Rämö, Jussi; Välimäki, Vesa; Vesa, Sampo; Väänänen, Riitta; Hämäläinen, Matti

Extracting music and noise from the ear canal using a headset

Published in:
Proceedings of the AES International Conference

Published: 21/08/2019

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Presented at the Conference on
Headphone Technology
2019 August 27 – 29, San Francisco, CA, US
This paper was presented at the AES Conference on Headphone Technology, as paper number 20. The full published version can be found at http://www.aes.org/e-lib/browse.cfm?elib=20513 .

Extracting Music and Noise from the Ear Canal using a Headset

Jussi Rämö[1], Vesa Välimäki[1], Sampo Vesa[2], Riitta Väänänen[2], and Matti Hämäläinen[2]

[1]*Acoustics Lab, Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland*
[2]*Nokia Technologies, Espoo, Finland*

Correspondence should be addressed to Jussi Rämö (`jussi.ramo@aalto.fi`)

ABSTRACT

An adaptive real-time filtering technique for separating music and ambient noise signals in the ear canal of a headset user is proposed. The system has been tested with a dummy head under laboratory conditions. This paper compares the responses given by the real-time system with those obtained with standard laboratory measurements. Both the headphone response estimate and the noise level measurements remain highly accurate even at a high ambient noise level of 86 $\mathrm{dB_{SPL}}$ inside the blocked ear canal. An efficient multirate version of the method is proposed for cases, where the ambient noise is dominated by low frequencies. The method has several useful applications for music listening in noise, including adaptive headphone target response equalization and unmasking of music.

1 Introduction

Many audio applications could benefit from having access to the information of the music and noise levels to which the user is exposed at any given moment. This is especially true when a user is wearing headphones in noisy environments, such as while commuting with public transport or in a noisy work place. Raised background noise levels often lead to louder music listening levels than in quiet surroundings, in order to unmask the background noise [1, 2, 3].

Estimation of the ambient noise level can be obtained by using an external microphone of a headset, which can be an in-wire microphone or a microphone installed directly to a headphone earpiece. This, however, does not directly indicate what the sound pressure level (SPL) inside user's ear canal is. Even though the transfer function from the external microphone to the ear canal of the user can be measured, the estimation is problematic, since the amount of leaking of the headset can vary, especially between different users and also over time.

An estimation of the total SPL inside the ear canal, when using an in-ear headset with a microphone installed on the ear-canal side of the headset, can be obtained from the in-ear microphone signal. Headsets with in-ear microphones have become common, e.g., in active noise cancelling headphones and in headphones that can calibrate themselves [4]. However, for many applications, it is important to discern between music and noise signals within the ear canal, e.g, when evaluating the auditory masking of music signals caused by ambient noise [3, 5], or loudness correction, which requires the knowledge of the listening level [6, 7, 8].

An important safety concern when using headphones is the listening level and total noise exposure users experience over an extended period of time [2, 9, 10, 11]. By using the in-ear microphone, it is possible to estimate the total amount of SPL inside the ear canal. However, in addition to the total exposure it may be useful to separate between the exposure caused by the background noise and music to get a better understanding what is causing the high noise level exposures users may experience.

This paper proposes an adaptive algorithm that separates ambient noise and music signals from an in-ear signal consisted of the mix of music and noise inside the ear canal of a user. This is conducted by using an adaptive least-mean-square (LMS) algorithm [12, Ch. 9], which takes the in-ear microphone signal and the unprocessed music signal as inputs, and outputs an estimate of the music signal at the in-ear microphone position along with an error signal, which, in this case, is the ambient noise signal. Furthermore, the adaptive filter coefficients, estimated by the LMS, form an estimate of the impulse response of the headset.

Obtaining the impulse response estimate of the headphone brings additional benefits and further possible use cases for the proposed algorithm, such as headphone magnitude response monitoring and evaluation of the perceived quality of the response [13, 14] and its real-time equalizing to a given target response [15, 16], or monitoring the fit and leakage of the headset [17] to avoid poor fitting of the headset, which can deteriorate the sound quality substantially and introduce even more noise into the ear canal.

The rest of this paper is organized as follows. Section 2 introduces the prototype headset used in this work, as well as measurements describing its acoustic properties. Section 3 describes the proposed adaptive algorithm. Section 4 shows the validation results, and Section 5 concludes the paper.

2 Prototype Headset

The headset used in this work was a commercially available in-ear headset with built-in in-ear microphones. The headset was modified in order to be able to obtain the in-ear microphone signal. A four-pin mini-plug connector was wired to the left and right loudspeaker elements of the headset, to the left in-ear microphone, and to ground. That is, the music played in stereo, but the sound inside the ear canal was estimated only from the left ear. This compromise was made in order to connect the headset directly to a laptop computer for easier testing and validation. Thus, all the further processing is applied to a mono signal and in one ear only. However, the signal processing for the right ear should be done exactly the same way as the left ear.

