An optimal high-pass filtering technique to improve the detectability of evoked otoacoustic emissions

G. Tognola, P. Ravazzani, F. Grandori
System Theory Centre - CNR, Polytechnic of Milan, I-20133 Milan, Italy

Abstract

Evoked otoacoustic emissions (EOAEs) are currently used as the most reliable test for hearing screening. The presence of emission is commonly detected by the visual inspection of the signals and the cross-correlation function between replicate responses recorded in the same conditions. Performances of these methods are greatly influenced by the residual noise level. Residual noise is often of low frequency and can be reduced by digital off-line filtering. For this purpose, an optimal high-pass filtering technique is proposed here. Cross-correlation between replicates of transient evoked otoacoustic emissions (TEOAEs) was used to determine the cut-off frequency that maximises the S/N ratio. This will be shown in a series of responses characterised by various S/N ratios. Data have been scored both visually and with quantitative methods before and after the use of high-pass filtering. The capability to detect the signal was increased with only a marginal decrease in the total power of the emissions.

1 Introduction

Evoked otoacoustic emissions (EOAEs) are acoustic signals produced by the inner ear [1]. Emissions can be evoked by transient acoustic stimuli (clicks and tone bursts) and can be recorded with a miniature microphone in the external auditory canal (for a review, see Grandori, Cianfrone & Kemp [2]). Since emissions can be reduced or even abolished by hearing losses of peripheral origin due to trauma and pathology, EOAEs are presently considered as the most sensitive clinical test for assessing the integrity of the end organ.

Presence of a transient-evoked otoacoustic emission (TEOAE) is usually detected by visual inspection, or by the use of simple criteria like cross-
correlation between two replicates, e.g. Lutman [3], Vohr [4], or by the analysis of their cross-spectrum, e.g. White [5]. These methods are obviously influenced by the residual noise content. In particular, noise reduction is one of the most outstanding problems in objective-classification methods [4].

There are typically three sources of noise: instrumentation (microphone included), environment, and the subject [6]. Instrumentation noise has stationary properties and is reduced with synchronous averaging. An effective reduction in the environmental noise is obtained with the use of a sound-proof cabin and/or with a noise-rejection threshold, and synchronous averaging. Instead, patient noise is more difficult to treat: it is often of low-frequency and is produced by swallowing, breathing, snoring, teeth grinding, heart beating, and cable rub. High-amplitude patient noise can be eliminated by noise rejection method, while low-level patient noise can be reduced by digital high-pass filtering [3]. Finally, it is should be noted that also some particular recording techniques, such as the derived non-linear method can intrinsically reduce the S/N ratio of the recordings; e.g. Kemp [6], Ravazzani & Grandori [7], Grandori & Ravazzani [8].

In this paper, an optimal high-pass filtering technique is proposed to maximise the S/N ratio in a given recording trial. Cross-correlation is used to determine an optimal high-pass frequency and to quantify the improvement obtained with the filtering technique. Finally, cross-power spectrum is used to estimate the loss of emission power due to filtering. A more comprehensive description of our procedure and results has been published in previous papers [9-10].

2 Methods

2.1 Materials
Click-evoked otoacoustic emissions were recorded binaurally from 16 normal hearing adults and 8 pathological subjects with moderate hearing losses due to otosclerosis. Emissions were collected both inside (3 subjects) and outside (21 subjects) a sound-proof cabin (estimated attenuation: 40 dBA) at various intensity levels to validate the proposed method under a wide variety of conditions and S/N ratios (for level of signal and noise). The background noise estimated outside the sound-proof cabin during a typical test session was -11 dB SPL. In most cases, the subjects were tested at 11 stimulus intensity levels from 50 to 80 dB SPL, step 3 dB SPL; during some other test sessions, subjects were tested at 9 stimulus intensity levels from 50 to 90.5 dB SPL. The equipment used for TEOAE measurements was an ILO88 (Otodynamics). The recordings were filtered with the ILO88 default procedure (second order high-pass set at 330 Hz, gain 1.57 and fourth order low-pass set at 10.6 kHz, gain 2.6) and digitised at a rate of 25000 samples/s with 260 repetitions of the non-linear four-click train. Finally, the averaged data were windowed using the default ILO window (2.5-20 ms) and filter (second order digital band pass set at 600-
6000 Hz). In practice, this last filtering stage has no relevant effects on the low-frequency content of the signal because the order of the filter is very small. Signals were stored and processed off-line on a host computer.

