SVD-BASED OPTIMAL FILTERING WITH APPLICATIONS TO NOISE REDUCTION IN SPEECH SIGNALS

Simon Doclo, Marc Moonen

K.U.Leuven, Dept. of Electrical Engineering (ESAT), SISTA,
Kardinaal Mercierlaan 94, 3001 Heverlee, Belgium
{simon.doclo,marc.moonen}@esat.kuleuven.ac.be

ABSTRACT

In this paper a class of SVD-based signal enhancement procedures is described, which amount to a specific optimal filtering technique for the case where the so-called 'desired response' signal cannot be observed. It is shown that this optimal filter can be written as a function of the generalized singular vectors and singular values of a so-called speech and noise data matrix. A number of simple symmetry properties of the optimal filter are derived, which are valid for the white noise case as well as for the coloured noise case. Also the averaging step of the standard one-microphone SVD-based noise reduction techniques is investigated, leading to serious doubts about the necessity of this averaging step.

When applying this technique for multi-microphone noise reduction, it is shown that for simple scenarios, where we consider localised sources and no multipath propagagtion, this technique exhibits some kind of beamforming behaviour. We further compare the performance of this technique with standard beamforming techniques, showing that for all reverberation times the performance of the SVD-based optimal filter is better than beamforming.

1. INTRODUCTION

In many speech communication applications, like audio-conferencing and hands-free mobile telephony, the recorded speech signals contain a considerable amount of acoustic noise. This is mainly due to the fact that the speaker is located at a certain distance from the microphones, which allows the microphones to record the noise sources too. Background noise causes a signal degradation which can lead to total unintelligibility of the speech.

Some techniques for noise reduction in speech have been proposed which are based on the singular value decomposition (SVD) [1][2][3]. Most of these techniques deal with the one-microphone case and have to rely on signal specific characteristics. The interpretation which is given to these SVD-based noise reduction techniques, is that they try to extract the most important formants from the noisy speech signal [3], thereby reducing the amount of noise.

When using a microphone array, the spatial configuration of the speech and noise sources and the microphone array constitutes an important aspect which should not be neglected. Therefore multi-microphone algorithms should also exploit the characteristics of the channel between the sources and the microphone array. The SVD-based multi-microphone extensions which have been proposed [4] don't yet exploit these characteristics.

Section 2 describes the SVD-based optimal filtering technique. A number of simple symmetry properties of the optimal filter are derived. Also the averaging step of the standard one-microphone SVD-based noise reduction techniques [2][3] is investigated, leading to serious doubts about the necessity of this averaging step. In section 3 SVD-based optimal filtering is applied to multi-microphone noise reduction. It is shown that for simple scenarios this technique exhibits some kind of beamforming behaviour. Section 4 further compares the performance of this technique with standard beamforming techniques, showing that the performance of the SVD-based optimal filter is better than standard beamforming.

2. SVD-BASED OPTIMAL FILTERING

2.1. Preliminaries

Consider the following optimal filtering problem (figure 1) : $\mathbf{u}_k \in \mathbb{R}^N$ is the filter input vector at time k, $\mathbf{y}_k$ is the filter output, $\mathbf{y}_k = \mathbf{u}_k^T \mathbf{W}$, with $\mathbf{W} \in \mathbb{R}^{N \times N}$ the optimal filter. The vector $\mathbf{d}_k \in \mathbb{R}^N$ is the desired response vector and $\mathbf{e}_k = \mathbf{d}_k - \mathbf{y}_k$ is the error vector. The MSE (mean square error) cost function for optimal filtering is

$$\begin{aligned} \mathbf{J}_{MSE}(\mathbf{W}) &= \mathcal{E}\{\|\mathbf{e}_k\|_2^2\} = \mathcal{E}\{\mathbf{d}_k^T \mathbf{d}_k\} - 2\mathcal{E}\{\mathbf{u}_k^T \mathbf{W} \mathbf{d}_k\} \\ &\quad + \mathcal{E}\{\mathbf{u}_k^T \mathbf{W} \mathbf{W}^T \mathbf{u}_k\}. \end{aligned} \tag{1}$$

The optimal filter is found by setting the derivative $\frac{\partial \mathbf{J}_{MSE}}{\partial \mathbf{W}}$ to zero. The optimal filter $\mathbf{W}_{WF}$ is the well-known Wiener filter,

$$\mathbf{W}_{WF} = \mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}^{-1} \cdot \mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{d}_k^T\right\}. \tag{2}$$

