Available online at www.sciencedirect.com

ScienceDirect

Journal of Ocean Engineering and Science 2 (2017) 172–185

www.elsevier.com/locate/joes

Improved signal de-noising in underwater acoustic noise using S-transform: A performance evaluation and comparison with the wavelet transform

Yasin Yousif Al-Aboosi[a,b,*], Ahmad Zuri Sha'ameri[a]

[a] *Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Skudai 81300, Johor, Malaysia*
[b] *Faculty of Engineering, University of Mustansiriyah, Baghdad, Iraq*

Received 15 March 2017; accepted 8 August 2017
Available online 13 August 2017

Abstract

Sound waves propagate well underwater making it useful for target locating and communication. Underwater acoustic noise (UWAN) affects the reliability in applications where the noise comes from multiple sources. In this paper, a novel signal de-noising technique is proposed using S-transform. From the time–frequency representation, de-noising is performed using soft thresholding with universal threshold estimation which is then reconstructed. The UWAN used for the validation is sea truth data collected at Desaru beach on the eastern shore of Johor in Malaysia with the use of broadband hydrophones. The comparison is made with the more conventionally used wavelet transform de-noising method. Two types of signals are evaluated: fixed frequency signals and time-varying signals. The results demonstrate that the proposed method shows better signal to noise ratio (SNR) by 4 dB and lower root mean square error (RMSE) by 3 dB achieved at the Nyquist sampling frequency compared to the previously proposed de-noising method like wavelet transform.
© 2017 Shanghai Jiaotong University. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: Underwater acoustic noise; Time–frequency analysis; Wavelet transforms; S-transforms; Signal de-noising.

1. Introduction

The capability to efficiently communicate and perform target locating underwater is important in various applications, including oceanographic studies, offshore oil prospecting, and defense operation [1]. Electromagnetic waves attenuate seriously in water; thus, sound waves are more suitable solution for underwater communication and target locating [1,2]. Underwater acoustic noise (UWAN) mainly comes from two sources: manmade (shipping, aircraft over the sea, and machinery sounds on the ship) and natural sources (rain, wind, marine lifeforms, and seismic). Considering that UWAN downgrades the acoustic signal quality [2,3], de-noising has to be implemented to improve signal quality [4].

Signal de-noising is important if the signal of interest is corrupted with noise, resulting in difficulty in recovering the information carried by the signal with minimum error. Noise is modeled as additive white Gaussian noise (AWGN) wherein the frequency components are distributed over all frequency range while the signal of interest lies within a specific range in frequency [5]. Different techniques for de-noising were reported, such as mean filtering [6], median filtering [7], Wiener filtering [8], and singular value decomposition (SVD) [9]. Wiener filter can be used to reduce the noise wherein the signal-to-noise ratio (SNR) is sufficiently high (usually higher than 4 dB) [8]. An SVD technique [9] represents a new time domain noise reduction approach. Recently, wavelet transform has emerged as a popular method in signal de-noising. Some of the methods proposed are wavelet correlation method [10], adaptive wavelet shrinkage [11], and dual-tree complex wavelet coefficient method [12]. The essence of these methods is the nonlinear processing on the wavelet coefficients and using the processed coefficients to reconstruct signals. Among

* Corresponding author.
E-mail address: ymyasin2@live.utm.my (Y.Y. Al-Aboosi).

http://dx.doi.org/10.1016/j.joes.2017.08.003
2468-0133/© 2017 Shanghai Jiaotong University. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

these methods, wavelet threshold method is widely used because of its suitability for many applications [13–15]. Given that UWAN is colored noise, two general methodologies can be used to de-noise a known signal. The first methodology performs pre-whitening of the signal before the de-noising operation is implemented using the same methods use for white noise [5]. Otherwise, the de-noising operation can be performed without using a pre-whitening filter but instead using a level-dependent threshold method [16].

In this study, a novel de-noising method based on time–frequency analysis is proposed using S-transform as an alternative to the wavelet transform [17]. The S-transform performs spectral localization derived in [18] and is closely related to the wavelet transform and short-time Fourier transform (STFT). Successful applications of the S-transform include the fields of geophysics [19], power quality analysis [20], and medicine [21]. A new filtering method uses the S-transform applied to the multichannel seismic data to selectively remove noise by applying an adaptive, time-dependent filter [22]. The results show that this technique is particularly effective to determine residual static corrections for data with a low SNR. A novel ECG signal de-noising technique is proposed using S-transform [23]. The experimental results demonstrate that the proposed method shows better SNR performance of 2 dB than the generally used ECG de-noising method. As its key feature, S-transform uniquely combines a frequency-dependent resolution and the localization of the real and imaginary spectra [24]. Furthermore, the S-transform uses time–frequency axis rather than the time-scale axis, resulting in the ease of directly interpreting the time-varying frequency characteristics of the signal.

The rest of this paper is organized as follows. Section 2 defines the signal de-noising problem. Section 3 describes the theoretical background of wavelet and S-transform. The methodologies of de-noising using discrete wavelet transform (DWT) and the proposed method using S-transform are also explained in this section. The results and discussion are discussed in Section 4. Finally, the conclusion of the paper is elaborated in Section 5.

2. Signal de-noising problem

The problem of interference because of noise is common in communication, radar, and sonar systems. In this section, the model for UWAN, which is colored noise, is presented for the signal de-noising in an additive noise channel.

2.1. Signal model

The signals used are single-frequency sinusoidal signal and linear frequency modulated (LFM) signal. They are used to represent fixed-frequency signals and time-varying signals that can be encountered in practical situations. An arbitrary sinusoidal signal can be defined as:

$$\begin{aligned} s(n) &= A\cos(\phi(n)) \qquad 0 \le n \le N-1 \\ &= 0 \qquad\qquad\quad \text{elsewhere} \end{aligned} \tag{1}$$

where N is the signal duration in samples, A is the signal amplitude, and $\phi(n)$ is the instantaneous phase. For a fixed-frequency signal, the instantaneous phase is defined as:

$$\phi(n) = 2\pi f_m n T_s \tag{2}$$

where f_m is signal frequency and T_s is the sampling period. The instantaneous phase for LFM signal is

$$\phi(n) = 2\pi\left(f_m + \frac{\propto}{2} n T_s\right) n T_s \tag{3}$$

where α is the frequency that is defined as $\propto = f_{BW}/NT_s$, and f_{BW} is the bandwidth of the signal.

For band-limited signals of finite energy that has no components higher than W Hz, the sampling frequency f_s must be equal or higher than $2W$ [25]. The sampling rate of $2W$ samples per second is known as the Nyquist rate. The underwater acoustic signals are mostly in the 0–2500 Hz frequency band, and the sampling frequency $f_s = 2\ W$ is the minimum requirement for digital sonar system [26]. In practice, the sampling frequency should be at least equal to $2.66\ W$ [26,27]. Therefore, the sampling frequency used in this work is 8000 Hz.

Given the combined impacts of the internal measurement system and the external environmental factors, measured signals are often contaminated by noise. Thus, the received signal can be defined as

$$x(n) = s(n) + v(n) \tag{4}$$

where $s(n)$ is the signal of interest and $v(n)$ is the UWAN. The assumptions of Gaussian distribution for UWAN are described in [28]. However, recent work suggested that the UWAN follows t-distribution [29] and stable alpha distribution [30]. Therefore, the purpose of de-noising is to reduce the degree of corruption to $s(n)$ by $v(n)$.

2.2. Characteristics of UWAN

Given that UWAN is frequency dependent [1,3], the assumption of additive white Gaussian noise (AWGN) is invalid and is appropriately modeled as colored noise [31–33]. The power spectrum of the colored noise is defined as [5,34]

$$S_{VV}\left(e^{j2\pi f}\right) = \frac{1}{f^{\beta}}\ \beta > 0,\ \frac{-f_s}{2} \le f \le \frac{f_s}{2} \tag{5}$$

The autocorrelation function for AWGN is characterized by the delta function means, and the adjacent samples are independent identically distributed. However, the autocorrelation function of the $\frac{1}{f^{\beta}}$ noise is no longer a delta function but rather takes the form of a *sinc*() function [25,34]. Contrary to AWGN, the noise samples are correlated [5].

