Filter Bank Design for Subband Adaptive Microphone Arrays

Jan Mark de Haan* Nedelko Grbić Ingvar Claesson Sven Nordholm

Abstract— **This paper presents a new method for the design of oversampled uniform DFT-filter banks for the special application of subband adaptive beamforming with microphone arrays. Since array applications rely on the fact that different source positions give rise to different signal delays, a beamformer alters the phase information of the signals. This in turn leads to signal degradations when perfect reconstruction filter banks are used for the subband decomposition and reconstruction. The objective of the filter bank design is to minimize the magnitude of all aliasing components individually, such that aliasing distortion is minimized although phase alterations occur in the subbands. The proposed method is evaluated in a car hands-free mobile telephony environment and the results show that the proposed method offers better performance regarding suppression levels of disturbing signals and much less distortion to the source speech.**

Keywords— **Speech Enhancement, Microphone Arrays, Beamforming, Filter Banks, Filter Bank Design.**

Proposed SAP EDICS: 1-ENHA

corresponding address:
Jan Mark de Haan
Department of Telecommunications and Signal processing
Blekinge Institute of Technology
372 25 Ronneby
Sweden
Tel: +46-457-385728
Fax: +46-457-27914
e-mail: jmd@bth.se

signature: Jan Mark de Haan ______________________________

I. Introduction

Filter banks have been of great interest in a number of signal processing applications. Two important applications of filter banks are subband coding [1], and subband adaptive filtering [2]. Examples where subband adaptive filtering is applied successfully are acoustic echo cancellation [3], [4], [5], [6], [7], single microphone speech enhancement [8], [9], signal separation [10], [11], and microphone arrays [12]. Modulated filter banks have been of special interest because of their simplicity and efficient implementations [13], [14], [15].

Perfect reconstruction decimated filter banks have been of great interest in subband coding. However, these filter banks are less suitable for subband adaptive filtering, since the perfect reconstruction property with aliasing cancellation is not maintained when the subband signals are modified by filters with arbitrary magnitude and phase response. This implies that aliasing is present in the reconstructed output of the subband adaptive filter. Design methods for filter banks for subband adaptive filtering have been proposed in [16], [17], [18].

Also, low filter bank delays are important in speech communications, since speech conversations are impaired by additional transmission delays, [7]. Design methods for perfect reconstruction filter banks exist where the delay can be chosen independently of the filter length, e.g. biorthogonal filter banks, [19]. Design of oversampled DFT filter banks with minimum delay and perfect reconstruction has been proposed in [20].

This paper proposes a new method for the design of analysis and synthesis uniform DFT-filter banks, using unconstrained quadratic optimization, where the main goal is that adaptive filtering in the subbands should cause minimal source signal degradations at the output. The design method aims at minimizing aliasing in the subband signals as well as minimizing magnitude, phase, and aliasing distortions in the reconstructed output signal.

The proposed method consist of two steps. In the first step the analysis filter bank is designed in such way that the aliasing terms in each subband are minimized individually, contributing to minimal aliasing at the output without aliasing cancellation. In the second step the synthesis filter bank is designed to match the analysis filter bank where the analysis-synthesis response is optimized while all aliasing terms in the output signal are individually suppressed, rather than aiming at aliasing cancellation. Both design steps include constraints on the signal delay.

The paper is organized as follows. In Section II, the uniform DFT-filter bank concept is outlined and discussed to some depth. Section III, the filter bank design method is presented. Further, Section IV gives performance relations of the filter banks relative to numerous parameter choices. In Section V, the filter banks are used in automobile hands-free subband adaptive microphone array application. Finally, Section VI concludes the paper.

