
There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/271656/

Deposited on: 7 June 2022
A Hybrid Approach of Wavelet-based Total Variation and Wiener Filter to Denoise Adventitious Lung Sound Signal for an Accurate Assessment

Chang Sheng Lee1,2, Minghui Li3, Yaolong Lou1, and Ravinder Dahiya1
1James Watt School of Engineering, University of Glasgow, Glasgow, United Kingdom
2Global Technology Integration Department, Hill-Rom Services Pte Ltd, Singapore
Emails: ChangSheng.Lee@hillrom.com, David.Li@glasgow.ac.uk, Yaolong.Lou@hillrom.com, Ravinder.Dahiya@glasgow.ac.uk

Abstract—Adventitious sounds and their characteristics are the critical indicators of lung dysfunctions. Unfortunately, the captured lung sound often contains noise interferences, which may hinder the accuracy assessment of lung health. This paper proposes a hybrid approach of wavelet-based empirical Wiener filter and wavelet-based total variation (WATV) to denoise adventitious lung sound signals. As an optimal filter, the wavelet-based empirical Wiener filter requires appropriate selections of two wavelet transform bases, whereas WATV indirectly eliminates the need to select the wavelet transform bases by modifying a single objective function to achieve a minimax optimal filter in the sense of mean-squared error. We combined the two approaches by using the improved signal estimation from WATV to design an empirical Wiener filter for suppressing noise and smoothing the denoised signal. The performance of our proposed technique is evaluated via root-mean-squared error (RMSE) and signal-to-noise ratio (SNR) on simulated lung sound containing crackle and wheeze transmitted out of the chest wall and being corrupted with white Gaussian noise at various power levels. In simulation studies, our proposed technique achieved optimal RMSE similarly to the WATV filter accomplishes as an optimal filter — not only preserving signal characteristics but also further improving SNR by 6–9 dB compared to the wavelet soft and hard threshold functions, total variation denoising filter and the WATV filter. Additionally, our proposed technique is less sensitive to the variation of SNR values of the input signal.

Keywords—Denoising, lung sound signal, signal estimation, Wavelets, Wiener filter

I. INTRODUCTION

Auscultation is frequently used by doctors and clinicians to ‘listen’ to weird lung sounds. Despite the wide adoption of auscultation, it is filled with various problems such as variability and uncertainty of inter-listeners. The computer-based lung sound technique eliminates the subjective nature and provides a more reliable approach to assess lung functions [1], [2]. However, in a lung sound recording, interference is an inevitable noise source. The accuracy of the computer-based lung sound technique is lowered with interference; thus, noise reduction or denoising is crucial in lung sound signal processing. In literature, the presence of adventitious lung sounds (crackle and wheeze) are indicators of lung dysfunctions and can be related to airway obstruction and various pulmonary diseases such as chronic obstructive pulmonary disease and sputum production [3]–[5]. Differentiating the adventitious sounds from healthy lung sound is a critical step for assessing lung functions.

Classical wavelet-based thresholding methods are a practical signal denoising approach when the actual noise-free signal is practically unknown [6], [7]. The limitation with classical wavelet transform is introducing artifacts such as spurious-Gibbs oscillations and noise spikes around discontinuities [8]. Additional improvement is to perform empirical Wiener filtering in the wavelet transform domain [9]. However, the limitation with wavelet-domain empirical Wiener filtering is that the approach requires two wavelet transform bases. The effect on denoising the signals differs with different combinations of wavelet bases [9]–[12]. It was proposed in [8] a unified wavelet-based total variation (WATV) approach to overcome the artifacts produced during denoising by modifying a single objective function and indirectly eliminates the need for selecting the appropriate wavelet transform bases required in the wavelet-based empirical Wiener filter. However, WATV still presents small artifacts after denoising the signal, particularly in the lung sound signal containing crackle [13].