Field trials were conducted at Tanjung Balau, Johor, Malaysia (latitude of 1°35.169′N and longitude of 104°16.027′E) on November 5, 2013 to collect signal samples and investigate the statistical properties of UWAN (Fig. 1). The signals were received at a frequency range of 7–22 KHz through a broadband hydrophone (Dolphin EAR 100 Series) located approximately 5 km offshore. The measurements were collected at depths from 1 m to 9 m with a sea floor at a depth

Fig. 1. Field trials conducted at Tanjung Balau, Johor, Malaysia on November 5, 2013.

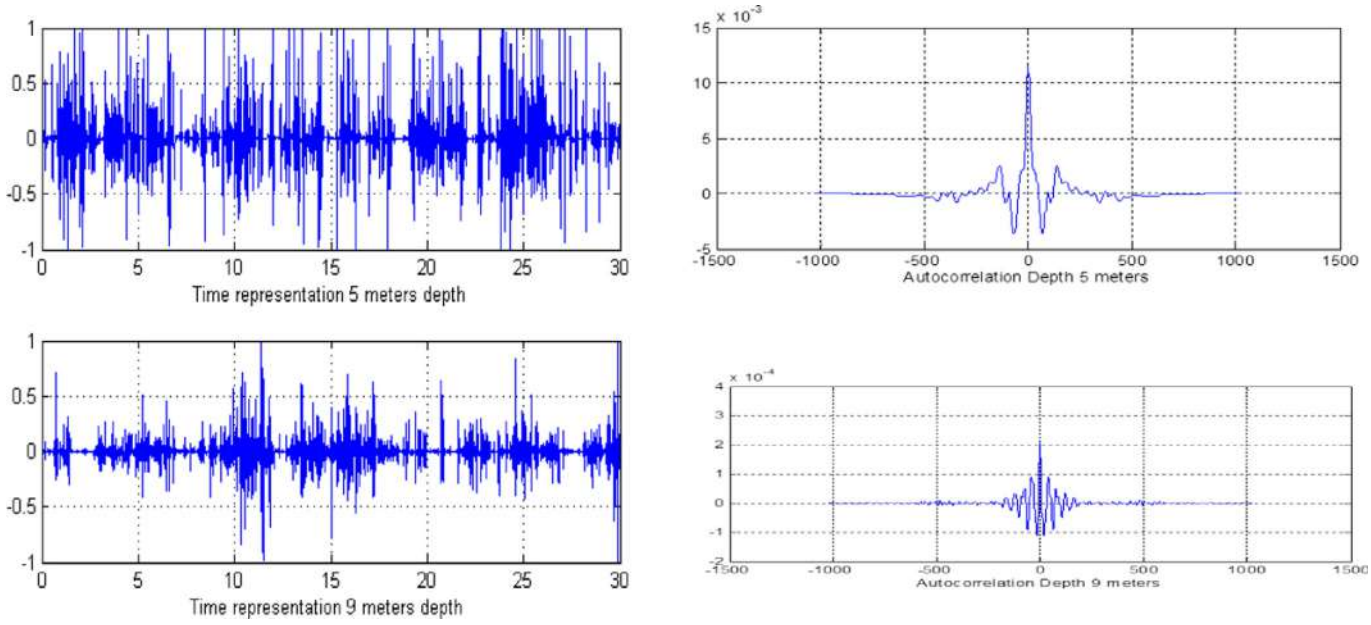


Fig. 2. Time representation and autocorrelation function of the UWAN for depths of 5 and 9 m.

of 10 m. The wind speed was nearly 7 knots, and the surface was at approximately 27 °C [35].

The UWANs were measured to determine the statistical properties such as the power spectral density (PSD), autocorrelation function, and probability density function (pdf). The sampling frequency used in this determination is 8000 Hz. Fig. 2 shows the time representation and the autocorrelation function at the depth of 5 and 9 m. The biased autocorrelation function is not an impulse function, indicating that the noise samples are correlated. The power spectrum is estimated using Welch's modified periodogram technique [36]. The settings for the power spectrum estimation are as follows: window type = Hanning, N−-point fast Fourier transform (FFT) = 2048, FFT window size = 256, and overlapping = 50%. Fig. 3 shows the PSD estimate at different depths. Clearly, the UWAN exhibited a decaying PSD with rates between ($1/f^2$) and ($1/f^3$). Therefore, the field trials confirm that UWAN is colored noise.

3. De-noising of signal in UWAN

3.1. De-noising process flow

Since the UWAN is colored noise, there are two general approaches as shown in Fig. 4 to de-noise a known signal. The first approach shown in Fig. 4(a) performs pre-whitening of the signal before the de-noising operation is implemented [5,25]. Alternatively, the de-noising operation can perform without using a pre-whitening filter but instead using the spectrum characteristic of the colored noise as shown in Fig. 4(b) [16]. This method minimizes the processing with the exclusion of the pre-whitening filter.

3.2. Whitening process and inverse whitening filter

The colored noise can be transformed into white noise by passing it into LTI whitening filter [25,37]. The prediction error filter (PEF) with transfer function $H_p(z)$ is used for this

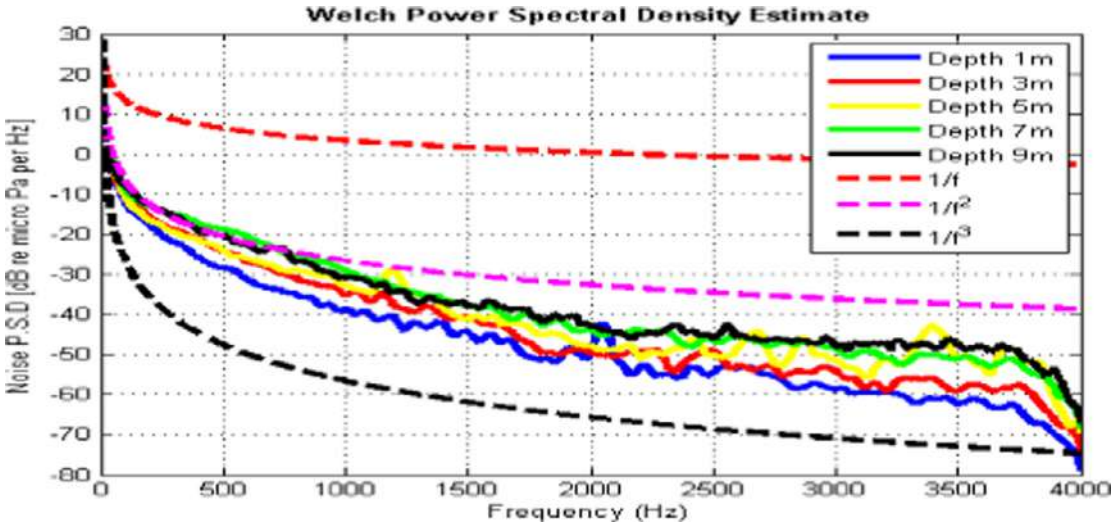
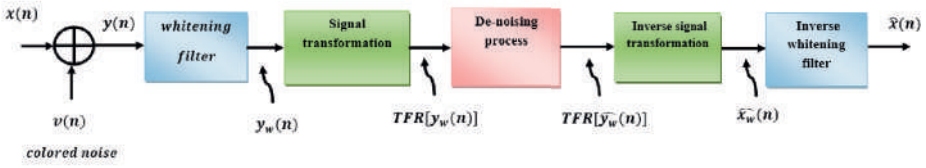
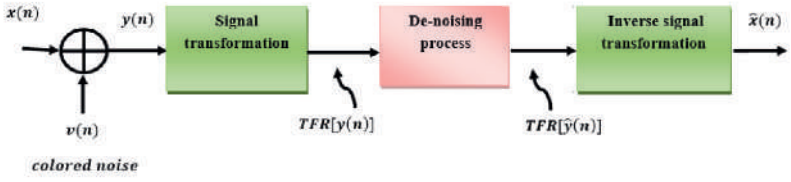


Fig. 3. Welch PSD estimate for the UWAN for depths of 1, 3, 5, 7, and 9 m.



(a) De-noising using a pre-whitening filter.



(b) Direct de-noising method.

Fig. 4. De-noising methods. [Whitening filter: prediction error filter (PEF), signal transformation: wavelet transform or S-transform, de-noising: single-level or multi-level threshold estimation and soft thresholding, inverse signal transformation: inverse wavelet transform or inverse S-transform and inverse whitening filter: inverse PEF].

purpose [5]. The output of the PEF is used as the difference between the estimate of the linear predictor and the actual sequence. The transfer function of the PEF can be expressed as

$$H_p(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_p z^{-p} \tag{5}$$

where p is the length of the forward predictor filter, and $a_p(n)$ is the filter coefficients and depends on the UWAN data recording. If the order p of the PEF is sufficiently large, then the forward prediction error is orthogonal with constant variance; hence, the output of filter is similar to AWGN [5].

