which assured the same sound quality as that from a CD. This was because the future goal of our research is to realize telemedicine and automated auscultation. Two types of noises were added: a sine wave of 800 Hz and a male voice with frequency range similar to that of the respiratory sound. The sine wave noise simulated stationary noise, while the male voice speaking Japanese sentences simulated non-stationary noise. To quantitatively evaluate the noise reduction results, the noises were adjusted to have the same maximum amplitude as the respiratory sounds. Table 3 lists the parameters of STFT and OCNMF. To reduce respiratory sound distortion, the cost function was set to the Euclidean distance ($\beta = 2$), which evaluates both positive and negative norms equally. The effectiveness of the proposed OCNMF method was evaluated by comparing the signal-to-noise ratio (SNR) and signal-to-distortion ratio (SDR) with the results of the conventional SCNMF method. SNR is commonly used to evaluate the sound quality after noise reduction and SDR is commonly used to evaluate the accuracy of noise separation [45–47]. The details of the evaluation method are described in the following subsection.

Table 2. Characteristics of subjects.

Target	Age [years]	Gender	Disease
A	73	Male	Asthma and chronic obstructive pulmonary disease
B	74	Male	Chronic obstructive pulmonary disease
C	73	Male	Asthma
D	72	Male	Chronic obstructive pulmonary disease
E	72	Female	Asthma
F	74	Male	Asthma
G	77	Male	Chronic obstructive pulmonary disease
H	87	Male	Chronic obstructive pulmonary disease
I	62	Female	Asthma
J	58	Female	Asthma and Chronic obstructive pulmonary disease
K	65	Female	Asthma
L	72	Female	Asthma and chronic bronchitis
M	72	Male	Chronic obstructive pulmonary disease
N	63	Male	Asthma and chronic bronchitis
O	56	Male	Asthma and chronic bronchitis
P	81	Female	Asthma
Q	57	Female	Chronic obstructive pulmonary disease
R	24	Male	No disease
S	24	Male	No disease
T	24	Male	No disease
U	22	Male	No disease

Table 3. Predetermined parameters in short-time Fourier transform (STFT), harmonic percussion sound separation (HPSS), CNMF, SCNMF, and OCNMF.

Sampling frequency	44.1 kHz
Bit depth	16 bits
Window function in STFT	Hann window
Window length in STFT	1024 points
Shift length in STFT	512 points
Input SNR	0 dB
Parameters in HPSS	$m_t = 23$
Parameters in CNMF	$R = 30, K = 10,$ $i_{CNMF} = 200, \beta = 2$
Parameters in SCNMF	$J = 15, K = 10,$ $i_{SCNMF} = 200, \beta = 2$
Parameters in OCNMF	$J = 15, K = 10,$ $i_{OCNMF} = 200, \beta = 2, \mu = 1.0 \times 10^6$

4.2. Evaluation Methodology

The measured biological sound signals were divided into one unit of test data and the remaining as pre-training data based on the leave-one-out method [48]. The leave-one-out method is a type of cross-validation that is used in cases where a large amount of data cannot be prepared. In this experiment, it was difficult to collect a large number of biological sounds because we used our own sensors. Hence, the leave-one-out method was adopted.

To verify the effectiveness of noise reduction using both methods, noise was added to the test data on the computer. The test data were adjusted such that 0 to 5 s consisted of biological sound only, 5 to 10 s consisted of the biological sound and noise, and 10 to 15 s consisted of noise only, for a total of 15 s. The SNR was calculated from the time signals of the first 5 s and the last 5 s, and the SDR was calculated from the time signal of the first