This paper, inspired by [8], [9], [13], proposes a hybrid approach of WATV and the wavelet-based empirical Wiener filter. WATV is used to achieve an adequate denoised signal, and the wavelet-based empirical Wiener filter smooths the artifacts produced from the WATV denoised signal to obtain a significantly improved signal-to-noise ratio (SNR) and root-mean-squared error (RMSE) of the denoised signal for an accurate assessment. SNR reflects the denoised signal strength in relation to noise without compromising the frequency of interest; clinicians can better assess lung functions [2], [7], [14]. RMSE results show the filter capability in denoising and retaining significant characteristics from the noise-free lung sound. To the best of our knowledge and the literature survey, the hybrid technique of both WATV and the wavelet-based empirical Wiener filter has not been reported, particularly in the acoustic lung signal domain. The reason could be that both filters are termed as optimal denoising filters in the RMSE/MS sense and achieve good results in their capability. To evaluate our proposed technique’s performance, the wavelet soft and hard thresholding, total variation (TV) denoising, and the state-of-the-art WATV denoising approach are also applied to our simulated noisy lung sound signal containing crackle and wheeze. We achieved better RMSE results by 0.2–0.4 V and higher SNR by 6–9 dB than the wavelet soft and hard thresholding and the TV denoising methods. Our technique achieved similar optimal RMSE performance compared to the optimal WATV filter — showing the capability in preserving signal characteristics and further improving SNR by another 5.5–7.5 dB.

This work is funded by the Singapore Economic Development Board (EDB).
We organized the paper in the following: First, our data model and the assumption, followed by the problem formulation, are presented in Section II. Next, we presented our proposed technique in Section III. Section IV presented the denoised synthesized adventitious lung sound signal results and discussions. Finally, we presented the conclusion and future work in Section V.

II. DATA MODEL AND PROBLEM FORMULATION

Our lung sound model is based on the airflow transmission to the chest wall by the techniques in the communication system and signal processing [15], [16]. The lung sound model contains crackle and wheeze.

The lung sound is modeled as the flow source (airflow) hitting the airway [15], [16]. When the airflow hits the airway, the lung sound is modulated by amplitude and frequency,

\[ x_a(t) = x_s(t)m_a(t)m_f(t), \]  

(1)

where \( x_a(t) \) is the airflow hitting the airway, \( x_s(t) \) is the airflow; the amplitude and frequency modulation functions are denoted as \( m_a(t) \) and \( m_f(t) \), respectively.

The modulated airflow is accompanied by noises when it penetrates the airway wall,

\[ x_f(t) = x_a(t) + v_a(t), \]  

(2)

where \( x_f(t) \) is the airflow with accompanying noises after the airflow hits on the airway, and \( v_a(t) \) is the accompanying noise when \( x_a(t) \) hits the airway.

The noise from the sensor was also transferred, as is customary when noise from electronic devices is fed into the recording system [15], [16],

\[ x(t) = x_f(t) + v_f(t), \]  

(3)

where \( x(t) \) is the airflow that is transmitted out of the chest wall or the modulated signal with noises, and \( v_f(t) \) is the noise transferred from the sensor, such as electronic stethoscope.

Noise is also produced by the ambient and other factors such as speech and cough during the lung sound recording,

\[ y(t) = x(t) + v_n(t), \]  

(4)

where \( y(t) \) is the airflow that is captured by the sensor with noise, and \( v_n(t) \) is the noise caused by ambient. Substituting (1)–(3) into (4), we will have our received lung sound containing noise,

\[ y(t) = x_a(t)m_a(t)m_f(t) + v_a(t) + v_f(t) + v_n(t). \]  

(5)

A reasonable assumption is that the noises are zero-mean process having a probability density distribution that can be defined with mean and variance, uncorrelated with the transmitted lung sound \( x(t) \), with varying SNR levels, similar to those classical signal denoising studies [6], [8], [10], [11]. Hence, we modeled the noises as white Gaussian noise (WGN) [13], [14] and combined \( v_a(t), v_f(t), \) and \( v_n(t) \). Therefore, (5) can be simplified to (6) similar to a linear system, where \( y(t) \) is the received lung sound signal (output) containing WGN (error) \( v(t) \) and the desired lung sound signal (input) \( x_a(t) \) as in (1),

\[ y(t) = x_a(t) + v(t). \]  

(6)

From (6), the desired signal \( x_a(t) \) is contaminated by noise \( v(t) \) from the collisions of the airflow onto the airway, electronic devices, and ambient noise, thus, we have to remove the noise from the captured lung sound signal \( y(t) \) through denoising. However, artifacts are introduced during lung sound signal denoising. The artifacts can lead to misinterpretation or affect the assessment [17]–[19].