The headset was first measured in order to obtain the frequency response of the in-ear microphone and the headphone driver, as well as the passive isolation properties of the in-ear headset.

2.1 Measurement Equipment

The measurements were conducted in the large anechoic chamber at the Aalto University Acoustic Labs. Figure 1 shows photos of the setup, where the leftmost subfigure shows a closeup of the G.R.A.S. KEMAR Type 45BC dummy head with an IEC 60318-4 ear simulator RA0045 and anthropometric pinna type KB5001 (left ear). The prototype headset was fitted into the ear of the KEMAR, and a 1/2" free-field G.R.A.S. measurement microphone (type 46AF) was placed next to the ear.

The G.R.A.S. microphone was used for calibrating the levels of the other microphones, as well as a reference during the measurements, e.g., to make sure the sound field stays unchanged throughout the measurements. The subfigure on the right hand side in Fig. 1 shows a broader view of the setup, including the placement of a Genelec loudspeaker at approximately 2.8 m in front of the KEMAR. The speaker was used as an external sound source for isolation measurements.

A MacBook Pro laptop computer was connected to an RME Fireface 800 audio interface, which was used to output sound to the Genelec loudspeaker and to record the signals of the G.R.A.S. microphone and KEMAR's left ear canal simulator. The headset was connected directly to a MacBook Pro's headphone connector, since the headset was equipped with the four-pin mini plug. The joint use of two separate sound interfaces was enabled by MacOS's Audio MIDI Setup, which allows the user to create aggregate devices, i.e., to combine different audio interfaces and use them as one.

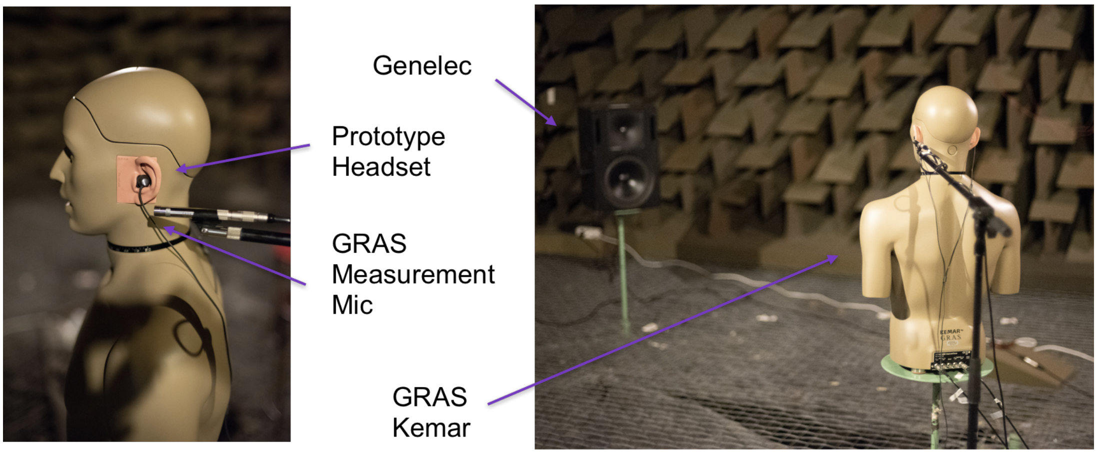

Fig. 1: Photos of the measurement setup in the anechoic chamber of the Aalto Acoustics Lab.

2.2 Level Calibration

The setup consisted of three microphones, namely the external G.R.A.S. reference microphone, KEMAR's left ear canal simulator, and the in-ear microphone inside the headset. The microphone in KEMAR's ear is located at the drum reference point (DRP), which correspond to the location of the ear drum.

First, the levels of the three microphones were aligned, so that the recorded signals could be compared against each other. The calibration procedure for the microphones was conducted as follows.

1. The reference G.R.A.S. microphone was calibrated using a calibrator (94 $\mathrm{dB}_{\mathrm{SPL}}$ at 1 kHz) in order to obtain the absolute level of the G.R.A.S. microphone.

2. A 1-kHz sine signal was played back from the external loudspeaker and simultaneously recorded by using the G.R.A.S. microphone and the left ear of the KEMAR dummy head. The absolute level of the sine signal was calculated using the calibrated G.R.A.S. microphone, and the level of the KEMAR mic was set to be that of the G.R.A.S. reference microphone.

3. A 1-kHz sine signal was played back from the headset, while inserted into the left ear of the KEMAR. The signal was recorded using the KEMAR and the in-ear microphone. Then, the headset's in-ear microphone level was set according to the known level of the KEMAR microphone.

2.3 Isolation Response

An isolation response of a headset illustrates how ambient noise is leaked around and through the headset into the ear canal of the user [18, 3]. The isolation is defined as the difference of the open ear response and the occluded ear response.