II. Uniform DFT-Filter Banks

A. Definitions and Notations

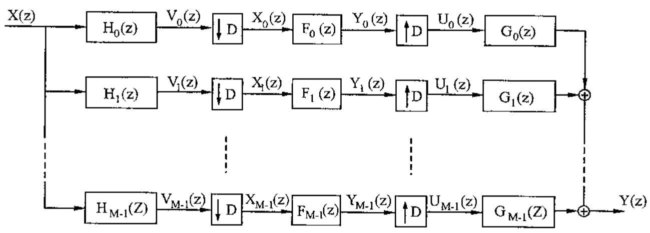

Fig. 1. *Analysis and synthesis filter banks with subband filtering.*

properties of the resulting filter. However, filter design techniques with complex approximation for filters with arbitrary phase exist [22]. These methods do not take the behavior of the filter bank into consideration. However, an extension of the filter design method to filter bank design has been proposed in [23].

An appropriate design criterion may be to minimize the following objective function

$$\epsilon_{\mathbf{h}} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega \qquad (12)$$

where $H_d(z)$ is a desired frequency response of the prototype analysis filter in the passband region $\Omega_p = [-\omega_p, \omega_p]$. The desired frequency response is defined as

$$H_d(e^{j\omega}) = e^{-j\omega\tau_H}, \qquad \omega \in \Omega_p \qquad (13)$$

where τ_H is the desired group delay of the prototype analysis filter and, consequently, the desired delay of the whole analysis filter bank.

By only specifying a complex specification for the passband region, a suitable prototype filter cannot be obtained. The stopband regions of the prototype analysis filters must also be defined, otherwise significant inband-aliasing distortion may occur in the subbands.

One approach to combat the undesired inband-aliasing is to minimize the energy in the stopband. This can be done with any filter design method where passband and stopband can be defined separately. A more appealing approach would be to address the inband-aliasing directly in the objective function. This can be done by complementing the design criterion with an *inband-aliasing distortion* term. The design criterion is complemented by adding the energies of all $D-1$ aliasing terms in subband $m = 0$ to the objective function in Eq. (12). This will address all aliasing in the analysis bank due to the modulated structure. The objective function becomes

$$\epsilon_{\mathbf{h}} = \alpha_{\mathbf{h}} + \beta_{\mathbf{h}}, \qquad (14)$$

with the *Passband Response Error*

$$\alpha_{\mathbf{h}} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega, \qquad (15)$$

and the *Inband-Aliasing Distortion*

$$\beta_{\mathbf{h}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{d=1}^{D-1} \left|H(e^{j\omega/D}W_D^d)\right|^2 d\omega. \qquad (16)$$

The passband response error can be rewritten as

$$\begin{aligned} \alpha_{\mathbf{h}} &= \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega \\ &= \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \left[|H(e^{j\omega})|^2 - 2\mathrm{Re}\left\{H_d^*(e^{j\omega})H(e^{j\omega})\right\} + \right. \\ &\quad \left. + |H_d(e^{j\omega})|^2\right] d\omega, \end{aligned} \qquad (17)$$

where $H_d(z)$ is the desired frequency response, defined in Eq. (13). The prototype analysis filter response $H(z)$ can be expressed in terms of its impulse response, $h(n)$, according to

$$H(z) = \sum_{n=0}^{L_{\mathbf{h}}-1} h(n)z^{-n} = \mathbf{h}^T\phi_{\mathbf{h}}(z), \qquad (18)$$

where $\mathbf{h} = [h(0), \ldots, h(L_{\mathbf{h}} - 1)]^T$ and $\phi_{\mathbf{h}}(z) = [1, z^{-1}, \ldots, z^{-L_{\mathbf{h}}+1}]^T$. Substituting Eqs. (18) and (13) into Eq. (17) yields

$$\begin{aligned} \alpha_{\mathbf{h}} &= \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \left[\mathbf{h}^T\phi_{\mathbf{h}}(e^{j\omega})\phi_{\mathbf{h}}^H(e^{j\omega})\mathbf{h} - \right. \\ &\quad \left. -2\mathrm{Re}\left\{e^{j\omega\tau_H}\mathbf{h}^T\phi_{\mathbf{h}}(e^{j\omega})\right\} + 1\right] d\omega, \end{aligned} \qquad (19)$$

