IMPROVED NOISE REDUCTION FOR HANDS-FREE CAR PHONES UTILIZING INFORMATION ON VEHICLE AND ENGINE SPEEDS

*Henning Puder and Frank Steffens**
Signal Theory, Darmstadt University of Technology
Merckstr. 25, D–64283 Darmstadt, Germany
Tel: +49 (0)6151 16 2815, Fax: +49 (0)6151 16 3778
e-mail: hpuder@nesi.tu-darmstadt.de

ABSTRACT

In this paper we present an improved method for the spectral estimation of car noise in order to enhance the performance of noise reduction systems. An algorithm is developed that allows us to track changes in the noise spectrum during speech activity. For this tracking, we use the knowledge of the speed of the car and the revolutions of the engine. The paper starts with a detailed analysis of the car noise. The proposed algorithm based on this analysis first removes the harmonic components of the engine noise by selective filtering in time. The remaining wind and tyre noise is predicted during speech activity, based on the last available estimate and the vehicle speed.

1 INTRODUCTION

In recent years, speech processing systems for cars, like hands-free telephones or voice controlled operations have become more and more popular. They all require noise reduction, increasing both the communication comfort and the recognition rate of voice controlled systems. A major problem inherent to all noise reduction methods which are based on the modified Wiener filter [1] or the weighting rule of Ephraim/Malah [2], is the estimation of the power spectral density of the car noise. To estimate this density, one can make use of minimum statistics [3] or voice activity detectors [4]. The disadvantage of each of these methods is that they follow changes in the noise power, e.g. due to an acceleration of the car, insufficiently when speech is present. This degrades the performance of the noise reduction.

In this contribution, we first present an analysis of the different car noise components and show the dependence of these components on car parameters like vehicle speed and revolutions per minute (rpm) of the engine.

Then we will develop methods using the knowledge of these parameters to enhance the estimation of the noise spectrum, especially for tracking during speech activity. For this, we use the information about the speed and the rpm of the car, being available for modern cars on the Controlled Area Network (CAN) bus.

* Now with Robert Bosch GmbH, Stuttgart, Germany

2 CLASSIFICATION OF THE DIFFERENT CAR NOISE COMPONENTS

The three major components of car noise which are weighted differently according to the type of car and engine, are **engine, wind, and tyre** noise.

These three components are examined in detail in the following and a dependence of these components on the characteristic car parameters is worked out.

2.1 Engine Noise

The rotating engine and the movement of the pistons generate a harmonic noise spectrum having narrow band spectral components. The frequency localization of these components depends directly on the rpm of the engine. A spectrogram for typical engine noise (run-up and run-down) is depicted in Fig. 1. The harmonic structure of the spectrum can be easily identified. To

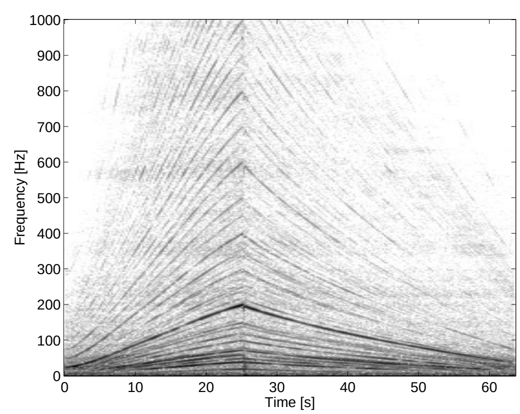

Figure 1: Spectrogram of a typical engine noise

underline the high power of the spectral components the power spectral density of an engine noise at 3485 rpm is depicted in Fig. 2. Engines normally exhibit spectral components at multiples of half the rpm frequency. These multiples are marked explicitly in Fig. 2.

2.2 Wind Noise

For modern cars with low aerodynamic drag, wind noise generally exhibits less power than tyre noise. Remarkable components normally only appear at higher speeds.

