MULTI-MICROPHONE NOISE REDUCTION USING GSVD-BASED OPTIMAL FILTERING WITH ANC POSTPROCESSING STAGE

Simon Doclo, Marc Moonen

K.U.Leuven, Dept. of Elec. Engineering, SISTA
K. Mercierlaan 94, 3001 Leuven, Belgium
{doclo,moonen}@esat.kuleuven.ac.be

Erik De Clippel

Philips ITCL
Interleuvenlaan 74-76, 3001 Leuven, Belgium
erik.de.clippel@philips.com

ABSTRACT

Recently a signal enhancement technique, based on a generalized singular value decomposition (GSVD), has been proposed for acoustic noise reduction in multi-microphone speech signals. In this paper a postprocessing stage for this GSVD-based signal enhancement technique is discussed. Using the GSVD-based optimal filtering technique it is possible to create a 'speech reference' and a 'noise reference'. These references can then be used in an adaptive noise cancellation (ANC) algorithm.

It is shown that this ANC postprocessing stage gives rise to an additional SNR-improvement. Simulations show that the GSVD-based optimal filtering technique with ANC postprocessing stage has a better noise reduction performance than standard adaptive beamforming algorithms for all reverberation times.

1. INTRODUCTION

In many speech communication applications, like hands-free mobile telephony and audio-conferencing, the recorded speech signals are corrupted by acoustic background noise and echo signals (see figure 4). This causes a signal degradation which can lead to total unintelligibility of the speech and which decreases the performance of speech coding and speech recognition devices. Therefore efficient noise and echo reduction techniques are called for.

Recently a signal enhancement technique, based on a generalized singular value decomposition (GSVD), has been proposed, which amounts to a specific optimal filtering technique for the case where the so-called 'desired response' signal cannot be observed. The optimal filter can be written as a function of the generalized singular vectors and singular values of a speech and noise data matrix [1][2][3]. In [4] it has been shown that the computational complexity of this technique can be reduced by using recursive and approximate GSVD-updating algorithms and by using downsampling techniques. The GSVD-based optimal filtering technique is briefly reviewed in section 2.

Although this GSVD-based optimal filtering technique reduces a considerable amount of noise, its noise reduction performance can be improved by adding a postprocessing stage. The postprocessing stage closely resembles the structure of a generalized sidelobe canceller, using a 'speech reference' and a 'noise reference' in an adaptive noise cancellation (ANC) algorithm. The output of the GSVD-based optimal filtering technique is used as 'speech reference', whereas different possibilities exist for creating a 'noise reference'. This ANC postprocessing stage is discussed in section 3.

Section 4 describes the used simulation environment and in section 5 the noise reduction performance of the ANC postprocessing stage is analyzed for different parameters of the algorithm. In section 6 the performance of the GSVD-based optimal filtering technique with the ANC postprocessing stage is compared with standard beamforming algorithms (delay-and-sum beamformer, generalized sidelobe canceller), showing that the SNR-improvement is better for all reverberation times.

2. GSVD-BASED OPTIMAL FILTERING

2.1. General case

The GSVD-based optimal filtering technique [2] considers problems where the observed signal vector $\mathbf{u}_k \in \mathbb{R}^N$ contains a signal-of-interest $\mathbf{s}_k \in \mathbb{R}^N$ (*e.g.* a speech signal) and an additive noise term $\mathbf{n}_k \in \mathbb{R}^N$, such that $\mathbf{u}_k = \mathbf{s}_k + \mathbf{n}_k$.