Figure 2 plots the results of an example measurement, where the magnitude response of the external speaker was measured using the KEMAR and the headset's in-ear microphone, while the KEMAR was facing the loudspeaker (0 degrees). The blue lines, in the top subfigure, show the response measured with the open ear (solid) and the occluded ear (dashed) of the KEMAR at DRP, while the black line shows the measured response at the in-ear microphone position.

The bottom subfigure in Fig. 2 shows the isolation response—the difference between the occluded ear response and the open ear response at DRP (blue curve)—and the difference between KEMAR's DRP microphone and the headset's in-ear microphone in the occluded case (dashed black curve). As can be seen, there is not much suppression at frequencies below 500 Hz, but between approximately 600 Hz and 8 kHz the attenuation provided by the headset is between 10 and 40 dB.

Figure 3 shows the calculated isolation responses for four different angles, namely 0°, 30°, 60°, and 90°, where at 0° the KEMAR was facing the loudspeaker and at 90° the ear of the KEMAR was pointed towards the loudspeaker. As can be seen, the isolation responses

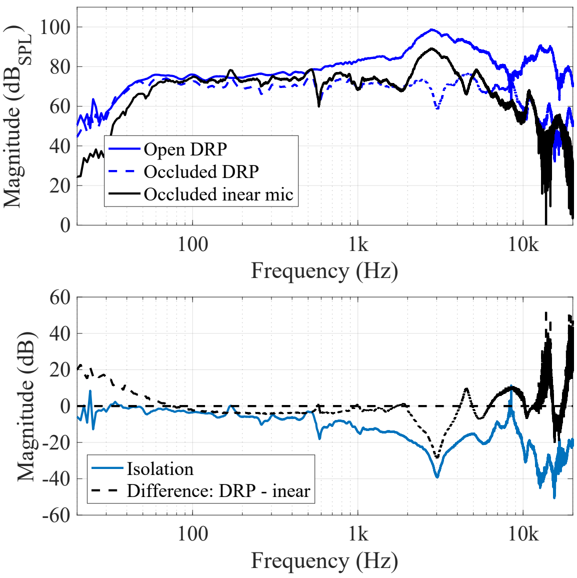

Fig. 2: (Top) Measurements of the external loudspeaker response with open and closed ear, and (bottom) the related headphone isolation curve and the difference of the responses at the in-ear and DRP microphones.

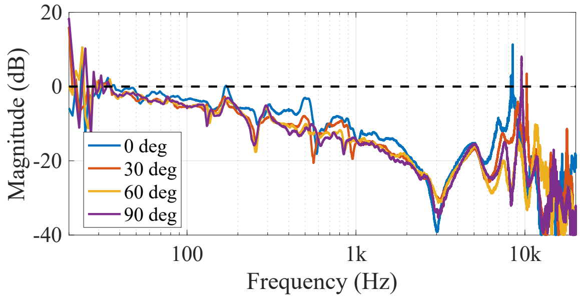

Fig. 3: Isolation responses with different incident angles of external sound.

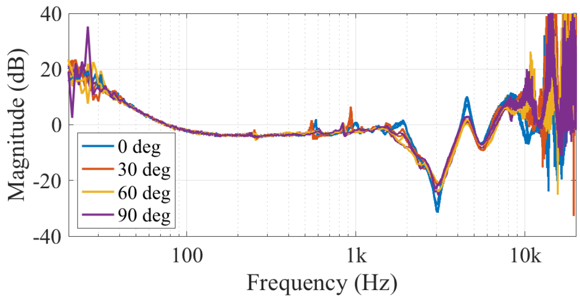

Fig. 4: Magnitude response differences between the in-ear microphone and KEMAR's DRP microphone, with the external sound source. Two measurements per angle (i.e., one repetition).

are (surprisingly) similar for different angles, indicating that ambient sounds from different directions (front and side) are attenuated quite uniformly.

Furthermore, Fig. 4 shows the transfer functions from the in-ear microphone to the DRP for the angles corresponding to the isolation curves in Fig. 3. The transfer function describing the change in magnitude response between the in-ear microphone and the DRP allows the estimation of the sound pressure level at the ear-drum position of the user by using the audio signal captured with the in-ear microphone. Similarly to the isolation responses, the transfer functions of different angles do not change much when the angle of the external sound source is varied.

2.4 Magnitude Response

The magnitude response of the headphone driver was measured using the in-ear microphone and KEMAR's DRP microphone. Figure 5 (top) shows the magnitude response measured at the in-ear (black) and DRP (blue) positions. There are total of eight measurements with four repositions of the headset. The bottom subfigure plots the magnitude difference between the DRP and in-ear microphones, i.e., the transfer function from the in-ear microphone to the "ear drum" of the simulator.

It is slightly surprising that there is such a big difference in the bass responses between the two microphones. The expectation was that there would not be any differences in the bass end of the spectrum, since there should not be any position-dependent resonances in the ear-canal at those frequencies [19], small electret microphones have typically no problem of capturing low frequencies [20], and the air in the sealed "ear canal" is operating under pressure chamber principle [21].

