



Article

A Two-Step Adaptive Noise Cancellation System for Dental-Drill Noise Reduction

Jitin Khemwong^a and Nisachon Tangsangiumvisai^{b,*}

Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand

E-mail: ^aJitin.14856@gmail.com, ^bNisachon.T@chula.ac.th (Corresponding author)

Abstract. This paper introduces a two-step dental-drill Noise Reduction (NR) technique based upon the Adaptive Noise Cancellation (ANC) system. The proposed technique is particularly designed for the NR headphone, which the patients should be wearing while having their dental treatment. Similarly, the dentists are also suggested to wear these NR headphones to prevent hearing impairment due to excessively high level of drill noise. In the first step of the proposed NR technique, a tone-frequency extraction algorithm is proposed to estimate the main spectral component of the dental-drill noise. A sinusoidal signal with the estimated tone frequency is generated and subsequently fed to the first adaptive filter of the ANC system to remove the main tone frequency from the dental-drill noise. In the second-step, another adaptive filter of the ANC system is then employed to eliminate the residual high-frequency components of the dental-drill noise. Computer simulations based on recorded dental-drill sounds and real speech signals demonstrate the effectiveness of the proposed two-step dental-drill NR (TSDNR) technique, both in terms of the noise attenuation performance and the speech quality of the enhanced speech signal. Moreover, results of a subjective listening test with 15 listeners are also given to guarantee satisfied speech quality of the enhanced speech signal employing the proposed TSDNR technique.

Keywords: Noise reduction, dental-drill sounds, tone-frequency extraction, adaptive noise cancellation, speech quality.

ENGINEERING JOURNAL Volume 22 Issue 4

Received 4 November 2017

Accepted 31 March 2018

Published 31 July 2018

Online at <http://www.engj.org/>

DOI:10.4186/ej.2018.22.4.51

1. Introduction

Generally, we should go to see the dentists every six months in order to have dental check-up. However, some people avoid going to see the dentist regularly because they feel scared to the dental-drill sounds [1]. In fact, the dental-drill sounds also affect hearing systems of the dentists who have been exposed to these high-pitched sounds throughout his career. Most high-speed dental drills normally contain spectral components lying in the range between 2 to 14 kHz [2]. The high-frequency components of the dental-drill sounds are due to the use of high-speed electromotor of the drill equipment. Besides, the sound pressure level (SPL) of dental-drill sounds in a dental room can reach up to 80 dBA [3] for overall spectrum within that range. Thus, the dentists can suffer a temporary hearing loss if they work continuously without any hearing-protection devices [4]. Furthermore, permanent hearing loss of the dentists can start to develop after they have been working for more than five years [5]. Therefore, if the SPL of the dental-drill sounds can be significantly reduced, the risk of hearing loss in the dentists can be minimized. In addition, the patients will be able to go for regular dental check-ups.

Several noise reduction (NR) techniques have been developed to reduce the effect of undesired sounds in working environment, such as factory noise, train noise, street noise, engine noise, etc. However, only a few NR techniques are proposed for dental-drill noise removal. Normally, the NR techniques are classified by their functions into two groups; Passive Noise Control (PNC) and Active Noise Control (AcNC) [6]-[9]. The PNC reduces the unwanted noise in a physical way, i.e. by using noise-isolating materials, such as ear muffs, ear plugs, etc. This type of NR techniques can reduce the medium-frequency and high-frequency components of the undesired noise. However, the PNC type can reduce the SPL up to about 20-30 dB and cause discomfort when wearing the devices for a long time. On the other hand, the AcNC approaches consume electrical energy and offer noise reduction at low frequencies [7], [8]. The noise-cancelling headphones with AcNC use an external microphone to detect the noise. An inverted-phase sound is then generated to remove the noise [9]. One example of this type of NR techniques is known as Adaptive Noise Cancellation (ANC) system, which usually employs an adaptive filter to estimate the unknown ambient noise [10]. This estimated noise signal is subtracted from the noisy speech signal afterwards.

From the spectral characteristics of the dental-drill noise, there are a few high spectral peaks in the low-frequency region and mask nearly all of the speech spectral components, while lower spectral components of the dental-drill noise are spreading in the high-frequency region, especially beyond the speech frequency range. A noise-cancelling headphone has been proposed to remove the dental-drill noise from the desired speech signal of the dentist [11]. A combination between PNC and AcNC is used for dental-drill noise reduction. It is shown in [11] that PNC attenuates the drill noise particularly in the mid-frequency and high-frequency range at least 20 dB. However, the effect of peaks in the low-frequency region is still audible by the patient who wears this noise-cancelling headphone. As for the low-frequency drill-noise reduction, the AcNC is used by applying the ANC system to further reduce the peak noises by additional 10 dB [11]. Nevertheless, the main drawback of this dental-drill noise-cancelling technique is that the PNC part causes discomfort to the wearers and also attenuates the target speech signal. As a result, this prevents effective communications between the dentist and the patient.

Therefore, an efficient dental-drill NR technique is proposed in this manuscript for noise-cancelling headphones. The proposed NR technique aims to alleviate the peaks and the high-frequency dental-drill noise without the use of PNC, while preserving the speech quality of the dentists. By employing the ANC system in two steps, thus, it will be referred to as the Two-Step Dental-Drill Noise Reduction (TSDNR) technique. First, the fundamental sinusoidal frequency of the dental-drill sounds is estimated and is removed by the use of the first adaptive filter of the single-microphone ANC system. In the second step, the residual high-frequency components of the dental-drill sounds are significantly removed via another adaptive filter of the ANC system. The proposed TSDNR technique is particularly designed for the dental-drill NR headphone without the use of PNC and does not need any Voice Activity Detector (VAD) to distinguish between speech and noise frames, while several conventional NR techniques do. Instead, the frequency characteristics of the dental-drill sounds are analysed and utilised to obtain the information about the dental-drill noise. Hence, the noise attenuation performance of the proposed TSDNR technique is independent to the accuracy of VADs. Based upon computer simulations, the noise attenuation performance and the distortion of the enhanced speech signal of the proposed TSDNR technique are investigated and compared with the conventional two-microphone ANC system. Furthermore, results of a

subjective listening test with 15 listeners are given to guarantee satisfied speech quality of the enhanced speech signal employing the proposed TSDNR technique.

In Section 2, some favourite existing NR techniques are summarized; including their advantages and disadvantages. The spectral characteristics of the dental-drill sounds are investigated in Section 3. The proposed TSDNR is presented in Section 4. Simulation results based on recorded dental-drill sounds and speech signals are given in Section 5, followed by the conclusions in Section 6.

2. Related Work

One of the most popular NR techniques for voice communication systems is the Spectral Subtraction (SS) method [12]. This method is very simple, spectral efficient, and employs only one microphone in the system as a sensor for the noisy input signal. An assumption that the desired speech and the additive noise signals being uncorrelated to each other is necessary. If the spectrum of the background noise can be estimated and is consequently subtracted from the noisy speech spectrum, the enhanced speech spectrum can therefore be obtained. It is clear that the quality of the enhanced speech signal by employing the SS method depends on the accuracy of the estimated background noise spectrum. In addition, the residual noise spectral components at random frequencies result in an artefact, which is known as the *musical noise* [13], and affects the quality of the enhanced speech signal. Various NR techniques have been proposed in order to reduce the musical noise effect. One SS-based method employs an adaptive gain function, which is obtained from averaging speech sub-spectrum ratio, instead of using direct subtraction [14]. Due to the noise characteristics that affect the speech spectrum differently along the frequency regions, a multi-band SS method is introduced [15]. In [16], the noise over-subtraction is proposed with spectral floor setting, however, this technique results in a trade-off between the residual noise, including the musical noise, and speech distortion. Instead of a single over-subtraction parameter, a non-linear SS method with multiple over-subtraction parameters is proposed for different interfering noises at each frequency region [17]. Alternatively, the *a priori* Signal-to-Noise Ratio (SNR) estimation is applied to various NR techniques without causing the musical noise effect [18], [19]. These NR techniques, however, require the VADs for the process of noise spectrum estimate. In fact, the performance of these NR techniques depends on the accuracy of the VADs. Based on perceptually motivated cost function, a Gaussian statistical distribution in Bayesian framework is applied for speech enhancement [20]. However, this technique introduces speech distortion for low input SNR level. Although, a number of noise power spectrum estimation techniques have been proposed without the use of any VADs, however, accurate estimation of the noise spectrum is not yet obtained, particularly when the noise signal is rapidly changing with time [21]-[25].

