COMPARISON OF DIFFERENT ADAPTIVE ALGORITHMS FOR STEREOPHONIC ACOUSTIC ECHO CANCELLATION

Peter Eneroth, Jacob Benesty†, Tomas Gänsler†, and Steven Gay†

Dept. of Applied Electronics, Signal Processing Group, Lund University, Sweden
† Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey, USA
peter.eneroth@tde.lth.se, {jbenesty,gaensler,slg}@bell-labs.com

ABSTRACT

In this paper, different adaptive algorithms for stereophonic acoustic echo cancellation are compared. The algorithms include the simple LMS algorithm and two specialized two-channel adaptive algorithms. Due to the high calculation complexity needed in stereophonic acoustic echo cancellation applications, the time domain algorithms are applied in a subband structure. The comparison include aspects as convergence rate, calculation complexity, signal path delay and memory usage. Real-life recordings are used in the evaluation.

1. INTRODUCTION

The increasing use of teleconferencing systems and desktop conferencing where the acoustic echo canceler (AEC) plays a central role, has led to the requirement of faster and better performing algorithms. In these applications there is a desire to have far better sound quality and sound localization than what has been provided previously. These quality improvements can be achieved by increasing the signal bandwidth and also by adding more audio channels to the system. This last fact spurred the need for multi-channel acoustic echo cancelers of which the two channel (stereo) AEC is the the most interesting since only complexity issues differ in the more general multichannel case. Figure 1 illustrates the concept of stereophonic echo cancellation between a transmission room and a receiving room. As is depicted in the figure, the echo is due to the acoustic coupling between the loud-speakers and the microphones in the receiving room. The solution is to estimate this coupling, and subtract an estimated echo from the return signal.

Stereophonic acoustic echo cancellation (SAEC) is fundamentally different from traditional mono echo cancellation. Four mono echo cancelers straightforwardly implemented in the stereo case not only would have to track changing echo paths in the receiving room *but also in the transmission room*! For example, the canceler has to reconverge if one talker stops talking and another starts talking at a different location in the transmission room. There is no adaptive algorithm that can track such a change sufficiently fast and this scheme therefore results in poor echo suppression. Thus, a generalization of the mono AEC in the stereo case does not result in satisfactory performance.

The theory explaining the problem of SAEC is described in [1]. The fundamental problem is that the two channels may carry linearly related signals which in turn may make the normal equations to be solved by the adaptive algorithm singular. This implies that there is no unique solution to the equations but an infinite number of solutions and it can be shown that all (but the physically true) solutions depend on the transmission room. As a result, intensive studies have been made of how to handle this properly.

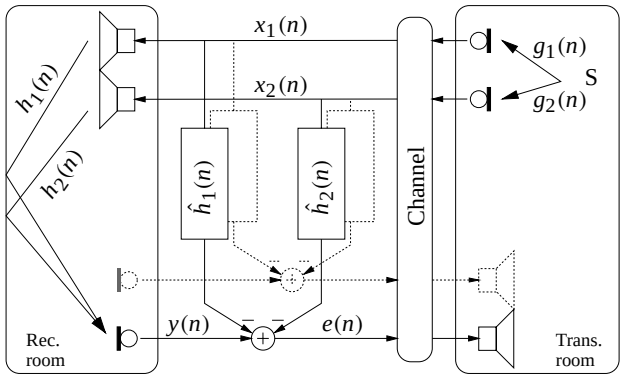

Figure 1: Schematic diagram of a stereophonic echo canceler.

A complete theory of non-uniqueness and characterization of the SAEC solution was presented in [2]. It was shown that the only solution to the non-uniqueness problem is to reduce the correlation between the stereo signals and an efficient low complexity method for this purpose was also given, [2]. Several decorrelation methods has since then been presented [3, 4, 5], and even if these decrease the correlation between the channels, the normal equations to be solved will still represent an ill-conditioned problem.

The performance of the SAEC is also more severely affected by the choice of algorithm than the monophonic counterpart. This is easily recognized since the performance of most adaptive algorithms depends on the condition number of the input signal. In the stereo case, the condition number is very high and algorithms that do not take the cross-correlation between the channels into account, such as the standard Least Mean Square (LMS) or Normalized LMS (NLMS), converge slowly to the true solution.