III. PROPOSED TECHNIQUE

In this paper, WATV is employed on noisy signal \( y(n) \) with wavelet transform \( W \) and a single objective function \( \tilde{\omega} \) to obtain \( x_t(n) \). Next, the denoised signal \( x_t(n) \) from WATV is employed to design an empirical Wiener filter \( H \) to smooth the denoised signal to reduce artifacts and obtain the desired signal \( x_d(n) \) [8], [9], [11]–[13]. \( n \) is denoted as the sample index, and the total number of samples \( N \) over a known time \( T \) is defined as \( N = F_sT \), where \( F_s \) is the sampling frequency, and set to \( F_s = 4000 \text{ Hz} \) in this work.

In WATV [8], [13], a 5-scale undecimated discrete wavelet transform \( W \) with two vanishing moments fulfilling the Parseval frame condition with Daubechies filter (due to its translation-invariant property in denoising) is used for denoising signal with a low- and high-pass analysis filter [8],

\[ W y(n) = W x_a(n) + W v(n), \quad n = 1, 2, \ldots, N. \]  

(7)

\( W \) is denoted as wavelet transform for denoising signal in (7). The ‘nonstationary’ region of the lung sound signal produces significant wavelet transform coefficients (amplitude) over many wavelet scales. Most of the significant coefficients at each wavelet scale correspond to the desired lung sound signals, whereas the insignificant wavelet coefficients with small values, typically noise, are shrunk during denoising. \( \omega \) is denoted as the wavelet coefficients containing our signal \( x_t \) required for the designing of the empirical wiener filter,

\[ \omega = W x_t. \]  

(8)

Thus, the estimation of signal \( x_t \) denoted as \( \hat{x}_t \) can be obtained by inverse wavelet transform \( W^T \) of wavelet coefficients \( \omega \) shown in (9) once the estimated wavelet coefficients \( \tilde{\omega} \) is available [11],

\[ \hat{x}_t = W^T \tilde{\omega}. \]  

(9)

The wavelet coefficients \( \tilde{\omega} \) can be identified in the following way. We index the terms \( j \) and \( k \) to represent the scale and time information of the signal in the wavelet coefficients \( \omega_j, k \) respectively. The \( \| Dw^T \tilde{\omega} \| \) can be defined as the total variation of signal estimation, where \( D \) is the first-order difference matrix. The single indexed normalized wavelet coefficient is represented as, e.g., \( \| x \|_1 = \sum_n |x_n|; \| x \|_2 = \sum_n |x_n|^2 \). Doubly indexed normalized wavelet coefficient is denoted as, e.g., \( \| \omega \|_2^2 = \sum_{j,k} |\omega_{j,k}|^2 \). To optimize the recovered signal, the choice of regularization parameters \( \lambda \), the threshold shape controller \( \alpha \), minimax concave penalty function \( \phi \) and TV parts
\( \beta \) are critical. We have chosen the parameters as suggested in [8].

The split augmented Lagrangian shrinkage algorithm (SALSA) was applied to solve the WATV denoising problem [8], [13] in (10).

\[
\hat{\omega}(n) = \arg\min_{\omega} \left\{ F(\omega) = \frac{1}{2}\|Wy - \omega\|_2^2 + \sum_{j,k} \rho \lambda\phi(\omega_{j,k}; \alpha_j) + \beta\|D\omega\|_2 \right\}.
\]

(10)

The following parameters have been optimized for denoising and reported in the literature to resist spurious noise spikes, resulting in lower and optimal RMSE [8], [13]. To achieve a balance between wavelet-based and TV denoising, they are controlled by a parameter \( \rho \), and it was chosen as a value of 0.9 [8], [13]. Thus, the regularization parameter is given as \( \lambda = 2.5\rho\sigma/2^{\beta/2} \) and TV parts \( \beta = (1-\rho)(\sqrt{N}/4)\sigma \) where \( \sigma = 3 \), is related to the noise variance \( \sigma^2 \) in each wavelet scale \( j \) [8]. From the regularization parameter \( \lambda \), above, we can identify the threshold shape controller as \( \alpha = 1/\lambda \).