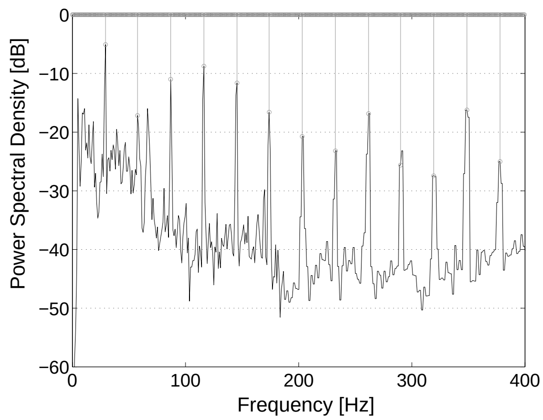

Figure 2: PSD of engine noise at 3485 rpm

Fig. 3 depicts the power spectral density of wind noise for a car measured in a wind tunnel at two different wind speeds. The higher noise power for the higher speed is obvious. However, it is especially interesting that the characteristic of the power spectral density stays nearly the same.

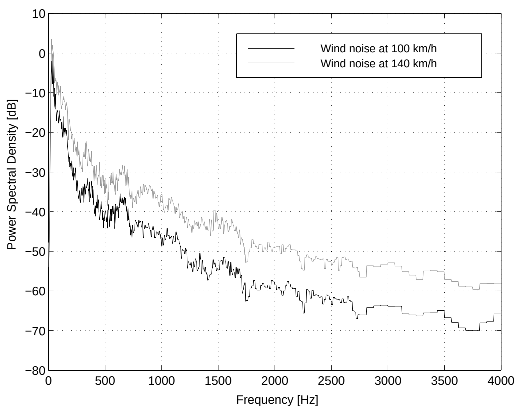

Figure 3: Wind noise at two different speeds

2.3 Tyre / Rolling Noise

For tyre noise, a correlation of the power spectral density of the noise and the speed of the car can also be discerned. To confirm this, two different noise spectra are

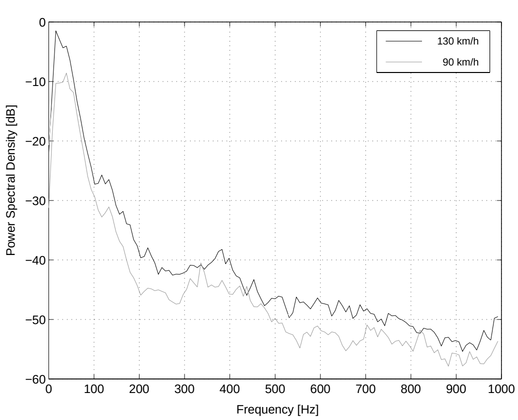

Figure 4: Tyre noise at different speeds

depicted in Fig. 4 for two different speeds. Even though the spectral characteristics are not identical, they are very similar. It is important to notice that frequency shifts of spectral components do not appear when the car speed changes.

3 ENHANCED NOISE ESTIMATION BY USING CAR PARAMETERS

The analyses of the previous chapter showed that car noise can be divided into two different classes: An engine noise that depends on the rpm, and the vehicle speed-dependent wind and tyre noises. The engine noise consists of harmonic spectral components at multiples of half the rpm. Wind and tyre noises have a different power depending on the speed of the car. However, their frequency characteristic stays nearly the same.

For the calculation of the weighting coefficients of the noise reduction,

$$G_{opt}(k,n) = \begin{cases} 1 - \sqrt{\frac{\kappa NPSD(k,n)}{XPSD(k,n)}} & G_{opt} > \beta_f \\ \beta_f & \text{otherwise} \end{cases}, \quad (1)$$

an estimation of the power spectral densities of the input signal,

$$\begin{aligned} XPSD(k,n) = & \lambda_{fast}\, XPSD(k,n-1) \\ & + (1-\lambda_{fast})\, |X(k,n)|^2 \end{aligned} \quad (2)$$

and the noise

$$NPSD(k,n) = \begin{cases} \lambda_{slow}\, NPSD(k,n-1) & \text{no voice} \\ +(1-\lambda_{slow})\, |X(k,n)|^2 & \text{detected} \\ NPSD(k,n-1) & \text{otherwise} \end{cases} \quad (3)$$

at time n and frequency k is necessary. $X(k,n)$ denotes the spectrally analyzed noisy input signal, whereas β_f and κ are the spectral floor and the overestimation factor, respectively.