10 s. The SNR was defined using the time signal $s(t)$ of only the biological sound and the time signal $n(t)$ of only noise, as shown in the following equation:

$$\text{SNR} = 10 \log_{10} \frac{\max(|s(t)|)}{\max(|n(t)|)} \quad [\text{dB}], \quad (24)$$

where the operator $|\cdot|$ denotes the absolute value. As shown in Equation (24), the SNR can evaluate sound quality using the amplitude ratio of the signal to noise in mixed sounds. Note that the higher the SNR, the higher the sound quality. The SNR of the test data before noise reduction in this experiment was set to 0 dB and the difference between this value and the SNR after noise reduction quantitatively evaluates the effectiveness of noise reduction. The SDR was calculated from the time signals before and after the noise reduction process; it indicates the level of signal distortion caused by the noise reduction process. The defining equation of SDR based on the time signal before the noise reduction process $b(t)$ and the time signal after the process $a(t)$ is as follows:

$$\text{SDR} = 10 \log_{10} \frac{\sum_t^{S-1} b^2(t)}{\sum_t^{S-1} \{b(t) - \lambda a(t)\}^2} \quad [\text{dB}], \quad (25)$$

where S represents the number of samples for the time signal and $S = 441 \times 10^3$ (equivalent to 10 s). λ is a parameter that adjusts the volume before and after noise reduction and is defined as the following equation:

$$\lambda = \sqrt{\frac{\sum_t^{S-1} b^2(t)}{\sum_t^{S-1} a^2(t)}}. \quad (26)$$

4.3. Results and Discussion

Figure 8 shows the average values of the SNR and SDR for all subjects for the conventional and proposed methods. Table A1 in Appendix A shows the SNR and SDR for each subject in the case of sine wave noise and Table A2 shows the SNR and SDR for each subject in the case of male voice noise. The SNR and SDR of our proposed OCNMF were significantly higher than those of the SCNMF for both noises ($p < 0.001$).

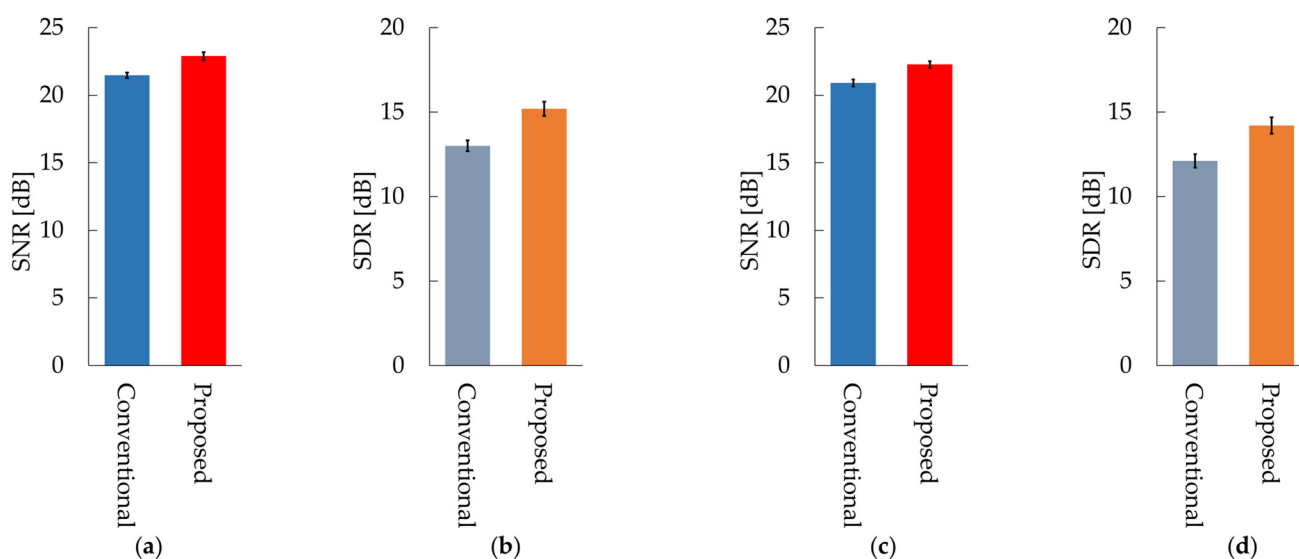


Figure 8. Results of the conventional method SCNMF and proposed method OCNMF: (a) signal-to-noise ratio (SNR) in 800 Hz sine wave noise; (b) signal-to-distortion ratio (SDR) in 800 Hz sine wave noise; (c) SNR in male voice noise; (d) SDR in male voice noise.

