SPARSE REPRESENTATION-BASED EXTRACTION OF PULMONARY SOUND COMPONENTS FROM LOW-QUALITY AUSCULTATION SIGNALS

Tomoya Sakai[1] , *Haruka Satomoto*[2], *Senya Kiyasu*[1], *Sueharu Miyahara*[2]

[1]Graduate School of Engineering, Nagasaki University, Nagasaki, Japan
[2]Faculty of Engineering, Nagasaki University, Nagasaki, Japan

ABSTRACT

Toward assistance of respiratory system diagnosis, sparse representation of auscultation signals is utilized to extract pulmonary sound components. This signal extraction is a challenging task because the pulmonary sounds such as vesicular sounds and crackles are overlapping each other in the time and frequency domains, and they are so faint that the quality of recorded signals is quite low in many cases. It is experimentally shown that the pulmonary sound components are successfully extracted from low-quality auscultation signals via the sparse representation. This extraction method is confirmed to be highly robust against random noise and digital quantization.

Index Terms— Respiratory system diagnosis, electronic auscultation, source separation, compressed sensing

1. INTRODUCTION

Respiratory auscultation using electronic stethoscopes enables us not only to record, store, and transmit lung sounds, but also to derive medical information by signal processing [1]. We present a substantial application of sparse signal representation to computer-aided respiratory auscultation.

The lung sound components associated with pulmonary disorders, a.k.a rales, are named according to their audio characteristics: wheezes, squawks, crackles, etc. The prior works have also focused on the characteristics of the signal components in the time and frequency domains. There is a body of literature reporting on signal analysis and classification of lung sounds [1, 2, 3, 4, 5, 6]. The classification techniques use frequency spectra, waveforms, and/or wavelet coefficients to describe the features of the lung sounds.

Such signal features, however, are not always discriminative. One would obtain mixture of the features of the signal components from the Fourier or wavelet coefficients, because the lung sound components are overlapping in the time and frequency domains. Separating the signal components is a challenging but essential approach to the improvement of classification performance. It is also noticeable that some types of pulmonary sounds such as vesicular sounds and crackles are so faint that internal noise of the electronic stethoscope is not negligible in many cases.

In this paper, we aim to extract the pulmonary sound components from low-quality auscultation signals. We exploit the fact that the signal components can be efficiently represented by a small number of suitable basis functions. We design the signal extraction method based on the so-called sparse representation, and demonstrate its robustness against noise.

The first author was partially supported by MEXT KAKEN 22700163.

2. SPARSE REPRESENTATION AND EXTRACTION OF LUNG SOUND SIGNALS

2.1. Sparse representation

Assume the following properties of a time signal $y(t)$.

Assumption 1 *$y(t)$ consists of signal components $y_k(t)$ ($k = 1, \ldots, K$) and noise $e(t)$.*

Assumption 2 *$y_k(t)$ can be expanded in a known basis $\mathcal{A}_k$ (card $\mathcal{A}_k = n_k$).*

Assumption 3 *A small number of basis functions in $\mathcal{A}_k$ can synthesize $y_k(t)$, and those in $\mathcal{A}_i$ ($i \neq k$) cannot.*

Let $\mathbf{y}$ be a d-dimensional vector containing d samples of $y(t)$ at $t = t_i \in \mathcal{T}$. Define $\mathbf{y}_k$ and $\mathbf{e}$ for $y_k(t)$ and $e(t)$ in the same manner. Let $\mathbf{A}_k$ be a $d \times n_k$ matrix whose columns are the vectors of d samples of the basis functions in $\mathcal{A}_k$. Then, Assumption 2 states

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{x}_k \tag{1}$$

where $\mathbf{x}_k$ is a n_k-dimensional vector containing the coefficients for $y_k(t)$ expanded in $\mathcal{A}_k$. Under Assumption 1, the discrete signal $\mathbf{y}$ can be represented as a linear combination of the basis functions:

$$\mathbf{y} = \sum_k \mathbf{y}_k + \mathbf{e} = \sum_k \mathbf{A}_k \mathbf{x}_k + \mathbf{e} = \mathbf{A}\mathbf{x} + \mathbf{e} \tag{2}$$

