21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

# Robust Detection of Adventitious Lung Sounds in Electronic Auscultation Signals

Tomoya Sakai, Madoka Kato, Sueharu Miyahara, Senya Kiyasu
*Nagasaki University*
*tsakai@cis.nagasaki-u.ac.jp*

## Abstract

*We present a sparse representation-based method for detecting adventitious lung sounds in low-quality auscultation signals. Since the noise cannot be represented sparsely by any bases, we can extract clear breath sounds and adventitious sounds from noisy electronic auscultation signals via the sparse representation. Using these clear sound components, we measure the level of abnormality, and robustly detect adventitious sounds with pulsating waveforms, a.k.a crackles. We have experimentally confirmed that our detection achieves an average precision of about 85 percents regardless of noise level.*

## 1. Introduction

Listening to the patient's body with a stethoscope is a simple but effective method for physical diagnosis. Chest auscultation, in particular, has been a fundamental approach to detect heart and pulmonary disorders. As the electronic stethoscope [1] becomes popular, pattern recognition techniques can contribute to deriving medical information from the electronic auscultation signals.

Pattern recognition of lung sounds is a challenging issue because of a variety of generation mechanisms of the lung sounds. 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 on the analysis of lung sound signals have focused on the characteristics of the signal components in the time and frequency domains [12, 8, 10, 16, 2, 11]. The classification techniques use frequency spectra, waveforms, and/or wavelet coefficients to describe the features of the lung sounds.

Such signal features computed from the lung sound signals, 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. 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 present a sparse representation-based method for the detection of adventitious lung sounds. Lung sound components can be efficiently represented by a small number of suitable basis functions. Exploiting this fact, we have achieved the separate extraction of breath sounds and adventitious sounds from low-quality auscultation signals [13]. Section 2 provides the summary of our extraction method. In section 3, we extend this work to the detection of adventitious sounds. We experimentally show that the detection of the crackles is highly robust against internal noise.

## 2. Sparse representation-based extraction of lung sound components

### 2.1. Sparse representation of lung sounds

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

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

**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 $\mathbf{w}$ for $y_k(t)$ and $w(t)$ 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

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{x}_k \quad (1)$$

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:

$$\mathbf{y} = \sum_k \mathbf{y}_k + \mathbf{w} = \sum_k \mathbf{A}_k \mathbf{x}_k + \mathbf{w} = \mathbf{A}\mathbf{x} + \mathbf{w} \quad (2)$$

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

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

$$\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_K] \quad (4)$$

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. Separate extraction by 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

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

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 [5, 6, 3, 4]. The solution to a convex minimization problem

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

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

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

also has a sparse solution equivalent to (6) with a suitable value of $\lambda$, which controls the sparsity of the solution. One can find efficient algorithms for (7) [9, 7, 15, 14].

## 3. Detection of adventitious components

The sparse representation-based extraction achieves detection and classification of the adventitious sounds, since the nonzero coefficients in $\mathbf{x}_k$ of the sparse solution indicate the existence of the $k$-th sound components. For an adventitious sound component $y_j(t)$ accompanied by respiration, we measure its level of abnormality $a_j(t)$ so that it is invariant to the amplitude scale of the auscultation signal $y(t)$ and the noise $w(t)$. Let $e(t) \geq 0$ be the envelope of the basic normal breath sound component of $y(t)$. We define the abnormality level by simply normalizing $y_j(t)$ by $e(t)$ as

$$a_j(t) = \begin{cases} y_j(t)/e(t) & \text{if } e(t) > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

The adventitious sounds are detected as the time segments with high abnormality levels. We have shown that the sparse representation-based extraction can recover noise-free lung sound components from a noisy auscultation signal [13]. We expect that the evaluation of the abnormality level as (8) using the extracted sound components is highly robust against noise.

## 4. Experimental evaluation

### 4.1. Detection of crackles

As a practical example, we address the detection of coarse and fine crackles from noisy auscultation signals. While normal breath sounds are 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 $\mathbf{A}_1 = \mathbf{A}_C$ and $\mathbf{A}_2 = \mathbf{A}_W$ 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

$$\mathbf{y} = [\mathbf{A}_C \mathbf{A}_W] \begin{bmatrix} \mathbf{x}_C \\ \mathbf{x}_W \end{bmatrix} + \mathbf{w} \quad (9)$$

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.