More sophisticated algorithms such as the APA (Affine Projection Algorithm) or RLS (Recursive Least Squares), that are less affected by a high condition number, handles the stereo case much better. This is even more true for algorithms that are specially derived for the two-channel situation. In the following we will study two different types of two-channel algorithms, the two-channel fast RLS algorithm, and a frequency domain adaptive algorithm.

Table 1 Two-channel FRLS algorithm for complex arithmetic. The transposition and the Hermitian transposition operators are denoted ^{T} and ^{H}, respectively, and the conjugate is denoted *. $\lambda \in (0, 1]$ is the forgetting factor and $k \in [1, 2.5]$ a stabilization parameter.

Input signals	**Matrix sizes**
$\boldsymbol{\chi}(n) = \begin{bmatrix} x_1(n) & x_2(n) \end{bmatrix}^T$	(2×1)
$\mathbf{x}(n) = \begin{bmatrix} \boldsymbol{\chi}^T(n) & \dots & \boldsymbol{\chi}^T(n-L+1) \end{bmatrix}^T$	$(2L \times 1)$
Prediction	
$\mathbf{e}_\mathrm{A}(n) = \boldsymbol{\chi}(n) - \mathbf{A}^H(n-1)\mathbf{x}(n-1)$	(2×1)
$\varphi_1(n) = \varphi(n-1) + \mathbf{e}_\mathrm{A}^H(n)\mathbf{E}_\mathrm{A}^{-1}(n-1)\mathbf{e}_\mathrm{A}(n)$	(1×1)
$\begin{bmatrix} \mathbf{M}(n) \\ \mathbf{m}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{G}(n-1) \end{bmatrix} + \begin{bmatrix} \mathbf{I}_2 \\ -\mathbf{A}(n-1) \end{bmatrix} \mathbf{E}_\mathrm{A}^{-1}(n-1)\mathbf{e}_\mathrm{A}(n)$	$((2L+2) \times 1)$
$\mathbf{e}_{\mathrm{B}_2}(n) = \boldsymbol{\chi}(n-L) - \mathbf{B}^H(n-1)\mathbf{x}(n)$	(2×1)
$\varphi(n) = \varphi_1(n) - \mathbf{e}_{\mathrm{B}_2}^H(n)\mathbf{m}(n)$	(1×1)
$\mathbf{A}(n) = \mathbf{A}(n-1) + \mathbf{G}(n-1)\mathbf{e}_\mathrm{A}^H(n)/\varphi(n-1)$	$(2L \times 2)$
$\mathbf{E}_\mathrm{A}(n) = \lambda[\mathbf{E}_\mathrm{A}(n-1) + \mathbf{e}_\mathrm{A}(n)\mathbf{e}_\mathrm{A}^H(n)/\varphi(n-1)]$	(2×2)
$\mathbf{G}(n) = \mathbf{M}(n) + \mathbf{B}(n-1)\mathbf{m}(n)$	$(2L \times 1)$
$\mathbf{e}_{\mathrm{B}_1}(n) = \mathbf{E}_\mathrm{B}(n-1)\mathbf{m}(n)$	(2×1)
$\mathbf{e}_\mathrm{B}(n) = k\mathbf{e}_{\mathrm{B}_2}(n) + (1-k)\mathbf{e}_{\mathrm{B}_1}(n)$	(2×1)
$\mathbf{B}(n) = \mathbf{B}(n-1) + \mathbf{G}(n)\mathbf{e}_\mathrm{B}^H(n)/\varphi(n)$	$(2L \times 2)$
$\mathbf{E}_\mathrm{B}(n) = \lambda[\mathbf{E}_\mathrm{B}(n-1) + \mathbf{e}_{\mathrm{B}_2}(n)\mathbf{e}_{\mathrm{B}_2}^H(n)/\varphi(n)]$	(2×2)
Filtering	
$e(n) = y(n) - \hat{\mathbf{h}}^H(n-1)\mathbf{x}(n)$	(1×1)
$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{G}(n)e^*(n)/\varphi(n)$	$(2L \times 1)$
Definition	
$\hat{\mathbf{h}}(n) = \begin{bmatrix} \hat{h}_{1,0}(n) & \hat{h}_{2,0}(n) & \cdots & \hat{h}_{1,L-1}(n) & \hat{h}_{2,L-1}(n) \end{bmatrix}^T$	

Table 2 Two-channel frequency domain adaptive filter, i denotes channel 1 or 2. The block size is denoted L, and the number of adaptive coefficients per frequency bin K, resulting in a total adaptive filter length of KL. Forgetting factor $0 < \beta < 1$, channel cross-correlation dependence $0 \le \rho \le 1$ and adaptive step size μ_b. $\mathbf{F}$ denotes the Fourier matrix.