If we consider speech applications and use a robust speech-noise detection algorithm [5], noise-only observations $\mathbf{n}_{k'}$ can be made during speech pauses. Our goal is to reconstruct the signal-of-interest $\mathbf{s}_k$ from $\mathbf{u}_k$ by means of a linear filter $\mathbf{W} \in \mathbb{R}^{N \times N}$, using $\hat{\mathbf{s}}_k = \mathbf{u}_k^T \mathbf{W}$. It can be shown that when optimizing a MMSE-criterion and making some statistical assumptions, the optimal filter $\mathbf{W}_{WF}^{[k]}$ at time k is equal to

$$\mathbf{W}_{WF}^{[k]}=\mathcal{E}\left\{\mathbf{u}_k\cdot\mathbf{u}_k^T\right\}^{-1}\left(\mathcal{E}\left\{\mathbf{u}_k\cdot\mathbf{u}_k^T\right\}-\mathcal{E}\left\{\mathbf{n}_k\cdot\mathbf{n}_k^T\right\}\right). \tag{1}$$

In practice this filter is computed by means of a generalized singular value decomposition (GSVD) [6][7] of a speech data matrix $\mathbf{A}_{[k]} \in \mathbb{R}^{p\times N}$, containing p speech vectors, and a noise data matrix $\mathbf{B}_{[k]} \in \mathbb{R}^{q\times N}$, containing q noise vectors,

$$\mathbf{A}_{[k]} = \begin{bmatrix} \mathbf{u}_{k-p+1}^T \\ \vdots \\ \mathbf{u}_{k-1}^T \\ \mathbf{u}_k^T \end{bmatrix} \quad \mathbf{B}_{[k]} = \begin{bmatrix} \mathbf{n}_{k-q+1}^T \\ \vdots \\ \mathbf{n}_{k-1}^T \\ \mathbf{n}_k^T \end{bmatrix}. \tag{2}$$

At time k, the GSVD of the two matrices $\mathbf{A}_{[k]}$ and $\mathbf{B}_{[k]}$ is defined as

$$\begin{cases} \mathbf{A}_{[k]} = U_{A[k]}\cdot\Sigma_{A[k]}\cdot X_{[k]}^T \\ \mathbf{B}_{[k]} = U_{B[k]}\cdot\Sigma_{B[k]}\cdot X_{[k]}^T, \end{cases} \tag{3}$$

with $\Sigma_{A[k]} = \text{diag}\{\sigma_{i[k]}\}$, $\Sigma_{B[k]} = \text{diag}\{\eta_{i[k]}\}$, $U_{A[k]}$ and $U_{B[k]}$ orthogonal matrices, $X_{[k]}$ an invertible (but not necessarily orthogonal) matrix and $\frac{\sigma_{i[k]}}{\eta_{i[k]}}$ the generalized singular values. Substituting these formulas into (1) gives

$$\mathbf{W}_{WF}^{[k]} = X_{[k]}^{-T}\cdot\text{diag}\{1-\frac{p}{q}\frac{\eta_{i[k]}^2}{\sigma_{i[k]}^2}\}\cdot X_{[k]}^T \tag{4}$$

In fact the filter $\mathbf{W}_{WF}^{[k]}$ belongs to a more general class of estimators, which can be described by

$$\mathbf{W}^{[k]} = X_{[k]}^{-T}\cdot\text{diag}\{f(\sigma_{i[k]}^2,\eta_{i[k]}^2)\}\cdot X_{[k]}^T. \tag{5}$$

In this paper we will use the following gain function

$$f(\sigma_{i[k]}^2,\eta_{i[k]}^2) = 1-\alpha\cdot\frac{p}{q}\frac{\eta_{i[k]}^2}{\sigma_{i[k]}^2}, \tag{6}$$

where the factor α is a noise overestimation factor (when $\alpha = 1$ we obtain the MMSE estimator). By increasing the factor α the SNR of the enhanced signal increases, but some signal distortion is introduced. Therefore the factor α will have to be limited.