An alternative approach for reducing the additive noise signal in voice communication systems is the ANC system, which employs an adaptive filter. The coefficients of the adaptive filter are adapted according to the error signal minimization. The performance of the ANC system is specified by the choice of the adaptive filtering algorithm. In contrast to the SS method, the ANC-based NR technique does *not* need any VAD to distinguish between speech and noise frames. Furthermore, the ANC system does not lead to the musical noise effect. Normally, the ANC system requires the use of two microphones. The first microphone signal contains the noisy speech signal and is known as the *primary* signal. The second microphone, on the other hand, is assumed to be located very close to the noise source and far away from the speech source so that it picks up mostly the additive noise signal, but not the desired speech signal, and is referred to as the *reference* signal. In fact, it is impossible to place the second microphone to pick up only the additive noise signal, without being contaminating by the desired speech signal.

The block diagram of two-microphone ANC system is illustrated in Fig. 1 [10]. The primary signal, $x(n)$, contains the desired speech signal, $s(n)$, and the noise signal, $d_1(n)$, which are assumed to be uncorrelated to each other. This additive noise signal, $d_1(n)$, travels through the signal path between the noise source and the first microphone, which is represented by a Finite Impulse Response (FIR) filter, $h_{21}(n)$. Similarly, the reference signal contains the noise signal, $d(n)$, and the additive speech signal which travels through another signal path between the desired source and the second microphone, representing by $h_{12}(n)$. Ideally, $h_{12}(n)$ should be zero. Thus, the adaptive filter, $w(n)$, attempts to estimate $h_{21}(n)$, resulting in the output of the adaptive filter approaching $\hat{d}_1(n)$. The error signal, $e(n)$, is used to control the adaptive filter, $w(n)$, by the choice of a suitable adaptive filtering algorithm. Based upon the Normalized Least Mean Square (NLMS) algorithm [10], the computational complexity of the two-

microphone ANC system depends on the order of the adaptive filter, $w(n)$.

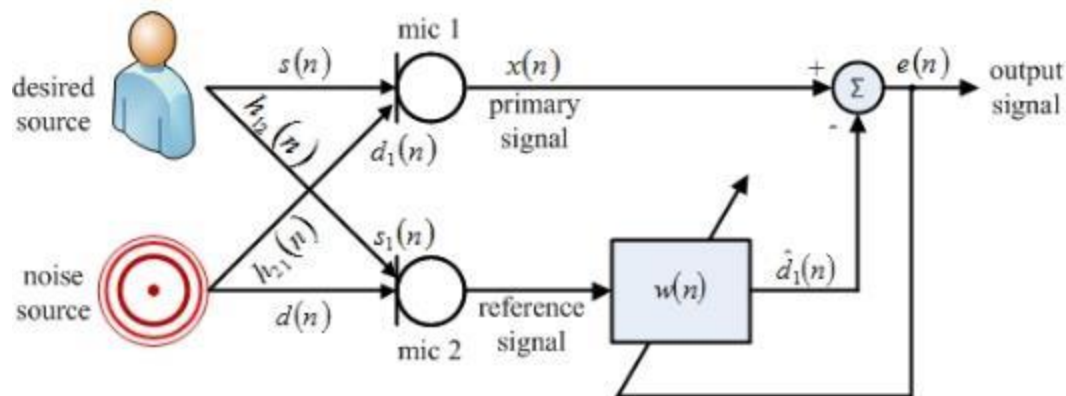


Fig. 1. The block diagram of the two-microphone ANC system. (mic1 denotes the first microphone, mic2 denotes the second microphone).

3. Spectral Characteristics of the Dental-Drill Sounds

To understand the dental-drill sounds, which is crucial to designing a noise reduction algorithm, its temporal and spectral characteristics, as well as noise levels over frequency range are studied in this Section. The magnitude spectra over consecutive frames of the recorded dental-drill sound are shown in Fig. 2. It is noted that the magnitude spectrum in Fig. 2(a) represents typical dental-drill sound when the dental equipment is turned on but with no dental treatment [26], whereas those in Fig. 2(b) and Fig. 2(c) indicate typical dental-drill sounds under dental treatment conditions [27], [28].

Considering Fig. 2(a), it is evident that the recorded dental-drill sound in [26] has dominant peak magnitude at a single frequency around 8 kHz. From Fig. 2(b) and (c), the dental-drill spectra from [27], [28], there are not only dominant peak magnitudes, but these peak frequencies also change over time, i.e. from around 7-8 kHz for the fundamental frequency component. In addition, there also exist their corresponding peak magnitudes at their harmonic frequency around 14-16 kHz. Note that, the frequency drift of the dental-drill sound under dental treatment can be caused by either the change in the motor speed (which is controlled by the dentist) or the vibrational resonant frequency of the dental drill when it is in contact with teeth. In addition to the dominant peak magnitudes and their harmonics, the dental-drill spectra in Fig. 2(a)-2(c) exhibit wide-band characteristics distributed over the audible frequency band, however, at lower spectral magnitudes. Fig. 3(a)-3(c) are spectrogram plots of the corresponding dental-drill noise spectra in Fig. 2, which provide more details of the temporal characteristics, particularly the change of the dominant tone-frequency over time.

From the above discussion, it can be deduced that the dental-drill sounds are of non-stationary type, where most of its power is concentrated at the dominant peak or tone frequency, while the rest of the noise power spreads over the entire audible frequency band. These spectral characteristics will be employed in the designing process of the proposed efficient dental-drill NR technique in the next Section.

For typical speech spectral characteristics, it is known that the spectra of vowel (or voiced) sounds in low frequency band up to 1 kHz carry most of the speech power, whereas the spectra of consonant (or unvoiced) sounds in high-frequency band from 1 kHz up to 5-6 kHz provides speech intelligibility. By comparing the dental-drill spectrum plots with a speech spectrum of the same sampling frequency in Fig. 4, it is clearly seen that most of the dental-drill noise spectral components during the speech bandwidth; i.e. 0-5 kHz, the power of their magnitude spectrum is typically much lower than that of vowel sounds under normal sound utterance. Hence, the dental-drill noises are totally masked by the speech spectral components during the speech bandwidth.

