

Fractality evaluation for pulmonary crackle sound using the Degree of Self-Similarity

Achmad Rizal^{1,2,*}, Risanuri Hidayat², and Hanung Adi Nugroho¹

¹ Universitas Gadjah Mada, Department of Electrical Engineering and Information Technology, 55281, Yogyakarta, Indonesia

² Telkom University, School of Electrical Engineering, 40287, Bandung, Indonesia

Abstract. Lung sound is a complex signal produced by the respiratory process. The complex signal has several properties including a chaotic behavior, fractality or self-similarity property. One of lung sounds that arise from abnormalities occurred in the respiratory tract is pulmonary crackle sound. In this study, we tested the degree of self-similarity of pulmonary crackle sound and examined whether the degree of similarity can be used as a feature to differentiate the pulmonary lung crackle sound with normal lung sound. The results showed the sufficient strength of the self-similarity nature of the pulmonary crackle sound. Meanwhile, a test using K-mean clustering produced an accuracy of 87.5% to differentiate between the pulmonary crackle sound and normal lung sound. It can be stated then that it is deemed important to take another feature to obtain higher accuracy. The high self-similarity degree indicates that a pulmonary crackle sound has fractals properties.

1 Introduction

Pulmonary crackle is an adventitious lung sound occurred in the respiratory sounds indicating several abnormalities in the airways. It might be caused by some diseases including lung fibrosis, pneumonia, and chronic bronchitis [1]. Lung sound could be heard using a stethoscope by doctors in diagnosing certain abnormalities in the respiratory system. It is commonly diagnosed in agreement with the expertise and experience of the physician. Along with the development of digital signal processing technology, many methods have been developed to analyze lung sounds automatically.

Various methods have been used to detect the pulmonary crackle sound automatically. In a study reported by Mendes and co-author [2], some Teager features such as energy, information entropy, and local Higuchi fractal dimension were calculated on the non-stationary part of the output of wavelet packed stationary transform-non-stationary transform filter (WPST-NST). In another study, some tests were conducted on the effects of the use of the window, wavelet, and machine learning for the pulmonary sound detection in the time-frequency domain and time-scale domain [3]. The results indicated that Support Vector Machine (SVM) as a classifier could produce the highest accuracy compared to multilayer perceptron and K-NN [3]. Rizal and co-workers used multi-order Tsallis entropy (TE) as a feature extraction method for pulmonary crackle [4]. The reported results showed that TE with the order of 2, 3, and 4 could produce accuracy up to 95.35%. Another study as reported took only three sub-bands (D3, D4, and D5) to analyze the

crackle sound using a wavelet analysis [5]. The suggested mother wavelets were Daubechies 7 and Symlet 7.

Lung sound is a complex biological signal. One characteristic that often appears on a complex signal is fractal properties. Most of the fractals have a self-similarity feature. The self-similarity of the signal can be tested using several methods such as R/S method, variance method, the absolute moment method, and variance ratio of residuals [6]. In this study, we measured the degree of self-similarity (H) using the variance method to observe the characteristics of pulmonary crackle. Self-similarity is one of the characteristics of the fractal signal. Using the degree of self-similarity (H), we could find the fractality nature of pulmonary crackle sound. The H value would be used to determine whether the fractal method could be used for a lung sound analysis in future studies. It also tested whether the value of H could be utilized as a feature to distinguish between crackle sound and normal lung sound.

2 Materials and Methods

The flowchart of the process in this paper is displayed in Fig.1. First, normalization process was applied in the lung sound signal. Then, we calculated the degree of self-similarity (H) of the lung sound signal. We used maximum block size $m = 100$ in this paper. We analyzed the H value of normal and crackle lung sound using ANOVA then tested using K-mean clustering.