In the following, we consider problems where only observations of $\mathbf{u}_k$ are available, and the observed signal $\mathbf{u}_k$ contains a signal-of-interest $\mathbf{s}_k$ (*e.g.* a speech signal) plus additive noise $\mathbf{n}_k$,

$$\mathbf{u}_k = \mathbf{s}_k + \mathbf{n}_k. \tag{3}$$

If we consider speech applications and use a robust speech/noise detection algorithm [5], noise-only observations can be made during speech pauses (time k'), $\mathbf{u}_{k'} = \mathbf{n}_{k'}$, which allows to estimate the spatial and temporal colour of the noise. Our goal is to reconstruct the signal-of-interest $\mathbf{s_k}$ from $\mathbf{u}_k$ by means of a linear filter

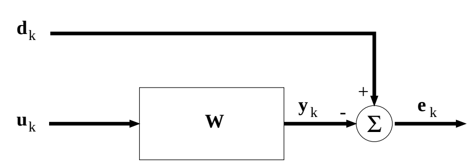

Figure 1: Optimal filtering problem with desired response $\mathbf{d}_k$

$\mathbf{W}$. In the optimal filtering context this means that the desired signal is in fact equal to the signal-of-interest, $\mathbf{d}_k = \mathbf{s}_k$, but that now the desired signal $\mathbf{d}_k$ is an unobservable signal.
We make two *assumptions* : short-term stationarity of the noise, $\mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} = \mathcal{E}\left\{\mathbf{n}_{k'} \cdot \mathbf{n}_{k'}^T\right\}$, and statistical independence of the speech and noise signals, $\mathcal{E}\left\{\mathbf{s}_k \cdot \mathbf{n}_k^T\right\} = 0$. The first assumption allows us to estimate the noise correlation matrix $\mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\}$ during speech pauses. From the second assumption it is easily verified that $\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\} = \mathcal{E}\left\{\mathbf{s}_k \cdot \mathbf{s}_k^T\right\} + \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\}$ and that $\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{s}_k^T\right\} = \mathcal{E}\left\{\mathbf{s}_k \cdot \mathbf{s}_k^T\right\}$, such that the optimal filter becomes

$$\mathbf{W}_{WF} = \mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}^{-1}\left(\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\} - \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\}\right) \quad (4)$$

An interesting simplification is derived from the joint diagonalization [6] of the symmetric matrices $\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}$ and $\mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\}$,

$$\left\{\begin{array}{lcl} \mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\} & = & X \cdot \text{diag}\{\sigma_i^2\} \cdot X^T \\ \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} & = & X \cdot \text{diag}\{\eta_i^2\} \cdot X^T \end{array}\right. \quad (5)$$

In practice, X, σ_i^2 and η_i^2 are computed by means of a *generalized singular value decomposition* (GSVD) of the speech data matrix $\mathbf{U}_k \in \mathbb{R}^{p \times N}$ and the noise data matrix $\mathbf{N}_k \in \mathbb{R}^{q \times N}$ (with p and q typically larger than N),

$$\mathbf{U}_k = \begin{bmatrix} \mathbf{u}_k^T \\ \mathbf{u}_{k+1}^T \\ \vdots \\ \mathbf{u}_{k+p-1}^T \end{bmatrix} \quad \mathbf{N}_k = \begin{bmatrix} \mathbf{n}_k^T \\ \mathbf{n}_{k+1}^T \\ \vdots \\ \mathbf{n}_{k+q-1}^T \end{bmatrix} \quad (6)$$

such that $\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\} \simeq \mathbf{U}_k^T \cdot \mathbf{U}_k$ and $\mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} \simeq \mathbf{N}_k^T \cdot \mathbf{N}_k$. The GSVD of the matrices $\mathbf{U}_k$ and $\mathbf{N}_k$ is defined as

$$\left\{\begin{array}{lcl} \mathbf{U}_k & = & U \cdot \text{diag}\{\sigma_i\} \cdot X^T \\ \mathbf{N}_k & = & V \cdot \text{diag}\{\eta_i\} \cdot X^T, \end{array}\right. \quad (7)$$

with U and V orthogonal matrices, X an invertable (but not necessarily orthogonal) matrix and $\frac{\sigma_i}{\eta_i}$ the generalized singular values. By substituting these formulas into formula (4), one obtains

$$\mathbf{W}_{WF} = X^{-T} \cdot \text{diag}\{\frac{\sigma_i^2 - \eta_i^2}{\sigma_i^2}\} \cdot X^T \quad (8)$$

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

$$\mathbf{W} = X^{-T} \cdot \text{diag}\{f(\sigma_i^2, \eta_i^2)\} \cdot X^T. \quad (9)$$

This formula can be interpreted as follows :

- X^{-T} is an analysis filterbank which performs a transformation from the time domain to a (signal-dependent) transform domain
- $f(\sigma_i^2, \eta_i^2)$ is a function which modifies the transform domain parameters
- X^T is a synthesis filterbank which performs a transformation from the transform domain back to the time domain.