The estimated denoised signal \( \hat{x}_t \) is applied into empirical Wiener filter design in (11)-(12) for smoothing and eliminates the artifacts by minimizing the RMSE to design an improved weighting profile in (12), where \( \sigma^2 \) is the noise variance from the wavelet transform [9],

\[
\hat{x}_u = W^THW\hat{x}_t, \quad (11)
\]

\[
H = \frac{\hat{\omega}^2}{\hat{\omega}^2 + \sigma^2}. \quad (12)
\]

Inspired by [8], [9], the proposed technique is summarized in Fig. 1. In [8], [13], a good denoised signal is achieved (low RMSE); however, small defects still exist, and performance on recovering SNR has not been discussed. Hence, we propose a hybrid WATV and wavelet-based empirical Wiener filtering to smoothen the denoised signal further to achieve a better-denoised signal in terms of SNR and RMSE.

\[
\begin{align*}
&\begin{array}{c}
W
\end{array} \begin{array}{c}
\begin{array}{c}
W
\end{array}
\end{array} \begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\hat{\omega}
\end{array}
\end{array}
\end{array} \begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
W^T
\end{array}
\end{array}
\end{array}
\end{array} \begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
\hat{x}_t
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}
\end{align*}
\]

Fig. 1. A hybrid technique of WATV and wavelet-based empirical Wiener filtering.

From Fig. 1 and (7)-(12), we applied the estimated denoised signal \( x_u(n) \) from WATV to obtain an adequate signal estimate instead of deciding on two wavelet transform bases to obtain an optimal empirical Wiener filter [8], [9], [11]-[13]. The pseudocode of the algorithm is shown below.

Input: Noisy data \( y \); Number of vanishing moment \( k \); Regularization parameter \( \lambda \); TV parts \( \beta \); Step size \( \mu \); Number of wavelet scale \( j \); Number of iteration \( I \).

Initialization: \( \omega = Wy; \alpha_j = 1/\lambda_j \).

//Identifying wavelet coefficient in (1) by iteratively minimizing with respect to \( \omega \) and \( u \) with variable splitting and augmented Lagrangian approach.

\[
u = \omega; d = \omega; v = 0;
\]

//Iteration till convergence between \( \omega \) and \( u \).

\[
\begin{align*}
\text{For} \quad i = 1: I \\
\quad p_{j,k} = [Wy + \mu(u - d)]/(1 + \mu) \\
\quad \text{//Finding the output threshold of } \omega \text{ for all } j,k \text{ with the information from } \phi, \rho, \lambda, \mu, \alpha_j = 1/\lambda_j \\
\quad \omega_{j,k} = \phi(p_{j,k}; \lambda_j/(1 + \mu); \alpha_j) \\
\quad v = d + \omega \\
\quad \text{//Total variation denoising (tvd) requires data input from } v, \text{ length of the data input } (N) \text{ and TV parts} \\
\quad d = W[W^Tv - tvd(W^Tv; N; \beta/\mu)] \\
\quad u = v - d \\
\quad d = d - (u - \omega) \\
\end{align*}
\]

End For

Preliminary Output: Denoised wavelet coefficient (\( \hat{\omega} \)), where signal \( \hat{x}_t = W^T\omega \).