Here, the vector $\mathbf{x}$ and the matrix $\mathbf{A}$ are the concatenations of $\mathbf{x}_k$ and $\mathbf{A}_k$, respectively.

$$\mathbf{x} = [\mathbf{x}_1^\top, \ldots, \mathbf{x}_K^\top]^\top \tag{3}$$

$$\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_K] \tag{4}$$

It follows from Assumption 3 that $\mathbf{x}$ has a small number of nonzero scalar components, say at most m nonzeros out of $n = \sum n_k$ components. We call (2) a sparse representation of $\mathbf{y}$.

2.2. Sparse solution

If a discrete lung sound signal $\mathbf{y}$ is represented as (2) by a sparse vector $\mathbf{x}$ and bases $\mathbf{A}_k$ of normal and adventitious sound signals $\mathbf{y}_k$, any signal component $\mathbf{y}_k$ can be recovered as (1). The recovered signal components are beneficial for the detection and classification of lung sound abnormalities.

Finding the sparse vector $\mathbf{x}$ can be formulated as

$$\min_{\mathbf{x}} ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2 \quad \text{s. t.} \quad ||\mathbf{x}||_0 \leq m \tag{5}$$

Here, l^0 norm $||\mathbf{x}||_0$ denotes the cardinality, or the number of nonzero scalar components of $\mathbf{x}$. This minimization is a combinatorial problem and hard to solve. A common approach to this problem is to relax (5) to a convex minimization problem with l^1 norm [7, 8, 9, 10]. The solution to a convex minimization problem

$$\min_{\mathbf{x}} ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2 \quad \text{s. t.} \quad ||\mathbf{x}||_1 \leq \delta \tag{6}$$

coincides with the sparse solution to (5) with overwhelming probability. An l^1-regularization problem

$$\min_{\mathbf{x}} \left(\frac{1}{2} ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 + \lambda ||\mathbf{x}||_1 \right) \tag{7}$$

also has a sparse solution equivalent to (6) with a suitable value of λ, which controls the sparsity of the solution. One can find efficient algorithms for (7) [11, 12, 13, 14].

2.3. Extraction of breath sounds and crackles

We attempt to extract different types of lung sound signals using the sparse representation. As a practical example, we address the signal extraction of vesicular breath sounds and discontinuous types of pulmonary adventitious sounds from noisy auscultation sounds. While the frequency spectrum of the vesicular sounds is confined under 500Hz, pulmonary adventitious sounds such as the coarse and fine crackles have wide-ranging frequency components because of their pulsating waveforms. Incorporating this prior knowledge, we adopt a sinusoidal basis and a wavelet basis for the vesicular sounds and adventitious sounds, respectively.

Let $\mathbf{A}_C$ and $\mathbf{A}_W$ be a discrete cosine transform matrix and a Daubechies wavelet transform matrix, respectively. The sparse representation of a lung sound signal $\mathbf{y}$ is written as

$$\mathbf{y} = [\mathbf{A}_C \mathbf{A}_W] \begin{bmatrix} \mathbf{x}_C \\ \mathbf{x}_W \end{bmatrix} + \mathbf{e} \tag{8}$$

Using the sparse solution to (6) or (7), we recover the vesicular sound and the crackles as $\mathbf{y}_C = \mathbf{A}_C\mathbf{x}_C$ and $\mathbf{y}_W = \mathbf{A}_W\mathbf{x}_W$, respectively. Note that the matrix multiplications by $\mathbf{A}_C$ and $\mathbf{A}_W$ can be performed via $\mathcal{O}(n \log n)$ and $\mathcal{O}(n)$ operations without storing $n \times n$ matrices.