Based on the results shown in Figs. 3, 4 and 5, it seems that the measurement results are consistent and reproducible up to about 6 kHz. With higher frequencies the measurement results vary quite much, which is often the case with headphone measurements, mainly due to slight variations on the placement of the headphones [22] and the limitations of the ear canal simulators [23]. The inconsistency of the measurements below 40 Hz in Figs. 3 and 4 is due to the external loudspeaker that

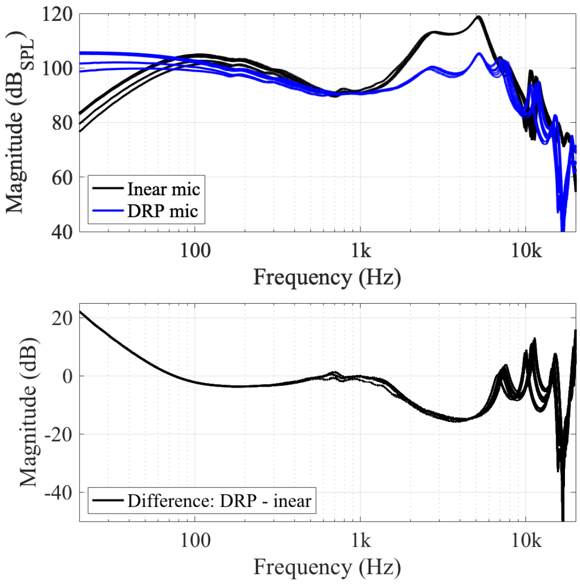

Fig. 5: (Top) Magnitude responses of the headset measured with the in-ear (black) and DRP (blue) microphones, and (bottom) the differences between the in-ear and DRP measurements.

was not able to reproduce frequencies lower than that. All in all, it is safe to assume that we can reliably use the in-ear microphone to estimate the leaked ambient noise at the ear drum at least up to 6 kHz.

3 Adaptive Filtering

The main idea in this work is to estimate the ambient noise inside the ear canal by removing the music signal from the noise-music mixture that is captured with the in-ear microphone of the headset. This is achieved by using an adaptive filter continuously in real time. Furthermore, the impulse response of the headphone is obtained with the adaptive processing, which can be highly useful for many applications.

Figure 6 shows the block diagram of the proposed adaptive filtering scheme that is used to estimate the ambient noise signal, as well as the music signal within the ear canal. It was decided to use the normalized LMS (NLMS) algorithm instead of a standard LMS algorithm, since the NLMS takes care of the normalization of the signal levels, which leads to a fast convergence and better stability, especially with nonstationary signals [24].

The clean music signal, which is fed to the headphone, is used as the reference signal for the adaptive filter. The in-ear microphone signal containing both ambient noise and music is the desired signal. This way, the purpose of the adaptive filter in Fig. 6 becomes to remove the music from the in-ear signal and leave the ambient noise in the error signal.

In general [12, Ch. 9][24], the filter coefficients $\mathbf{w}(n)$ of an LMS algorithm form an adaptive finite impulse response (FIR) filter producing an output

$$y(n) = \mathbf{w}^\top(n)\mathbf{x}(n), \tag{1}$$

where $\mathbf{x}(n)$ is the vector of reference signal samples, $\mathbf{w}(n)$ is the vector of adaptive filter coefficients, and $^\top$ refers to the transpose operation. The error signal $e(n)$ is calculated as a difference between the desired signal $d(n)$ and the filtered signal $y(n)$:

$$e(n) = d(n) - y(n). \tag{2}$$

The filter coefficients $\mathbf{w}(n)$ within the NLMS algoritm (in Matlab) are updated with

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu e(n)\mathbf{x}(n)}{\varepsilon + \mathbf{x}^\top(n)\mathbf{x}(n)}, \tag{3}$$

where μ is the adaptation step size and ε is a small positive constant to overcome possible numerical instability.

As illustrated in Fig. 6, in this work, the inputs of the NLMS algorithm are the clean music signal $\mathbf{x}(n)$ and the music and ambient noise mix $d(n)$ captured with the in-ear microphone. With this data, the NLMS algorithm continuously estimates the adaptive filter coefficients $\mathbf{w}(n)$ that approximate the impulse response from the headset driver to the in-ear microphone, after which it can effectively subtract the music signal estimate $y(n)$ from the captured in-ear signal. This leaves the error signal $e(n)$, which in this case becomes the ambient noise signal.