Input signals	**Matrix sizes**
$\mathbf{X}_i(m) = \mathrm{diag}\left\{\mathbf{F}\begin{bmatrix} x_i(mL-L) & \cdots & x_i(mL+L-1) \end{bmatrix}^T\right\}$	$(2L \times 2L)$
$\mathbf{y}(m) = \begin{bmatrix} \mathbf{0}_{1\times L} & y(mL) & \cdots & y(mL+L-1) \end{bmatrix}^T$	$(2L \times 1)$
Power spectrum estimation with regularization	
$\mathbf{S}_{i,j}(m) = \beta\mathbf{S}_{i,j}(m-1) + (1-\beta)\mathbf{X}_i^*(m)\mathbf{X}_j(m)$	$(2L \times 2L)$
$\tilde{\mathbf{S}}_{i,i}(m) = \mathbf{S}_{i,i}(m) + \mathrm{diag}\left\{\mathbf{e}_\mathrm{reg}\right\}$	$(2L \times 2L)$
$\mathbf{S}_i(m) = \tilde{\mathbf{S}}_{i,i}(m)\left[\mathbf{I}_{2L\times 2L} - \rho^2\mathbf{S}_{1,2}^*(m)\mathbf{S}_{1,2}(m)\left\{\tilde{\mathbf{S}}_{1,1}(m)\tilde{\mathbf{S}}_{2,2}(m)\right\}^{-1}\right]$	$(2L \times 2L)$
Filtering	
$\underline{\hat{\mathbf{y}}}_i'(m) = \sum_{k=0}^{K-1} \mathbf{X}_i(m-k)\underline{\hat{\mathbf{h}}}_{i,k}(m-1)$	$(2L \times 1)$
$\mathbf{e}(m) = \mathbf{y}(m) - \mathbf{W}_1\mathbf{F}^{-1}\left[\underline{\hat{\mathbf{y}}}_1'(m) + \underline{\hat{\mathbf{y}}}_2'(m)\right]$	$(2L \times 1)$
$\underline{\hat{\mathbf{h}}}_{1,k}(m) = \underline{\hat{\mathbf{h}}}_{1,k}(m-1) + \mu_\mathrm{b}\mathbf{F}\mathbf{W}_2\mathbf{F}^{-1}\mathbf{S}_1^{-1}\cdot \left[\mathbf{X}_1^*(m-k) - \rho\mathbf{S}_{1,2}\tilde{\mathbf{S}}_{2,2}^{-1}\mathbf{X}_2^*(m-k)\right]\mathbf{F}\mathbf{e}(m)$	$(2L \times 1)$
$\underline{\hat{\mathbf{h}}}_{2,k}(m) = \underline{\hat{\mathbf{h}}}_{2,k}(m-1) + \mu_\mathrm{b}\mathbf{F}\mathbf{W}_2\mathbf{F}^{-1}\mathbf{S}_2^{-1}\cdot \left[\mathbf{X}_2^*(m-k) - \rho\mathbf{S}_{2,1}\tilde{\mathbf{S}}_{1,1}^{-1}\mathbf{X}_1^*(m-k)\right]\mathbf{F}\mathbf{e}(m)$	$(2L \times 1)$
Definitions	
$\mathbf{e}(m) = \begin{bmatrix} \mathbf{0}_{1\times L} & e(mL) & \cdots & e(mL+L-1) \end{bmatrix}^T$	$(2L \times 1)$
$\underline{\hat{\mathbf{h}}}_{i,k}(m) = \mathbf{F}\begin{bmatrix} \hat{\mathbf{h}}_{i,k}^T(m) & \mathbf{0}_{1\times L} \end{bmatrix}^T$	$(2L \times 1)$
$\hat{\mathbf{h}}_{i,k}(m) = \begin{bmatrix} \hat{h}_{i,kL}(m) & \cdots & \hat{h}_{i,kL+L-1}(m) \end{bmatrix}^T$	$(L \times 1)$
$\mathbf{W}_1 = \mathrm{diag}\left\{\begin{bmatrix} \mathbf{0}_{1\times L} & \mathbf{1}_{1\times L} \end{bmatrix}\right\}, \mathbf{W}_2 = \mathrm{diag}\left\{\begin{bmatrix} \mathbf{1}_{1\times L} & \mathbf{0}_{1\times L} \end{bmatrix}\right\}$	