3. EXPERIMENTS

We demonstrate the sparse representation-based signal extraction of pulmonary sounds. For coarse and fine crackles, we set $\mathbf{A}_W$ to be the Daubechies tap-10 (db10) wavelet basis with three decomposition levels. We have found in preliminary experiments that $\mathbf{x}_W$ tends to be less sparse with a higher decomposition level, and the signal extraction is insensitive to the tap number. We have also confirmed that the relative residual $\varepsilon = ||\mathbf{y} - (\mathbf{y}_C + \mathbf{y}_W)||_2 / ||\mathbf{y}||_2$ is no more than -20dB if one second of clear auscultation sounds recorded at 44.1kHz (d =44,100) is represented as (2) with $m = \mathcal{O}(10^3)$-sparse vector $\mathbf{x}$. In the following experiments we set $m = 1,600$, which amounts to 51.2kbps with single-precision vector $\mathbf{x}$.

For the numerical solution to (7), we employ GPSR [13] with the initial values $\mathbf{x} = \mathbf{0}$ and $\lambda = 10^{-2} ||\mathbf{A}^\top \mathbf{y}||_\infty$. We repeat the minimization with a doubled parameter λ using the sparse solution as a new initial value of $\mathbf{x}$ until $||\mathbf{x}||_0 \leq m$.

3.1. Extraction of crackles

We show in Fig. 1 an example of the extraction of coarse crackles in a case of pneumonia showing soft infiltrative shadows on chest radiography. A clear signal $\mathbf{y}_C$ of vesicular lung sound and $\mathbf{y}_W$ of the crackles were extracted from the original signal $\mathbf{y}$ as in Fig. 1(a). One can observe in Fig. 1 that $\mathbf{y}_C$ and $\mathbf{y}_W$ are overlapping both in time and frequency domains, and these signals cannot be extracted by filtering in the frequency and/or wavelet domains. We remark that the extracted signals $\mathbf{y}_C$ and $\mathbf{y}_W$ have significantly less magnitudes of frequency components above 1kHz, which indicates noise rejection capability of our method. We have experimentally confirmed the similar performance of the extraction on various samples of normal and abnormal lung sounds of different subjects.

3.2. Robustness against noise

We demonstrate and evaluate the robustness of the extraction of lung sound signal components from noisy signals. In practice, internal thermal noise and quantization noise are considerable in pulmonary auscultation by a digital stethoscope. We simulate the noise-contaminated signal by adding white noise to an original signal and digital quantization.

Figure 2 shows an example of the extraction from the degraded version of a real auscultation sound accompanied by fine crackles in a case of interstitial pneumonitis (idiopathic pulmonary fibrosis). We degraded the original signal by adding white noise of 3dB SN ratio and 4-bit digital quantization. We could extract the vesicular sound $\mathbf{y}_C$ and the