where $(\cdot)^H$ denotes conjugate transpose. Eq. (19) can be rewritten in the quadratic form

$$\alpha_{\mathbf{h}} = \mathbf{h}^T\mathbf{A}\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + 1, \qquad (20)$$

where the $L_{\mathbf{h}} \times L_{\mathbf{h}}$ hermitian matrix $\mathbf{A}$ is

$$\mathbf{A} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \phi_{\mathbf{h}}(e^{j\omega})\phi_{\mathbf{h}}^H(e^{j\omega})d\omega, \qquad (21)$$

and the $L_{\mathbf{h}} \times 1$ vector $\mathbf{b}$ is

$$\mathbf{b} = \frac{1}{2\omega_p}\int_{-\omega_p}^{\omega_p} \mathrm{Re}\left\{e^{j\omega\tau_H}\phi_{\mathbf{h}}(e^{j\omega})\right\} d\omega. \qquad (22)$$

Calculating the integrals in matrix $\mathbf{A}$ and vector $\mathbf{b}$, the matrix entries $A_{i,j}$ are given by

$$A_{i,j} = \frac{\sin(\omega_p(j-i))}{\omega_p(j-i)}, \qquad (23)$$

and the vector entries b_i are given by

$$b_i = \frac{\sin(\omega_p(\tau_H - i))}{\omega_p(\tau_H - i)}. \qquad (24)$$

where $i, j = 0, \ldots, L_{\mathbf{h}} - 1$.

The inband-aliasing distortion term of the objective function, Eq. (16), can be rewritten as

$$\beta_{\mathbf{h}} = \frac{1}{2\pi}\sum_{d=1}^{D-1}\mathbf{h}^T\left[\int_{-\pi}^{\pi}\phi_{\mathbf{h}}(e^{j\omega/D}W_D^d)\phi_{\mathbf{h}}^H(e^{j\omega/D}W_D^d)\, d\omega\right]\mathbf{h}, \qquad (25)$$

which can be written in quadratic form

$$\beta_{\mathbf{h}} = \mathbf{h}^T\mathbf{C}\mathbf{h}, \qquad (26)$$

where the $L_{\mathbf{h}} \times L_{\mathbf{h}}$ hermitian matrix $\mathbf{C}$ is defined as

$$\mathbf{C} = \frac{1}{2\pi}\sum_{d=1}^{D-1}\int_{-\pi}^{\pi}\phi_{\mathbf{h}}(e^{j\omega/D}W_D^d)\phi_{\mathbf{h}}^H(e^{j\omega/D}W_D^d)\, d\omega. \qquad (27)$$

Calculating the integral in matrix $\mathbf{C}$, the matrix entries $C_{i,j}$ are given by

$$C_{i,j} = \frac{\varphi(j-i)\sin(\pi(j-i)/D)}{\pi(j-i)}, \qquad (28)$$

where

$$\varphi(n) = D \sum_{k=-\infty}^{\infty} \delta(n - kD) - 1. \tag{46}$$

The optimal prototype analysis filter, in terms of minimal total response error and minimal energy in the aliasing components, is found by minimizing the weighted objective function

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \gamma_{\mathbf{g}}(\mathbf{h}) + v\delta_{\mathbf{g}}(\mathbf{h}). \tag{47}$$

A weighting factor v is introduced in order to emphasize on either the total response error $(0 < v < 1)$ or the residual aliasing distortion $(v > 1)$. Inserting Eq. (36) and Eq. (43) into (47) yields

$$\epsilon_{\mathbf{g}}(\mathbf{h}) = \mathbf{g}^T(\mathbf{E} + v\mathbf{P})\mathbf{g} - 2\mathbf{g}^T\mathbf{f} + 1, \tag{48}$$