The output of the PEF is referred to as $x_w(n)$ and represents the convolution process between the input noisy signal $x(n)$ and the impulse response of the whitening filter $h_w(n)$. Thus, the resulting PEF is the transform of the signal with the AWGN:

$$x_w(n) = x(n) * h_w(n) = s(n) * h_w(n) + v(n) * h_w(n) \tag{6}$$

After the filter coefficients are determined, the de-noising process minimizes the noise term $v(n)^* h_w(n)$ to produce a clean estimate of the transform signal:

$$\hat{s}_w(n) = s(n) * h_w(n) \tag{7}$$

The original signal $s(n)$ can be recovered with an inverse whitening filter [38] with impulse response $h_{IWF}(n)$ and transfer function $1/H_p(z)$. The estimate of the original signal is

$$\hat{s}(n) = \hat{s}_w(n) * h_{IWF}(n) = s(n) * h_w(n) * h_{IWF}(n) \tag{8}$$

In the z-domain, the estimation of the original signal can be express as

$$\hat{S}(z) = S(z).H_P(z) \cdot \frac{1}{H_P(z)} = S(z) \tag{9}$$

3.3. Signal transformation

Signal transformation is used in the de-noising process. The two adopted methods are wavelet transform and S-transform.

3.3.1. Wavelet transform

The wavelet transform is a linear time–frequency distribution that decomposes the signal into a family of functions localize in time and frequency. The continuous wavelet transform can be expressed as

$$X(t,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(\tau) h\left(\frac{\tau - t}{a}\right) d\tau \tag{10}$$

where t is the time shift, a is the scale (known as dilation) factor, and $h(t)$ is the basis function (known as mother wavelet). The choice of basis function is signal dependent, and examples of basis functions are Debauchies, Coiflet, Symlet, and Biorthogonal. To find a suitable basis function, cross correlation is performed between the original signal $s(n)$, and the signal after the reverse wavelet transformation is applied. The highest cross correlation is then used to select the suitable basis function [39].

The signal is processed in discrete time. Thus, using the DWT is more appropriate. The DWT is defined as

$$X(n,k) = \frac{1}{\sqrt{k}} \sum_{m=0}^{N-1} x(m) h\left(\frac{m-n}{k}\right) \tag{11}$$

where n is the time shift and k is the scale factor. The DWT is computed by passing the time domain signal $x(n)$ successively through L level of high pass and low pass filters with decimation by 2. The procedure formed from hierarchical set of approximation and detail analysis is known as multiresolution analysis, which can be implemented using computationally efficient algorithms [40].

3.3.2. S-transform

The S-transform is a special case of the STFT by replacing the window function with a frequency-dependent Gaussian window [18,41]. The Gaussian window width is inversely proportional to the frequency, and its height is scaled linearly to the frequency. Given the behavior of the window scaling, the S-transform possesses good time resolution for high-frequency components and good frequency resolution for low-frequency components. The S-transform can be expressed as [18]

$$X(t,f) = \int_{-\infty}^{\infty} x(\tau)\ g(\tau - t, f) e^{-j2\pi f\tau} d\tau \tag{12}$$

where $x(t)$ is the signal and $g(t, f)$ is the frequency-dependent Gaussian window. The window is given as [18]

$$g(t,f) = \frac{|f|}{\sqrt{2\pi}} e^{\left(\frac{-t^2 f^2}{2}\right)} \tag{13}$$

The presence of the variable f makes the window spread frequency dependent. Given that $X(t, f)$ is complex valued, the modulus $|X(t, f)|$ is usually plotted in practice to derive the time–frequency representation. The S-transform represents the local spectrum; thus, averaging of the local spectrum over time produces the Fourier spectrum that can be expressed as [18]

$$\int_{-\infty}^{\infty} X(t,f) dt = X(f) \tag{14}$$

The signal in time representation $x(t)$ is exactly recoverable from $X(t, f)$ based on the following expression [18,42]:

$$x(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} X(t,f) dt \right\} e^{j2\pi ft} df \tag{15}$$

In this study, the de-noising is performed in the time–frequency representation and the signal is recovered from noise using Eq. (15).

The discrete S-transform is used in this study to allow processing in the continuous S-transform. $x(n)$, $n = 0, 1, \ldots (N-1)$ denotes a discrete time series corresponding to $x(t)$ with a time sampling interval of T_s. The S-transform of the discrete time series $x(n)$ is given by

$$X(n,k) = \sum_{m=0}^{N-1} x(m-n) e^{\frac{-2\pi^2 m^2}{k^2}} e^{\frac{-j2\pi mk}{N}} \tag{16}$$

The inverse of the discrete S-transform is given by [43]:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} \left\{ \sum_{n=0}^{N-1} X(n,k) \right\} e^{\frac{j2\pi nk}{N}} \tag{17}$$

3.4. De-noising process

The de-noising methodology using wavelet transform and S-transform based on the block diagram in Fig. 4 is discussed in this section.

3.4.1. Wavelet-based method

In wavelet de-noising, the algorithm to de-noise a signal $s(n)$ corrupted by a noise signal $v(n)$ can be summarized by the following three steps [44,45]:

(1) Decomposition: Choose a wavelet and a decomposition level L. Compute the wavelet decomposition of the noisy signal at level L.
(2) Threshold detailed coefficient: For each level from 1 to L, select a threshold value and apply hard or soft thresholding on the detailed noisy coefficients.
(3) Reconstruction: Compute wavelet reconstruction using the original approximation coefficient of level L and modified detailed coefficient of levels 1 to L.

The de-noising method that applies thresholding in wavelet domain was proposed by [46]. The threshold value applied to the wavelet coefficients estimated at the kth decomposition level using adaptive universal threshold estimation [46] is given by

$$\gamma_k = c.\sigma_{v,k}\sqrt{2\log(N)} \tag{18}$$

where N is signal length at each level, $\sigma_{v,\,k}$ is the kth noise standard deviation, and c is adaptive universal threshold factor $0 < c < 1$. The kth noise standard deviation is expressed as

$$\sigma_{v,k} = \frac{\text{median}(|X_D(n,k)|)}{0.6745} \tag{19}$$

where $X_D(n,\ k)$ represents all the wavelet detail coefficients in level k [16]. For white noise, the estimation of the standard deviation from the median absolute value of the detail coefficients estimated for the first level alone is sufficient and can be used for the other levels. The reason is that the magnitude of the noise is the same for all frequencies. However, the noise standard deviation for colored noise has to be calculated for all levels [44].

Threshold value γ_k is used to remove noise and recover the original signal efficiently. The adaptive threshold factor c is introduced to further improve the de-noising performance [47]. The value of c is determined by gradually changing its value for each level with steps of 0.1. Through this procedure, the best results are found at the highest SNR at the output of de-noising process.

Thresholding is then applied on the coefficients $X_D(n,\ k)$ using the threshold value γ_k to separate the signal from the noise. Hard thresholding is the simplest, and the thresholded values of the detail coefficients, $X_{D,\,\gamma}(n,\ k)$, are obtained according to the following:

$$X_{D,\gamma}(n,k) = \begin{cases} X_D(n,k) & \text{if } |X_D(n,k)\ | > \gamma_k \\ 0 & \text{if } |X_D(n,k)| \le \gamma_k \end{cases} \tag{20}$$

Hard thresholding zeroes out all the signal values smaller than γ_k. Alternatively, soft thresholding can be used and the signal after thresholding is expressed as

$$X_{D,\gamma}(n,k) = \begin{cases} sgn(X_D(n,k))(|X_D(n,k)| - \gamma_k) & \text{if } |X_D(n,k)| > \gamma_k \\ 0 & \text{if } |X_D(n,k)| \le \gamma_k \end{cases} \tag{21}$$

In soft thresholding, all the coefficients with magnitudes smaller than the threshold value γ_k are set to zero; all the remaining coefficients are also reduced in magnitude by the amount of the threshold value [46]. Contrary to hard thresholding, soft thresholding causes no discontinuities in the de-noise signal.