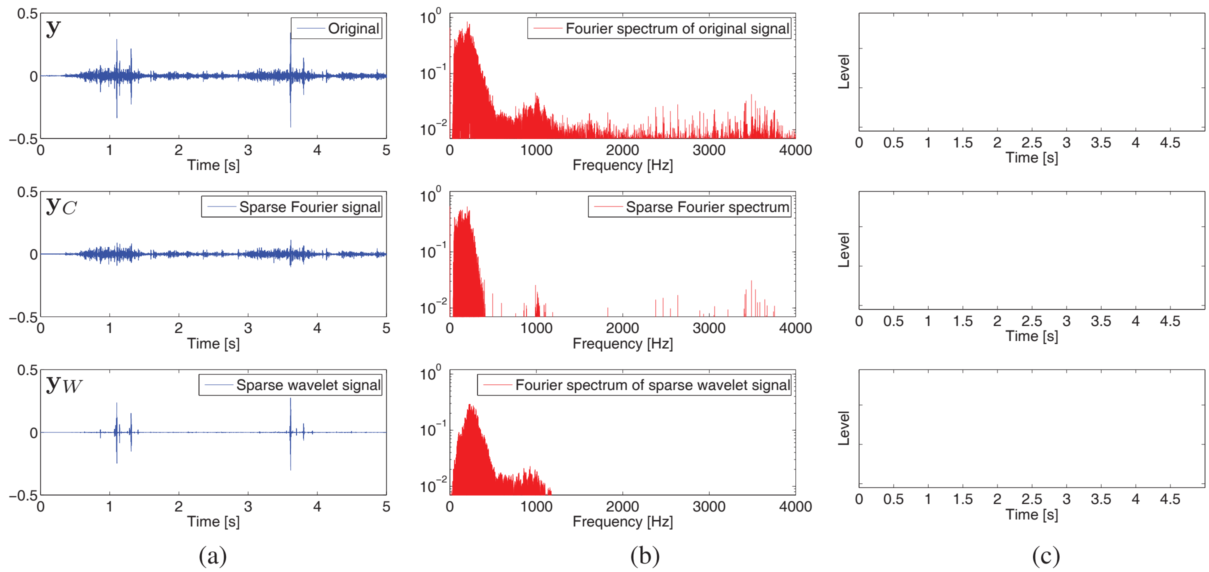

Fig. 1. Extraction of vesicular sound and coarse crackles. First row: original signal $\mathbf{y}$. Second row: extracted signal $\mathbf{y}_C$ with sparse frequency spectrum. Third row: extracted signal $\mathbf{y}_W$ with sparse wavelet coefficients. (a) Waveform in time domain, (b) amplitude spectrum in frequency domain, and (c) wavelet coefficients.

fine crackles $\mathbf{y}_C$ from the degraded signal $\mathbf{y}$ as shown in Fig. 2(a). Since the noise component $\mathbf{e}$ is not sparsely represented both in the frequency and wavelet domains, the lung sound signal components $\mathbf{y}_C$ and $\mathbf{y}_W$ can be recovered from the low-quality signal $\mathbf{y}$.

The reconstructed signal $\hat{\mathbf{y}} = \mathbf{y}_C + \mathbf{y}_W$ is an estimate of the original signal $\mathbf{y}_o$ before the degradation. We evaluate the reconstruction error $\delta = ||\mathbf{y}_o - \hat{\mathbf{y}}||_2/||\mathbf{y}_o||_2$ with respect to the additive noise level. Figure 3 plots the reconstruction error δ as a function of the ratio of the original signal to the additive noise. The stepwise dependence indicates that the quality of the reconstructed signal tends to be sustained despite the noise increase. The reconstruction error abruptly changes when some nonzeros are misplaced in the sparse solution.

4. CONCLUDING REMARKS

Sparse representation has a great potential for lung sound signal processing. Clear pulmonary sounds can separately be recovered from low-quality auscultation signals. The sparse representation of auscultation signals can also play a role of the signal classification because the nonzero coefficients indicate the types of lung sounds. Further research should address this classification capability for various types of lung sounds. Sparse coding techniques would be of great help in learning bases for the signal classification of lung sounds.

5. REFERENCES

[1] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, "Respiratory sounds: advances beyond the stethoscope," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, no. 3, pp. 974–987, 1997.

[2] A. Kandaswamy, C. S. Kumar, Rm. Pl. Ramanathan, S. Jayaraman, and N. Malmurugan, "Neural classification of lung sounds using wavelet coefficients," *Computers in Biology and Medicine*, vol. 34, no. 6, pp. 523–537, 2004.

[3] A. Marshall and B. Boussakta, "Signal analysis of medical acoustic sounds with applications to chest medicine," *Journal of the Franklin Institute*, vol. 344, pp. 230–242, 2007.

[4] M. Yeginer and P. Y. Kahya, "Feature extraction for pulmonary crackle representation via wavelet networks," *Computers in Biology and Medicine*, vol. 39, no. 8, pp. 713–721, 2009.

[5] M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes," *Computers in Biology and Medicine*, vol. 39, no. 9, pp. 824–843, 2009.

[6] T. Masada, S. Kiyasu, and S. Miyahara, "Unmixed spectrum clustering for template composition in lung sound classification," in *Proceedings of PAKDD '08, Lecture Notes in Computer Science*, 2008, vol. 5012, pp. 964–969.