2.2 The proposed filtering procedure

It is well known that the frequency content of both the noise and the TEOAE overlap; this implies that high-pass filtering removes not only the noise but also some signal components. We used an optimal high-pass filtering technique to remove residual low-frequency noise. The filtering technique was optimal in the sense that the high-pass cut-off frequency maximises the cross-correlation between the two replicate TEOAEs. It will be shown that the analysis of cross-correlation coefficients can be a useful tool to determine the particular cut-off frequency that reduces noisy components, but maintains the most significant features of emissions.

The entire procedure is schematically depicted in Figure 1.

Figure 1. Schematic representation of the optimal high-pass filtering procedure. STEP is the sample width in the frequency domain; Fo is the optimal high-pass cut-off frequency that is, the frequency at which the cross-correlation reaches its maximum; FTh is the high-pass cut-off frequency at which the loss in emission power reaches the threshold level Th; Fc is the final value of the cut-off frequency.
Cross-correlation between the two replicate recordings (A and B, respectively, in the ILO88 system) has been used as an estimate of the S/N ratio. Recordings were digitally high-pass filtered (zero-phase-shift, finite impulse response, 63rd order) at varying cut-off frequencies, $F_c$, from 400 to 1600 Hz, in steps empirically set to 200 Hz. For each replicate, cross-correlation coefficients were computed as a function of the cut-off frequency in a window 7-16 ms to avoid the influence of the acoustical artefact and the tapering procedure. For each recording, the optimal high-pass cut-off frequency $F_o$ was defined as the frequency at which the correlation reached its maximum. The cross-power spectrum magnitude between two replicates was computed together with the normalised cumulative power (for details, see below). The normalised cumulative power has been used to quantify the loss of signal due to filtering. To prevent an undesirable reduction in the signal components, responses were filtered at the cut-off frequency $F_o$ only when the corresponding decrease in total power was less than a threshold level $T_h$ empirically set to 30%. In the remaining cases, responses were filtered at the cut-off frequency $F_{Th}$ (lower than $F_o$) at which the corresponding power loss reaches the threshold level.

**Cross-power spectrum and Coherence function** An estimate of the magnitude of cross-power spectrum $G_{xy}$ of two replicate recordings $x$ and $y$ was obtained from the relationship:

$$|G_{xy}(k)| = \frac{1}{N} |W_{xy}(k)|$$

where $W_{xy}(k) = X(k)^* Y(k)$; $X(k)$ and $Y(k)$ are the Fourier transform of $x$ and $y$; $N$ is the total number of data points; $X(k)^*$ is the complex conjugate of $X(k)$, and $k$ is the sampled frequency. To reduce the presence of large side-lobes in the equivalent filter response, the leading edges of the time series $x$ and $y$ were tapered by a cosine window and the complex quantity $W_{xy}(k)$ was smoothed using a five-point Blackman routine [11].

The similarity between two signals can be expressed with a quantitative figure in normalised form. We have done this by computing a Coherence Function [11-12] given by:

$$C_{xy}^2(k) = \frac{|G_{xy}(k)|^2 N^2}{|T_x(k)||T_y(k)|N^2}$$

where $|T_x(k)|$ and $|T_y(k)|$ are $|X(k)|^2$ and $|Y(k)|^2$ after smoothing with the five-point Blackman routine. The coherence function can assume values between 0 and 1; higher values imply a higher similarity. In the computation of the cumulative power (see below) only the components with a coherence factor greater than 0.5 were considered; see, e.g. Beauchamp [11].
Normalised Cumulative Power  The cumulative power is defined here as the power of the emission in the frequency band [0-\(k\) Hz]:

\[
CP(k) = \sum_{j=1}^{k} |G_{xy}(j)| \cdot \Delta
\]

\[\text{(3)}\]

where \(\Delta\) is the sample width in the frequency domain [0-\(k_{\text{MAX}}\) Hz], and \(k_{\text{MAX}}\) is the maximum signal frequency.