* Corresponding author: rizal.s3te14@mail.ugm.ac.id

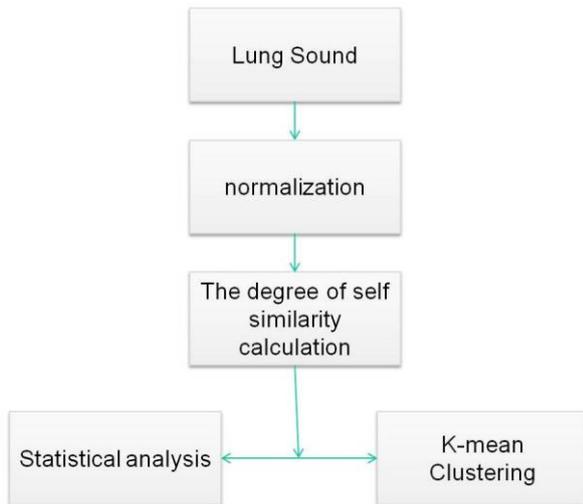


Fig. 1. Flowchart of the process in the research

K-mean clustering was used to test whether the H value could be used as feature for lung sound classification. The detail of each process will be described in next subsection.

2.1 Lung Sound Data

We used the same lung sound data set were collected from the internet as did in a previous paper [4]. The data consisted of 20 crackle sound data and 20 normal bronchial sounds. Crackle sounds were taken from cystic fibrosis and interstitial pulmonary fibrosis patient. Meanwhile, normal data was recorded from 26-year-old man [7]. Data had a single respiratory cycle with a sampling frequency of 8000 Hz. To overcome the differences in the recording process, we did a normalization process as shown in equation (1).

$$y(n) = x(n) - \frac{1}{N} \sum_{i=0}^{N-1} x(n) \quad (1)$$

where $x(n)$ is the input signal, and $y(n)$ is the result of normalization. This process was intended to eliminate the DC component in the signal. Amplitude normalization process was then performed as in equation (2).

$$y(n) = \frac{x(n)}{\max|x|} \quad (2)$$

where $|x|$ is a maximum absolute value of signal x . By the normalization process, the input signal would have a range from -1 to +1 and a zero mean.

2.2 The Degree of Self-Similarity

The degree of self-similarity (H) was proposed by Kalden and Ibrahim to measure self-similarity on GPRS signal [6]. By using H value, the characteristic of the traffic on the GPRS network could be predicted. If a sequence of a signal $x(i)$ is stated $y^m(k)$ as the aggregated process. $y^m(k)$ is formed from a non-overlapping block with the size m as in (3).

$$y^m(k) = \frac{1}{m} \sum_{l=0}^{m-1} x(km - l) \quad (3)$$

If the signal $x(i)$ is stated as self-similar, the variance of the aggregated process will decay based upon the increase of block size m . The variance of y^m is expressed as in (4).

$$\text{var}(y^m) \approx m^{-\beta} \quad (4)$$

where $0 < \beta < 1$, hence the degree of self-similarity (H) can be expressed as in (5)

$$H = 1 - \beta/2 \quad (5)$$

The high value of H shows a strong self-similarity nature [8]. The H value can be an indication of the fractal properties of the signal.

3 Results and Discussion

Typical pulmonary crackle sound and normal bronchial are shown in Fig. 1 indicating that the duration of the crackle sound tends to be short and to have the largest frequency below 500 Hz. On the other hand, the normal bronchial sound has a clear pause between inspiration and expiration phase. Normal bronchial sound tends to continue.

Fig. 2 shows the variance decay as a function of block size m , as we used $m = 1$ to 100. Normal lung sound has slower decay than pulmonary sound so that the H value of normal lung sound is higher than pulmonary crackle as explained in (4) and (5). The boxplot of H value for pulmonary crackle sound and normal lung sound is shown in Fig. 3. Some H value of normal lung sound has a lower value than other so make it overlapped with H value of pulmonary crackle. In another hand, some H value of pulmonary crackle has a higher value to make it have the same value with normal lung sound.

To test the H value separation between normal bronchial sounds crackle, we used Analysis of Variance (ANOVA). Using single-factor ANOVA test, the F-value of 31.93 with a p-value < 0.001 were obtained. This indicated that the value of H in normal bronchial sounds and crackle sound statistically significant differences.