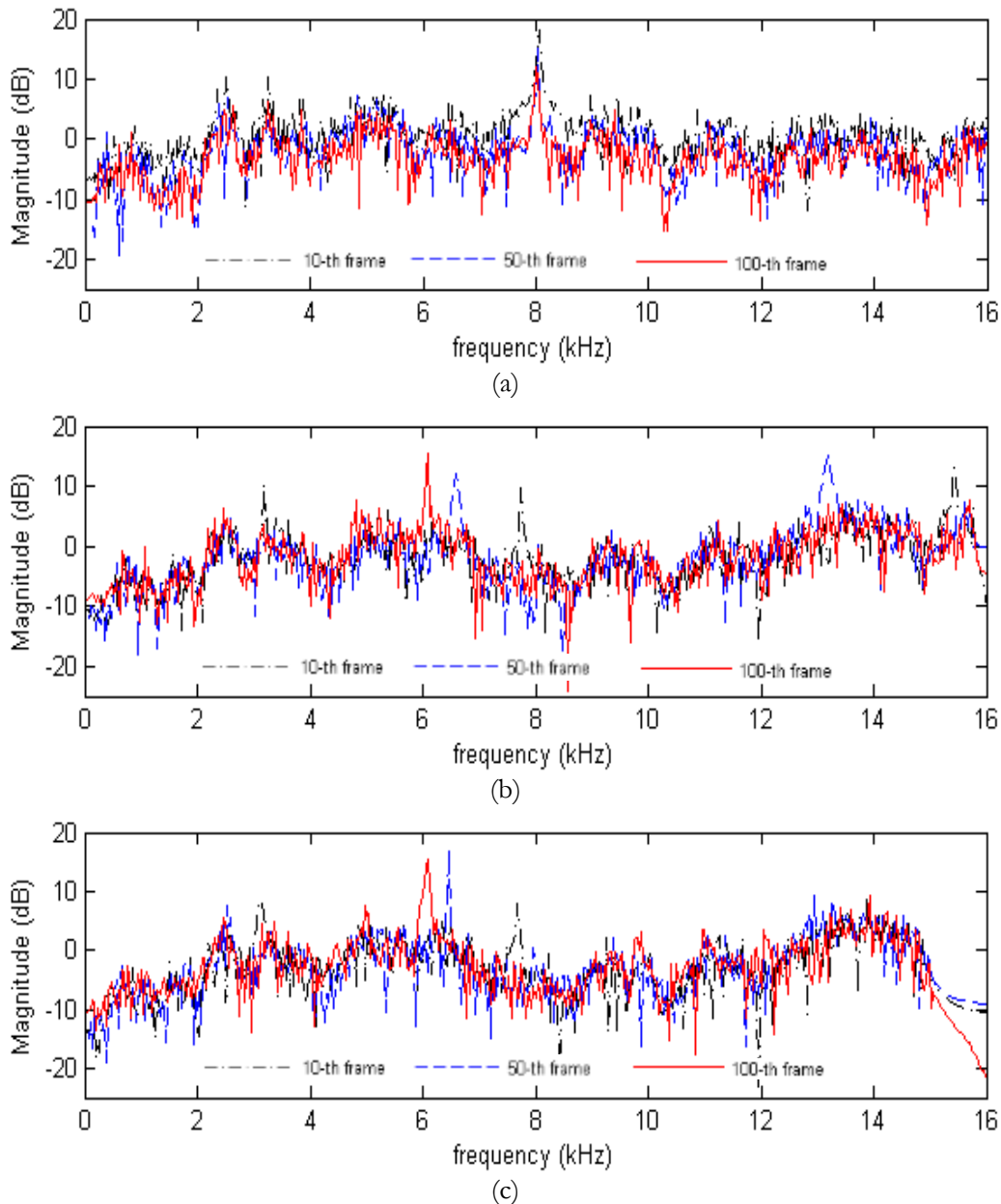


Fig. 2. Overlay plots of the magnitude spectrum of (a) dental-drill sound from [26], (b) dental-drill sound from [27], and (c) dental-drill sound from [28]. (Sampling rate of 32 kHz).

On the other hand, during the frequency range beyond the speech bandwidth; i.e. 5-16 kHz, the peak magnitude spectrum of dental-drill noises is located within the frequency range of consonant sounds. In fact, its peak power is much higher than the speech spectral ones. Thus, in this frequency region, the speech spectrum is negligible and only the dental-drill noise spectrum is present.

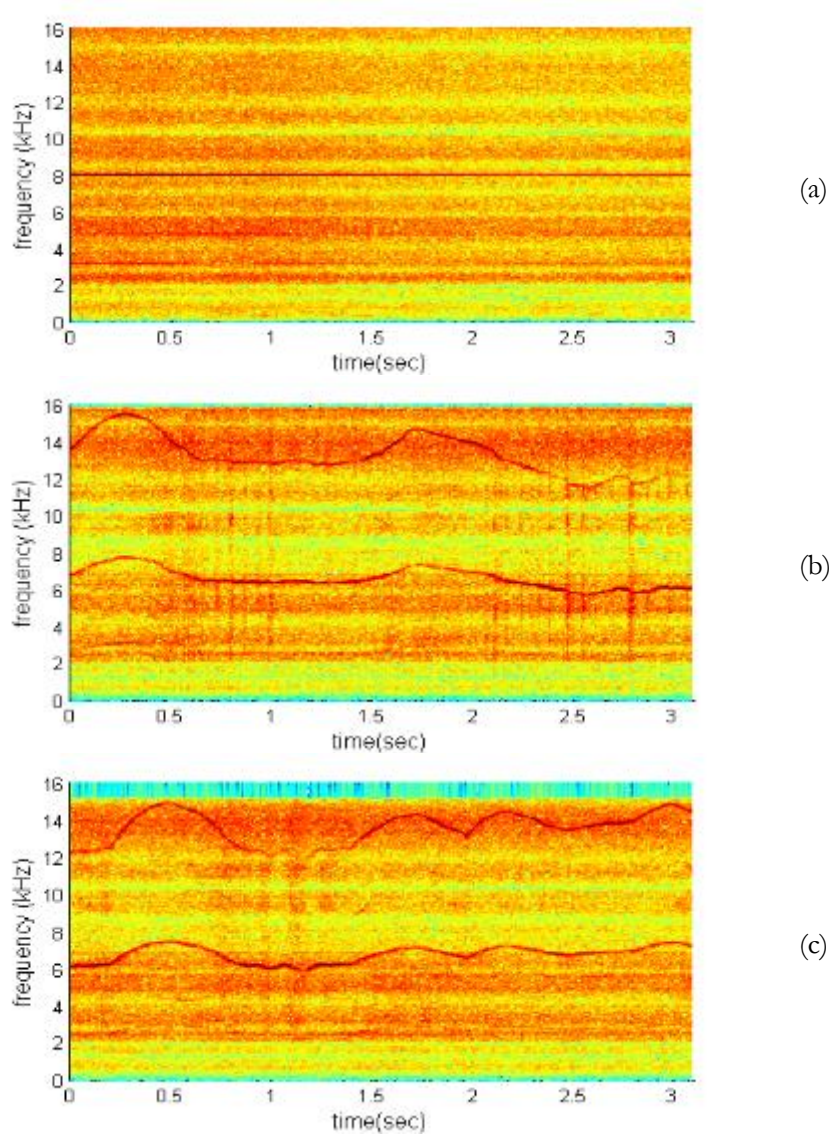


Fig. 3. Spectrogram plot of dental-drill sounds: (a) fixed tone-frequency from [26], (b) and (c) time-varying tone-frequency from [27] and [28], respectively. (Sampling rate of 32 kHz).