The solution

$$\mathbf{g}_{\text{opt}} = \arg\min_{\mathbf{g}} \{\epsilon_{\mathbf{g}}(\mathbf{h})\}, \tag{49}$$

can be found by solving the set of linear equations

$$(\mathbf{E} + v\mathbf{P})\mathbf{g} = \mathbf{f}. \tag{50}$$

IV. Design Parameters

A. Pre-specified Parameters

In order to use the design method, a number of parameters are set. These parameters are the number of subbands, M, the filter lengths of the analysis and synthesis filters, $L_{\mathbf{h}}$ and $L_{\mathbf{g}}$, and the decimation factor, D. The filter lengths are commonly set to multiples of the decimation factor because it gives rise to the efficient polyphase implementation, [25]. The parameters mentioned here influence both performance and complexity, as will be seen in the following sections. Other design parameters are the passband cut-off frequency, ω_p, optional weighting in the synthesis filter bank optimization, v, the delay of the analysis filter bank τ_H, and the total delay τ_T. For the sake of simplicity and clearness, the results for filter banks with $M = 4$ subbands are presented in this section. The results in this section show that an increment in degrees of freedom, which is obtained obtained by reducing the passband cut-off frequency or decreasing the decimation factor, generally yields higher performance. However, decreasing the decimation factor is at the expense of increased computational complexity. The results also show that the balance between the group-delays of the filter banks can have a large impact on different design measures.

B. Passband Cut-Off Frequency

The bandwidth of the prototype analysis filter is set by selecting the passband cut-off frequency ω_p. The parameter determines the bandwidth of the analysis filters. Prototype analysis filters, obtained from Eq. (31), with different passband cut-off frequencies are shown in Fig. 2. The figure illustrates that a lower stop band level, and thus lower inband-aliasing distortion, can be obtained by choosing a lower passband cut-off frequency.

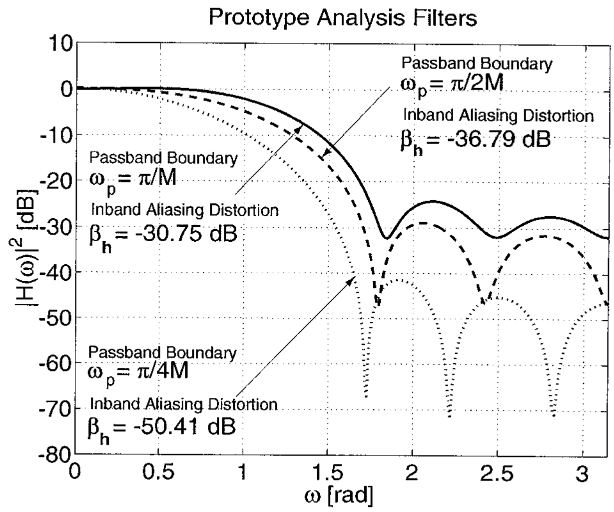

Fig. 2. *Prototype analysis filter responses with filter lengths $L_{\mathbf{h}} = 8$, no. of subbands $M = 4$ and decimation factor $D = 2$. The frequency response of the filters is shown for passband boundaries $\omega_p = \frac{\pi}{M}$, $\omega_p = \frac{\pi}{2M}$ and $\omega_p = \frac{\pi}{4M}$. The corresponding inband-aliasing distortion, $\beta_{\mathbf{h}}$, is also shown.*