2.2. Multi-channel time series filtering

Consider M microphones where each microphone signal $m_j(k)$, $j = 1\ldots M$, consists of a filtered version of the speech signal and an additive noise term,

$$m_j(k) = s_j(k) + n_j(k) = h_j^s(k)\otimes s(k) + n_j(k). \tag{7}$$

The vector $\mathbf{u}_k \in \mathbb{R}^N$ with $N = LM$ now takes the form

$$\mathbf{u}_k = \begin{bmatrix} \mathbf{m}_{1k} & \mathbf{m}_{2k} & \ldots & \mathbf{m}_{Mk} \end{bmatrix}^T \tag{8}$$
$$\mathbf{m}_{jk} = \begin{bmatrix} m_j(k) & m_j(k-1) & \ldots & m_j(k-L+1) \end{bmatrix}. \tag{9}$$

The enhanced speech signal $\hat{s}(k)$ is then computed as

$$\hat{\mathbf{s}}_k = \begin{bmatrix} \hat{s}(k-p+1) & \ldots & \hat{s}(k-1) & \hat{s}(k) \end{bmatrix}^T = \mathbf{A}_{[k]}\cdot\mathbf{w}_{WF[k]}^i,$$

with $\mathbf{w}_{WF[k]}^i$ the i^{th} column of $\mathbf{W}_{WF}^{[k]}$. This can be considered a multi-channel filtering operation, where each of the M microphone signals is filtered with an L-taps FIR-filter (figure 1).

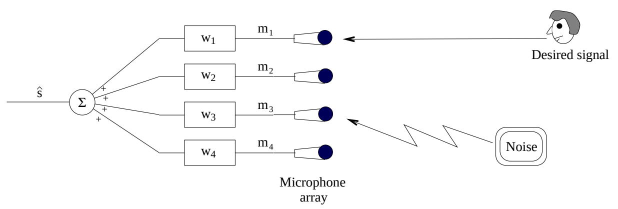

Figure 1: GSVD-based optimal filtering technique

When filtering with $\mathbf{w}_{WF[k]}^i$, the enhanced signal $\hat{s}(k)$ is in fact an optimal estimate for the speech component $s_l(k-\Delta)$ in the l^{th} microphone signal, with

$$l = mod(i-1,L)+1 \tag{10}$$
$$\Delta = rem(i-1,L) \tag{11}$$

The elements on the diagonal of the error covariance matrix $\{\mathcal{E}\{\mathbf{e}_k\cdot\mathbf{e}_k^T\}\}_{ii}$, with $\mathbf{e}_k = \mathbf{s}_k - \mathbf{u}_k^T\mathbf{W}_{WF}^{[k]}$, indicate how well the i^{th} component of $\mathbf{s}_k$ is estimated. The smallest element on the diagonal of this matrix therefore corresponds to the best estimator, which is the corresponding column of $\mathbf{W}_{WF}^{[k]}$. However simulations indicate that e.g. taking $i = \frac{L}{2}$ instead of the best value does not decrease the performance considerably.

2.3. Computational complexity

Since in each time step new data vectors $\mathbf{u}_k$ or $\mathbf{n}_k$ are appended to the speech or noise matrix, the GSVD of $\mathbf{A}_{[k]}$ and $\mathbf{B}_{[k]}$ and the optimal filter $\mathbf{W}_{WF}^{[k]}$ need to be recomputed. Instead of recomputing the GSVD from scratch at each time step, recursive GSVD-updating algorithms can be used which compute the GSVD at time k using the decomposition at time $k-1$ [8][9]. The total computational complexity of the algorithm can be further reduced by using a square root-free implementation for the GSVD-updates and by using downsampling techniques [4]. In this context downsampling means that the GSVD of $\mathbf{A}_{[k]}$ and $\mathbf{B}_{[k]}$ and the optimal filter $\mathbf{W}_{WF}^{[k]}$ are not updated for every sample, but only every d samples (d typically $10\ldots20$).