2. ADAPTIVE ALGORITHMS

A few adaptive algorithms, which exploit the correlation between the two channels in SAEC, have been derived. In this section, we will investigate important properties of such algorithms. The algorithms that have been chosen are the two-channel fast recursive least mean squares (FRLS), Table 1 [6], applied in a subband scheme and the unconstrained versions of a two-channel frequency domain adaptive filter, Table 2 [7]. For comparison, the normalized least mean square (NLMS), is also studied. In Table 3 calculation complexity, signal path delay and memory usage of these algorithms are exemplified, and finally in Sec. 3, convergence and tracking properties are shown in simulations. All complexity and memory usage numbers are calculated for the full SAEC, i.e. with *four* adaptive filters and *two* return signals, as depicted in Fig. 1.

2.1. Normalized Least Mean Square

The NLMS is without competition the most common adaptive filter, and its strengths include robust behavior, and its structure allows for simple implementation. The error signal and the filter updates are calculated as

$$e(n) = y(n) - \mathbf{x}_1^H\hat{\mathbf{h}}_1(n-1) - \mathbf{x}_2^H\hat{\mathbf{h}}_2(n-1) \quad (1)$$

$$\hat{\mathbf{h}}_i(n) = \hat{\mathbf{h}}_i(n-1) + \frac{\mu}{\mathbf{x}_1^H\mathbf{x}_1 + \mathbf{x}_2^H\mathbf{x}_2}\mathbf{x}_i e(n), \quad i = 1, 2 \quad (2)$$

where $\mathbf{x}_i$ is the transmission room signal vector containing the last L samples for channel i, and $\hat{\mathbf{h}}_i(n)$ the filter estimate vector. The main disadvantage with this algorithm is the slow convergence rate on signals with large correlation matrix eigenvalue spread [8], as speech signals, and with correlated channels in the two channel situation. For the long adaptive filters needed in SAEC, calculation complexity is also significant. The number of real-valued multiplications per sample needed for the four adaptive filters depicted in Fig. 1 is $8L$ for real-valued signals and $32L$ for complex-valued signals. The calculation complexity can be reduced by applying the adaptive filters in a subband structure, but this will also introduce signal transmission delay to the otherwise delayless NLMS algorithm. This is described in Sec. 2.4.

2.2. Two-channel Fast Recursive Least Mean Squares

A complete analysis of the two-channel FRLS algorithm, Table 1, is beyond the scope of this paper. However, a general analysis of

the RLS algorithm can be found in [8] and stabilized two-channel versions are described in [6, 9, 10].
In contrast to the NLMS algorithm, estimates of the signal correlation and the cross-correlation between the two channels are incorporated in the Kalman gain vector **G**, and used in the update of the adaptive filter $\hat{\mathbf{h}}$. This improves the convergence rate significantly over the NLMS algorithm, but it is also the cause of the well-known instability problem of RLS type algorithms.
In this version, stability is improved with the stability parameter k. But for operation on non-stationary signals, like speech-signals, further enhancements are needed. First of all, by monitoring φ, it is possible to detect if the algorithm is about to become unstable. If this is the case, the parameters in the prediction part are reset to a suitable start value, while the adaptive filter estimate, $\hat{\mathbf{h}}$, can be left unchanged. Secondly, the algorithm can be applied in a two-path structure [11], where the adaptive filter $\hat{\mathbf{h}}$ is copied to a static filter, if it is a better estimate than a previous copy. The static filter is then used for the actual filtering.
Another disadvantage with the two-channel FRLS in most situations is the high calculation complexity. The number of real-valued multiplications per sample needed for the four adaptive filters depicted in Fig. 1 is $32L$ and $128L$ for complex-valued signals. Most likely this algorithm will be used in subband structures, in order to decrease the complexity, but also to improve stability by having shorter adaptive filters. Like for the NLMS, this will introduce signal transmission delay.