Further, we tested the possibility of using the value of H as a feature for lung sound classification using K-mean clustering. K-mean clustering groups the data into K clusters. The centroid of each cluster was calculated from the average of data grouped in it. Centroid and the distance calculation process were repeated until no data moved to another cluster. We used accuracy for performance evaluation parameter. We defined accuracy as:

$$\text{accuracy}(\%) = \frac{\text{number of data grouped in correct cluster}}{\sum \text{data}} \times 100\% \quad (6)$$

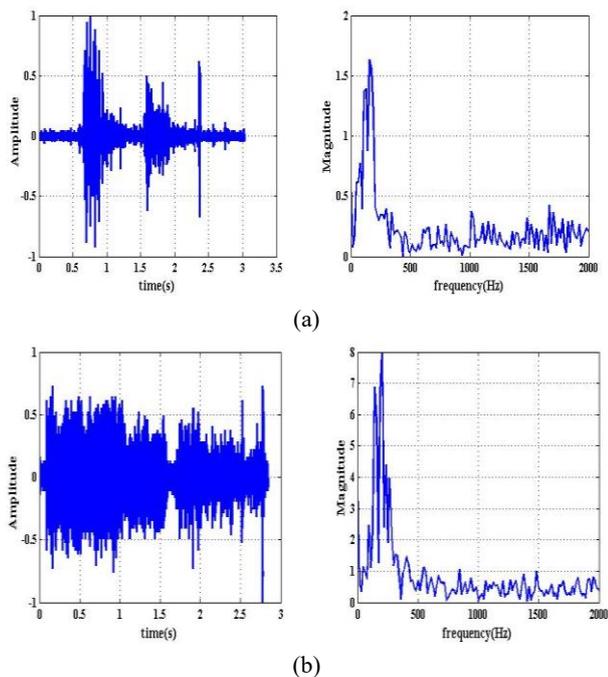


Fig. 1. (a) Crackle sound and the spectral (b) Normal bronchial sound and the spectral

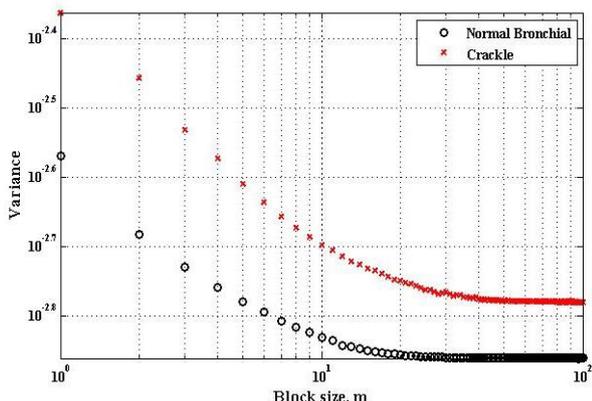


Fig. 2. Variance decay as function of block size (m) for pulmonary crackle sound and normal bronchial sound

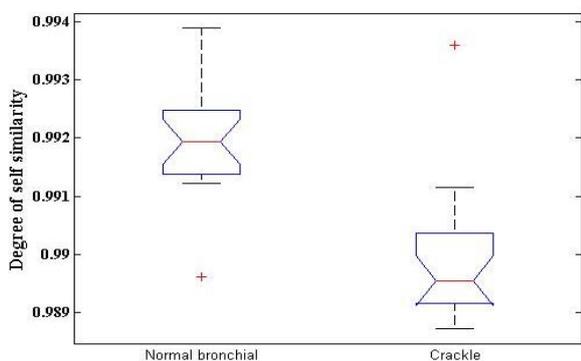


Fig. 3. Box-plot of the degree of self-similarity for normal bronchial sound and crackle sound