//Empirical Wiener filter design for smoothing: \( H \)

\[
H = \frac{\hat{\omega}^2}{\hat{\omega}^2 + \sigma^2}
\]

//Denoised output:

\[
\hat{x}_u = W^THW\hat{x}_t
\]

IV. SIMULATION STUDIES

This paper performed 500 adventitious lung sounds denoising simulation runs at each noise level generated and analyses on MATLAB R2019b. The simulated adventitious lung sound signal shown in Fig. 2 is fed into the proposed technique and established denoising filters such as the wavelet soft and hard threshold functions [13], [20], the TV denoising filter [13], and the WATV filter [8], [13], [20] for comparing signal denoising performance. The wavelet soft and hard threshold functions are widely used filters in the medical-signal process for achieving better-denoised signal in terms of SNR [1], [7], while the TV denoising filter and WATV have achieved good and excellent RMSE results, respectively, in denoising noisy lung sound signal [13]. The mean RMSE and SNR results are presented and discussed in sub-section D.

To obtain adventitious lung sound shown in (1), The airflow source \( x_s(n) \) is first modulated by the frequency modulation \( f \) cosine wave with an amplitude of 1 V and frequency of 400 Hz, followed by the amplitude modulation \( m_s \) sawtooth wave with amplitude of 1 V amplitude and frequency of 400 Hz.

A. Synthesis of Lung Sound with Crackles

Employing the equations proposed in [21], we simulated adventitious airflow (crackle) transmitted to the airway using (13)-(14). We represent the crackling signal \( x_c(n) \) as two periods, and the crackle modulation function \( m_c(n) \) is employed to shift the energy of \( x_s(n) \) to the initial part of the shape. Fig. 2(a) presented the simulated crackle, with initial deflection width (IDW) = 1.2 ms and two cycle duration (2CD) = 9.8 ms [21].

\[
x_s(n) = [\sin(4\pi n^\alpha)]m_c(n), \quad \alpha = \frac{\log(0.25)}{\log(0.12)} \quad (13)
\]

\[
m_c(n) = 0.5[1 + \cos[2\pi(n^{0.5} - 0.5)]] \quad (14)
\]

B. Synthesis of Lung Sound with Wheezes

Synthesis of wheeze as airflow source \( x_s(n) \) [16] and then transmitted to the airway \( x_w(n) \) is presented in (15). The airflow
source \( x_s(n) \) for wheeze was simulated as a pure sine wave with WGN \( v_w(n) \) power at 50 \( \mu \)W, 1 V amplitude, and \( F = 100 \) Hz for the duration of 100 ms [16]. The simulated wheeze is presented in Fig. 2(c),

\[
x_s(n) = \sin(2\pi(F/F_s)n) + v_w(n).
\]

(15)

C. Synthesis of Noises

The modulation’s accompanying noises \( v_n(n) \) were inserted into the acoustic signals (13) and (15) that penetrate to the airwall shown in (2), with WGN power level and SNR at 0.6 dBm and 0.01 dB [16], respectively. The parameters chosen demonstrated that the proposed communication model corresponds with the physiological characteristics of the actual lung sounds [16]. Finally, the microphone received sound combined with the WGN \( v_f(n) \), power at \( 10^{-6} \) dBm, as is usually the case in electronic communication [15], [16].

![Fig. 2. Simulated adventitious lung sound containing crackle and wheeze transmitted onto the chest wall, corrupted with additive WGN as the noise component \( v_f(t) \): (a) Simulated airflow source crackle; (b) Crackle transmitted onto chest wall with additive WGN; (c) Simulated airflow source wheeze; and (d) Wheeze transmitted onto chest wall with additive WGN.](image)

D. Simulation Results and Discussion

WGN having various SNR values were generated and employed as the noise component \( v_n(t) \) in (4)–(5), similar to the literature [13], [22]. We varied the noise SNR values between 0 dB and 20 dB with a 2 dB increment rate resulting in 11 noise levels. From Fig. 2, we can observe the similarity between our simulated noisy lung sound signals and the actual noisy lung sound signals captured in an uncontrolled environment with microphones in the literature [6], [7].

Equations (16) and (17) showed RMSE, and SNR calculation comparison between our proposed technique, the wavelet soft and hard threshold functions, the TV denoising filter, and the WATV filter in denoising simulated noisy respiratory signals,

\[
\text{RMSE} = \sqrt{\text{mean}[(\|d\| - \|x\|)^2]}.
\]

(16)

where \( d \) is the denoised lung sound amplitude and \( x \) is the noise-free lung sound signal amplitude.