To obtain the de-noised signal, the new threshold detail coefficients $X_{D,\,\gamma}(n,\ k)$ and original approximation coefficient of level L, $X_{A,\,L}(n,\ k)$, are used for the signal reconstruction process as in Eq. (22) and this process consists of up-sampling by 2 and filtering.

$$x(n) = A_L(n,k) + \sum_{\ell=1}^{L} D_\ell(n,k) \tag{22}$$

3.4.2. S-transform-based method

If the received signal is AWGN, then it can be considered a combination of narrow-band Gaussian random processes wherein each process is centered at a frequency f_l. If the random process is stationary, then the narrow-band Gaussian random process can be expressed as [5,48]

$$x(n) = v_I \cos(2\pi f_l n) - v_Q \sin(2\pi f_l n) \tag{23}$$

where v_I is the in-phase component and v_Q is the quadrature component. v_I and v_Q are independent Gaussian random processes with zero mean and variance σ_v^2. Combining all the narrow-band noise components in Eq. (23) results in the AWGN that is defined as

$$x(n) = \sum_{\ell=-\infty}^{\infty} v_{I,\ell} \cos(2\pi f_\ell n) - v_{Q,\ell} \sin(2\pi f_\ell n) \tag{24}$$

By limiting the signal duration within $0 \le n \le N-1$, the frequency representation obtained from the discrete Fourier transform (DFT) of $x(n)$ is:

$$\begin{aligned} X(k) &= \sum_{\ell=-\infty}^{\infty} v_{I,\ell}\ \text{DFT}\big[\cos(2\pi f_\ell n)\big] - v_{Q,\ell}\ \text{DFT}\big[\sin(2\pi f_\ell n)\big] \\ &= \sum_{\ell=-\infty}^{\infty} v_{I,\ell} \mathrm{X}_{I,\ell}(k)\ - v_{Q,\ell} \mathrm{X}_{Q,\ell}(k) \end{aligned} \tag{25}$$

Segregating the $X(k)$ into real and imaginary part results in:

$$\begin{aligned} X(k) &= \sum_{\ell=-\infty}^{\infty} \left(v_{I,\ell}\ re\{\mathrm{X}_{I,\ell}(\mathrm{k})\}\ - v_{Q,\ell} re\{\mathrm{X}_{Q,\ell}(\mathrm{k})\}\right) \\ &\quad + j\ \left(v_{I,\ell}\ im\{\mathrm{X}_{I,\ell}(\mathrm{k})\}\ - v_{Q,\ell} im\{\mathrm{X}_{Q,\ell}(\mathrm{k})\}\right) \\ &= X_{re}(k) + \ jX_{im}(k) \end{aligned} \tag{26}$$

For the, ℓth terms in $X_{re}(k)$ and $X_{img}(k)$, the magnitude is a linear combination of $v_{I,\,\ell}$ and $v_{Q,\,\ell}$. Since both the pdf of $v_{I,\,\ell}$ and $V_{Q,\,\ell}$ are independent Gaussian random process, than the resulting magnitude pdf is a convolution of their two pdfs which result in a Gaussian pdf [5]. Similarly, a colored noise where the power spectrum decays at the rate of *1/f*, the frequency representations becomes:

$$\begin{aligned} X(k) = \frac{1}{f_\ell^{1/2}} \Bigg\{ &\sum_{\ell=-\infty}^{\infty} \left(v_{I,\ell}\ re\{\mathrm{X}_{I,\ell}(\mathrm{k})\} - v_{Q,\ell} re\{\mathrm{X}_{Q,\ell}(\mathrm{k})\}\right) \\ &+ j\left(v_{I,\ell}\ im\{\mathrm{X}_{I,\ell}(\mathrm{k})\}\ - v_{Q,\ell} im\{\mathrm{X}_{Q,\ell}(\mathrm{k})\}\right)\Bigg\} \end{aligned} \tag{27}$$

By using the properties of the discrete Fourier transform (DFT) and the convolution theorem, the S-transform in Eq. (12) can be considered as a convolution process in the frequency domain between the signal $X(k)$ and the scalable localizing Gaussian window (k). The S-transform can be

expressed as:

$$X(n,k) = \sum_{m=0}^{N-1} x(m-n) e^{\frac{-2\pi^2 m^2}{k^2}} e^{\frac{-j2\pi mk}{N}} = \sum_{m=0}^{N-1} x(m-n) w(m) e^{\frac{-j2\pi mk}{N}} = X(k) \underset{(k)}{*} W(k) e^{-j\frac{2\pi nk}{N}} \quad (28)$$

By considering the signal in terms of the real and imaginary parts, Eq. (28) can be written as:

$$X(n,k) = \left[X_{re}(k) + jX_{im}(k)\right] \underset{(k)}{*} W(k) e^{-j\frac{2\pi nk}{N}} = \left[X_{re}(k) \underset{(k)}{*} W(k) e^{-j\frac{2\pi nk}{N}} + jX_{im}(k) \underset{(k)}{*} W(k) e^{-j\frac{2\pi nk}{N}} \right] \quad (29)$$

The complex exponential denoted by $(e^{-j\frac{2\pi nk}{N}})$ represents the shift of the window in time domain.

From Eq. (29), the de-noising process in frequency domain requires thresholding in the real and imaginary parts of the spectrum. Thus, the adaptive universal threshold estimation described in Eq. (18) has to include real and imaginary components. The threshold value is given by

$$\gamma_{k,re} = c.\sigma_{v,k,re}\sqrt{2\log(N)} \qquad \gamma_{k,im} = c.\sigma_{v,k,im}\sqrt{2\log(N)} \quad (30)$$

where $\sigma_{v,\,k,\,re}$ and $\sigma_{v,\,k,\,im}$ are the noise standard deviations for the real and imaginary parts, respectively. The kth noise standard deviations are calculated as

$$\sigma_{v,k,re} = \frac{\text{median}(|X_{re}(n,k)\ |)}{0.6745} \qquad \sigma_{v,k,im} = \frac{\text{median}(|X_{im}(n,k)|)}{0.6745} \quad (31)$$

The median for AWGN and colored noise is calculated in the same way as explained in Section 3.4.1.

After determining the threshold values for real and imaginary parts $\gamma_{k,\,re}$ and $\gamma_{k,\,im}$, the time–frequency representations of the real and imaginary parts after hard thresholding are

$$X_{\gamma,re}(n,k) = \begin{cases} X_{re}(n,k) & \text{if } |X_{re}(n,k)\ | > \gamma_{k,re} \\ 0 & \text{if } |X_{re}(n,k)| \le \gamma_{k,re} \end{cases} \qquad X_{\gamma,im}(n,k) = \begin{cases} X_{im}(n,k) & \text{if } |X_{im}(n,k)\ | > \gamma_{k,im} \\ 0 & \text{if } |X_{im}(n,k)\ | \le \gamma_{k,im} \end{cases} \quad (32)$$

Alternatively, the time–frequency representations of the real and imaginary parts after soft thresholding are

$$X_{\gamma,re}(n,k) = \begin{cases} sgn(X_{re}(n,k))\left(|X_{re}(n,k)| - \gamma_{k,re}\right) & \text{if } |X_{re}(n,k)| > \gamma_{k,re} \\ 0 & \text{if } |X_{re}(n,k)| \le \gamma_{k,re} \end{cases}$$

$$X_{\gamma,im}(n,k) = \begin{cases} sgn(X_{im}(n,k))\left(|X_{im}(n,k)\ | - \gamma_{k,im}\right) & \text{if } |X_{im}(n,k)| > \gamma_{k,im} \\ 0 & \text{if } |X_{im}(n,k)| \le \gamma_{k,im} \end{cases} \quad (33)$$

The time–frequency representation obtained by combining real and imaginary parts is

$$X_{\gamma}(n,k) = X_{\gamma,re}(n,k) + jX_{\gamma,im}(n,k) \quad (34)$$

The de-noised signal $x(n)$ in time representation can be recovered using the inverse discrete S-transform in Eq. (17).

3.5. Performance measures

Many quantitative parameters can be used to evaluate the performance of the de-noising procedure for reconstructing signal quality. The two most widely used performance measures are [23,44] SNR and root mean square error (RMSE). The SNR is defined as

$$\text{SNR} = 10\log\left[\frac{\sum_{n=1}^{N}[s(n)]^2}{\sum_{n=1}^{N}[\hat{s}(n) - s(n)]^2}\right] \quad (35)$$

where $s(n)$ is the original signal, $\hat{s}(n)$ is the de-noised signal, and N is the length of the signal. The de-noising is successful if post SNR is higher than pre SNR.