The cumulative power at a given frequency can be expressed in the so called normalised form:

\[
NCP_{\text{norm}}(k) = \frac{CP(k)}{CP(k_{\text{MAX}})} \cdot 100
\]

\[\text{(4)}\]

3 Results

Typically, optimal cut-off frequencies were found to be in the range 600-1200 Hz. In some cases, however, final values as low as 400 Hz or as high as 1400 Hz were found.

An illustrative example of cross-correlation as a function of the cut-off frequency is given in Figure 2 for the response evoked at 68 dB SPL. For this response, the maximum is at 1400 Hz. For cut-off frequencies higher than the optimal value, cross-correlations decrease rapidly and take values well below 50%. It is observed that the cross-correlation values can be improved only if the reduction of noisy components is greater than that of the signal. On the contrary, when the cut-off frequency is above the optimal value, signal components are filtered out thus producing a progressive decrease in the correlation values.

Fig. 2. Cross-correlation of two replicates from subject DNLL as a function of the cut-off frequency. The curves were computed from responses evoked at 68 dB SPL. Cross-correlation coefficients were computed in a time window 7-16 ms. For a sake of completeness, the high-pass filtering was performed up to 3000 Hz. The dashed line corresponds to the correlation value before filtering. TEOAEs from this subject are given in Figure 3.
Some examples of TEOAEs obtained from normal ears are given in Figures 3 and 4. Responses were recorded at various stimulus intensities in a soundproof cabin. Even a simple visual inspection reveals that filtering improves the reproducibility of the recordings and does not alter the basic features of the responses.

Fig. 3. (A) Examples of emissions evoked at decreasing stimulus levels from a normal hearing adult; stimulus intensities range from 50 to 90.5 dB SPL (intensity values are given on the left of each trace). Two averaged responses of 260x4 sweeps each are superimposed. Responses have been windowed from 2.5 to 20 ms post-stimulus time. Left: recordings before optimal high-pass filtering (raw data). Right: filtered responses. The high-pass cut-off frequencies $F_c$ are given on the right of each trace. (B) Cross-correlation from raw data (dashed line) and optimal high-pass filtered data (solid line) as a function of the stimulus intensity level for the same subject. Cross-correlation coefficients were calculated in a time window 7-16 ms.

As expected, the correlation is an increasing function of the stimulus intensity level for both raw and filtered data. Assuming that the noise level is about constant during the whole recording trial, the lower the intensity level, the lower are the emission amplitudes (and the S/N ratio). After high-pass filtering, the correlation was increased for all the responses. For some responses, only after filtering, the correlation values fall well above 60%. Note that in the
objective-classification methods, emission is considered absent if the cross-correlation coefficient is below 50-60%.

Despite the improvement in the cross-correlation values, the average loss in the total power was typically around 15-20%.

![Figure 4](image)

Figure 4. (A) Examples of emissions evoked at decreasing stimulus levels from a normal hearing subject. (B) Cross-correlation from raw data and optimal high-pass filtered data as a function of the stimulus intensity. Details as in previous Figure 3.

4 Conclusions

In general, high-pass filtering can improve the reproducibility of the recordings without a significant loss in the main features of the evoked responses. We propose a filtering procedure capable of improving the signal detectability from each recording trial. Cross-correlation was used to determine the optimal cut-off frequency that maximises the S/N ratio of a given response. In a sample of more than 350 TEOAEs from normal and hearing impaired adults, the optimal cut-off frequencies were typically found to be in the range 400-1400 Hz, with an average value around 700-800 Hz. This optimum filtering method was shown to increase the detectability of the evoked emissions in cases of low and high S/N ratio; nevertheless, this filtering algorithm was found to create no false negatives when it is applied to data collected from hearing impaired patients.
References