3. ANC POSTPROCESSING STAGE

Although the GSVD-based optimal filtering technique reduces a considerable amount of noise, its noise reduction performance can be improved by adding a postprocessing stage (see figure 2). The postprocessing stage is a structure widely used in adaptive beamformers, using a 'speech reference' and a 'noise reference' in an adaptive noise cancellation (ANC) algorithm. The adaptive filter will remove the correlation between the noise reference and the speech reference. For this algorithm to work properly, the noise reference therefore should be (highly) correlated with the noise still present in the speech reference. In order to avoid signal cancellation, signal leakage into the noise reference also has to be limited. Since in most cases this signal leakage cannot be completely avoided, the adaptive filter is only allowed to adapt during periods when no speech is present [10][11].

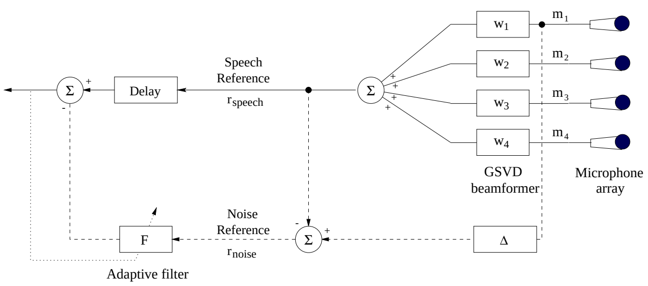

Figure 2: GSVD-based optimal filtering technique with ANC postprocessing stage

A well known adaptive beamformer, depicted in figure 3, is the generalized sidelobe canceller (GSC) [12][13][14]. A GSC uses the output of a fixed delay-and-sum beamformer as speech reference and creates noise references by combining the delayed microphone signals using a blocking matrix. A multi-channel adaptive filter (NLMS, APA, RLS) then removes the correlation between the noise references and the speech reference.

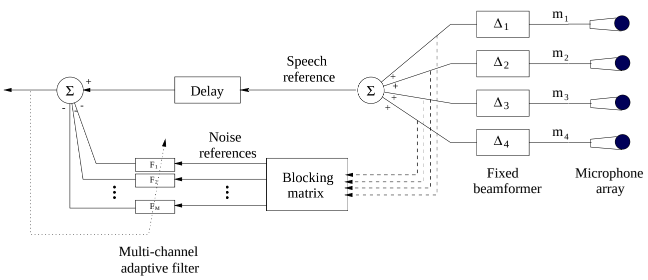

Figure 3: Generalized Sidelobe Canceller

In the ANC postprocessing stage of the GSVD-based optimal filtering technique (figure 2), the output of the GSVD-based optimal filter is used as speech reference. For the creation of a noise reference different possibilities exist. Define the speech reference $r^{l,\Delta}_{speech}(k)$ as the optimal estimate of the speech component in the l^{th} microphone signal (delayed with Δ), such that

$$r^{l,\Delta}_{speech}(k) = \hat{s}_l(k-\Delta). \tag{12}$$

This speech reference is created by using $\mathbf{w}^i_{WF[k]}$, with $i = (l-1)L + \Delta$. An obvious choice for creating a noise reference is simply subtracting the speech reference $r^{l,\Delta}_{speech}(k)$ from the delayed l^{th} microphone signal,

$$r^1_{noise}(k) = m_l(k-\Delta) - r^{l,\Delta}_{speech}(k). \tag{13}$$

Indeed, if $\mathbf{W}^{[k]}_{WF}$ is the optimal filter matrix for estimating the signal component $\mathbf{s}_k$, then $(I - \mathbf{W}^{[k]}_{WF})$ is the optimal filter matrix for estimating the noise component $\mathbf{n}_k$. This noise reference is depicted in figure 2. In the rest of the paper we will use this kind of noise reference and we will take $i = \frac{L}{2}$ (such that $l = 1$ and $\Delta = \frac{L}{2} - 1$)
However different possibilities for creating noise references exist. Instead of only calculating a 'noise reference' for the l^{th} microphone signal, one can calculate and use the 'noise references' for all the microphone signals, *i.e.*

$$r^2_{noise}(k) = \begin{bmatrix} m_1(k-\Delta) - r^{1,\Delta}_{speech}(k) \\ m_2(k-\Delta) - r^{2,\Delta}_{speech}(k) \\ \vdots \\ m_M(k-\Delta) - r^{M,\Delta}_{speech}(k) \end{bmatrix} \tag{14}$$