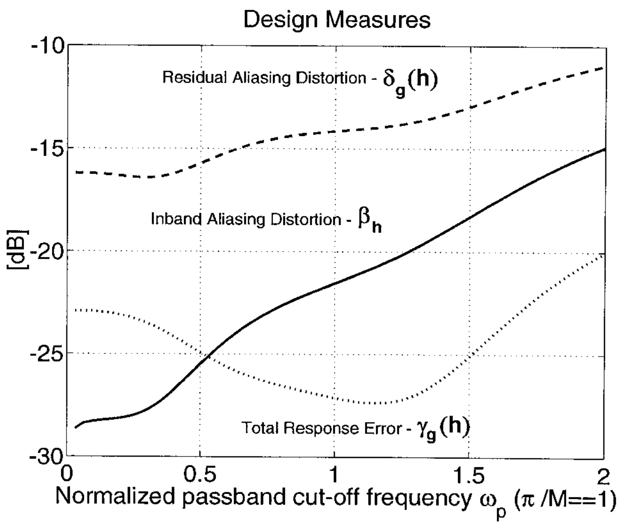

Fig. 3. *Filter bank performance as a function of the passband cut-off frequency. The solid line is the inband-aliasing distortion, $\beta_{\mathbf{h}}$, the dashed line is the residual aliasing distortion, $\delta_{\mathbf{g}}(\mathbf{h})$, and the dotted line is the total response error, $\gamma_{\mathbf{g}}(\mathbf{h})$. The number of subbands is $M = 4$ and the decimation factor is $D = 2$. The filter lengths are set to $L_{\mathbf{h}} = L_{\mathbf{g}} = 4$.*

The passband cut-off frequency also affects the total response error and the residual aliasing distortion since the synthesis filter bank design has a dependency on the analysis filter bank design. Figs. 3 and 4 show how the inband-aliasing distortion, $\beta_{\mathbf{h}}$ from Eq. (16), total response error, $\gamma_{\mathbf{g}}(\mathbf{h})$ from Eq. (34), and the residual aliasing distortion, $\delta_{\mathbf{g}}(\mathbf{h})$ from Eq. (35), are affected by the passband cut-off frequency parameter in the analysis filter bank design.

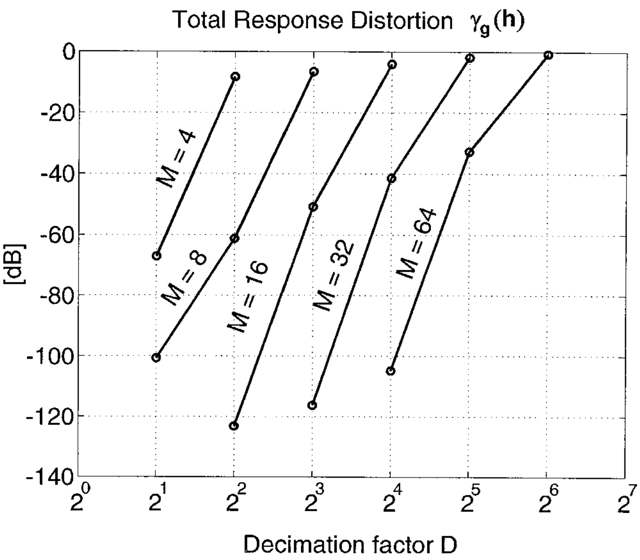

Fig. 7. *Minimum total response error, $\gamma_\mathbf{g}(\mathbf{h})$ as a function of the decimation factor D. The filter lengths are set to $L_\mathbf{h} = L_\mathbf{g} = 2M$. Each curve corresponds to a specific number of subbands, $M = 4, \ldots, 64$. In all cases, the passband cut-off frequency is set to $\omega_p = \frac{\pi}{8M}$ and the prototype filters have linear phase. The circles denote decimation factors chosen as powers of two.*