The RMSE between the original signal $s(n)$ and the de-noise signal $\hat{s}(n)$ is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left[\hat{s}(n) - s(n)\right]^2} \quad (36)$$

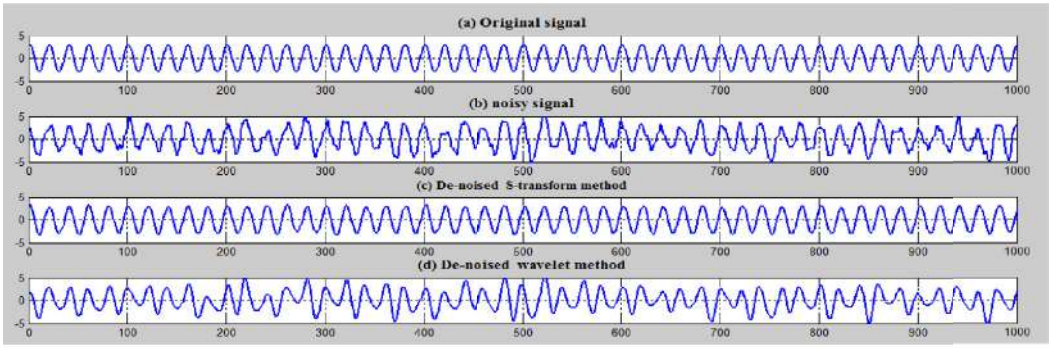
The RMSE gives a measure of how well the de-noised signal is similar to the original signal. The performance of the various methods in de-noising the signals can be compared using a suitable criterion, that is, identifying the technique that results in the highest SNR with the lowest possible RMSE.

4. Results and discussion

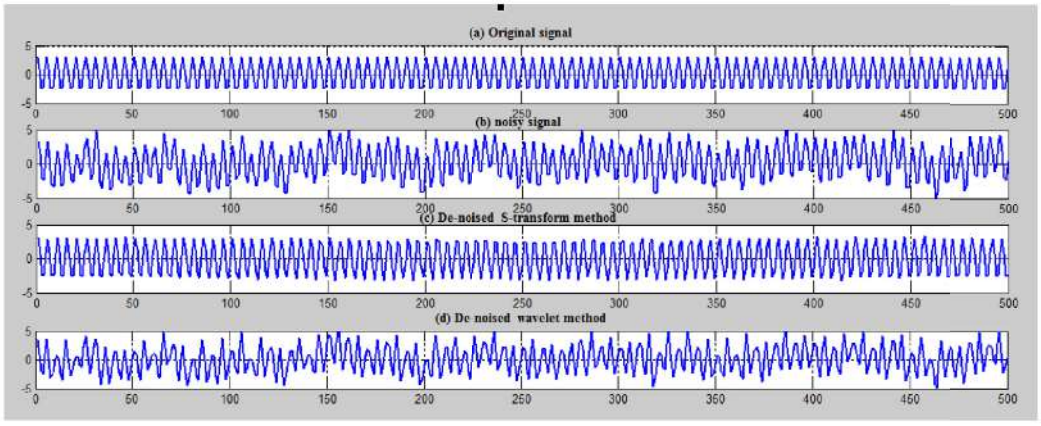
The different de-noising methods are tested for signals in UWAN. The UWAN used for the validation is sea truth data collected at Tanjung Balau, Johor, Malaysia using broad band hydrophones. The signals are defined in Eqs. (1)–(4) for fixed-frequency signal as well as time-varying signal. The characteristics of these signals are as follows:

(1) Single-tone signal with frequency of 400 Hz and length of 1000 samples.
(2) Single-tone signal with frequency of 1500 Hz and length of 1000 samples.
(3) LFM signal with initial frequency of 400 Hz and final frequency of 1500 Hz.

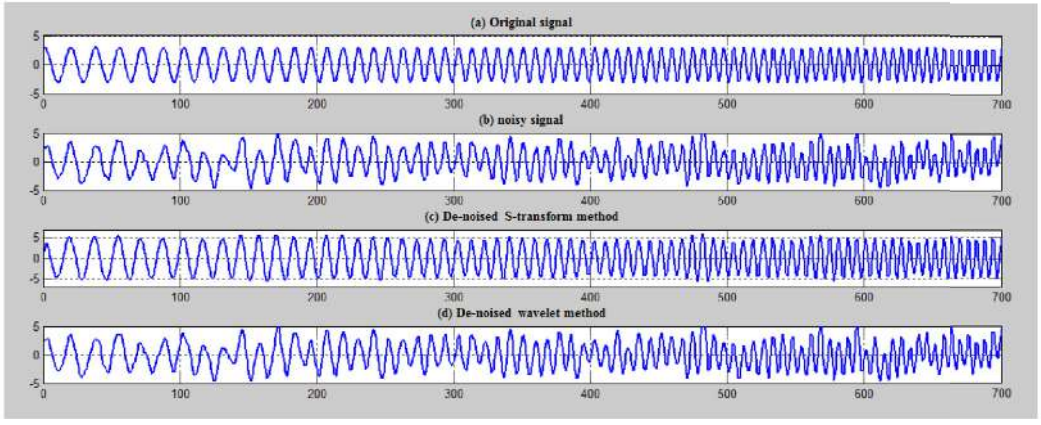
All these signals have sampling frequency of 8000 Hz. The simulations were performed at SNR from −6 dB to 8 dB by varying the signal power while keeping the noise



(a) Single tone signal with frequency 400 Hz.



(b) Single tone signal with frequency 1500 Hz.

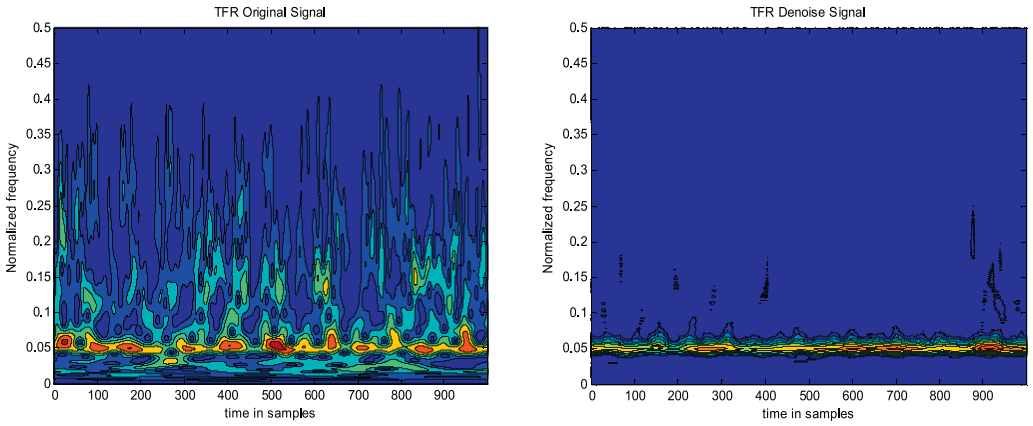


(c) LFM signal.

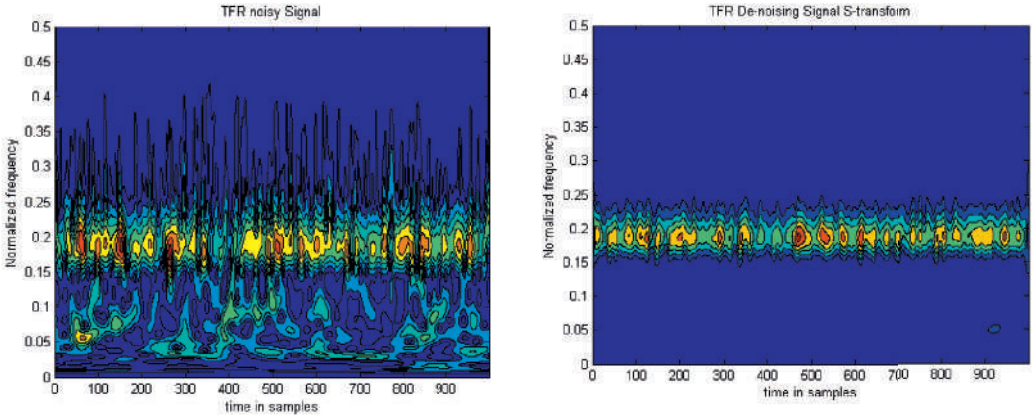
Fig. 5. Time representation of (a) original signal, (b) noisy signal with SNR of 6 dB, (c) de-noised signal using S-transform method, and (d) de-noised signal using wavelet method.