However since different 'speech references' are needed for the calculation of $r^2_{noise}(k)$ (only $r^{l,\Delta}_{speech}(k)$ is readily available), the microphone signals have to be filtered with different filters $\mathbf{w}^i_{WF[k]}$, implying increased computational complexity.
Another possibility for creating a noise reference would consist of using a blocking matrix on the filtered microphone signals (analogous to the generalized sidelobe canceller).

4. SIMULATION ENVIRONMENT

The used simulation environment is depicted in figure 4. It consists of a microphone array, a speech source $s(k)$ and background noise source $n(k)$. In our simulations the linear equi-spaced microphone array has a maximum number of $M = 6$ microphones and the distance between two adjacent microphones is 5 cm. The

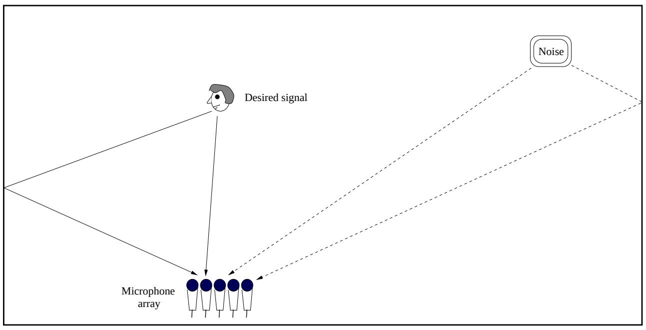

Figure 4: Simulation environment

speech source is located in front of the microphone array. The signals used are an 8 kHz clean speech signal and a temporally white noise source. In these simulations no far-end echo source is present, but in [15] it is shown how a far-end echo source can be easily incorporated into the algorithms.

Each microphone signal $m_j(k)$, $j = 1 \dots M$, consists of a filtered version of the speech and noise signal,

$$m_j(k) = h_j^s(k) \otimes s(k) + h_j^n(k) \otimes n(k), \qquad (15)$$

with $h_j^s(k)$ the impulse response between the speech source and the j^{th} microphone and $h_j^n(k)$ similarly defined for the noise source. The room impulse responses are obtained using the image method [16], with a filterlength of 1500 taps and for different reverberation times T_{60}. The reverberation time T_{60} can be expressed as a function of the reflection coefficient γ of the walls, according to Eyring's formula,

$$T_{60} = \frac{0.163V}{-S\log(1-\gamma)}, \qquad (16)$$

with V the volume of the room and S the total surface of the room. Simulations have been performed for a different number of microphones at different signal-to-noise ratios.

Since we are using simulations (and hence know the speech and noise components at each stage of the algorithm), the unbiased signal-to-noise ratio (SNR) and noise reduction (NR) can be computed as

$$\text{SNR} = 10\log_{10}\frac{\sum \tilde{s}^2(k)}{\sum \tilde{n}^2(k)}, \quad \text{NR} = 10\log_{10}\frac{\sum \tilde{n}^2(k)}{\sum n_1^2(k)}, \qquad (17)$$

where $\tilde{s}(k)$ and $\tilde{n}(k)$ correspond to the speech and noise component of the considered signal and $n_1(k)$ is the noise component of the first microphone signal.