The estimation error $\mathbf{e}_k$ is defined as $\mathbf{e}_k = \mathbf{s}_k - \mathbf{y}_k = \mathbf{s}_k - \mathbf{W}_{WF}^T \mathbf{u}_k$, such that error covariance matrix can be written as

$$\mathcal{E}\left\{\mathbf{e}_k \cdot \mathbf{e}_k^T\right\} = \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} \cdot \mathbf{W}_{WF}. \quad (10)$$

In particular, we are interested in the diagonal elements of the error covariance matrix $\{\mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} \cdot \mathbf{W}_{WF}\}_{ii}$, since these elements indicate how well $\{\mathbf{s}_k\}_i$ (the i^{th} component of $\mathbf{s}_k$) is estimated.
In the *white noise* case, the noise correlation matrix has the form $\mathcal{E}\{\mathbf{n}_k \cdot \mathbf{n}_k^T\} = \eta^2 \cdot I$, with η^2 the power of the noise process. The GSVD then reduces to an SVD with X an orthogonal matrix, such that $\mathbf{W}_{WF}$ is a symmetric matrix.

2.2. Time series filtering

When we apply the SVD-based optimal filtering to one-microphone noise reduction, the vector $\mathbf{u}_k$ is taken from a time series $u(k)$, *i.e.*

$$\mathbf{u}_k = \begin{bmatrix} u(k) & u(k-1) & u(k-2) & \ldots & u(k-N+1) \end{bmatrix}^T. \quad (11)$$

The vectors $\mathbf{s}_k$ and $\mathbf{n}_k$ are similarly defined. The data matrices $\mathbf{U}_k$ and $\mathbf{N}_k$, as defined in equation (6), now are Toeplitz matrices, such that the correlation matrices $\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}$ and $\mathcal{E}\left\{\mathbf{s}_k \cdot \mathbf{s}_k^T\right\}$ are *symmetric Toeplitz matrices*. Symmetric Toeplitz matrices belong to the class of double symmetric matrices, which are symmetric about both the main diagonal and the secondary diagonal. The eigenvectors of such matrices are known to have special symmetry properties [7]. Using these properties, one can prove the following symmetry property for the filter $\mathbf{W}_{WF}$.

Theorem 1 *If $\mathbf{W}_{WF}$ is constructed according to equation (9), then $\mathbf{W}_{WF}$ satisfies the symmetry properties*

$$\mathbf{W}_{WF} = J \cdot \mathbf{W}_{WF} \cdot J \quad (12)$$
$$\mathbf{W}_{WF}^T = J \cdot \mathbf{W}_{WF}^T \cdot J, \quad (13)$$

with J the reverse identity matrix. These properties hold in the white noise case as well as in the coloured noise case.

Proof : Considering the joint diagonalization of $\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}$ and $\mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\}$ (see equation (5)), one can easily verify that

$$\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}^{-1} \cdot \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} = X^{-T} \cdot \text{diag}\{\frac{\eta_i^2}{\sigma_i^2}\} \cdot X^T \quad (14)$$

is the eigenvector decomposition. Because

$$\mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}^{-1} \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} = J\, \mathcal{E}\left\{\mathbf{u}_k \cdot \mathbf{u}_k^T\right\}^{-1} \mathcal{E}\left\{\mathbf{n}_k \cdot \mathbf{n}_k^T\right\} J, \quad (15)$$

the eigenvectors (columns of X^{-T}) are known to have symmetry properties, in particular

$$J \cdot X^{-T} = X^{-T} \text{diag}\{\pm 1\}. \quad (16)$$

With this, one obtains

$$\begin{array}{rcl} J \cdot \mathbf{W}_{WF} \cdot J & = & J \cdot X^{-T} \cdot \text{diag}\{f(\sigma_i^2, \eta_i^2)\} \cdot X^T \cdot J \\ & = & X^{-T} \cdot \text{diag}\{f(\sigma_i^2, \eta_i^2)\} \cdot X^T = \mathbf{W}_{WF} \end{array}$$

These properties mean that the i^{th} row/column of $\mathbf{W}_{WF}$ is equal to the $(N+1-i)^{th}$ row/column in reverse order. For N odd, the middle column in $\mathbf{W}_{WF}$ is symmetric, and hence represents a linear phase filter. Note that a zero phase property has already been attributed to an SVD and rank truncation based estimator for the white noise case, if an additional averaging step is included [8]. For the coloured noise case [2][3], a similar linear phase property has apparently not been derived yet.