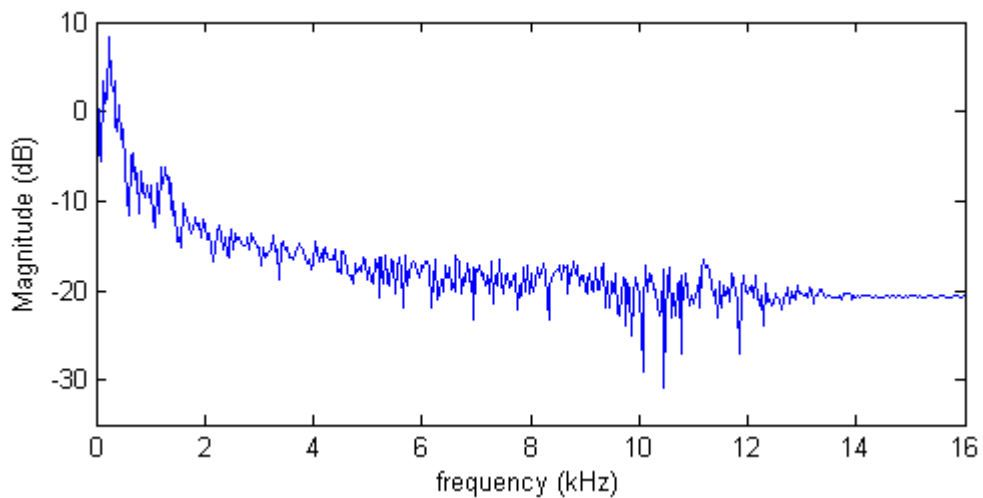


Fig. 4. The magnitude spectrum of a speech signal in [29].

4. The Proposed Two-Step ANC System for Dentist-Drill Noise Reduction Technique

In this work, it is supposed that the patient wears the noise-cancelling headphone that employs the investigated NR techniques on both sides of the earphones during dental treatment, as illustrated in Fig. 5. Hence, all the explanation is given for one side of the earphone and it will also be applied to the other side. In fact, similar explanation serves for the case when the dentist also wears the noise-cancelling headphone with the investigated NR techniques too. Based upon the operation of the two-microphone ANC system, the first microphone is located at the inner side of the loudspeaker of the earphone, enabling the detection of the noisy speech signal. The second microphone should be attached to the outer side of the loudspeaker or somewhere close to the drilling device in order to detect mainly the undesired dental-drill sound. Consequently, the dental-drill noise entering the ears will be removed from the noisy speech signal, albeit under the assumption that the reference signal contains only the noise. However, such an ANC system, suffers from the fact that the required assumption is entirely *impractical*. This is because the noise source, namely the dentist-drilling tool, is always close to the patient's ears during dental treatment, thereby making it difficult to allocate the reference microphone to pick up mainly the dental-drill noise. As a consequence, the noise suppression based upon the conventional two-microphone ANC system results in significant degradation of the enhanced speech quality.

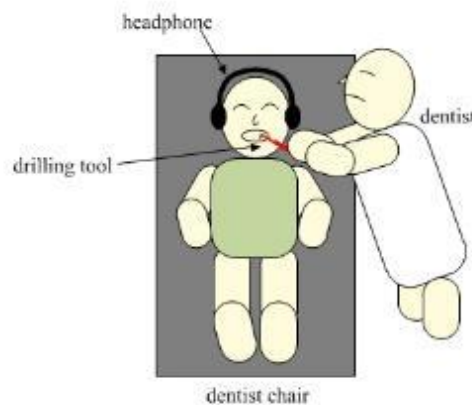


Fig. 5. Layout of the dental-drill NR headphone.

4.1. The single-microphone Two-Step ANC system: System and Operation

As illustrated by the block diagram in Fig. 6, the proposed TSDNR technique employs the “**single-microphone**” ANC system in two steps. As compared to the conventional two-microphone ANC system in Fig. 2, the proposed TSDNR system makes use of two parallel adaptive filters, $w_1(n)$ and $w_2(n)$, and only the *primary* microphone is used to detect the noisy speech signal. With **no** use of the second microphone, the reference signal is extracted from the noisy speech signal in order to obtain different spectral parts of the dental-drill noise (as will be explained later in this sub-section), and these are subsequently fed to the parallel adaptive filters for the first-step and the second-step noise reduction process, hence the so-called “**single-microphone two-step ANC system**”. The use of “**single-microphone**” ANC system offers improved feasibility and efficient implementation of the proposed TSDNR technique.

Due to the fact that the peak frequency magnitude of the dental-drill noise contains most of the power, and that the low-frequency spectrum of the dental-drill noise can be masked by vowel spectrum during speech activity, the “**first-step**” of the proposed TSDNR technique is set to mainly remove the peak tone-frequency from the primary signal. From the block diagram of the proposed TSDNR technique, as shown in Fig. 6, this is accomplished by extracting the peak tone-frequency from the noisy speech spectrum, which is detected by the primary microphone, and uses it as the reference signal for the first adaptive filter, $w_1(n)$. It is important to note that, for this first step of dental-drill noise reduction, the removal of the dental-drill noise within the consonant spectral band is avoided in order to maintain speech intelligibility.

This is justified by subjective test results performed in Section 5, where the residual noise showed minimum impact on the speech quality.

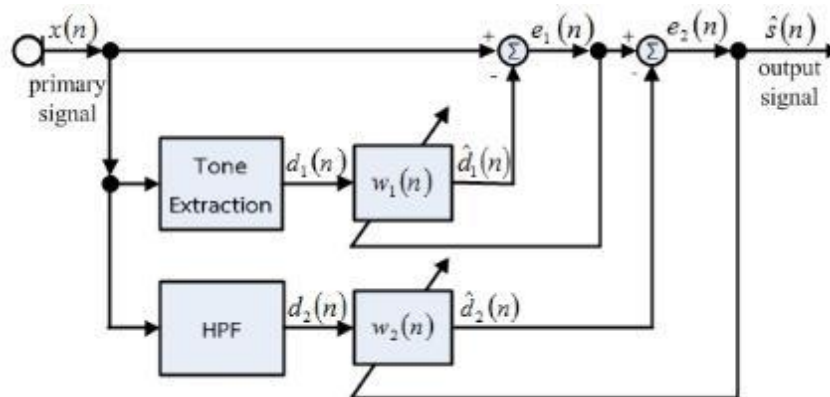


Fig. 6. The block diagram of the proposed TSDNR technique.

The **“second-step”** of the proposed TSDNR technique is set to eliminate the dental-drill spectral components beyond the speech frequency band, including the harmonic of the peak tone-frequency. This is achieved by employing the second adaptive filter, $w_2(n)$, where the reference signal to $w_2(n)$ is extracted from the noisy speech signal by the high-pass filter, HPF, as shown in Fig. 6. For a given filter order, N_b , the cut-off frequency (F_c) of the HPF must be sufficiently low to enable large reduction of high-frequency dental-drill noise spectral components, but sufficiently high to avoid interfering with speech spectrum, particularly during the consonant spectral band, which may result in speech degradation. It is noted that although the filter order of the HPF, N_b , should be kept small for low implementation complexity, a lower-order HPF exhibits a wider transition band, resulting in a higher level of speech degradation. Thus, there exist a trade-off among noise reduction, speech quality, and implementation complexity. From the aforementioned discussion, it can also be deduced that the use of the first-step of the proposed TSDNR technique considerably helps relax the trade-off by removing the peak tone-frequency of the dental-drill noise located near the consonant spectrum, thereby removing it from the performance trade-off space.

In summary, the peak tone-frequency and the residual high-frequency dental-drill noises are therefore removed from the primary signal, $x(n)$, by employing the first-step and the second-step of the proposed TSDNR technique, respectively. The enhanced speech signal, $\hat{s}(n)$, is therefore obtained at its output. In the next subsection, the proposed tone-frequency extraction algorithm is explained in details.