The clean music signal is delayed before it is sent to the NLMS, otherwise the NLMS would also estimate the “zeros” before the actual impulse response, caused by the delay occurring mainly in the audio interface, illustrated with the digital-to-analog (DA) and analog-to-digital (AD) blocks in Fig. 6. This allows the use of a shorter NLMS filter than without the delay. In this work, the delay is estimated from a sine sweep measurement by detecting the largest peak of the impulse

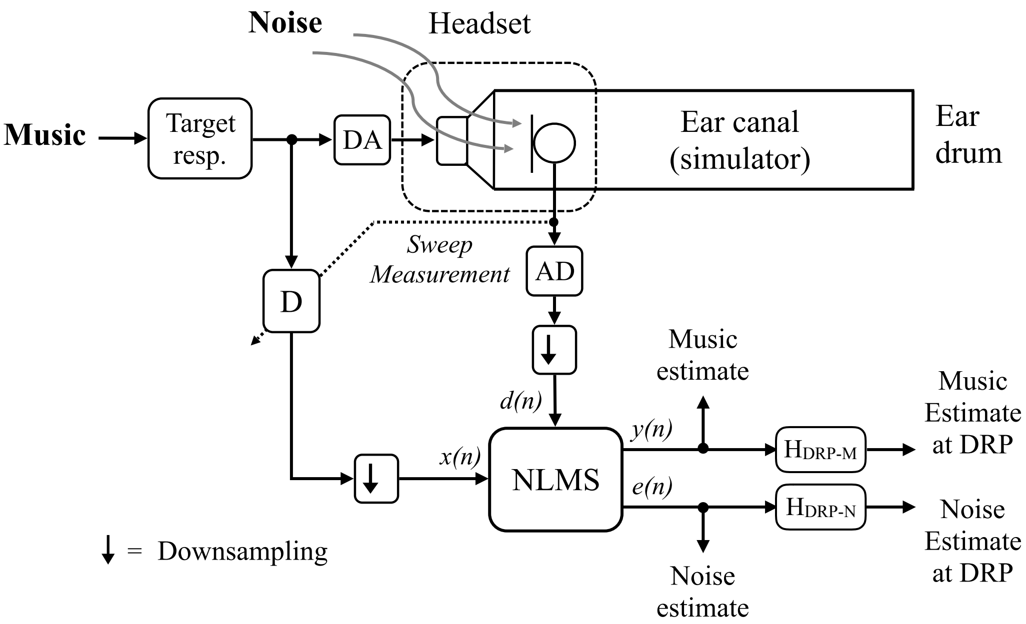

Fig. 6: Block diagram of the proposed adaptive filtering technique to separate music and noise.

response and subtracting 20 samples from that index value. With the hardware used during this work, the delay is about[1] 870 samples with the sample rate of 48 kHz, which corresponds to about 20 ms.

Typically, headphones isolate high frequencies better than low frequencies (see Fig. 3). Furthermore, many naturally occurring ambient noises have a spectral structure near to that of a pink noise, meaning there are more low frequencies, which pass the headphones' structure more easily. This is why low frequencies have a pronounced importance in many noise related applications.

In case of low frequencies being the main interest for an application, downsampling can be considered. For such applications downsampling can be advantageous, since it limits the frequency range at the high end of the spectrum, according to the Nyquist-Shannon theorem. The idea is to downsample the NLMS input signals before processing them, see Fig. 6. The benefits are that higher frequency components will not take any computing power, and thus, the NLMS algorithms can aim its focus on the frequency range that is of importance, and that the NLMS filter itself becomes considerably shorter, which reduces the computational complexity of the system.

Furthermore, $H_{\text{DRP-N}}$ and $H_{\text{DRP-M}}$ shown in Fig. 6 are the measured transfer functions of the prototype headset, from the in-ear microphone to DRP, for noise and music, respectively, see Fig. 7. They can be utilized to get music and noise estimates at the DRP position instead of the in-ear microphone position. However, the exact properties of these transfer functions may change between users.

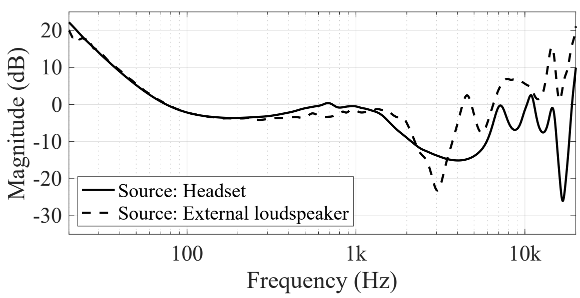

Fig. 7: Magnitude response differences, with 1/12-octave smoothing, between in-ear and DRP microphones averaged over multiple measurements (and angles in the external sound source case). See Figs. 5 and 4.

[1]The delay varies slightly every time the sound interface is reinitialized in Matlab.

4 Validation

The proposed adaptive system was implemented using Matlab with the help of Playrec [25] that allows non-blocking soundcard access from within Matlab. The validation measurements were conducted using the prototype headset and a G.R.A.S 45CA Headphone Test Fixture with an IEC 60318-4 ear simulator.

The ambient noise used in the experiments was pink noise and it was reproduced using a pair of Sennheiser HD 599 headphones placed on top of the prototype headset in the headphone test fixture. This proved to be a compact way of emulating a noisy situation while wearing an in-ear headset.