The performance of the following algorithms is compared: delay-and-sum beamformer, generalized sidelobe canceller (GSC) and recursive GSVD-based optimal filtering technique (with and without ANC postprocessing stage). The adaptive filter used in the GSC as well as in the ANC postprocessing stage of the GSVD-based optimal filtering technique is a time-domain NLMS algorithm. The filterlength of the adaptive filter is denoted by L_{NLMS} and the step size is $\mu = 0.1$. The desired signal of the adaptive filter is delayed with $\frac{L_{NLMS}}{2}$ in order for the adaptive filter to model some acausal taps. The noise reference for the GSC is calculated as

$$r_{noise}^{GSC}(k) = (M-1)\cdot m_1(k-\Delta_1) - \sum_{j=2}^{M} m_j(k-\Delta_j). \qquad (18)$$

The filterlength for the GSVD-based optimal filtering technique is denoted by L_{GSVD}.

5. NOISE REDUCTION PERFORMANCE OF THE ANC POSTPROCESSING STAGE

In figure 5 the unbiased SNR of the enhanced signal (output of the ANC postprocessing stage) is plotted in function of L_{GSVD} and L_{NLMS} for 2 values of the factor α (see equation 6). This figure shows that the SNR of the enhanced signal improves with increasing L_{GSVD} and L_{NLMS}. Since the total computational complexity is $\mathcal{O}(L_{GSVD}^2) + \mathcal{O}(L_{NLMS})$, it is sometimes better to use relatively small filterlengths L_{GSVD} for the GSVD-based optimal filter technique and relatively large filterlengths L_{NLMS} for the adaptive filter in the ANC postprocessing stage.

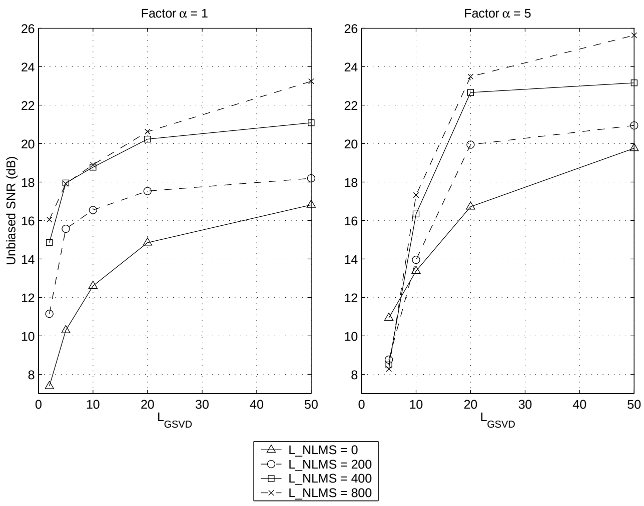

Figure 5: Dependence of SNR of enhanced signal on GSVD-filterlength L_{GSVD}, NLMS-filterlength L_{NLMS} and factor α ($M = 4$, SNR = 0 dB, $T_{60} = 200$ ms)

This figure also shows that the SNR of the enhanced signal improves with increasing factor α. This factor α can be considered a noise overestimation factor. For higher values of the factor α the SNR of the enhanced signal improves, but also signal distortion is introduced since the MMSE-criterion is not optimized any more. Therefore the factor α has to be limited (value depends on the number of channels M and on the filterlength L_{GSVD}), otherwise the speech intelligibility drops.

6. COMPARISON WITH STANDARD BEAMFORMING TECHNIQUES

Figure 6 compares the performance (unbiased SNR of the enhanced signal) for the GSVD-based optimal filtering technique (with and without ANC postprocessing stage) and standard beamforming algorithms (delay-and-sum beamformer and GSC). We use $M = 4$ microphones, the SNR of the noisy microphone signals is 0 dB, the filterlength L_{NLMS} of the adaptive NLMS-filter is 800 (for the GSC as well as for the ANC postprocessing stage), and for the GSVD-based optimal filtering technique the filterlength $L_{GSVD} = 20$ and $\alpha = 1$. The comparison is performed for different reverberation times T_{60} of the room. Low reverberation times correspond to highly correlated noise, while high reverberation times correspond to highly uncorrelated (diffuse) noise.
Figure 6 shows that for low T_{60} the GSC performs better than the GSVD-based optimal filtering technique (without ANC postprocessing stage), while for high T_{60} the GSVD-based optimal filtering technique performs better than than the GSC. However, when adding the ANC postprocessing stage, the GSVD-based optimal filtering technique clearly outperforms the GSC for all reverberation times.
Figure 7 shows the same comparison for different signal-to-noise ratios of the noisy microphone signals. One can observe that for higher SNRs, the difference in performance between the GSC and the GSVD-based optimal filtering technique with ANC postprocessing stage becomes smaller. However the same conclusion still holds that GSVD-based optimal filtering technique with ANC postprocessing stage outperforms the GSC for all reverberation times.