4.2. Tone-Frequency Extraction Algorithm

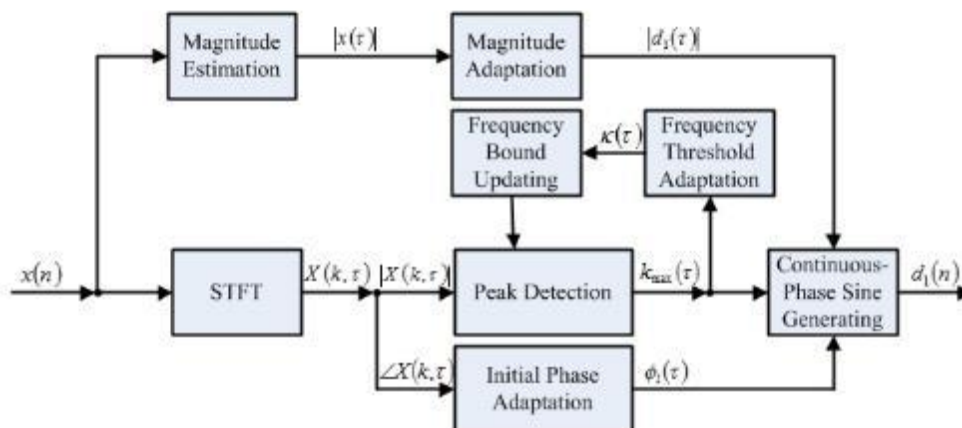


Fig. 7. The block diagram of the tone-frequency extraction algorithm.

In the first-step of the proposed TSDNR technique, the dominant sinusoidal frequency of the dental-drill noise is estimated. The “tone extraction” module, whose block diagram is illustrated in Fig. 7, is the main part of this first step. The noisy signal is partitioned into 30-*ms* frames, $x(\tau)$, which are analyzed via the Short-Time Fourier Transform (STFT) in order to detect for its peak, as given by

$$X(k, \tau) = \sum_{i=0}^{L-1} [x(\tau L + i) - \overline{x(\tau)}] \cdot e^{-j(2\pi/L)ik} \quad (1)$$

where k represents the frequency bin index, τ stands for the time frame index, L denote the length of framed signals, and $\overline{x(\tau)}$ is the DC component of τ -th frame signal.

The dominant tone-frequency of the dental-drill noise is then estimated by detecting its peak, as shown by the “peak detection” module. Once the peak is detected, the frequency bin index of each τ -th frame peak, $k_{max}(\tau)$, is therefore identified. For the “frequency threshold adaptation” module, the threshold of the frequency bin index, $\kappa(\tau)$, is adapted at each frame index by using $k_{max}(\tau)$ as follows.

$$\kappa(\tau) = \lfloor \gamma \cdot \kappa(\tau - 1) + (1 - \gamma) \cdot k_{max}(\tau) \rfloor \quad (2)$$

The parameter γ denotes the weighted factor and the $\lfloor \cdot \rfloor$ represents the floor operator. Then, in the “frequency bound updating” module, the obtained $\kappa(\tau)$ is used to update bounds for ensuring whether the detected frequency bin index of peak in the next frame is the frequency bin index of the dental-drill sounds, **not** the speech signal, i.e., $(\tau) - \delta \leq k_{max}(\tau + 1) \leq \kappa(\tau) + \delta$, whereas δ is a small positive value chosen appropriately from the observation of noise variation under the assumption that the dental-drill noise gradually change with time. Next, the concatenation of frame signals being generated must be smoothed by using phase-concatenating equation provided in the “initial phase adaptation” module, as given below.

$$\phi_i(\tau) = [k_{max}(\tau) - k_{max}(\tau - 1)] \cdot (\tau - 1) \cdot L + \phi_i(\tau - 1) \quad (3)$$

Finally, the magnitude of the sinusoidal signal being generated for the τ -th frame, $|d_1(\tau)|$, needs to be approximated and updated with smooth adaptation to avoid discontinuities from signal concatenations. The “magnitude estimation” module estimates $|d_1(\tau)|$ with the mean of the peak magnitude of framed noisy signal, $|x(\tau)|$. The estimated magnitude is also smoothed by the “magnitude adaptation” module via the smoothing magnitude equation, as given by

$$|d_1(\tau)| = \lambda \cdot |d_1(\tau - 1)| + (1 - \lambda) \cdot |x(\tau)| \quad (4)$$

where λ represents the smoothing magnitude factor. Due to the fact that the tone-frequency is slightly shifting with time. It is therefore necessary to generate the sinusoidal signal with continuous phase, otherwise there will be spectral leakage. This is obtained by employing the “continuous-phase sine generating” module. The signal at the τ -th frame can then be achieved:

$$d_1(\tau) = |d_1(\tau)| \cdot \sin(2\pi f_{max}(\tau)t + \phi_i(\tau)) \quad (5)$$

where $t = 0, \frac{1}{f_s}, \frac{2}{f_s}, \dots, \frac{(L-1)}{f_s}$, whereas f_s is the sampling frequency of the investigated signal. Then, the noise signal, $d_1(n)$, is estimated by concatenating of generating frame signals, $d_1(\tau)$.

In practice, once the frequency component of this pure tone at about 8 kHz is detected, we can therefore generate the sinusoidal signal with the same frequency of the main spectral component of the dental-drill noise. Consequently, this estimated sinusoidal signal is used as the reference signal of the first adaptive filter of the proposed TSDNR technique. As a result, the main tone frequency of the dental-drill noise, can be removed from the primary signal.

4.3. Computational Complexity of the Proposed TSDNR Technique

Finally, the computational complexity of the proposed TSDNR technique is investigated. By considering the case of the two-microphone ANC system, the adaptive filter, $w_1(n)$, needs to identify the unknown signal path, $h_{21}(n)$, between the dental-drill equipment and the first microphone (mic1). The order, N_1 , of the adaptive filter, $w_1(n)$, depends on the length of the impulse response of the signal path, $h_{21}(n)$. Based on the FIR structure, the length of the adaptive filter, $w_1(n)$, can generally be up to thousands of filter coefficients in order to model the signal path, $h_{21}(n)$ [30], [31]. This results in high computational complexity and slow convergence of the adaptive filter. Moreover, since the dental-drill equipment and the dentist's speech are very close to each other, this makes the reference signal contain a lot of speech signal, yielding the conventional two-microphone ANC system ineffective. Thus, a high-pass filter is also necessary to remove the residual dental-drill noise afterwards.

On the other hand, the proposed TSDNR technique detects the peak tone-frequency of the dental-drill noise in its first-step. Hence, a sinusoidal signal with this detected frequency is then generated and used as the reference signal of the adaptive filter, $w_1(n)$. The second microphone signal, or the reference signal is not required. The function of the adaptive filter, $w_1(n)$, is to handle with the magnitude and phase of this estimated sinusoidal signal. Thus, the order, N_1 , of the adaptive filter, $w_1(n)$, of the proposed TSDNR technique can be much shorter than that of the conventional two-microphone ANC system. Furthermore, a high-pass and another adaptive filters, $w_2(n)$, should also be used to remove the residual dental-drill noise beyond the speech frequency band. In particular, the use of single-microphone ANC system of the proposed TSDNR technique avoids the need of the second microphone and reduces the implementation cost.

5. Simulation Results

In this research work, the proposed TSDNR technique was investigated with 30 noisy speech signals, where 10 speech signals from IEEE Corpus database [29] and 3 recorded dental-drill noises were used [26]-[28]. All investigated signals were having the sampling rate of 32 kHz with various input Signal-to-Noise Ratio (SNR) levels at the primary microphone as -10, -5, 0, 5 and 10 dB. The orders of the adaptive filters in both steps were selected as 127. The cut-off frequency of the high-pass filter of order $N_b = 128$ in the second step of the proposed TSDNR technique was chosen to be $F_c = 10$ kHz so as not to affect the speech spectral components at low frequencies, thus, the speech distortion is minimized. The proposed TSDNR technique was investigated and compared with the conventional two-microphone ANC system in realistic scenario.