The music SPL was set to a moderate/high listening level (around 90 $\mathrm{dB_{SPL}}$). The noise was played back with three different levels of amplitudes: 1, 0.5, and 0.1, i.e., 0 dB, -6 dB, and -20 dB when normalized in dB units. The reference noise level was measured using the above described setup, when there was no music playing. In other words, the reference noise is the correct noise level that the adaptive algorithm is trying to estimate while the music is playing.

Figure 8 shows the results from the validation test, when the noise was played back at the middle level (-6 dB). The validation results were evaluated at the in-ear microphone position. The black line with markers, in the top subfigure, plots the impulse response of the headset measured using a sine sweep, while the green line with markers plots the adaptive filter coefficients $\mathbf{w}(n)$ of the NLMS, i.e., the impulse response estimated by the NLMS algorithm. The estimated impulse response is a calculated average of 40 frames, where each frame was 400 ms long. As can be seen, the impulse responses align well on top of each other.

Furthermore, the bottom subfigure in Fig. 8 shows the magnitude responses of both of the impulse responses depicted in the top subfigure. The main differences between these two magnitude responses are occurring at the very lowest and highest frequency regions. The inaccuracy at the bass end of the spectrum may be due to the fact that the NLMS filter length is only 250 samples and that is not enough to estimate the lowest frequencies. Another reason might be that the music signal does not have enough content at such a low frequency band to estimate it correctly. Lacking of high-frequency content in the music signal is probably the reason for high-frequency differences as well. All

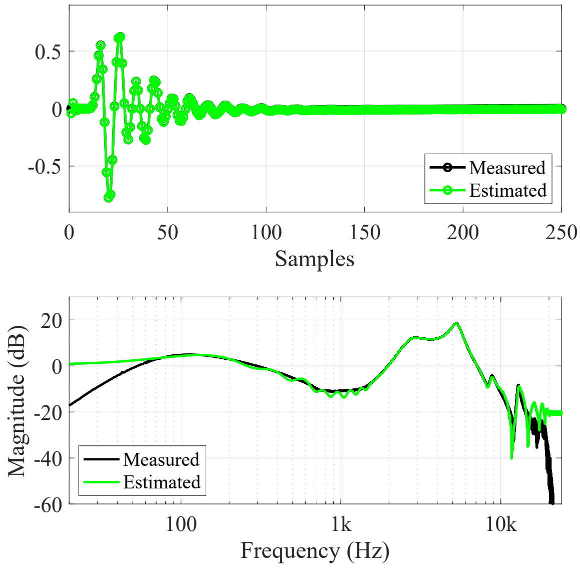

Fig. 8: NLMS performance with full 48-kHz sampling rate. (Top) Impulse responses and (bottom) magnitude responses of a sweep measurement (black) and NLMS adaptive filter coefficients (green).

in all, the NLMS estimation of the headset magnitude response is highly accurate.

Table 1 shows the results for noise estimation for three different noise levels, calculated from the NLMS output $e(n)$. Again, they are averages of 40 frames of adaptive processing. The middle row shows the corresponding noise level estimation result for the example case presented in Fig. 8. All three noise levels were accurately estimated. When the noise level was low, see the 66 $\mathrm{dB_{SPL}}$ case, there was a 2-dB difference between the estimate and the reference. This may be due to the fact that the surrounding ambient noise in the measurement situation starts to affect the measurement results, since the actual played back noise is not dominating anymore. The validation measurements were conducted in a moderately quiet office, but there might have been some extra noise, especially at low frequencies, that could interfere with the measurements when the target noise was at low level.

Figure 9 shows the magnitude spectra of the recorded in-ear signal (thick black line), the estimated noise (blue), and the estimated music (red). The spectra are averaged over 40 frames and 1/6-octave smoothed. As

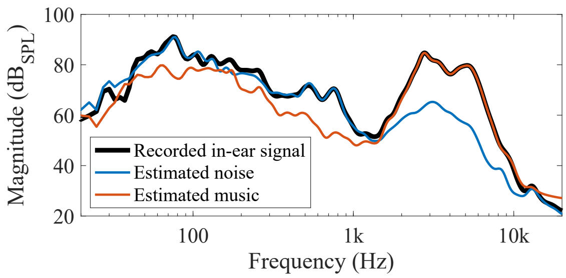

Fig. 9: Example of music and noise separation. Magnitude spectra of the recorded in-ear signal (thick black line), estimated noise (blue), and estimated music (red) signals. Average of 40 frames of adaptive processing at 48-kHz sampling rate is shown. The spectra were smoothed with a 1/6-octave filter.

can be seen, Fig. 9 demonstrates the fact that at low frequencies there is more ambient noise in the in-ear signal than music, while towards high frequencies the energy of the noise reduces and the music starts to dominate.