The following is a discussion of the experimental results. In the case of sine wave noise, the SNR of the proposed OCNMF method exceeded that of the conventional SCNMF method for almost all subjects. The reason for this result is that the noise basis matrix contains a sine wave with a different frequency pattern from the signal basis vector owing to the orthogonality constraint imposed on the noise basis matrix. This result suggests that OCNMF is effective in reducing stationary noise. The SDR of OCNMF exceeded that of SCNMF for all subjects, with a minimum difference of 0.2 dB and a maximum difference of 5.5 dB. This was because OCNMF was able to solve the problem of SCNMF, wherein the respiratory sound in the noise mixture signal was incorrectly contained in the noise basis matrix. This finding indicates that the orthogonality constraint is useful for reducing respiratory sound distortion. Likewise, the SNR of OCNMF exceeded the SNR of SCNMF for almost all subjects in the case of male voice noise. A previous study reported that an SNR of approximately 20 dB is sufficient for calculating the respiratory rate from respiratory sounds [24]. The SNRs of the experimental results were mostly above 20 dB, suggesting that our proposed method is sufficiently reliable for biological measurements. However, the SNR of subject L was 1.34 dB lower. The reason for this result could be that the frequency pattern of the added noise was similar to that of subject L's respiratory sound. The order of the parameter μ in the signal and noise similarity constraint was set on the basis of the number of elements in the input spectrogram in the experiments. For commercialization as a wearable healthcare device, the parameter μ should be carefully set according to the quality of the pre-training data and applications such as the detection of abnormal respiratory sounds. For example, in the case of pre-training with normal respiratory sounds that are relatively easy to collect and utilize in the abnormal respiratory sound detection system, a very high value of the parameter μ will cause the abnormal respiratory sounds to be removed as noise. The results of SDR using OCNMF in male voice noise exceeded those of the conventional method in almost all subjects. These results support the premise of our proposed method that orthogonality constraints can emphasize respiratory sounds without distortion.

Incidentally, the experimental results should be interpreted with caution, as they may be influenced by the similarity of the subjects' respiratory sounds. Most of the subjects in this experiment were patients with respiratory diseases, such as asthma or COPD. It has been reported that the spectral distribution of respiratory sounds differs for each individual [49,50]; thus, it is necessary to further verify whether the spectral distribution added to the signal basis matrix by pre-training is a generalized respiratory sound. Future studies are required to analyze the effects of individual differences in respiratory sounds by examining a large sample size, including people with other diseases.

5. Conclusions

The aging population has resulted in increasing mortality due to cardiovascular and respiratory diseases. To detect the early symptoms of these diseases, a device that can constantly evaluate biological sounds in a user's daily life is required. A noise reduction method based on SCNMF for biological sound measurement was proposed in our previous study. This method may distort the biological sound signal because the correlation between the basis matrices of the signal and noise was not considered. In this paper, we proposed a novel noise reduction system based on OCNMF, in which each vector of the basis matrix imposed the constraint of maximizing the cosine distance. The effectiveness of the proposed method was verified by experimentally comparing the SNR and SDR with the conventional method. The experimental results indicated that the SNR was significantly improved by 1.4 dB on average over the conventional method and the SDR by 2.1 dB on average. The findings proved that the proposed OCNMF-based system is advantageous for biological sound measurement in noisy environments. The proposed method can be applied to systems that assist medical professionals in diagnosis and automatically evaluate a user's health condition.

To utilize our proposed method for automatic diagnosis, it is necessary to reconstruct the phase of the biological sound signal and improve the speed of the model. In the future, we will extend the model to include phase estimation and attempt to improve its speed based on the sparsity of the matrix. In addition, we will study the optimization of the number of bases to cope with a large number of noise species.