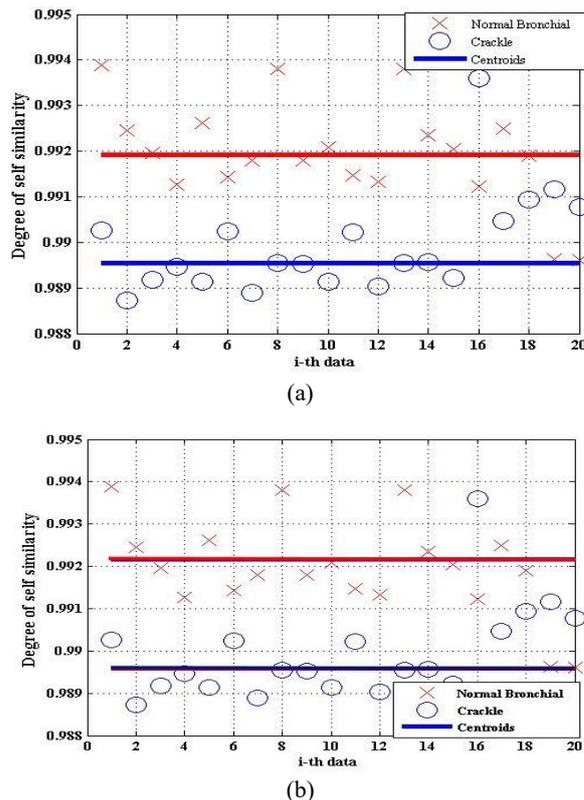


Fig. 4. (a) K-mean clustering result using city-block distance (b) K-mean clustering results using Euclidean distance

We used K-mean clustering with two distance measurement techniques as presented in Fig. 4. As seen in this Figure, some crackle sound data had a value closer to the normal bronchial centroid and vice versa. By using city-block distance, six data went to the wrong cluster, so the accuracy became 85%. On the other hand, by using Euclidean distance, five data were grouped into a wrong cluster in which the accuracy became 87.5%. The results showed that although the H value for pulmonary crackle was statistically different from the normal bronchial sound, it still requires another feature for enhancing the classification accuracy.

Some evaluations of lung sound characteristics had been done in the earlier study. The chaotic dynamics of respiration sound was reported in [9]. Lyapunov spectra, correlation dimension, and Kaplan-Yorke dimension indicated that respiratory sound had a chaotic behavior. Meanwhile, Gnitecki and Mousavi tested fractality of lung sounds using the Katz fractal dimension (KFD), Sevcik fractal dimension (SFD), and the variance fractal dimension (VFD)[10]. The results indicated that the measurement KFD and SFD could lead to defining respiratory sound as a fractal. With more lung sounds properties revealed, it would be easier for researchers to conduct further research. From these results, in future, it will be explored a fractal dimension as a feature for the classification of the various lung sound classes.

4. Conclusion

This paper presented the degree of self-similarity measurement for pulmonary crackle sound. The results

showed that pulmonary crackle sound was very strong indicated as a self-similarity signal. Pulmonary crackle sound also had a different degree of self-similarity compare with a normal bronchial sound. Using K-mean clustering, H value achieved a classification accuracy of 87.5% for normal and crackle sound data. This result indicated a need for other features to enhance the accuracy. Based on this result we can state that fractal dimension can be used for the extraction of crackle sound features for the next study

References

1. A. Bohadana, G. Izbicki, S. S. Kraman. *N. Engl. J. Med.*, **370**, 8 (2014)
2. L. Mendes, P. Carvalho, C. A. Teixeira, R. P. Paiva, J. Henriques. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2014)
3. G. Serbes, C. O. Sakar, Y. P. Kahya, N. Aydin. *Digit. Signal Process. A Rev. J.* **23**, 3(2013)
4. A. Rizal, R. Hidayat, H. A. Nugroho. *The 1st 2016 International Conference on Biomedical Engineering (iBioMed)* (2016)
5. V. I. Quandt, E. R. Pacola, S. F. Pichorim, H. R. Gamba. *Res. Biomed. Eng.* **31**, 2 (2015)
6. R. Kalden, S. Ibrahim. *Lect. Notes Comput. Sci.* **3015** (2004)
7. "The R.A.L.E Repository." [Online]. Available: <http://www.rale.ca/Repository.htm>. [Accessed: 22-Jul-2015].
8. D. K. Kumar, S. P. Arjunan, B. Aliahmad, *Fractal: Application in Biological Signaling and Image Processing* (CRC Press Book, New York, 2016)
9. C. Ahlstrom, A. Johansson, P. Hult, P. Ask, *Chaos, Solitons Fractals*, **29**, 5 (2006)
10. J. Gnitecki, Z. Moussavi. *Chaos, Solitons Fractals*, **26**, 4 (2005)