A number of performance indices were observed. The noise attenuation performance of all the investigated NR techniques was given by Segmental SNR Improvement ($\Delta SegSNR$), which was defined by the difference between the output $SegSNR$ and the input $SegSNR$, i.e.

$$SegSNR (dB) = \frac{10}{M} \cdot \sum_{m=0}^{M-1} \log_{10} \left(\frac{\sum_{n=Lm}^{Lm+L-1} s^2(n)}{\sum_{n=Lm}^{Lm+L-1} (s(n) - \hat{s}(n))^2} \right) \quad (6)$$

Moreover, the Log Spectral Distance (LSD) [32] between the enhanced signal and the original speech signal was used to measure the speech distortion, as given by

$$LSD (dB) = \frac{1}{M} \cdot \sum_{m=1}^M \sqrt{\frac{1}{\frac{K}{2} + 1} \sum_{k=0}^{K/2} \left(10 \cdot \log_{10} \left(\frac{|S(k, m)|}{|\hat{S}(k, m)|} \right) \right)^2} \quad (7)$$

where the parameter M represents the number of voice frames in the time domain and K represents the number of frequency bins.

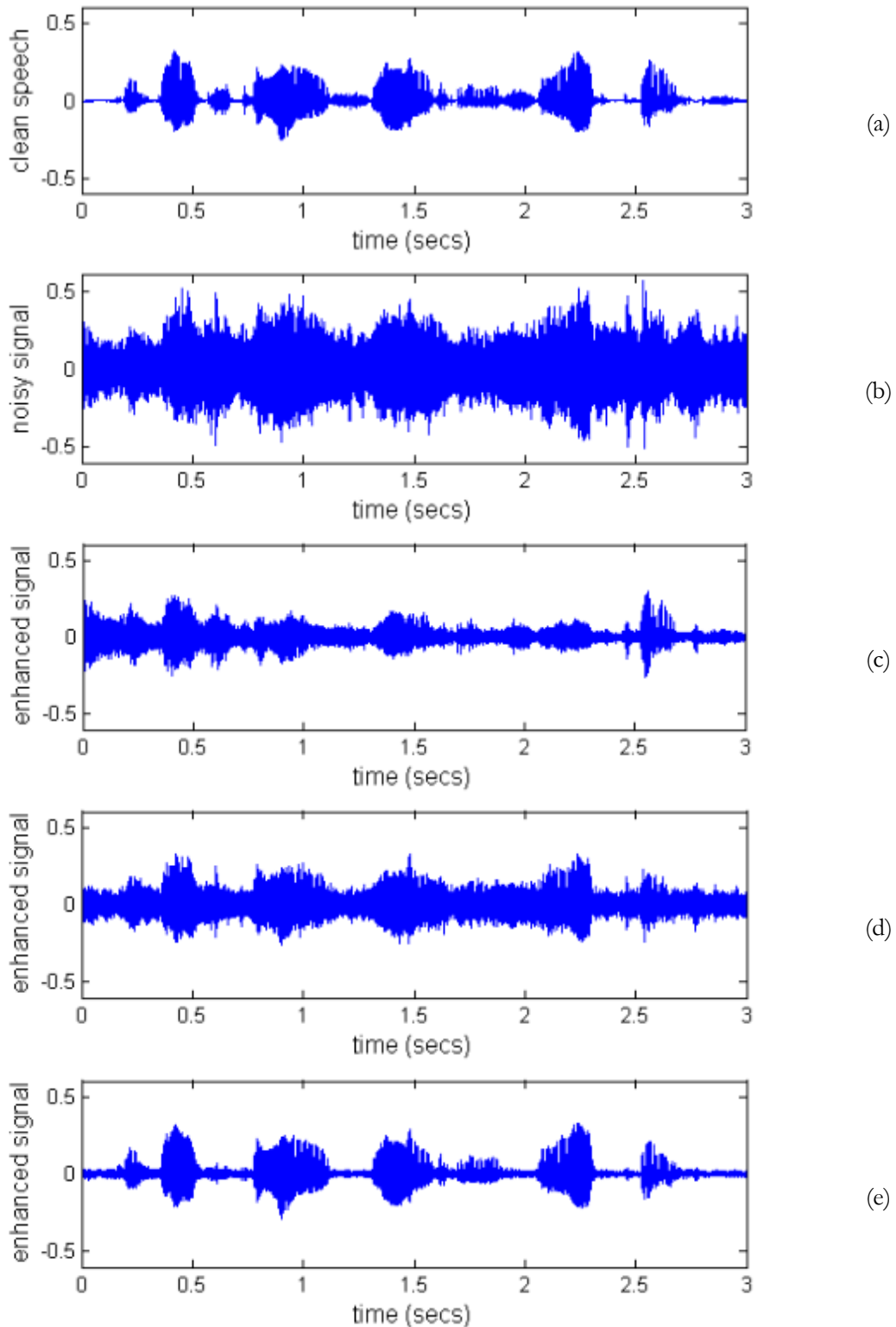


Fig. 8. Waveform plots of (a) clean speech signal, (b) noisy speech signal (input SNR of -5 dB), (c) the enhanced speech signal by using the conventional two-microphone ANC system, (d), (e) the enhanced speech signals by using the first-step and the second-step of the proposed TSDNR technique.