Author Contributions: Conceptualization, N.M. and S.N. (Shota Nakashima); methodology, N.M.; software, N.M. and S.M.; validation, N.M. and S.N. (Shota Nakashima); formal analysis, N.M.; investigation, K.D., X.L., T.H. and K.M.; resources, S.N. (Shota Nakashima), S.M., K.F., K.D., X.L., T.H. and K.M.; data curation, K.D., X.L., T.H. and K.M.; writing—original draft preparation, N.M., S.N. (Shota Nakashima), S.M., S.N. (Seiji Nishifuji) and K.F.; writing—review and editing, K.D., X.L., T.H. and K.M.; visualization, N.M. and K.F.; supervision, S.N. (Shota Nakashima) and S.N. (Seiji Nishifuji); project administration, N.M.; funding acquisition, S.N. (Shota Nakashima). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by JSPS KAKENHI grant number JP21K12797.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Yamaguchi University (ID number: H2021-031).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1 shows the SNR and SDR for each subject in the case of sine wave noise and Table A2 shows the SNR and SDR for each subject in the case of male voice noise.

Table A1. SNR and SDR for each subject in the case of sine wave noise.

Subject	SNR [dB]		SDR [dB]	
	SCNMF	OCNMF	SCNMF	OCNMF
A	20.5	20.5	15.9	18.1
B	23.4	24.3	14.8	15.0
C	22.7	24.2	11.3	16.7
D	21.1	22.8	12.4	13.2
E	22.3	25.7	13.2	18.7
F	21.8	21.6	11.5	11.9
G	20.9	22.3	11.2	12.3
H	21.2	22.4	15.7	17.2
I	21.1	24.1	11.2	14.2
J	22.6	23.2	14.1	16.7
K	22.0	24.9	12.1	14.8
L	21.4	22.3	11.5	13.9
M	23.2	24.1	12.3	12.6
N	20.1	21.2	14.2	17.1
O	20.3	21.5	14.8	15.7
P	20.6	22.4	12.0	13.4
Q	21.2	24.3	11.9	14.8
R	21.4	22.6	13.1	16.2
S	21.2	22.9	14.1	17.1
T	20.7	22.1	12.6	15.3
U	21.3	21.4	13.2	14.2

Table A2. SNR and SDR for each subject in the case of male voice noise.

Subject	SNR [dB]		SDR [dB]	
	SCNMF	OCNMF	SCNMF	OCNMF
A	19.6	20.8	13.8	16.0
B	21.7	22.4	14.1	13.3
C	21.5	23.5	9.0	16.1
D	18.9	23.7	10.2	11.6
E	20.9	22.9	13.9	18.5
F	20.4	20.9	9.1	12.4
G	21.6	22.4	10.2	10.1
H	21.2	23.5	15.1	17.7
I	21.3	22.8	10.6	14.0
J	22.5	22.5	12.7	14.7
K	22.3	23.4	11.2	13.4
L	22.5	24.1	11.7	12.4
M	22.3	21.0	10.4	12.6
N	21.2	20.8	13.1	14.7
O	18.6	20.7	13.5	15.8
P	19.4	22.9	11.0	12.0
Q	20.6	21.5	12.0	13.0
R	20.1	22.0	13.2	16.9
S	19.8	22.3	14.3	17.1
T	22.0	23.3	11.5	12.8
U	20.6	20.4	13.8	13.1