For coarse and fine crackles, we set $\mathbf{A}_W$ to be the Daubechies tap-10 (db10) wavelet basis. We have found

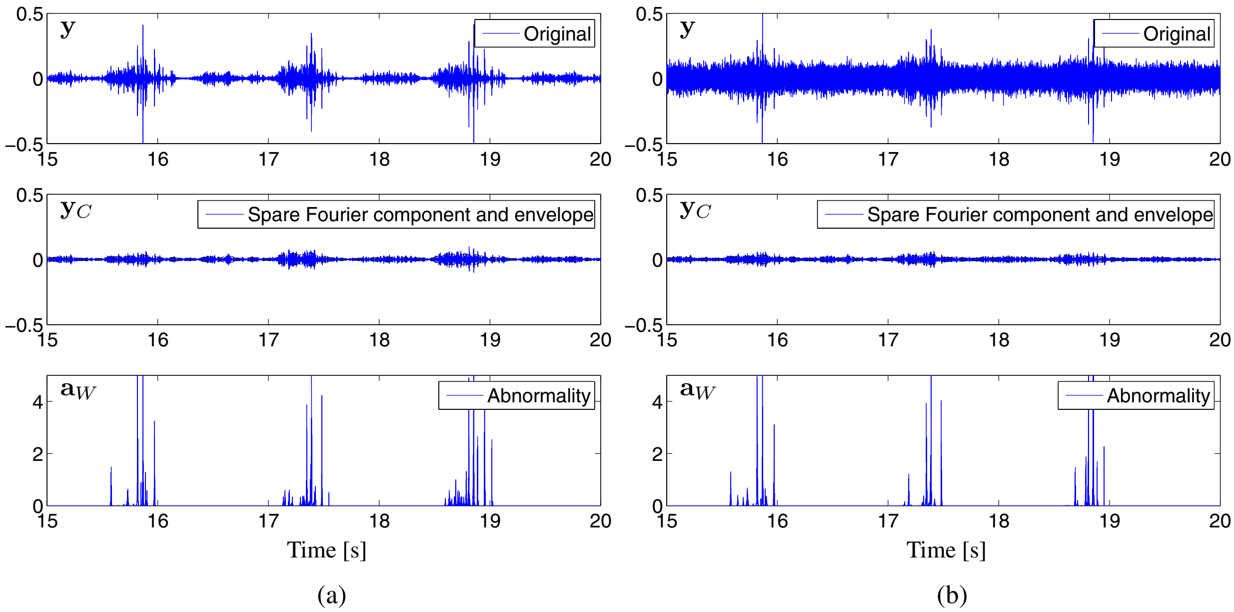


**Figure 1. Estimation of abnormality levels of (a) real auscultation signal of breath sound with fine crackles, and (b) degraded version of (a) by adding white noise of 3dB SN ratio. First row: original signal $\mathbf{y}$. Second row: extracted signal $\mathbf{y}_C$ with sparse frequency spectrum. Third row: abnormality level of extracted signal $\mathbf{y}_W$ with sparse wavelet coefficients.**

in preliminary experiments that 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 $-20$dB if one second of clear auscultation sounds recorded at $44.1$kHz ($d$ =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 $51.2$kbps with single-precision vector $\mathbf{x}$.

For the numerical solution to (7), we employ GPSR [15] 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 $\lambda$ using the sparse solution as a new initial value of $\mathbf{x}$ until $||\mathbf{x}||_0 \leq m$.

### 4.2. Abnormality estimate

We show in Fig. 1 an example of the estimation of abnormality levels for fine crackles in a case of interstitial pneumonitis (idiopathic pulmonary fibrosis). A clear signal $\mathbf{y}_C$ of breath sound is extracted from the original signal $\mathbf{y}$ as in Fig. 1(a). The estimated abnormality levels successfully indicate the accompanying crackles. One can detect the time segments containing crackles by thresholding the abnormality levels. The abnormality rarely shows nonzero levels in the time segments without crackles, which results in avoiding false positive detections.