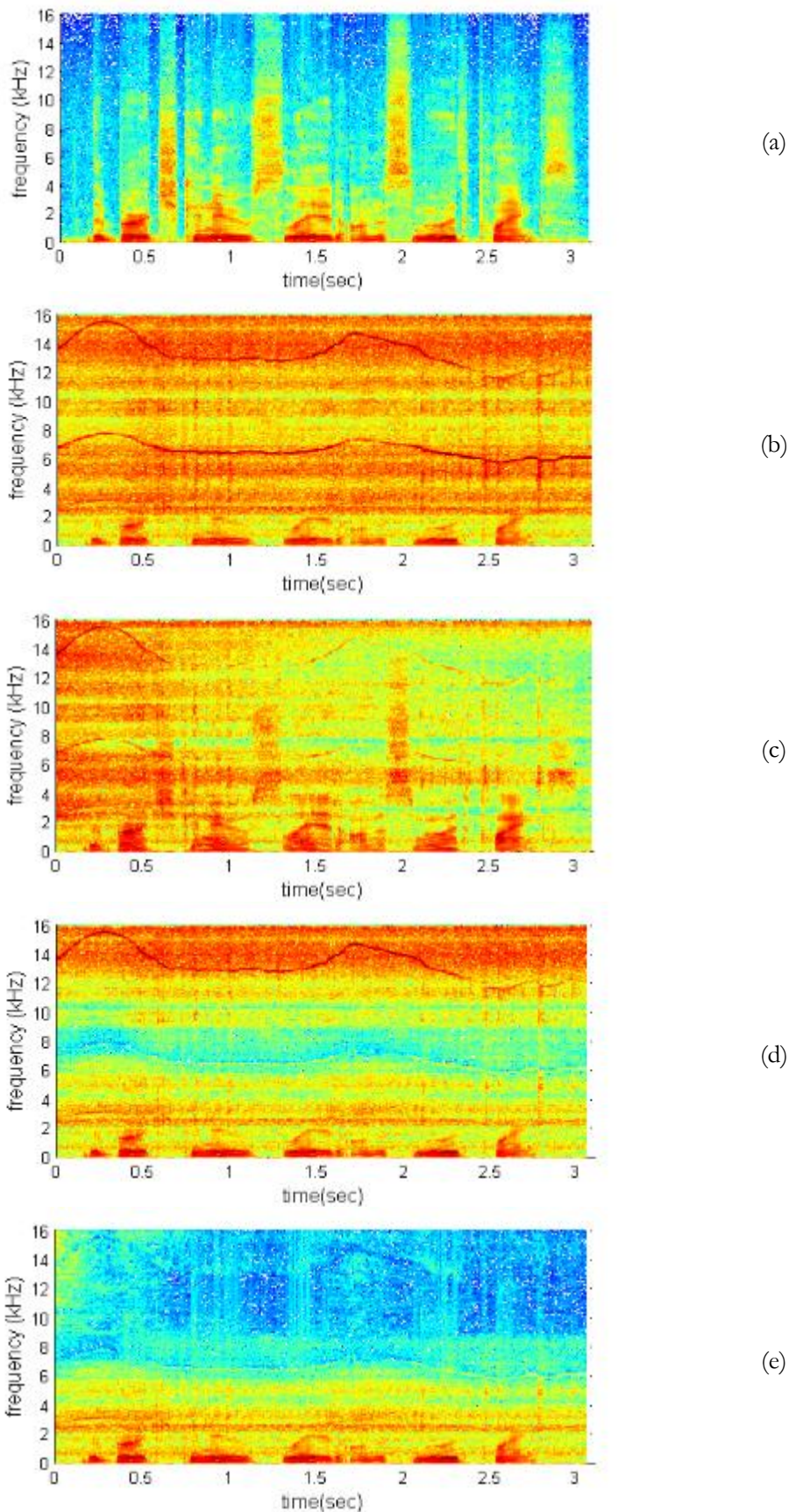


Fig. 9. Spectrogram plots of (a) clean speech signal, (b) noisy speech signal (input SNR of -5 dB), (c) the enhanced speech signal by using the conventional two-microphone ANC system, (d), (e) the enhanced speech signals by using the first-step and the second-step of the proposed TSDNR technique.

By considering at one particular case of a noisy speech signal when the input SNR was -5 dB, the clean speech and noisy speech signals are shown in Fig. 8(a) and Fig. 8(b), respectively. It is obviously shown in Fig. 8(d) and Fig. 8(e) that a large amount of the additive dental-drill noise is significantly removed by using the proposed TSDNR technique. On the contrary, as illustrated in Fig. 8(c), an inferior performance is obtained by using the conventional two-microphone ANC system, where the dental-drill noise spectrum cannot be removed effectively.

These comparisons was also observed via the spectrogram plots in order to illustrate in the frequency domain that the spectral components of dental-drill noise were drastically eliminated, by using the first-step and the second-step of the proposed TSDNR technique, as given in Fig. 9. Furthermore, the averaged magnitude spectrum plots of clean, noisy, and enhanced speech signals over 50 consecutive frames are also given in Fig. 10 to guarantee the noise attenuation performance of the proposed TSDNR technique.

It is clearly presented by observing at the averaged magnitude spectrum plots over 50 consecutive frames in Fig. 10 that the dental-drill spectral components are removed sufficiently after the first-step and the second-step of the proposed TSDNR technique. However, it is noticed that there exist residual dental-drill noise components in the frequency range beyond 9 kHz.

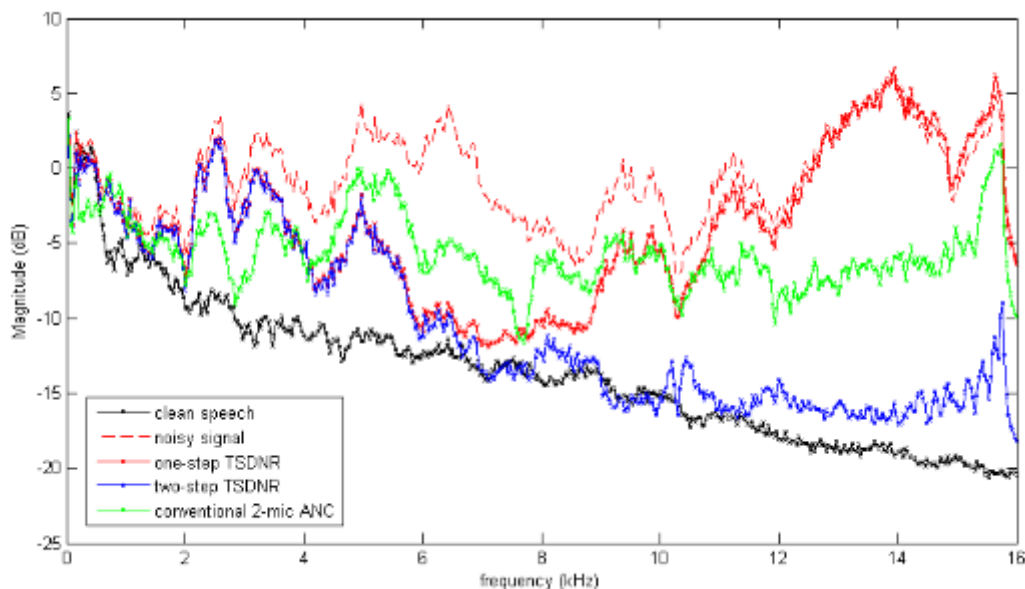


Fig. 10. The comparison of averaged magnitude spectrum plots of the clean, noisy, and enhanced speech signals of the conventional two-microphone ANC and the proposed TSDNR technique (after the first-step and the second-step), over 50 consecutive frames.

By observing over 30 noisy speech signals, the averaged Segmental SNR Improvement of the proposed TSDNR technique is presented for various input SNR levels in Table 1. It is illustrated that once the main spectral component of the dental-drill noise is removed by employing the first step of the proposed TSDNR technique, very large value of *SegSNR* improvement can be obtained. Moreover, the second step of the proposed TSDNR technique improved further the noise attenuation performance about 3-5 dB. Therefore, the overall performance of the TSDNR technique is far better than that of the conventional two-microphone ANC system.

By considering at the LSD, as given in Table 2, the proposed TSDNR technique provides lower level of speech distortion than the conventional two-microphone ANC system. It is also observed that the speech distortion level is increased slightly at very low input SNR levels.

Table 1. The SegSNR improvement of the investigated NR techniques.

Input SegSNR (dB)	ΔSegSNR (dB)		
	The conventional two-microphone ANC system	The proposed TSDNR technique (the 1 st step)	The proposed TSDNR technique (the 2 nd step)
-10.00	9.26	8.66	13.36
-5.00	6.26	6.69	10.89
0.00	3.34	4.96	9.32
5.00	1.54	3.68	8.63
10.00	0.68	2.96	7.39

Table 2. The LSD of the investigated NR techniques.

Input SegSNR (dB)	LSD (dB)		
	The conventional two-microphone ANC system	The proposed TSDNR technique (the 1 st step)	The proposed TSDNR technique (the 2 nd step)
-10.00	3.63	3.16	1.46
-5.00	3.41	2.97	1.45
0.00	3.10	2.76	1.39
5.00	2.90	2.60	1.36
10.00	2.51	2.36	1.28

Table 3. The MOS values of the investigated NR techniques.