7. CONCLUSION

In this paper we have described how a GSVD-based optimal filtering technique can create a 'speech reference' and 'noise reference', which can then be used in an adaptive noise cancellation (ANC) algorithm. This ANC postprocessing stage gives rise to an additional SNR-improvement for the GSVD-based optimal filtering technique. We have also shown that this GSVD-based optimal filtering technique with ANC postprocessing stage outperforms the standard generalized sidelobe canceller for all reverberation times and signal-to-noise ratios.

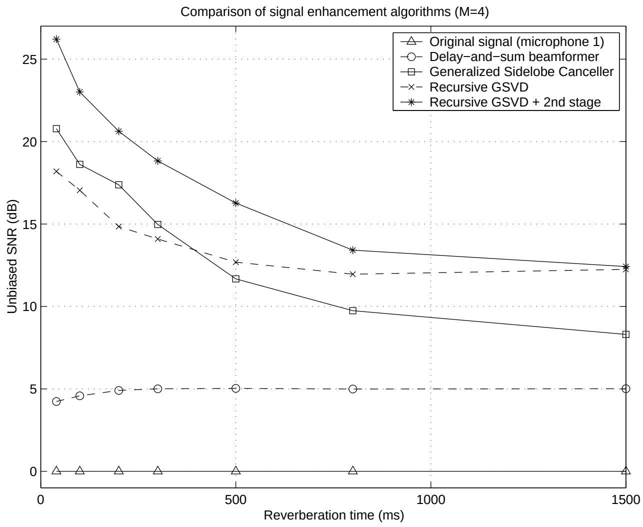

Figure 6: Comparison of unbiased SNR of enhanced signal for different signal enhancement algorithms for different reverberation times ($M = 4$, SNR = 0 dB, $L_{GSVD} = 20$, $L_{NLMS} = 800$, $\alpha = 1$)

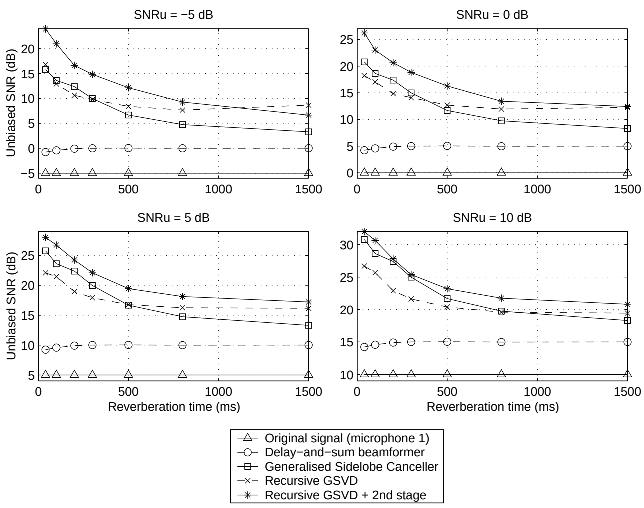

Figure 7: Comparison of unbiased SNR of enhanced signal for different signal enhancement algorithms for different reverberation times and signal-to-noise ratios ($M = 4$, $L_{GSVD} = 20$, $L_{NLMS} = 800$, $\alpha = 1$)