2.3. Time series filtering and averaging

The question now arises which of the N columns of $\mathbf{W}_{WF}$ is the best estimator. The answer is given by the error covariance matrix (see equation (10)). The smallest element on the main diagonal of the error covariance matrix corresponds to the best estimator. The best estimator, which is the corresponding column of $\mathbf{W}_{WF}$, will be denoted as $\mathbf{w}_{WF}^{min} \in \mathbb{R}^N$. The question remains if perhaps an even better estimator can be obtained by linearly combining the N columns of $\mathbf{W}_{WF}$. An obvious choice could be averaging, a technique which is often applied to rank truncation based estimation [1][2][3][8]. If w_i^j denotes the (i,j)-element of $\mathbf{W}_{WF}$, then the filter $\tilde{\mathbf{w}}$, obtained by averaging, can be written as

$$\tilde{\mathbf{w}} = \begin{bmatrix} \frac{1}{N} & \frac{1}{N} & \dots & \frac{1}{N} & \frac{1}{N} \end{bmatrix} \cdot \mathcal{W}_{WF}^T, \qquad (17)$$

$$\mathcal{W}_{WF}^T = \begin{bmatrix} 0 & 0 & \dots & 0 & w_1^1 & \dots & w_{N-1}^1 & w_N^1 \\ 0 & 0 & \dots & w_1^2 & w_2^2 & \dots & w_N^2 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & w_1^{N-1} & \dots & w_{N-2}^{N-1} & w_{N-1}^{N-1} & \dots & 0 & 0 \\ w_1^N & w_2^N & \dots & w_{N-1}^N & w_N^N & \dots & 0 & 0 \end{bmatrix}$$

The filter $\tilde{\mathbf{w}}$ has length $(2N-1)$. When we compare the performance of the filters $\mathbf{w}_{WF}^{min}$ and $\tilde{\mathbf{w}}$, simulations show that the filter $\mathbf{w}_{WF}^{min}$ outperforms the filter $\tilde{\mathbf{w}}$. Hence averaging does not seem to be a well-founded operation, while on the other hand it certainly increases computational complexity, since it requires $(2N-1)$-taps filtering instead of N-taps filtering.

2.4. Multichannel filtering

Consider M microphones where each microphone signal $m_j(k)$, $j = 1 \dots M$, consists of a filtered version of the desired signal $s(k)$ and an additive noise term, $m_j(k) = h_j(k) \otimes s(k) + n_j(k)$. The vector $\mathbf{u}_k \in \mathbb{R}^{MN}$ now takes the form

$$\mathbf{u}_k = \begin{bmatrix} \mathbf{m}_{1k} & \mathbf{m}_{2k} & \dots & \mathbf{m}_{Mk} \end{bmatrix}^T, \qquad (18)$$

$$\mathbf{m}_{jk} = \begin{bmatrix} m_j(k) & m_j(k-1) & \dots & m_j(k-N+1) \end{bmatrix}. \qquad (19)$$

Using the same formulas as for the one-channel case, the optimal filter $\mathbf{W}_{WF}$ and best (MN)-taps estimator $\mathbf{w}_{WF}^{min}$ can be computed. The estimated (enhanced) signal $\hat{s}(k)$ can be computed as

$$\hat{\mathbf{s}}(k) = \begin{bmatrix} \hat{s}(k) & \hat{s}(k+1) & \dots & \hat{s}(k+p-1) \end{bmatrix}^T = \mathbf{U}_k \cdot \mathbf{w}_{WF}^{min}. \qquad (20)$$

This operation can be considered as multichannel filtering, where each of the M channels is filtered with an N-taps filter A_j, with $\mathbf{w}_{WF}^{min} = \begin{bmatrix} A_1 & A_2 & \dots & A_M \end{bmatrix}^T$.