power constant with variance equals one. The de-noising using Daubechies wavelet of order 9 with soft thresholding [49] is used and compared with the proposed S-transform de-noising method. The pre-whitening filter of order 10 was used [24]. The different de-noising methods are referred to as follows:

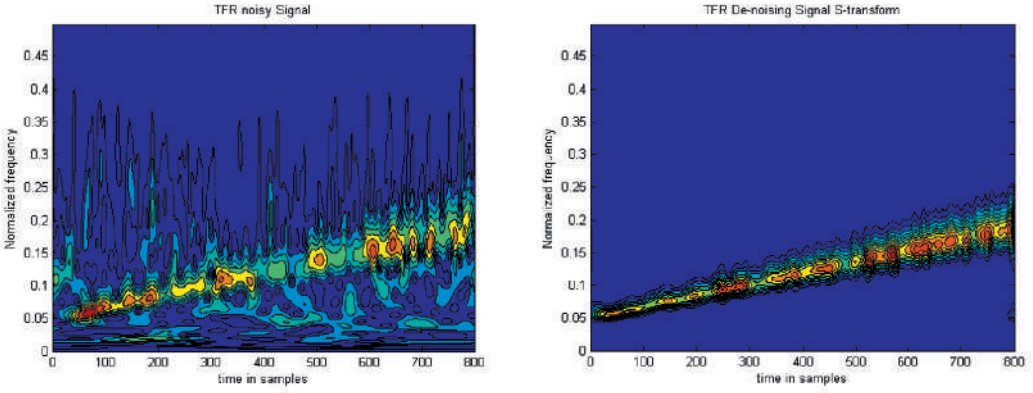
1. S-transform direct method without a pre-whitening filter and multi-level noise standard deviation estimation, as shown in Fig. 4(b).
2. S-transform method with pre-whitening filter and single-level noise standard deviation estimation, as shown in Fig. 4(a).



(a) Single tone signal with frequency 400 Hz.



(b) Single tone signal with frequency 1500 Hz.

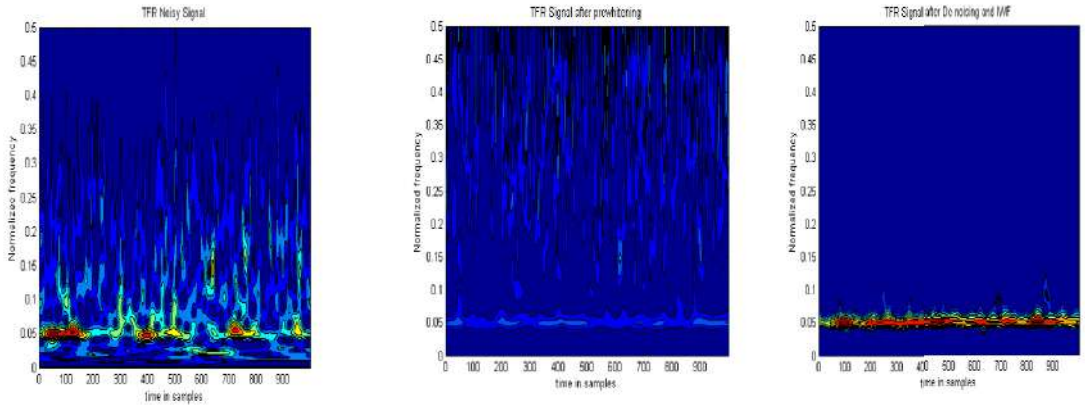


(c) LFM signal.

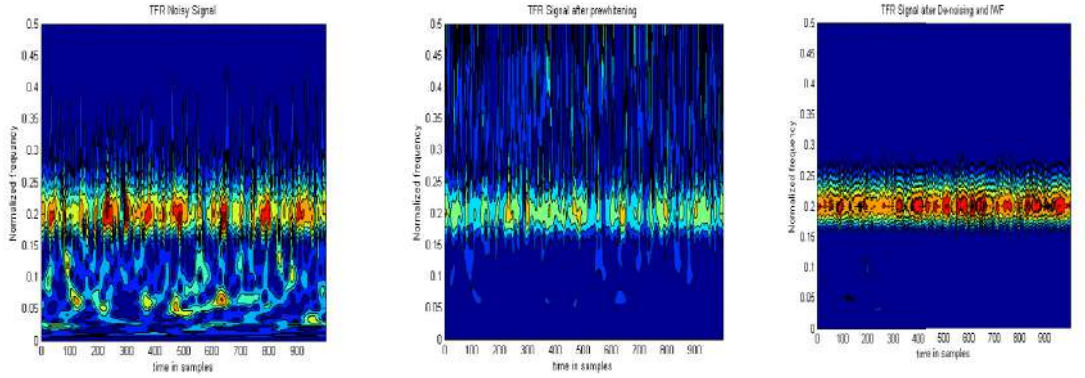
Fig. 6. Time–frequency representation of noisy signal and de-noised signal using S-transform direct method [the SNR of the signal is 6 dB].

3. Wavelet direct method without a pre-whitening filter and multi-level noise standard deviation estimation, as shown in Fig. 4(b).
4. Wavelet with pre-whitening filter and single-level noise standard deviation estimation, as shown in Fig. 4(a).

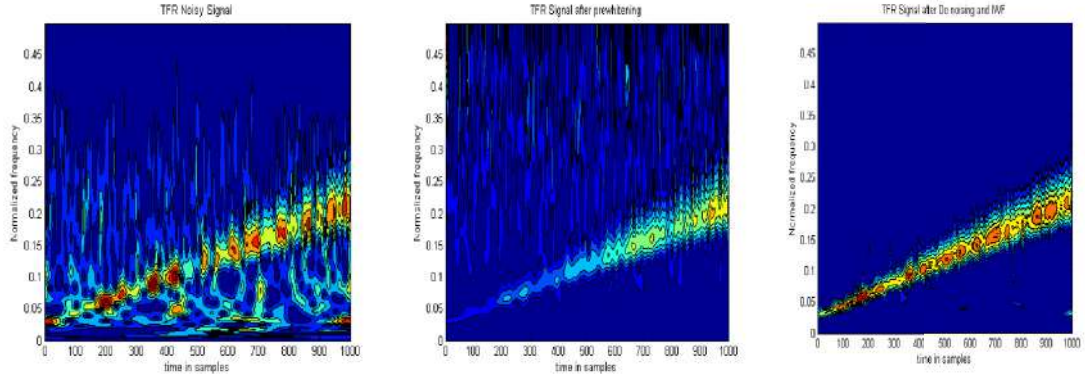
The time representation of original, noisy, and de-noised signals using S-transform direct method and wavelet transform direct method is shown in Fig. 5. The difference between the results produced by wavelet de-noising and S-transform de-noising can be easily compared with the help of this figure.



(a) Single tone signal with frequency 400 Hz.



(b) Single tone signal with frequency 1500 Hz.

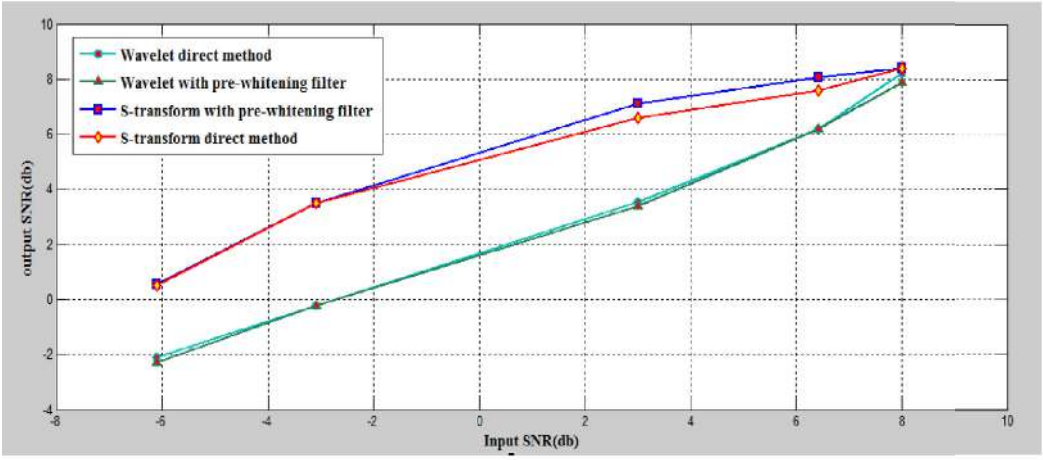


(c) LFM signal.