ACKNOWLEDGEMENTS

Simon Doclo is a Research Assistant supported by the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry). Marc Moonen is a Research Associate with the F.W.O.-Vlaanderen (Fund for Scientific Research-Flanders). This research work was carried out in the frame of the Concerted Research Action *Mathematical Engineering Techniques for Information and Communication Systems* of the Flemish Government, the F.W.O. Research Project *Design and implementation of adaptive digital signal processing algorithms for broadband applications*, the Interuniversity Attraction Pole IUAP P4-02, *Modeling, Identification, Simulation and Control of Complex Systems*, and the IT-project *Multimicrophone Signal Enhancement Techniques for handsfree telephony and voice controlled systems (MUSETTE)* of the I.W.T., and was partially sponsored by Philips ITCL.

REFERENCES

[1] S. H. Jensen, P. C. Hansen, S. D. Hansen, and J. A. Sørensen, "Reduction of Broad-Band Noise in Speech by Truncated QSVD," *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 6, pp. 439–448, Nov. 1995.

[2] S. Doclo and M. Moonen, "SVD-based optimal filtering with applications to noise reduction in speech signals," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, USA, Oct. 1999, pp. 143–146.

[3] S. Doclo and M. Moonen, "Robustness of SVD-based Optimal Filtering for Noise Reduction in Multi-Microphone Speech Signals," in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Pocono Manor, Pennsylvania, USA, Sept. 1999, pp. 80–83.

[4] S. Doclo and M. Moonen, "Noise Reduction in Multi-Microphone Speech Signals using Recursive and Approximate GSVD-based Optimal Filtering," in *Proc. of the IEEE Benelux Signal Processing Symposium (SPS2000)*, Hilvarenbeek, The Netherlands, Mar. 2000.

[5] S. Van Gerven and F. Xie, "A Comparative Study of Speech Detection Methods," in *Proc. EUROSPEECH*, Rhodos, Greece, Sept. 1997, vol. 3, pp. 1095–1098.

[6] F. T. Luk, "A parallel method for computing the generalized singular value decomposition," *Journal of Parallel and Distributed Computing*, vol. 2, pp. 250–260, 1985.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, MD : John Hopkins University Press, Baltimore, 3rd edition, 1996.

[8] M. Moonen, P. Van Dooren, and J. Vandewalle, "A systolic algorithm for QSVD updating," *Signal Processing*, vol. 25, pp. 203–213, 1991.

[9] M. Moonen, P. Van Dooren, and J. Vandewalle, "A Singular Value Decomposition Updating Algorithm for Subspace Tracking," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 4, pp. 1015–1038, Oct. 1992.

[10] D. Van Compernolle, "Switching Adaptive Filters for Enhancing Noisy and Reverberant Speech from Microphone Array Recordings," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Albuquerque, USA, Apr. 1990, pp. 833–836.

[11] S. Haykin, *Adaptive Filter Theory*, Information and system sciences series. Prentice Hall, Englewood Cliffs, New Jersey, 3rd edition, 1996.

[12] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. 30, no. 1, pp. 27–34, Jan. 1982.

[13] K. M. Buckley, "Broad-Band Beamforming and the Generalized Sidelobe Canceller," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 5, pp. 1322–1323, Oct. 1986.

[14] J. Bitzer, K. U. Simmer, and K.-D. Kammeyer, "Theoretical Noise Reduction Limits of the Generalized Sidelobe Canceller (GSC) for Speech Enhancement," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Phoenix, Arizona, USA, May 1999, vol. 5, pp. 2965–2968.

[15] S. Doclo, E. De Clippel, and M. Moonen, "Combined Acoustic Echo and Noise Reduction using GSVD-based Optimal Filtering," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Istanbul, Turkey, June 2000, vol. 2, pp. 1061–1064.

[16] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, vol. 65, pp. 943–950, Apr. 1979.