Commonly used estimation methods for the power spectral density of the noise cannot follow changes during speech activity (Eqn. 3) leading to a decrease in performance of the noise reduction, especially when the car accelerates.

In the following, we will develop methods for noise reduction that can follow changes in the PSD of the noise during speech activity. The first step in Section 3.1 is to cancel the harmonics of the engine noise that depend on the rpm. The second step in Section 3.2 is to track during speech activity the estimation of the remaining noise spectrum, i.e. wind and tyre noise, that depend on the velocity.

3.1 Engine Noise

Engine noise mainly occurs at multiples of half the engine frequency:

$$f_{harm,i} = i \cdot \frac{f_{engine}}{2} \qquad i \in \mathbb{N}. \quad (4)$$

The engine harmonics have a remarkable power up to an order of 15-17, or up to 400-500 Hz, respectively.

A specific property of the harmonics is that they have a very narrow-band character and that their frequency localization can be determined very precisely based on the rpm-signal. Nevertheless, it is not possible to predict their power which varies strongly and in an undetermined way with the rpm and time.

Therefore, our approach is to attenuate the engine noise harmonics up to 400-500 Hz

$$G_{opt}(k_i, n) = L, \tag{5}$$

continuously, i.e. also during speech activity. Here L denotes an attenuation value, e.g. -20 dB, n the time, and k_i the discrete frequency values of the harmonics. This procedure requires a very selective attenuation in frequency in order not to attenuate the speech components too much which would cause speech distortions.

Tests have shown that it is not possible to obtain the required selectivity in the frequency domain, e.g. by a STFT or filterbank analysis, without causing large speech distortions. This is due to a long time period of the signal that has to be analyzed in order to obtain the required frequency resolution of a few Hertz. This is equivalent to a large smoothing in time of the output signal and leads to distortions of the instationary speech components. For this reason we examine if it is possible to achieve the required selective attenuation by filtering in the time domain.

One possibility is to apply a notch filter of second order to attenuate every engine harmonic separately.

The transfer function of the notch filter is given by:

$$H_N(z) = \frac{1+\alpha}{2}\,\frac{1-2\beta z^{-1}+z^{-2}}{1-\beta(1+\alpha)z^{-1}+\alpha z^{-2}} \tag{6}$$

with:
$$\beta = \cos\left(\frac{f_0}{8000Hz}\,2\pi\right) \tag{7}$$

$$\alpha = \frac{1}{\cos\left(\Delta\omega_{3\,\mathrm{dB}}\right)} - \sqrt{\frac{1}{\cos^2\left(\Delta\omega_{3\,\mathrm{dB}}\right)} - 1} \tag{8}$$

where f_0 denotes the stop-band frequency and $\Delta\omega_{3\,\mathrm{dB}}$ the 3 dB cutoff frequency.

Due to the second order of the filter, the attenuation that can be obtained is lower the more the filter is chosen to be selective in frequency.

As a remedy to this problem, we developed a modified notch filter of higher order that can be designed in two steps:

1. Design of a higher order highpass filter $H(z)$. The zeros and poles of a possible filter of order two are depicted in Fig. 5 and the corresponding transfer function is shown in Fig. 6.

2. Shift and multiply of poles ($z_P = a_P + j\,b_P$) and zeros ($z_0 = a_0 + j\,b_0$) which is equivalent to oversampling by the factor U according to

$$z_{P,0i} = \sqrt[U]{a_{P,0}^2 + b_{P,0}^2}\; e^{j\left(\frac{\phi_{P,0}}{U} + i\frac{2\pi}{U}\right)}, \quad i \in [-K, K]. \tag{9}$$