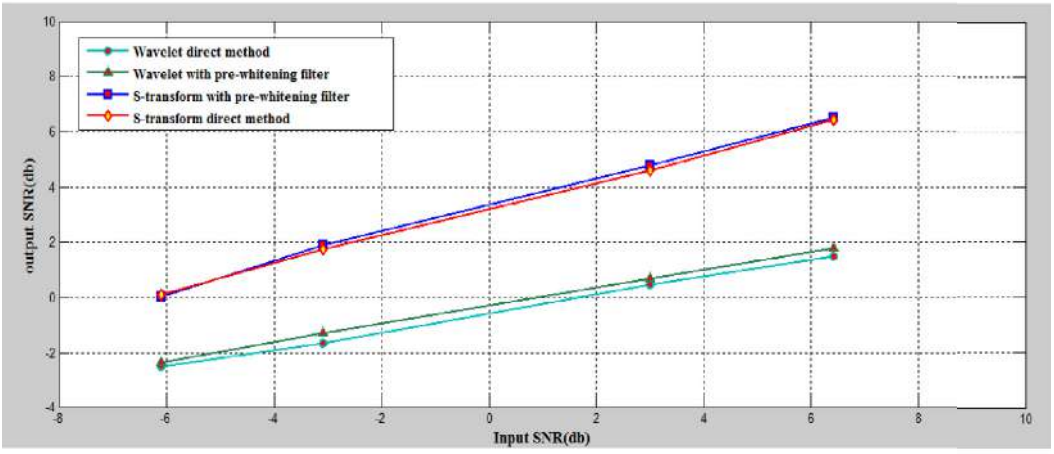
Fig. 7. Time–frequency representation for noisy signal and de-noised signal using S-transform method with pre-whitening filter [the SNR of the signal is 6 dB].

Fig. 6 shows the time–frequency representation for the noisy and de-noised signals using the S-transform direct method. Given that UWAN is colored noise, the power concentration for all signals is in the low frequency region ($f < 0.25$ Hz) rather than in the high frequency region ($f > 0.25$ Hz). Thus, the noise standard deviation at each level in frequency axis needs to be estimated before calculating the threshold value and applying the soft thresholding. Clearly, significant differences exist between the noisy signal and the de-noised signal.

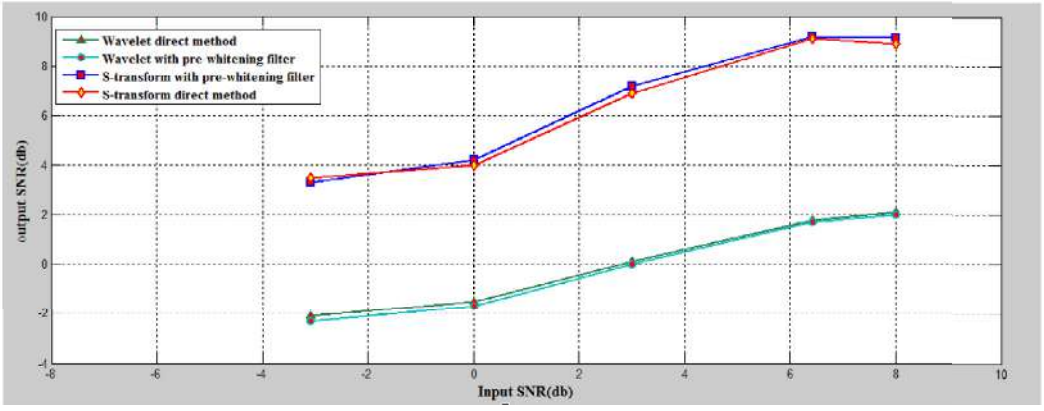
Fig. 7 shows the time–frequency representation for the noisy and de-noised signals using the S-transform method with pre-whitening filter. The pre-whitening filter transforms the colored noise in the noisy signal to white noise. Given that



(a) Single tone signal with frequency 400 Hz.



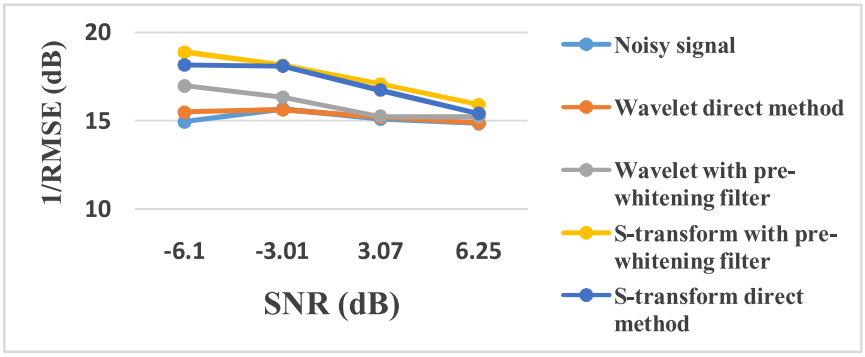
(b) Single tone signal with frequency 1500 Hz.



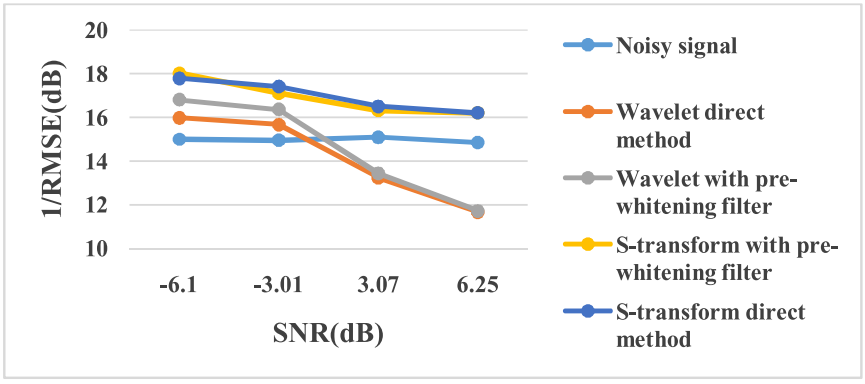
(c) LFM signal.

Fig. 8. Comparison of the output SNR obtained using DWT methods and the proposed S-transform de-noising methods.

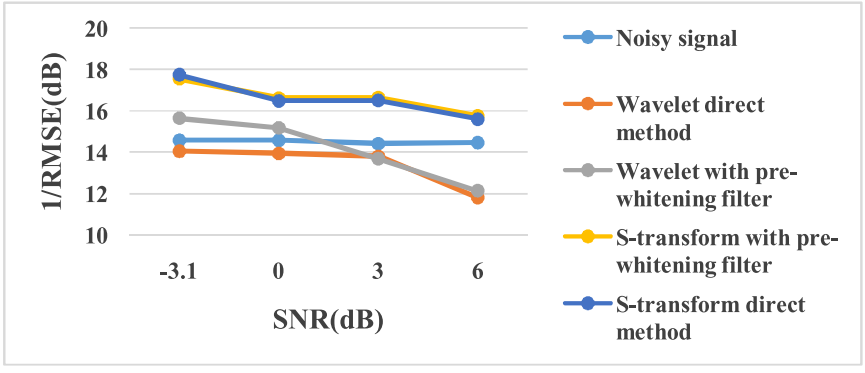
the power concentration is equally distributed in frequency, the noise standard deviation can be estimated at fixed frequency level to calculate the threshold value and then apply the soft thresholding. After de-noising, the inverse whitening filter is applied to the recovered original signal. The difference between the noisy signal and the de-noised signal is observed in this figure.



(a) Single tone signal with frequency 400 Hz.



(b) Single tone signal with frequency 1500 Hz.



(c) LFM signal.

Fig. 9. Comparison of the RMSE obtained using DWT and the proposed S-transform de-noising methods.

Fig. 8 shows the output SNR of the proposed S-transform de-noising methods with wavelet de-noising methods for input SNR from −6 dB to 8 dB. The de-noising is successful if output SNR has the highest possible value. In all cases shown in this figure, the output SNR using S-transform de-noising methods performs better than using the wavelet de-noising methods. No significant difference is found if the pre-whitening filter is used or not between either the S-transform de-noising methods or wavelet de-noising methods. However, the difference in the output SNR of approximately 3 dB is observed between the S-transform de-noising methods and wavelet de-noising methods in case of a single-tone signal with frequency of 400 Hz, nearly 4 dB difference in the case of a single-tone signal with frequency of 1500 Hz, and around 6 dB difference in the case of LFM signal.