3. BEAMFORMING BEHAVIOUR

When applying this technique for multi-microphone noise reduction, it is shown by simulations that when we consider localised sources and no multipath propagagtion, this technique exhibits some kind of beamforming behaviour. We will study the spatial directivity pattern $H(f,\theta)$ of the SVD-based optimal filter, defined as

$$H(f,\theta) = \sum_{l=1}^{M} H_l(f) \cdot \exp\left(j\ 2\pi f \frac{(l-1)d\cos\theta}{c}\right). \qquad (21)$$

$H(f,\theta)$ is a function of frequency f and angle θ. $H_l(f)$ is the frequency response of the filter A_l, d is the distance between the microphones and c is the speed of sound ($c \simeq 340\frac{m}{s}$).

In the first case we consider spatio-temporal white noise, *i.e.* the noise $n_j(k)$ present in every microphone signal $m_j(k)$ is temporal white noise and is uncorrelated with the noise in every other microphone signal. We consider the situation where the speech source impinges on the microphone array at an angle $\theta = 45°$. The distance $d = 5cm$, the number of microphones $M = 5$ and the filterlength $N = 10$. Figure 2 shows the noisy microphone signal $m_1(k)$, enhanced signal $\hat{s}(k)$, the frequency response $|H_j(f)|$ for the M filters A_j and the spatial directivity pattern $|H(f,\theta)|$ for all frequencies f and for one specific frequency $f = 1000$ Hz. The SNR of the noisy signal is 3.1 dB, while the SNR of the enhanced signal is 13.73 dB. As can be seen from the directivity pattern, the gain is maximal for the direction $\theta = 45°$. This is even better illustrated in figure 3 where the directivity pattern is plotted for the frequences $f = i \cdot 100, i = 1 \dots 40$. For most frequencies the directivity gain is maximal for the direction $\theta = 45°$. However for low frequencies the spatial selectivity is very poor.

In the second case we consider a localised white noise source which impinges on the microphone array at an angle $\theta = 150°$, such that the noise signals $n_j(k)$ are delayed versions of each other. The speech source is located in front of the microphone array ($\theta = 90°$). The number of microphones $M = 2$ and the filterlength $N = 10$. Figure 4 shows the directivity pattern for the frequences $f = i \cdot 100, i = 1 \dots 40$. As can be seen, for most frequencies the directivity gain is approximately zero for $\theta = 150°$, the direction of the noise source. Only for low frequencies the spatial selectivity is very poor. We can conclude that *the SVD-based filtering technique has the desired beamforming behaviour*.

4. COMPARISON WITH STANDARD BEAMFORMING

We further compare the performance (SNR of enhanced signal $\hat{s}(k)$) of the SVD-based optimal filtering technique and standard beamforming techniques (delay-and-sum beamforming and Generalised Sidelobe Canceller (GSC)) [9]. This comparison is done 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 5 compares the performance of the delay-and-sum beamformer and the GSC-beamformer with the SVD-based optimal filtering technique (filterlength $N = 10, 20, 50$). As can be seen, for small T_{60} the GSC-beamformer performs much better than for high T_{60}. This is normal because the GSC-beamformer is designed for correlated noise, not for diffuse noise. Unlike the GSC-beamformer, the SVD-based optimal filtering technique still performs well for high T_{60}. As can be seen, *for all reverberation times, the SVD-based optimal filtering technique performs better than the GSC-beamformer*, if the filterlength N is high enough.

5. CONCLUSIONS

In this paper we have discussed a class of SVD-based signal enhancement procedures, which amount to a specific optimal filtering technique. A number of simple symmetry properties have been derived for the optimal filter. When this SVD-based optimal filtering technique is applied to multi-microphone noise reduction, it is shown that it exhibits some kind of beamforming behaviour and that it outperforms standard beamforming techniques.

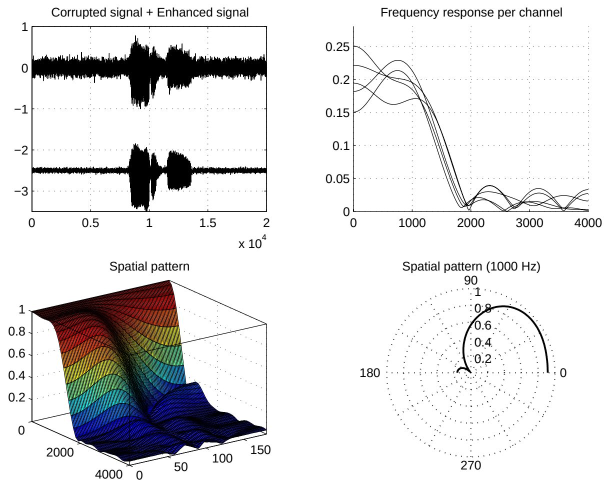

Figure 2: Noisy and enhanced signal, frequency response of the filters A_j and directivity pattern for spatio-temporal white noise