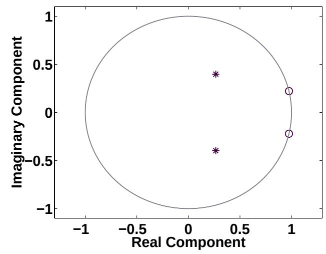

Figure 5: Poles (*) and zeros (o) of the second order highpass $H(z)$

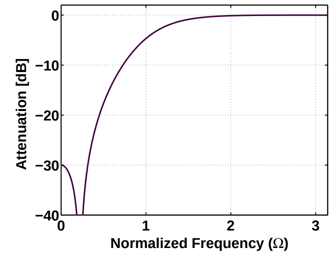

Figure 6: Transfer function $H(z)$ of the second order highpass

A filter with these zeros and poles exhibits K stop-band frequencies at $\Omega_s = i\,\frac{2\pi}{U}$ with $i \in [1, K]$ and can be denoted as:

$$G(\Omega) = \frac{\prod_i (1 - z_{0i}\,z^{-1})}{\prod_i (1 - z_{Pi}\,z^{-1})}. \tag{10}$$

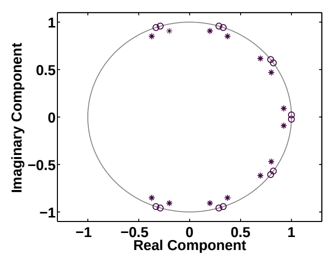

Figure 7: Poles (*) and zeros (o) of the modified notch filter $G(z)$

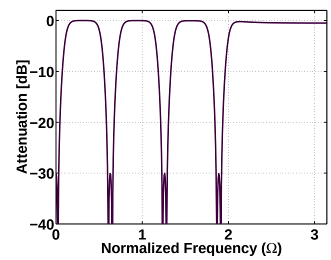

Figure 8: Transfer function of the modified notch filter $G(z)$

The position of the resulting poles and zeros is shown in Fig. 7 and the corresponding transfer function in Fig. 8.

In comparison with comb-filters $G_K(z)$, which can be generated by an oversampling of a highpass by the factor M, i.e. $G_K(z) = H(z^M)$, the modified notch filters have the following advantages:

- They offer the possibility to continuously adjust the stop-band frequency to the rpm, as $U \in \mathbb{R}^+$ in contrast to $M \in \mathbb{N}$.

- The number of harmonics to attenuate can be determined with the constant K.

A selective filtering in frequency can now be performed as follows: With the rpm f_{engine} the factor $U = \frac{8000Hz}{f_{engine}}$ can be determined and then, based on a fixed prototype highpass (s. Fig. 6), the position of the poles and the zeros of the modified notch filter can be calculated whenever U changes. Combining successively two conjugate-complex poles and zeros, we get IIR filters of second order which are - one after the other - utilized to filter the noisy speech signal.

Generally, only the engine harmonics up to 500 Hz exhibit a remarkable power such that it is worth applying a selective attenuation. Thus, it is reasonable to first decompose the noisy signal with an 8-channel filterbank in subsampled subband signals. Then the filtering with the modified notch filters has only to be applied to the lowest band covering the frequency range up to 500 Hz (s. Fig. 9). The advantage of this procedure is that the poles and zeros of the modified notch filter are further apart. Thus, the designed modified notch filter is less sensitive to quantization errors of the filter coefficients. Additionally, computational power is saved as the signal $x_1(l)$ is subsampled. The first stage of the frequency de-

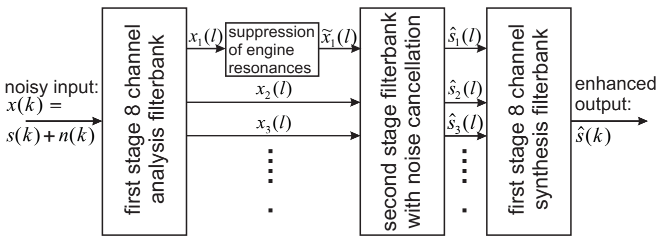

Figure 9: Structure of the noise cancellation algorithm

composition is followed by a second, which is required by the noise reduction algorithm to remove the other noise components. We obtained satisfying results with this filtering in time, especially without speech distortions.

3.2 Wind and Tyre Noise

The analyses of wind and tyre noise in Section 2 showed that changes in speed affects all frequency components in nearly the same manner.