[7] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

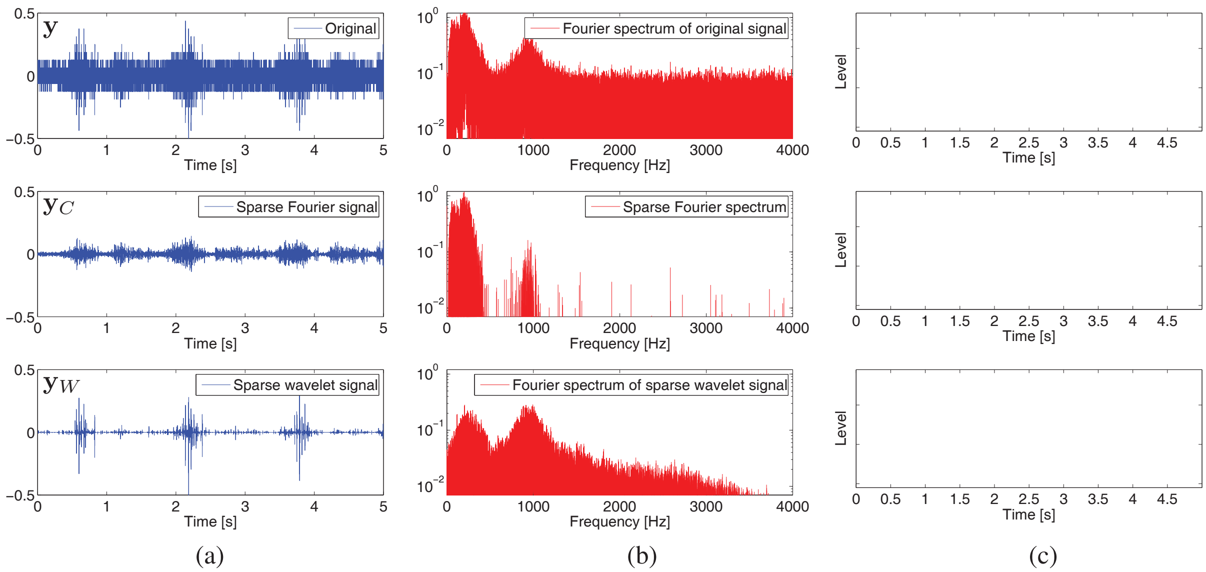

Fig. 2. Recovery of breath sound and fine crackles from degraded signal. First row: degraded version of the original signal by adding white noise of 3dB SN ratio, followed by 4-bit digital quantization. Second row: extracted signal $\mathbf{y}_C$ with sparse frequency spectrum. Third row: extracted signal $\mathbf{y}_W$ with sparse wavelet coefficients. (a) Waveform in time domain, (b) amplitude spectrum in frequency domain, and (c) wavelet coefficients.

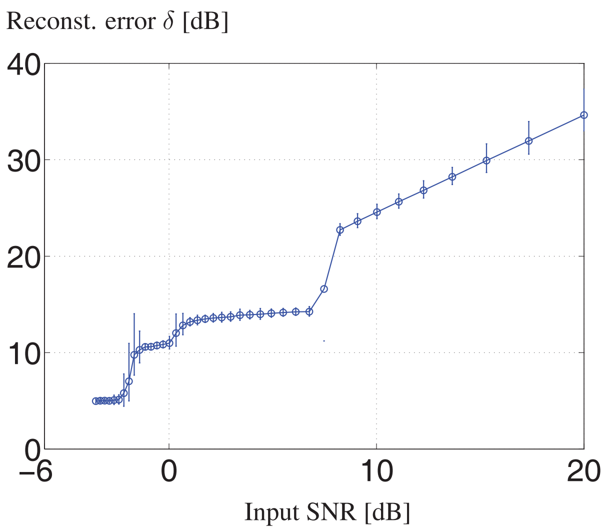

Fig. 3. SN ratios of degraded and reconstructed signals.

[8] D. L. Donoho and M. Elad, "Maximal sparsity representation via l_1 minimization," in *Proceedings of the National Academy of Sciences*, 2003, vol. 100, pp. 2197–2202.

[9] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. on Pure and Applied Math*, vol. 59, no. 8, pp. 1207–1223, 2006.

[10] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, pp. 21–30, March, 2008.

[11] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D Gorinevsky, "An interior-point method for large-scale l_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.

[12] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[13] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.

[14] R. Tomioka and M. Sugiyama, "Dual augmented Lagrangian method for efficient sparse reconstruction," *IEEE Signal Processing Letters*, vol. 16, no. 2, pp. 1067–1070, 2009.