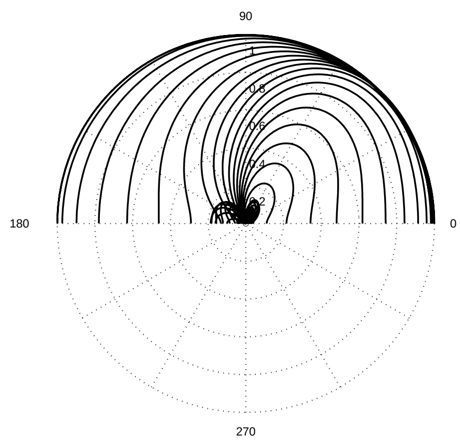

Figure 3: Spatial directivity pattern $|H(f,\theta)|$ for spatio-temporal white noise and speech source at $\theta = 45^\circ$

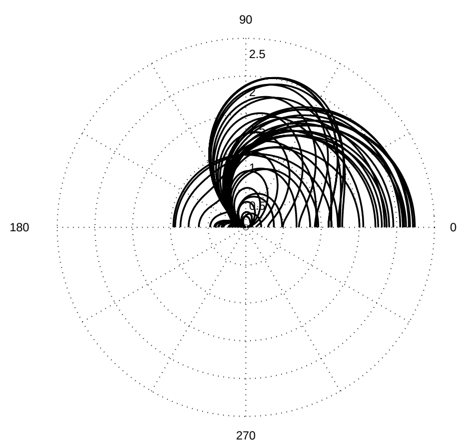

Figure 4: Spatial directivity pattern $|H(f,\theta)|$ for localised white noise source at $\theta = 150^\circ$ and speech source at $\theta = 90^\circ$

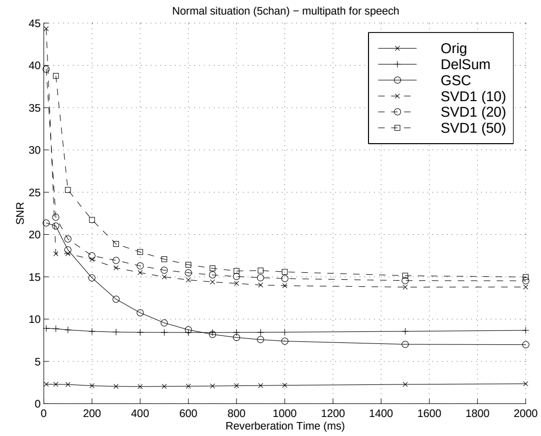

Figure 5: Comparison with standard beamforming techniques

6. 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 framework of 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.

7. REFERENCES

[1] M. Dendrinos, S. Bakamidis, and G. Carayannis, "Speech enhancement from noise : A regenerative approach," *Speech Communication*, vol. 10, no. 2, pp. 45–57, Feb. 1991.

[2] 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, Audio Processing*, vol. 3, no. 6, pp. 439–448, Nov. 1995.

[3] P. S. K. Hansen, *Signal Subspace Methods for Speech Enhancement*, Ph.D. thesis, Technical University of Denmark, Lyngby, Denmark, Sept. 1997.

[4] I. Dologlou, J.-C. Pesquet, and J. Skowronski, "Projection-based rank reduction algorithms for multichannel modelling and image compression," *Signal Processing*, vol. 48, pp. 97–109, Jan. 1996.

[5] F. Xie and S. Van Gerven, "Comparative study of 3 speech detection methods," Tech. Rep. MI2-SPCH-95-8, ESAT, K.U.Leuven, Belgium, Oct. 1995.

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

[7] P. Butler and A. Cantoni, "Eigenvalues and eigenvectors of symmetric centrosymmetric matrices," *Linear Algebra and its Applications*, vol. 13, pp. 275–288, Mar. 1976.

[8] I. Dologlou and G. Carayannis, "Physical Representation of Signal Reconstruction from Reduced Rank Matrices," *IEEE Trans. Signal Processing*, vol. 39, no. 7, pp. 1682–1684, July 1991.

[9] B. D. Van Veen and K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, pp. 4–24, Apr. 1988.