References

- Şahin, B.; İlgün, G. Risk factors of deaths related to cardiovascular diseases in World Health Organization (WHO) member countries. *Health Soc. Care Community* **2020**. [CrossRef] [PubMed]
- Nattel, S. New ideas about atrial fibrillation 50 years on. *Nature* **2002**, *415*, 219–226. [PubMed]
- Dai, H.; Zhang, Q.; Much, A.A.; Maor, E.; Segev, A.; Beinart, R.; Adawi, S.; Lu, Y.; Bragazzi, N.L.; Wu, J. Global, regional, and national prevalence, incidence, mortality, and risk factors for atrial fibrillation, 1990–2017: Results from the Global Burden of Disease Study 2017. *Eur. Heart J. Qual. Care Clin. Outcomes* **2020**, *7*, 574–582. [CrossRef]
- Williams, B.A.; Chamberlain, A.M.; Blankenship, J.C.; Hylek, E.M.; Voyce, S. Trends in atrial fibrillation incidence rates within an integrated health care delivery system, 2006 to 2018. *JAMA Netw. Open* **2020**, *3*, e2014874. [CrossRef] [PubMed]
- Zulkifly, H.; Lip, G.Y.H.; Lane, D.A. Epidemiology of atrial fibrillation. *Int. J. Clin. Pract.* **2018**, *72*, e13070. [CrossRef]
- Wijesurendra, R.S.; Casadei, B. Mechanisms of atrial fibrillation. *Heart* **2019**, *105*, 1860–1867. [CrossRef] [PubMed]
- Harju, J.; Tarniceriu, A.; Parak, J.; Vehkaoja, A.; Yli-Hankala, A.; Korhonen, I. Monitoring of heart rate and inter-beat intervals with wrist plethysmography in patients with atrial fibrillation. *Physiol. Meas.* **2018**, *39*, 065007. [CrossRef]
- MacNee, W. Pathogenesis of chronic obstructive pulmonary disease. *Proc. Am. Thorac. Soc.* **2005**, *2*, 258–266. [CrossRef] [PubMed]
- The Top 10 Causes of Death. Available online: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (accessed on 22 October 2021).
- World Health Organization. *Global Status Report on Noncommunicable Diseases 2010*; World Health Organization: Geneva, Switzerland, 2010. Available online: https://www.who.int/nmh/publications/ncd_report_full_en.pdf (accessed on 22 October 2021).
- Miravittles, M.; Ribera, A. Understanding the impact of symptoms on the burden of COPD. *Respir. Res.* **2017**, *18*, 67. [CrossRef]
- Pelter, M.M.; Mortara, D.; Badilini, F. Computer Assisted Patient Monitoring: Associated Patient, Clinical and ECG Characteristics and Strategy to Minimize False Alarms. *Hearts* **2021**, *2*, 459–471. [CrossRef]
- Asmare, M.H.; Filtjens, B.; Woldehanna, F.; Janssens, L.; Vanrumste, B. Rheumatic Heart Disease Screening Based on Phonocardiogram. *Sensors* **2021**, *21*, 6558. [CrossRef] [PubMed]
- Torii, K.; Nakashima, S.; Nakamura, H.; Ooyagi, T.; Tanaka, K.; Yagami, H. Distinction of heart sound and respiratory sound using body conduction sound sensor based on HPSS. In Proceedings of the 7th ACIS International Conference on Applied Computing & Information Technology, Honolulu, HI, USA, 31 May 2019; pp. 178–183.
- Leng, X.; Chen, J.; Cohen, I.; Benesty, J. Study of widely linear multichannel Wiener filter for binaural noise reduction. In Proceedings of the 2017 25th European Signal Processing, Kos, Greece, 28 August–2 September 2017.
- Yu, L.; Antoni, J.; Wu, H.; Jiang, W. Reconstruction of cyclostationary sound source based on a back-propagating cyclic wiener filter. *J. Sound Vibr.* **2019**, *442*, 787–799. [CrossRef]
- Balaji, V.R.; Maheswaran, S.; Babu, M.R.; Kowsigan, M.; Prabhu, E.; Venkatachalam, K. Combining statistical models using modified spectral subtraction method for embedded system. *Microprocess. Microsyst.* **2020**, *73*, 102957. [CrossRef]