Observed Signals	MOS values
Clean speech signal	4.9
Noisy speech signal (input SNR of -5 dB)	2.1
The conventional two-microphone ANC system	2.8
The proposed TSDNR technique (after the first step)	3.4
The proposed TSDNR technique (after the second step)	4.1

Furthermore, the subjective listening test was also carried out. The Mean Opinion Score (MOS) value was obtained based on 15 listeners, at the input SNR of -5 dB. By the informal listening, the high-pitched dental-drill sounds were no longer audible in the enhanced speech signals of the first-step and the second-step of the proposed TSDNR technique. It can be clearly seen from Table 3 that the MOS value of the enhanced speech signal employing two steps of the proposed TSDNR technique is better than that of the conventional two-microphone ANC system.

6. Conclusions

The TSDNR technique has been proposed for dental-drill noise reduction with the use of two separate single-microphone ANC systems in two steps. In the first step, the proposed tone-frequency extraction algorithm and an adaptive filter are employed together to removed significantly the main tone-frequency of the dental-drill noise. Then, in the second step, with the use of the second adaptive filter and a high-pass filter, the residual high-frequency dental-drill noise is satisfactorily removed. Although, the proposed TSDNR technique employs two adaptive filters, its computational complexity can be a lot lower than that of the conventional two-microphone ANC system due to its lower order of the adaptive filter, especially in the first step of the proposed TSDNR technique. In addition, the use of a single microphone offers

improved feasibility and efficient implementation of the proposed TSDNR technique. It has been demonstrated both by objective and subjective results that the proposed TSDNR technique can remove the dental-drill noise effectively, whereas the speech quality of the enhanced speech signal is reasonably preserved.

Acknowledgement

This research work is financially supported by the Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University and EECU Master Honours Scholarship, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University.

References

- [1] C. M. Mak, H. M. Wong, and Y. J. Chu, "Effect of the sound of dental equipment on dental anxiety and noise control techniques," in *Int. Commission on Biological Effects of Noise*, Zurich, Switzerland, 2017.
- [2] C. E. Wilson, "Hearing-damage risk and communication interference in dental practice," *Journal of Dental Research*, vol. 69, no. 2, pp. 489-493, Feb. 1990.
- [3] A. Dutta, K. Mala, and S. R. Acharya, "Sound levels in conservative dentistry and endodontics clinic," *Journal of conservative dentistry*, vol. 16, no. 2, pp. 121, Mar. 2013.
- [4] N. Bali, S. Acharya, and N. Anup, "An assessment of the effect of sound produced in a dental clinic on the hearing of dentists," *Oral Health and Preventive Dentistry*, vol. 5, no. 3, pp. 187, Jun. 2007.
- [5] F. Gijbels, "Potential occupational health problems for dentists in Flanders," *Clinical oral investigations*, vol. 10, no. 1, pp. 8-16, Mar. 2006.
- [6] W. S. Gan and S. M. Kuo, "An integrated audio and active noise control headset," *IEEE Trans. on Consumer Electronics*, vol. 48, no. 2, pp. 242-247, Aug. 2002.
- [7] S. M. Kuo, S. Mitra, and W. S. Gan, "Active noise control system for headphone applications," *IEEE Trans. on Control Systems Technology*, vol. 14, no. 2, pp. 331-335, Feb. 2006.
- [8] E. Kaymak, M. A. Atherton, K. R. G. Rotter, and B. Millar, "Real-time adaptive filtering of dental drill noise using a digital signal processor," in *Research and Education in Mechatronics*. Stockholm, Sweden, 2006.
- [9] E. Kaymak, M. Atherton, K. Rotter, and B. Millar, "Dental drill noise reduction using a combination of active noise control, passive noise control and adaptive filtering," in *Inter-noise*, Istanbul, Turkey, 2007, p. 110.
- [10] K. R. G. Rotter, M. A. Atherton, E. Kaymak, and B. Millar, "Noise reduction for dental drill noise," in *Proceedings of the 11th Mechatronics*, University of Limerick, Ireland, 2008, pp. 1-5.
- [11] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice-Hall, 2002.
- [12] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Acoust., Speech Signal Process.*, vol. 27, no. 2, pp. 113-120, Apr. 1979.
- [13] T. Esch and P. Vary, "Efficient musical noise suppression for speech enhancement systems," in *IEEE Int. Conf. on Acoust., Speech, Signal Processing (ICASSP)*, Taipei, Taiwan, 2009, pp. 4409-4412.
- [14] H. Gustafsson, S. Nordholm, and I. Claesson, "Spectral subtraction using reduced delay convolution and adaptive averaging," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 799-807, Nov. 2001.
- [15] S. Kamath and P. C. Loizou, "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," in *IEEE Int. Conf. on Acoust., Speech Signal Process.*, 2002, pp. 44164-44164.
- [16] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in *IEEE Int. Conf. on Acoust., Speech, Signal Processing (ICASSP)*, Washington, WA, USA, 1979, vol. 4, pp. 208-211.
- [17] P. Lockwood and J. Boudy, "Experiments with a nonlinear spectral subtractor (NSS), Hidden Markov models and the projection, for robust speech recognition in cars," *Speech Communication*, vol. 11, pp.215-228, 1992.

- [18] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Trans. on Acoust., Speech, and Signal Process.*, vol. 32, no. 6, Dec. 1984.
- [19] P. Scalart, "Speech enhancement based on a priori signal to noise estimation," in *IEEE Int. Conf. on Acoust., Speech, Signal Processing (ICASSP)*, Atlanta, GA, 1996, pp. 629-632.
- [20] P. C. Loizou, "Speech enhancement based on perceptually motivated Bayesian estimators of the magnitude spectrum," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 5, Sep. 2005.
- [21] R. Martin, "Spectral subtraction based on minimum statistics," in *Proc. Eur. Signal Process.*, 1994, pp.1182-1185.
- [22] H. G. Hirsch and C. Ehrlicher, "Noise estimation techniques for robust speech recognition," in *IEEE Int. Conf. on Acoust., Speech, Signal Process.*, Detroit, MI, 1995, pp.153-156.
- [23] C. Ris and S. Dupont, "Assessing local noise level estimation methods: Application to noise robust ASR," *Speech Communication*, vol. 34, no. 1, pp. 141-158, Feb. 2001.
- [24] R. Martin, "Speech enhancement using MMSE short time spectral estimation with Gamma distributed speech priors," in *Proc. Int. Conf. Speech, Acoustics, Signal Processing*, 2002, vol. 1, pp. 253-256.
- [25] R. Martin, "Speech enhancement based on minimum mean-square error estimation and supergaussian priors," *IEEE Trans. on Speech and Audio Process.*, vol. 13, no. 5, pp. 845-856, Aug. 2005.
- [26] N. Tangsangiumvisai, *A Dentist Drill Noise Reduction Technique* [Sound recording]. Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University, Report, 2016.
- [27] nBeats. (2014). *Dentist Drill Sound Effect [HD]* [Online]. Available: <https://youtube.com> [Accessed: 1 May 2017]
- [28] Sound Effects. (2015). *Dentist Drill Sound Effect* [Online]. Available: <https://youtube.com> [Accessed: 1 May 2017]
- [29] E. H. Rothauser, "IEEE Recommended practice for speech quality measurements," *IEEE Trans. Audio and Electroacoustics*, vol. 17, pp. 225-246, 1969.
- [30] A. Kar and M. N. S. Swamy, "Tap-length optimization of adaptive filters used in stereophonic acoustic echo cancellation," *Signal Processing*, vol. 131, pp. 422-433, 2017.
- [31] C. Breining, P. Dreiscitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, "Acoustic Echo Control—An application of very high order adaptive filters," *IEEE Signal Processing Magazine*, vol. 16, no.4, pp. 42-69, 1999.
- [32] I. Cohen, "Speech enhancement using super-Gaussian speech models and non-causal a priori SNR estimation," *Speech Communication*, vol. 47, no. 3, pp. 336-350, Apr. 2005.