2.3. Two-channel Frequency Domain Algorithm

Frequency domain algorithms perform better than the NLMS on signals with large correlation matrix eigenvalue spread, like speech signals. In the adaptive filter update of the NLMS, (2), the same normalization factor, $\mathbf{x}_1^H\mathbf{x}_1+\mathbf{x}_2^H\mathbf{x}_2$, is used for all frequencies, independent of the spectral characteristic of the excitation signals. In the two-channel frequency domain filter, Table 2, each frequency bin has an independent normalization factor. In the update of $\hat{\underline{\mathbf{h}}}_{1,k}$, in Table 2, this normalization is represented by the spectrum estimate vector $\mathbf{S}_1^{-1}$. In the same equation, the channel cross-correlation is taken into account by the term $\rho\mathbf{S}_{1,2}\tilde{\mathbf{S}}_{2,2}^{-1}\mathbf{X}_2^*(m-k)$, where the parameter ρ controls the amount of cross-correlation in the adaptation. This improves the performance when the two channels x_1 and x_2 are correlated. $\tilde{\mathbf{S}}_{i,i}$, $i=1,2$, are regularized versions of $\mathbf{S}_{i,i}$, since this greatly decreases the misalignment in poorly excited frequency bands.
The algorithm given in Table 2 is, except for the regularization, an exact derivative of a time-domain block LMS, and is denoted the constrained version. Large complexity reduction is possible by replacing $\mathbf{F}\mathbf{W}_2\mathbf{F}^{-1}$ with the identity matrix, without significant reduction of the performance [7], and this version is denoted unconstrained. For the unconstrained version, the optimal step-size parameter is $\mu_b = 2(1-\beta)$. In the algorithm in Table 2, the input data blocks have no overlap ($\alpha_o = 1$) . By overlapping the input data, it is possible to increase the convergence rate, and in the simulation section, we use only 25% new input-data for each iteration ($\alpha_o = 4$). It is also possible to have several adaptive filter taps per frequency bin, however in this paper, we will only consider the situation with 1 tap per frequency bin, i.e. $K = 1$.
For $\alpha_o = 1$, the unconstrained version of the algorithm in Table 2, with two return signals (e_1,e_2), requires 3 FFTs on vectors with $2L$ elements plus $148L$ real-valued multiplications and $8L$ real-valued divisions per L samples. The fact that 2 real-valued FFTs can be calculated with one complex-valued FFT, have been used. For the overlapped version, the complexity is approximately increased with the factor α_o. The signal transmission delay is equal to the block size L. In contrast to the FRLS, this algorithm is very robust. A complete analysis of the algorithm is given in [7].

Table 3 Calculation complexity as the number of real-valued multiplication per sample, the transmission delay in ms and the memory usage in words, for the algorithms used in the Simulation section.

	NLMS	FRLS	NLMS in subbands	FRLS in subbands	Frequency LMS
complex.	8.2k	33k	0.72k	2.6k	1.0k
delay (ms)	0	0	64	64	64
memory	6.1k	18k	17k	40k	66k

2.4. Adaptive Filtering In Filterbank Structures

In subband structures, an analysis filterbank divides a signal into M down-sampled subband signals, which each represent a frequency region. Then the adaptive filtering is performed on the subband signals, and the fullband residual echo signal $e(n)$ is finally reconstructed with a synthesis filterbank. Since down-sampling aliasing has a very negative effect on the convergence of the adaptive filter [12], the downsampling factor r is usually less than M. For FFT based filterbanks we will only need to have $M/2+1$ complex-valued adaptive filters since the $M/2-1$ upper subbands differ only by the conjugate from the lower counterpart. The calculation complexity reduction comes from the fact that each adaptive filter is r times shorter and only updated once for each r fullband sample.
The biggest disadvantage with a filterbank is the introduced signal path delay. For linear phase filterbanks, this delay is equal to the length of the prototype filter, which depends on several factors, including: aliasing, attenuation, and the ratio r/M. For 60 dB stopband suppression and $r/M = 0.75$ we need approximately $11M$ filter coefficients. In addition to this we need approximately 5 non-causal taps per subband [13]. If the the inherent flat delay in the receiving room is less, we can artificially delay the signal y.
Four analysis filterbanks, decomposing x_1, x_2, y_1, y_2 in Fig. 1, and two analysis filterbanks, e_1, e_2, are needed. The number of real-valued multiplication for the four analysis filterbanks is $\frac{1}{r}(4K + 4M\log_2 M - 14M + 24)$ per sample, and $\frac{1}{r}(2K + 2M\log_2 M - 7M + 12)$ for the two synthesis filterbanks, where K is the filter length [10].