Thus, we predict the wind and tyre noise spectrum during speech activity by taking its last estimate, which was calculated during the last speech pause and raise or lower it according to a change in the speed of the car,

$$NPSD(k,n) = NPSD(k,n_0) \cdot P(v(n), v(n_0)), \quad (11)$$

where n_0 denotes the time of the last estimation and $P(x,y)$ a scalar function of two parameters.

We examined the dependence of the noise power on the car speed. A typical result is depicted in Fig. 10 which visualizes the power of wind and tyre noises for different speeds referred to the power at a speed of 80 km/h. The curve may be approximated to a straight line being equivalent to the following relation:

$$NPSD(k,n) = NPSD(k,n_0) \cdot e^{K\,(v(n)-v(n_0))}. \quad (12)$$

In normal driving situations, wind and tyre noise is only measurable together with engine noise. As we apply the cancellation of the engine harmonics before the estimation of the PSD of the noise (s. Sec. 3.1), the input spectrum $X(k,n)$ is already free of these harmonics.

The new method for estimating the power spectral density of the noise is then given by the following:

$$NPSD(k,n) = \begin{cases} \lambda_{slow}\, NPSD(k,n-1) + (1-\lambda_{slow})\,|X(k,n)|^2 & \text{no voice detected} \\ NPSD(k,n_0) \cdot e^{K\,(v(n)-v(n_0))} & \text{otherwise} \end{cases} \quad (13)$$

K denotes a constant corresponding to the gradient of the approximating line of the curve in Fig. 10. This

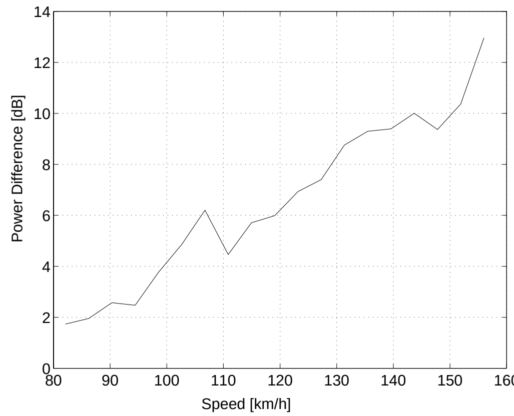

Figure 10: Power of wind and tyre noise dependent on the car speed referred to the power at 80 km/h

parameter depends on the car, the tyres, etc. For our examinations, we determined the parameter with the help of measured data. A real-time system could estimate this parameter by evaluating noise measurements at different speeds.

4 RESULTS AND CONCLUSIONS

In this paper we presented methods to enhance noise reduction by predicting and selectively attenuating car noise components. The algorithm was tested with a frequency decomposition using a two stage filterbank, like depicted in Fig. 9 and described in [5]. Thus, besides the advantages for the cancellation of the engine harmonics, it is possible to perform a high frequency resolution for the low frequency bands and a high time resolution for the high frequency bands adapted to the human hearing.

The performance of the algorithm was examined with different speech and noise signals recorded in cars while accelerating. In comparison to the methods for noise estimation known so far, we were able to obtain satisfying and promising results.

References

[1] S. Boll: *Suppression of Acoustic Noise in Speech Using Spectral Subtraction,* IEEE Trans. Acoustics, Speech, and Signal Processing ASSP-27, pp. 113-120, 1979

[2] Y. Ephraim, D. Malah: *Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator,* IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-32, No.6, 1984

[3] R. Martin: *Spectral Subtraction Based on Minimum Statistics,* in Proc. Seventh European Signal Processing Conference, pp. 1182-1185, 1994

[4] H. Puder: *Single Channel Noise Reduction Using Time-Frequency Dependent Voice Activity Detection,* in Proc. IWAENC-99, pp. 68-71, USA, 1999

[5] P. Dreiseitel, H. Puder: *Speech Enhancement For Mobile Telephony Based on Non-Uniformly Spaced Frequency Resolution,* in Proc. EUSIPCO-98, vol.2, pp. 965-968, Greece, 1998