18. Wang, L.; Odani, K.; Kai, A. Dereverberation and denoising based on generalized spectral subtraction by multi-channel LMS algorithm using a small-scale microphone array. *EURASIP J. Adv. Signal Process.* **2012**, *2012*, 1–11. [[CrossRef](#)]
19. Salman, A.H.; Ahmadi, N.; Mengko, R.; Langi, A.Z.R.; Mengko, T.L.R. Empirical Mode Decomposition (EMD) Based Denoising Method for Heart Sound Signal and Its Performance Analysis. *Int. J. Electr. Comput. Eng. Syst.* **2016**, *6*, 1–8.
20. Wang, H.; Wang, M.; Li, J.; Song, L.; Hao, Y. A Novel Signal Separation Method Based on Improved Sparse Non-Negative Matrix Factorization. *Entropy* **2019**, *21*, 445. [[CrossRef](#)]
21. Peng, C.; Kang, Z.; Hu, Y.; Cheng, J.; Cheng, Q. Nonnegative matrix factorization with integrated graph and feature learning. *ACM Trans. Intell. Syst. Technol.* **2017**, *8*, 42. [[CrossRef](#)]
22. Ishizuka, R.; Nishikimi, R.; Nakamura, E.; Yoshii, K. Tatum-level drum transcription based on a convolutional recurrent neural network with language model-based regularized training. In Proceedings of the APSIPA2020, Auckland, New Zealand, 7–10 December 2020; pp. 359–364.
23. Selim, I.M.; Keshk, A.E.; El Shourbugy, B.M. Galaxy image classification using non-negative matrix factorization. *IJCA* **2016**, *137*, 4–8. [[CrossRef](#)]
24. Murakami, N.; Torii, K.; Nakashima, S. Biological sound sensor robust to air conduction noise. In Proceedings of the International Symposium on Computer, Consumer and Control, Online, 13–16 November 2020; pp. 537–540.
25. Shuvo, S.B.; Ali, S.N.; Swapnil, S.I.; Al-Rakhami, M.S.; Gumaei, A. CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings. *IEEE Access* **2021**, *9*, 36955–36967. [[CrossRef](#)]
26. Wu, J.; Wei, W.; Zhang, L.; Wang, J.; Damaševičius, R.; Wang, H.; Wang, G.; Zhang, X.; Yuan, J.; Woźniak, M. Risk Assessment of Hypertension in Steel Workers Based on LVQ and Fisher-SVM Deep Excavation. *IEEE Access* **2019**, *7*, 23109–23119. [[CrossRef](#)]
27. Choi, S.; Jiang, Z. Comparison of envelope extraction algorithms for cardiac sound signal segmentation. *Expert Syst. Appl.* **2008**, *34*, 1056–1069. [[CrossRef](#)]
28. Choi, S.; Jiang, Z. Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique. *Comput. Biol. Med.* **2010**, *40*, 8–20. [[CrossRef](#)] [[PubMed](#)]
29. Reichert, S.; Gass, R.; Brandt, C.; Andrès, E. Analysis of respiratory sounds: State of the art. *Clin. Med. Insights: Circ. Respirat. Pulm. Med.* **2008**, *2*, 45–58. [[CrossRef](#)]
30. Kompis, M.; Pasterkamp, H.; Oh, Y.; Wodicka, G.R. Distribution of inspiratory and expiratory respiratory sound intensity on the surface of the human thorax. In Proceedings of the 19th International Conference—IEEE/EMBS, Chicago, IL, USA, 30 October–2 November 1997; pp. 2047–2050.
31. Hagiwara, K.; Goto, M.; Iguchi, Y.; Yasuno, Y.; Kodama, H.; Kidokoro, K.; Tajima, T. Soft X-ray charging method for a silicon electret condenser microphone. *Appl. Phys. Exp.* **2010**, *3*, 091502. [[CrossRef](#)]
32. Petaros, A.; Sholts, S.B.; Slaus, M.; Bosnar, A.; Wärmländer, S.K. Evaluating sexual dimorphism in the human mastoid process: A viewpoint on the methodology. *Clin. Anat.* **2015**, *28*, 593–601. [[CrossRef](#)]
33. Nakano, M.; Kameoka, H.; Le Roux, J.; Kitano, Y.; Ono, N.; Sagayama, S. Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with β -divergence. In Proceedings of the 2010 IEEE International Workshop on Machine Learning for Signal Processing, Kittila, Finland, 29 August–1 September 2010; pp. 283–288.
34. Févotte, C.; Cemgil, A.T. Nonnegative matrix factorizations as probabilistic inference in composite models. In Proceedings of the 2009 17th European Signal Processing Conference, Glasgow, Scotland, 24–28 August 2009; pp. 1913–1917.
35. Lee, D.D.; Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791. [[CrossRef](#)] [[PubMed](#)]
36. Chan, T.K.; Chin, C.S.; Li, Y. Semi-supervised NMF-CNN for sound event detection. *IEEE Access* **2021**, *9*, 130529–130542. [[CrossRef](#)]
37. Lefevre, A.; Bach, F.; Févotte, C. Semi-supervised NMF with time-frequency annotations for single-channel source separation. In Proceedings of the 13th International Society for Music Information Retrieval Conference, Porto, Portugal, 8–12 October 2012; pp. 115–120.
38. Joder, C.; Weninger, F.; Eyben, F.; Virette, D.; Schuller, B. Real-time speech separation by semi-supervised nonnegative matrix factorization. In Proceedings of the 10th International Conference on Latent Variable Analysis and Signal Separation, Tel Aviv, Israel, 12–15 March 2012; pp. 322–329.
39. Smaragdīs, P. Convolutional speech bases and their application to supervised speech separation. *IEEE Trans. Audio Speech Lang. Process.* **2007**, *15*, 1–12. [[CrossRef](#)]
40. Wang, W.; Cichocki, A.; Chambers, J.A. A multiplicative algorithm for convolutional non-negative matrix factorization based on squared Euclidean distance. *IEEE Trans. Signal Process.* **2009**, *57*, 2858–2864. [[CrossRef](#)]
41. Yagi, K.; Takahashi, Y.; Saruwatari, H.; Shikano, K.; Kondo, K. Music signal separation by orthogonality and maximum-distance constrained nonnegative matrix factorization with target signal information. In Proceedings of the 45th International Conference on Applications of Time-Frequency Processing in Audio, Helsinki, Finland, 1–4 March 2012; pp. 142–147.
42. Fitzgerald, D. Harmonic/percussive separation using median filtering. In Proceedings of the 13th International Conference on Digital Audio Effects, Graz, Austria, 6–10 September 2010; pp. 10–13.
43. Oyabu, S.; Kitamura, D.; Yatabe, K. Linear multichannel blind source separation based on time-frequency mask obtained by harmonic/percussive sound separation. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021.