3. SIMULATIONS

The convergence rate and the ability of the algorithm to track echo path changes highly effects the performance of SAECs. In this section we will exemplify these properties with real-life data recorded in a quiet office-like room. Both the recordings and the simulations are performed with 16 kHz sampling rate. In order to reduce the correlation between the two channels, the data was processed with a non-linear function, [2] $\alpha_n = 0.5$. As performance index the mean square error (MSE) energy of the residual is used. The MSE is given by,

$$\text{MSE} = \frac{P_e}{P_y}, \qquad P_e = \text{LPF}\Big[e^2(n)\Big], \tag{3}$$

where LPF denotes a lowpass filter; in this case it has a single pole in 0.9996. Four different systems are used in the simulation. The classical NLMS algorithm, with each adaptive filter having a length of 1024 taps. Both the NLMS and the two-channel FRLS algorithms in a subband structure, with 64 subbands and a downsample factor of 48. Each adaptive filter have 28 taps, corresponding to 1024 fullband taps plus 320 non-causal adaptive filter taps. Finally, the frequency

based algorithm has a block-size $L = 1024$, $K = 1$ filter taps per frequency bin and the input-data overlap factor $\alpha_o = 4$.

In Fig. 2 the initial convergence and the tracking ability of the algorithms are shown. There is an instantaneously change of *the receiving room impulse responses*, $h_i(t)$, after 20 s. Both the FRLS in subbands and the frequency domain algorithm perform very well.

In Fig. 3 there is an instantaneously change of *the transmission room impulse responses*, $g_i(t)$, after 5.1 s. As can be seen in the figures, the two algorithms specially derived for the two-channel situation are less affected than the other two.

Finally in Fig. 4, it is shown that even if the fullband NLMS appear to perform reasonably well in Fig. 2 and 3, it does only perform well in the region with the most signal energy, 200–2500 Hz. As the human ear is sensitive also for higher frequencies, Fig. 4 tells us more about the perceptual quality than previous figures.

In order to study the stability of the FRLS algorithm, 24 hours of real-life data was processed. On average each subband was restarted every 18 s. These restarts are handled with little effect on the performance. In Fig. 2 there are 33 restarts, on average one per subband, and 13 of them occur in the time-interval 16–19 s.

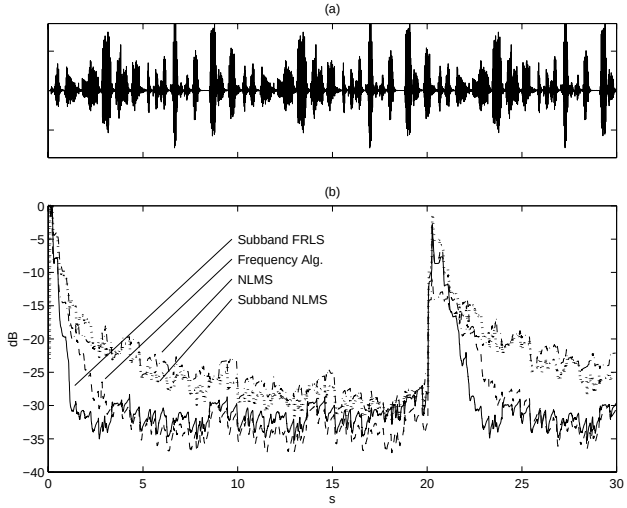

Figure 2: (a) Transmission room speech signal (left channel). Instantaneous receiving room impulse responses, $h_i(t)$, change after 20 s. (b) Mean square error performance for four adaptive algorithms: NLMS (dash-dotted line), subband NLMS (dotted line), two-channel FRLS (solid line), two-channel frequency algorithm (dashed line).