4.1 Downsampling

As discussed in Sec. 3, downsampling can be useful in certain situations, since the spectrum of a typical noise signal has more energy in the low-frequency range compared to high frequencies, and the isolation properties of a typical headset attenuates high frequencies better than low frequencies. Figure 10 illustrates the effect of downsampling, where the signals were downsampled with a factor of six, in order to have a sampling frequency of 8 kHz.

The top part of Fig. 10 shows the impulse responses of the same example as the one in Fig. 8, but downsampled. The main difference is that now the NLMS filter

Table 1: Estimated noise SPLs with two different sample rates at various reference noise levels. Music level was approx. 90 dB$_{SPL}$ in all cases.

Normalized level (dB)	Reference noise (dB$_{SPL}$)	Estimation $f_s = 48$ kHz (dB$_{SPL}$)	Estimation $f_s = 8$ kHz (dB$_{SPL}$)
0	85.8	85.9	85.8
−6	80.1	80.4	80.3
−20	66.2	67.7	67.2

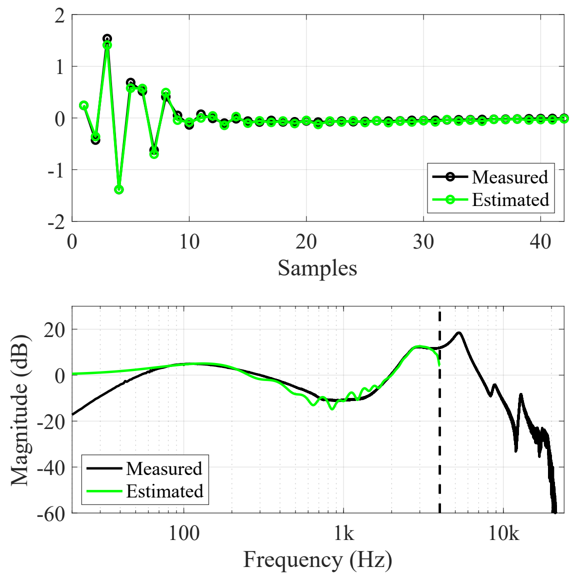

Fig. 10: NLMS performance at the downsampled 8-kHz sample rate, where black lines show the sweep-measured results and green lines show an average of 40 frames of adaptive processing. The vertical dashed line indicates the Nyquist limit at 4 kHz.

length is only 42 samples, i.e., one sixth of the original length of 250 samples. The bottom subfigure shows the estimated magnitude response (green) against the full-band response of the headset measured with a sine sweep (black). As can be seen, the magnitude response is estimated up to 4 kHz, and up to about 3.5 kHz the estimation is as accurate as that of the 48-kHz case.

Table 1 shows the noise level estimates for the downsampled cases as well. In fact, the results are as good as in the 48 kHz case, which emphasizes the fact that most of the noise energy leaking inside the ear canal is below 4 kHz.

5 Conclusions

This paper proposed an adaptive algorithm that extracts music and ambient noise signals from inside the ear canal using a single in-ear microphone of a headset. The proposed adaptive filtering technique was implemented using an NLMS algorithm, which was able to adaptively extract the ambient noise signal, as well as estimate the impulse response of the headset at the in-ear microphone position.

The algorithm takes the raw music signal as a reference signal and the in-ear microphone signal as the desired signal. Based on these, it adaptively estimates the filter coefficients, which correspond to the impulse response of the headset. The NLMS uses these filter coefficients to process the raw music signal to obtain an estimate for the music signal at the in-ear microphone position. It then subtracts the music estimate from the desired signal leaving an error signal, which corresponds to the ambient noise in the ear canal.

Validation measurements showed that the adaptive system works as expected in real time, and it can accurately estimate the impulse response of the headset and the ambient noise level while the headphone is being used. It was demonstrated that the algorithm also works at the sampling rate of 8 kHz, while still producing accurate ambient noise level estimations, since most of the energy of the ambient noise that leaks through the headset into the ear canal is below 4 kHz.

For many headphone signal processing applications, it is essential to obtain the information about the headphone response, as well as the levels of the music and ambient noise signals. The proposed adaptive system is a useful building block for sophisticated headset signal processing for music listening in a noisy environment.

6 Acknowledgment

This research was part of the EarEQ project (Aalto University project no. 410752) funded by Nokia Technologies.

References

[1] Rämö, J., Välimäki, V., Alanko, M., and Tikander, M., "Perceptual Frequency Response Simulator for Music in Noisy Environments," in *Proc. AES 45th Int. Conf.*, pp. 1–10, Helsinki, Finland, 2012.

[2] Oksanen, S., Hiipakka, M., and Sivonen, V., "Estimating Individual Sound Pressure Levels at the Eardrum in Music Playback over Insert Headphones," in *Proc. AES 47th Int. Conf.*, pp. 1–8, Chicago, IL, 2012.

[3] Rämö, J., Välimäki, V., and Tikander, M., "Perceptual Headphone Equalization for Mitigation of Ambient Noise," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 724–728, Vancouver, Canada, 2013.