44. Cavalcanti, Y.C.; Oberlin, T.; Dobigeon, N.; Févotte, C.; Stute, S.; Ribeiro, M.J.; Tauber, C. Factor analysis of dynamic PET images: Beyond Gaussian noise. *IEEE Trans. Med. Imaging* **2019**, *38*, 2231–2241. [[CrossRef](#)]
45. Chuan, T.C.; Darsono, A.M.; Saat, M.S.M.; Isa, A.A.M.; Hashim, N.M.Z. Blind source separation on biomedical field by using nonnegative matrix factorization. *ARPN J. Eng. Appl. Sci.* **2016**, *11*, 8200–8206.
46. Vincent, E.; Gribonval, R.; Févotte, C. Performance measurement in blind audio source separation. *IEEE Trans. Audio Speech Lang. Process.* **2006**, *14*, 1462–1469. [[CrossRef](#)]
47. Jaiswal, A.; Fitzgerald, D.; Barry, D.; Coyle, E.; Rickard, S. Clustering NMF basis functions using shifted NMF for monaural sound source separation. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 245–248.
48. Cawley, G.C. Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. In Proceedings of the 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 16–21 July 2006; pp. 1661–1668.
49. Badjatiya, S.; Mhetre, M.; Kodgule, R.; Abhyankar, H. Breathe sound analysis: A new approach for diagnosis. *IJCT* **2017**, *4*, 1–8.
50. Sanchez, I.; Vizcaya, C. Tracheal and lung sounds repeatability in normal adults. *Respir. Med.* **2003**, *97*, 1257–1260. [[CrossRef](#)]