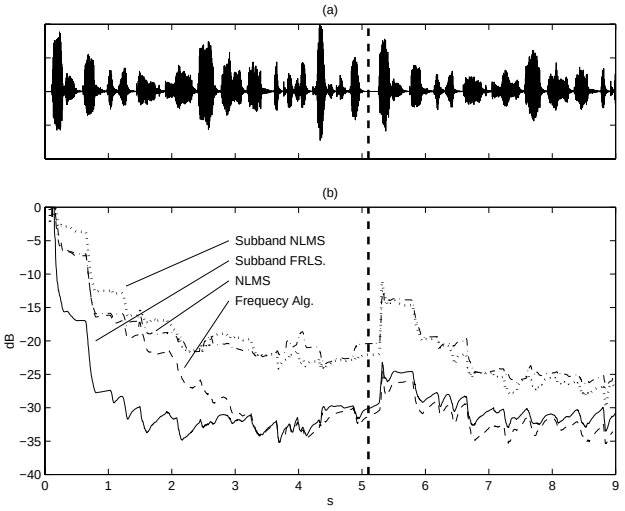

Figure 3: Instantaneous transmission room impulse response, $g_i(t)$, change after 5.1 s. Other conditions as in Fig. 2.

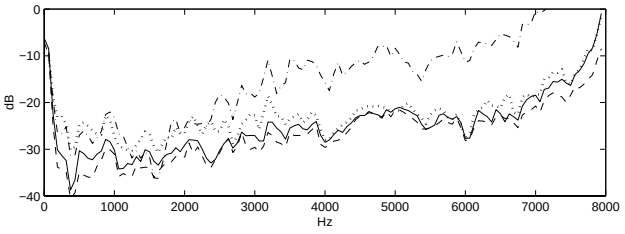

Figure 4: Echo canceler suppression in the frequency domain. Line types as in Fig. 2.

4. REFERENCES

[1] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic acoustic echo cancellation—an overview of the fundamental problem," *IEEE Signal Processing Lett.*, vol. 2, no. 8, pp. 148–151, August 1995.

[2] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. Speech and Audio Processing*, vol. 6, no. 2, pp. 156–165, March 1998.

[3] T. Gänsler and P. Eneroth, "Influence of audio coding on stereophonic acoustic echo cancellation," in *Proc. IEEE ICASSP*, 1998, pp. 3649–3652.

[4] A. Gilloire and V. Turbin, "Using auditory properties to improve the behavior of stereophonic acoustic echo cancellers," in *Proc. IEEE ICASSP*, 1998, pp. 3681–3684.

[5] S. Shimauchi, Y. Haneda, S. Makino, and Y Kaneda, "New configuration for a stereo echo canceller with nonlinear preprocessing," in *Proc. IEEE ICASSP*, 1998, pp. 3685–3688.

[6] J. Benesty, F. Amand, A. Gilloire, and Y. Grenier, "Adaptive filtering algorithms for stereophonic acoustic echo cancellation," in *Proc. IEEE ICASSP*, 1995, pp. 3099–3102.

[7] J. Benesty, A. Gilloire, and Y. Grenier, "A frequency domain stereophonic acoustic echo canceler exploiting the coherence between the channels," *The Journal of the Acoustical Society of America*, vol. 106, pp. L30–L35, Sept 1999.

[8] S. Haykin, *Adaptive Filter Theory*, chapter 9, Prentice Hall International, 3rd edition, 1996.

[9] M. G. Bellanger, *Adaptive Digital Filters and Signal Analysis*, Marcel Dekker, 1987.

[10] S. L. Gay and J. Benesty, Eds., *Acoustic Signal Processing For Telecommunication*, vol. 551, chapter 8, Kluwer Academic Publishers, 2000, ISBN 0-7923-7814-8.

[11] K. Ochiai, T. Araseki, and T. Ogihara, "Echo cancellation with two path models," *IEEE Trans. on Communications*, vol. COM-25, no. 6, pp. 589–595, June 1977.

[12] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. on Signal Processing*, vol. 40, no. 8, pp. 1862–1875, August 1992.

[13] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustical echoes," in *Proc. of ICASSP*, 1988, pp. 2570–2573.