The RMSE values of all methods when input SNR ranges from −6 dB to 6.25 dB are compared, as shown in Fig. 9. In all cases, the output RMSE using the S-transform de-noising methods exhibits lower error compared with using the wavelet de-noising methods. Similar to the output SNR,

Table 1
Output SNR and RMSE of the various de-noising methods and the input SNR of 3 dB and input 1/RMSE of 15 dB.

Performance measure	S-transform direct method	S-transform with pre-whitening filter	Wavelet direct method	Wavelet with pre-whitening filter
		Single tone signal with frequency 400 Hz		
Output SNR (dB)	6.8	7.2	3.31	3.28
1/RMSE (dB)	16.77	17.1	15.2	15.25
		Single tone signal with frequency 1500 Hz		
Output SNR (dB)	4.9	5.1	0.6	0.71
1/RMSE (dB)	15.52	15.13	13.44	13.24
		LFM signal		
Output SNR (dB)	6.84	7.15	0.1	1.65
1/RMSE (dB)	16.51	16.65	14.4	13.82

the pre-whitening filter does not affect the performance of both de-noising methods. Between the two de-noising methods, a difference of 2 dB is observed for single-tone signals with frequencies of 400 and 1500 Hz and around 3 dB difference in the case of LFM signal. The output SNR and RMSE for the various de-noising methods for a given input SNR of 3 dB and input 1/RMSE of 15 dB are summarized in Table 1.

5. Conclusion

This study focuses on the de-noising of acoustic signals in colored noise, specifically UWAN. From the time–frequency representation generated by the S-transform, de-noising is performed using soft thresholding with universal threshold estimation. The proposed method is compared with the more conventionally used wavelet transform de-noising method. The de-noising methods are applied to simulated and real fixed-frequency signal and time-varying signal. The S-transform-based de-noising method showed better performance than the wavelet-based de-noising method in terms of the SNR and RMSE at 4 and 3 dB, respectively. The S-transform can resolve frequency components better than the wavelet transform. The S-transform is also useful to analyze unknown signals before performing de-noising to recover the signal.

Acknowledgments

The authors would like to thank the Universiti Teknologi Malaysia (UTM) and Ministry of Higher Education (MOHE) Malaysia for supporting this work.

References

[1] M. Stojanovic, J. Preisig, IEEE Commun. Mag. 47 (1) (2009) 84–89.
[2] P.H. Dahl, et al., Acoust. Today 3 (1) (2007) 23–33.
[3] R.J. Urick, Ambient Noise in the Sea, 1984 DTIC Document.
[4] S.S. Murugan, V. Natarajan, R.R. Kumar, Acoust. Australia 40 (2) (2012) 111.
[5] C.W. Therrien, Discrete Random Signals and Statistical Signal Processing, Prentice Hall PTR, 1992.
[6] X. Zhang, Y. Xiong, IEEE Signal Process. Lett. 16 (4) (2009) 295–298.
[7] R. Lukac, B. Smolka, K.N. Plataniotis, Signal Process. 87 (9) (2007) 2085–2099.
[8] J. Chen, et al., IEEE Trans. Audio Speech Lang. Process. 14 (4) (2006) 1218–1234.
[9] H. Hassanpour, A. Zehtabian, S. Sadati, Digital Signal Process. 22 (5) (2012) 786–794.
[10] Y. Xu, et al., IEEE Trans. Image Process. 3 (6) (1994) 747–758.
[11] S.M. Govindan, P. Duraisamy, X. Yuan, Digital Signal Process. 33 (2014) 180–190.
[12] P.R. Hill, et al., Signal Process. 105 (2014) 464–472.
[13] D.L. Donoho, J.M. Johnstone, Biometrika 81 (3) (1994) 425–455.
[14] G. Kalpana, V. Rajendran, S.S. Murugan, J. Mar. Eng. Technol. 13 (3) (2014) 29–35.
[15] M. Mamun, M. Al-Kadi, M. Marufuzzaman, J. Appl. Res. Technol. 11 (1) (2013) 156–160.
[16] I.M. Johnstone, B.W. Silverman, J. R. Stat. Soc. Ser. B 59 (2) (1997) 319–351.
[17] S. Roopa, S. Narasimhan, Signal Process. 105 (2014) 207–215.
[18] R.G. Stockwell, L. Mansinha, R. Lowe, IEEE Trans. Signal Process. 44 (4) (1996) 998–1001.
[19] C.R. Pinnegar, L. Mansinha, Geophysics 68 (1) (2003) 381–385.
[20] T. Nguyen, Y. Liao, Electr. Power Syst. Res. 79 (4) (2009) 569–575.
[21] S. Assous, et al., IEEE Trans. Biomed. Eng. 53 (6) (2006) 1032–1037.
[22] C.R. Pinnegar, D.W. Eaton, J. Geophys. Res. Solid Earth 108 (B9) (2003).
[23] M. Das, S. Ari, IRBM 34 (6) (2013) 362–370.
[24] S. Ventosa, et al., IEEE Trans. Signal Processing 56 (7) (2008) 2771–2780.
[25] A.V. Oppenheim, G.C. Verghese, Signals, Systems, and Inference, 2010 Class notes for, 6.
[26] Q. Li, Digital Sonar Design in Underwater Acoustics: Principles and Applications, Springer Science & Business Media, 2012.
[27] W.C. Knight, R.G. Pridham, S.M. Kay, Proc. IEEE 69 (11) (1981) 1451–1506.
[28] R.P. Hodges, Underwater Acoustics: Analysis, Design and Performance of Sonar, John Wiley & Sons, 2011.
[29] J. Panaro, et al., in: Brazilian Telecommunication Symposium, 2012.
[30] M. Chitre, J.R. Potter, S.-H. Ong, IEEE J. Oceanic Eng. 31 (2) (2006) 497–503.
[31] T. Melodia, et al., Mobile Ad Hoc Networking: Cutting Edge Directions, 2013, pp. 804–852.
[32] G. Burrowes, J.Y. Khan, Short-Range Underwater Acoustic Communication Networks, INTECH Open Access Publisher, 2011.
[33] M. Chitre, J. Potter, O.S. Heng, in: OCEANS'04. MTTS/IEEE TECHNO-OCEAN'04, IEEE, 2004.
[34] M.L.A.W. Zhao, Math. Probl. Eng. 2012 (2012) Article ID 673648.
[35] A.Z. Sha'ameri, Y. Al-Aboosi, N.H.H. Khamis, in: 2014 International Conference on Computer and Communication Engineering (ICCCE), IEEE, 2014.
[36] H.R. Gupta, R. Mehra, Int. J. Sci. Res. Eng. Technol. 2 (6) (2013) 389–392.
[37] J.P.C. Da Costa, et al., Signal Process. 93 (11) (2013) 3209–3226.
[38] B. Widrow, E. Walach, Adaptive Inverse Control: A Signal Processing Approach, Reissue ed., John Wiley & Sons, Inc, 2008.
[39] I. McLoughlin, Applied Speech and Audio Processing: with Matlab Examples, Cambridge University Press, 2009.

[40] S.G. Mallat, IEEE Trans. Pattern Anal. Mach. Intell. 11 (7) (1989) 674–693.
[41] S.-C. Pei, et al., Signal Process. 91 (6) (2011) 1466–1475.
[42] C.R. Pinnegar, D.W. Eaton, J. Geophys. Res. Solid Earth 108 (B9) (2003) (1978–2012).
[43] M. Schimmel, J. Gallart, IEEE Trans. Signal Process. 53 (11) (2005) 4417–4422.
[44] J. Baili, et al., NDT & E Int. 42 (8) (2009) 696–703.
[45] H. Olkkonen, Discrete Wavelet Transforms: Algorithms and Applications, InTech, 2011 August.
[46] D.L. Donoho, IEEE Trans. Inf. Theory 41 (3) (1995) 613–627.
[47] R. Aggarwal, et al., Int. J. Comput. Appl. 20 (5) (2011) 14–19.
[48] G. Wei, C. Yang, F.-J. Chen, Signal Process. 91 (4) (2011) 841–851.
[49] A.E. Villanueva-Luna, et al., De-Noising Audio Signals Using MATLAB Wavelets Toolbox, INTECH Open Access Publisher, 2011.