This property is preserved under noisy conditions. We degraded the original signal by adding white noise of 3dB SN ratio. We could extract the lung sound components from the degraded signal $\mathbf{y}$, and estimate the abnormality levels as shown in Fig. 1(b). We have experimentally confirmed the similar performance of the extraction and the abnormality level estimation on various samples of normal and abnormal lung sounds of different subjects.

### 4.3. Robustness against noise

We evaluate the performance and robustness of the detection of crackles from noisy signals. We manually segmented the signals of lung sounds with fine and coarse crackles, and assigned the labels of ‘normal’ and ‘abnormal’ under the supervision of a doctor of medicine. We added noise to the original signals and calculated the precision and recall values at different thresholds on the abnormality levels for 433 segments with the ground truth labels.

Figure 2 shows the precision-recall curves. We have achieved an average precision of about 85% regardless of noise level. Since the noise cannot be represented sparsely by any bases, our approach with the extracted lung sound signal components can avoid false positive detections due to noise at low threshold levels of abnor-

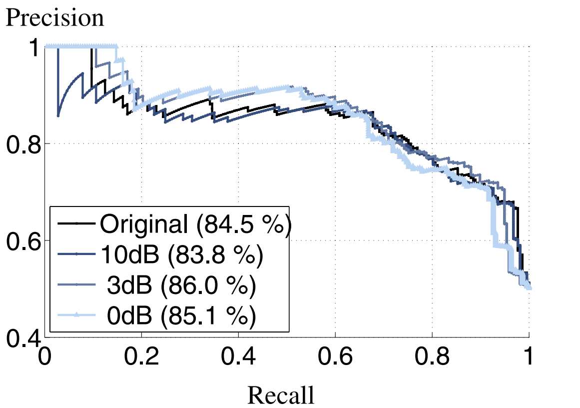


**Figure 2. Precision-recall curves by varying a threshold on the abnormality levels for the detection of crackles under different noise levels. The percentages are the average precisions for the signals having the indicated SN ratios.**

mality to achieve high recall. As the noise increases, the detection performs with slightly higher precision at low recall and lower precision at high recall.

## 5. Concluding remarks

Sparse representation has a great potential for the detection of adventitious lung sounds. Using the clear breath sounds and adventitious sounds separately extracted from noisy auscultation signals, we can measure the level of abnormality and robustly detect the adventitious sounds.

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 and enhancing the robustness not only against internal but also against external noises.

## References


[1] I. 3M Health Care. Electronic stethoscope. Electronic stethoscope model 4100, 2005.

[2] M. Bahoura. Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. *Computers in Biology and Medicine*, 39(9):824–843, 2009.

[3] E. J. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. on Pure and Applied Math*, 59(8):1207–1223, 2006.

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

[5] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[6] D. L. Donoho and M. Elad. Maximal sparsity representation via $l_1$ minimization. In *Proceedings of the National Academy of Sciences*, volume 100, pages 2197–2202, 2003.

[7] 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*, 1(4):586–597, 2007.

[8] A. Kandaswamy, C. S. Kumar, R. P. Ramanathan, S. Jayaraman, and N. Malmurugan. Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine*, 34(6):523–537, 2004.

[9] 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*, 1(4):606–617, 2007.

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

[11] 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*, volume 5012, pages 964–969, 2008.

[12] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka. Respiratory sounds: advances beyond the stethoscope. *American Journal of Respiratory and Critical Care Medicine*, 156(3):974–987, 1997.

[13] T. Sakai, H. Satomoto, S. Kiyasu, and S. Miyahara. Sparse representation-based extraction of pulmonary sound components from low-quality auscultation signals. *Proceedings of ICASSP'12*, 2012.

[14] R. Tomioka and M. Sugiyama. Dual augmented Lagrangian method for efficient sparse reconstruction. *IEEE Signal Processing Letters*, 16(2):1067–1070, 2009.

[15] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Trans. Signal Processing*, 57(7):2479–2493, 2009.

[16] M. Yeginer and P. Y. Kahya. Feature extraction for pulmonary crackle representation via wavelet networks. *Computers in Biology and Medicine*, 39(8):713–721, 2009.