We defined SNR by finding the ratio of the peak amplitude of the denoised signal to peak amplitude of noise signal and expressed the ratio using the logarithmic decibel scale in (17),

\[
\text{SNR} = 20 \log \left( \frac{d}{y - x} \right),
\]

(17)

where \( x \) is the noise-free simulated signal, \( y \) is the simulated noisy signal, and \( d \) is the denoised signal.

From Fig. 3 and Fig. 4, our proposed technique performed better than the wavelet soft and hard thresholding and the TV denoising approach by about 0.3 V in terms of RMSE. Referring to Fig. 3, the WATV filter and our proposed technique achieved an RMSE of 0.44 V and 0.45 V, respectively. From Fig. 4, our proposed technique achieved an RMSE of 0.45 V, and the WATV filter achieved 0.47 V. From the simulation results, our proposed technique RMSE results are estimated to be within ±0.01 V of the WATV filter.

Referring to Fig. 5 and Fig. 6, our proposed technique performed better than the wavelet soft and hard threshold functions, the TV denoising filter, and the WATV filter in terms of SNR. From Fig. 5, our proposed technique improves SNR by about 6.5 dB compared to the other widely used filters in the literature [8], [13], [20]. Similarly, in Fig. 6, our proposed technique improves SNR by about 8.5 dB compared to the other filters. The WATV filter is known as an optimal filter in the RMSE sense, and simulations showed that our technique can achieve the best (optimal) RMSE performance as well, while further achieving higher noise removal in terms of SNR by another 5.5–7.5 dB comparing to the WATV filter [8], [13], [20]. From the RMSE and SNR results, our proposed technique showed its robustness to severe noise in denoising noisy lung sound signals while achieving optimal RMSE.

The performance benefits could be achieved due to the wavelet-based empirical wiener filter smoothing [9]–[12] of the already minimized artifacts [8], [13] and denoised signal with the complementing design of diagonal weighting matrix \( H \) from WATV. Both techniques use the diagonal matrix to design the filter, e.g., translation-invariant denoising matrix in WATV and diagonal weighting matrix in wavelet-based empirical wiener filter.

WATV estimates the wavelet coefficients \( \hat{\omega} \) by considering both insignificant (noise) and significant (signal) coefficients, we used the estimated signal estimates from WATV to design an empirical Wiener filter \( H \) to smooth and reduce the artifacts on the denoised signal. The empirical Wiener filter scales the coefficients by minimizing the MSE to design an improved weighting profile \( H \approx 1 \), with a WATV coefficient more significant than the noise variance, \( \hat{\omega}^2 \gg \sigma^2 \), with the improved weighting profile, our proposed hybrid technique can decrease the denoised signal’s bias and achieve a minimax optimal filter in RMSE. If the noise variance \( \sigma^2 \) is greater than the estimated wavelet coefficient \( \hat{\omega}^2 \), the weighting profile will contribute to the gain in RMSE.
V. CONCLUSION AND FUTURE WORK

Lung sound signal contains unwanted noise, which hinders the assessment of lung function; hence, denoising is critical in lung sound signal processing. Typically artifacts are introduced during lung sound signal denoising; therefore, we propose a hybrid of WATV and wavelet-based empirical Wiener filtering on suppressing noise and further smoothing the signal to achieve high SNR denoised lung sound signal for an accurate lung health assessment. WATV is recognized as an optimal filter in the RMSE sense in the literature. Our proposed technique obtained similar optimal RMSE performance compared to the WATV filter while outperforming other commonly used wavelet-based denoising functions such as the wavelet soft and hard thresholding filters and the TV denoising filter in our simulation studies. In addition, our proposed technique also enhances the SNR of denoised lung sound signals containing crackle and wheeze by about 5.5–7.5 dB compared to the WATV filter. Furthermore, denoising lung sounds in [6]-[9] are directly based on collecting actual respiratory data in a noisy environment. Following the same way, further validation on actual respiratory sound signals containing crackle or wheeze is preferred to show the consistent denoising performance between our simulation studies and the real diagnostic cases.

REFERENCES