[4] Backman, J., Campbell, T., Kleimola, J., and Hiipakka, M., "A Self-Calibrating Earphone," in *Proc. AES 142nd Conv.*, pp. 1–5, Berlin, Germany, 2017.

[5] Välimäki, V., Franck, A., Rämö, J., Gamper, H., and Savioja, L., "Assisted Listening Using a Headset: Enhancing Audio Perception in Real, Augmented, and Virtual Environments," *IEEE Signal Process. Mag.*, 32(2), pp. 92–99, 2015.

[6] Holman, T. and Kampmann, F., "Loudness Compensation: Use and Abuse," *J. Audio Eng. Soc.*, 26(7/8), pp. 526–536, 1978.

[7] Nielsen, S. B., "A Loudness Function for Analog and Digital Sound Systems Based on Equal Loudness Level Contours," in *Proc. AES 140th Conv.*, pp. 1–10, Paris, France, 2016.

[8] Fierro, L., Rämö, J., and Välimäki, V., "Adaptive Loudness Compensation in Music Listening," in *Proc. 16th Sound and Music Computing Conference*, pp. 135–142, Málaga, Spain, 2019.

[9] Moore, B. C. J., "Effects of Sound-Induced Hearing Loss and Hearing Aids on the Perception of Music," *J. Audio Eng. Soc.*, 64(3), pp. 112–123, 2016.

[10] Chasin, M., "Musicians and the Prevention of Hearing Loss," in *Proc. AES Int. Conf. Music-Induced Hearing Disorders*, pp. 1–6, Chicago, IL, 2018.

[11] Sinnott, L. M. and Weinstein, B., "Risk of Sound-Induced Hearing Disorders for Audio Post Production Engineers: A Preliminary Study," in *Proc. AES 145th Conv.*, pp. 1–10, New York, NY, 2018.

[12] Haykin, S., *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, NJ, third edition, 1996.

[13] Olive, S. E., Welti, T., and Khonsaripour, O., "A Statistical Model That Predicts Listeners' Preference Ratings of In-Ear Headphones: Part 1—Listening Test Results and Acoustic Measurements," in *Proc. AES 143st Conv.*, pp. 1–12, New York, NY, 2017.

[14] Olive, S. E., Welti, T., and Khonsaripour, O., "A Statistical Model That Predicts Listeners' Preference Ratings of In-Ear Headphones: Part 2—Development and Validation of the Model," in *Proc. AES 143st Conv.*, pp. 1–10, New York, NY, 2017.

[15] Olive, S. E., Welti, T., and McMullin, E., "Listener Preference For Different Headphone Target Response Curves," in *Proc. AES 134st Conv.*, pp. 1–12, Rome, Italy, 2013.

[16] Liski, J., Välimäki, V., Vesa, S., and Väänänen, R., "Real-Time Adaptive Equalization for Headphone Listening," in *Proc. 25th European Signal Process. Conf. (EUSIPCO)*, pp. 608–612, Kos, Greece, 2017.

[17] Tikander, M., "Modeling the Attenuation of a Loosely-Fit Insert Headphone for Augmented Reality Audio," in *Proc. AES 30th Int. Conf.*, pp. 1–9, Saariselkä, Finland, 2007.

[18] Rämö, J. and Välimäki, V., "Signal Processing Framework for Virtual Headphone Listening Tests in a Noisy Environment," in *Proc. AES 132nd Conv.*, pp. 1–6, Budapest, Hungary, 2012.

[19] Rämö, J. and Välimäki, V., "Digital Augmented Reality Audio Headset," *J. Electrical and Computer Engineering*, 2012, pp. 1–13, 2012.

[20] Zahn, R., "Analysis of the Acoustic Response of Circular Electret Condenser Microphones," *J. Acoust. Soc. Am.*, 69(4), pp. 1200–1203, 1981.

[21] Gido, S. P., Schulein, R. B., and Ambrose, S. D., "Sound Reproduction within a Closed Ear Canal: Acoustical and Physiological Effects," in *Proc. AES 130th Conv.*, pp. 1–19, London, UK, 2011.

[22] International Telecommunication Union, "Electro-Acoustic Measurements on Headsets," in *Telecommunication Standardization Sector of ITU, ITU-T Recommendation P.380*, 2003.

[23] International Electrotechnical Comission, "Electroacoustics—Simulators of Human Head and Ear—Part 4: Occluded-Ear Simulator for the Measurement of Earphones Coupled to the Ear by Means of Ear Inserts," in *International Electrotechnical Comission, IEC 60318-4*, Geneva, Switzerland, 2010.

[24] Morgan, D. R. and Kratzer, S. G., "On a Class of Computationally Efficient, Rapidly Converging, Generalized NLMS Algorithms," *IEEE Signal Process. Lett.*, 3(8), pp. 245–247, 1996.

[25] Humphrey, R., "Playrec: Multi-Channel Matlab Audio," *http://www.playrec.co.uk*, Accessed: Feb. 28, 2019.