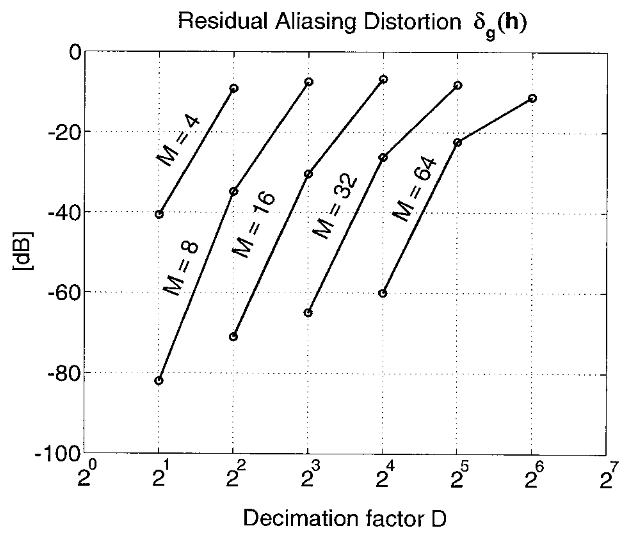

Fig. 8. *Minimum residual aliasing distortion, $\delta_\mathbf{g}(\mathbf{h})$, as a function of the decimation factor D. The filter lengths are set to $L_\mathbf{h} = L_\mathbf{g} = 2M$. Each curve corresponds to a specific number of subbands, $M = 4, \ldots, 64$. In all cases, the passband cut-off frequency is set to $\omega_p = \frac{\pi}{8M}$ and the prototype filters have linear phase. The circles denote decimation factors chosen as powers of two.*

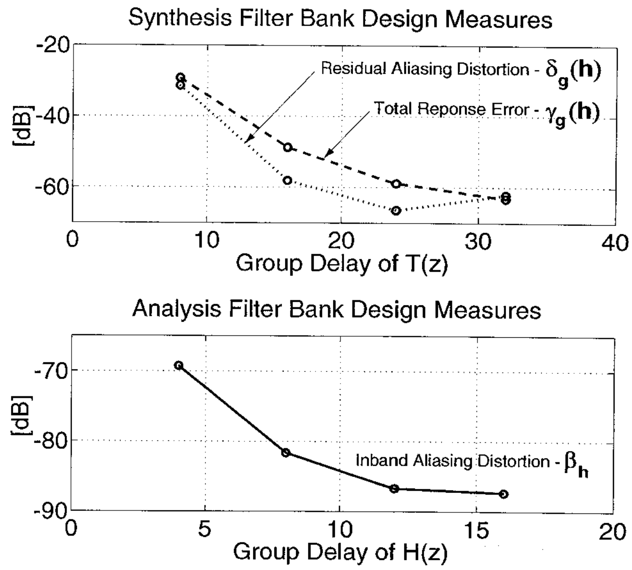

Fig. 9. *Total response error and residual alisiasing distortion as a function of τ_H and τ_T. The number of subbands is $M = 8$. The decimation factor is set to $D = 4$. The filter lengths are set to $L_\mathbf{h} = L_\mathbf{g} = 32$. The delay of the analysis filter-bank is set to $\tau_H = \frac{1}{2}L_\mathbf{h}$.*

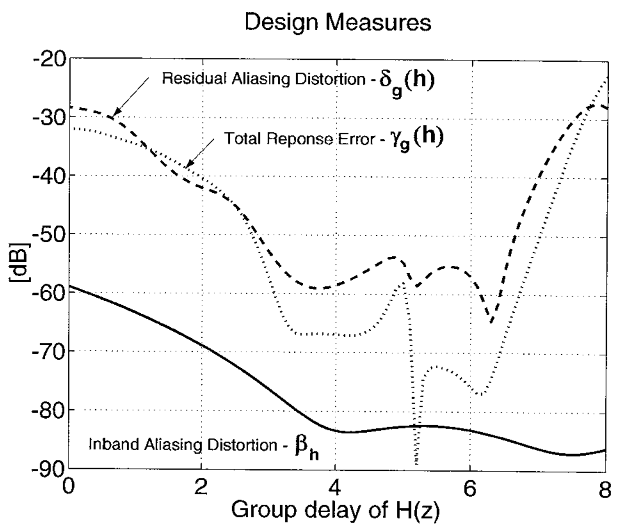

Fig. 10. *Design errors as a function of the analysis prototype filter group-delay. The solid line is the inband aliasing distortion, the dashed line is the residual aliasing distortion, and the dotted line is the total response error. The number of subbands is $M = 4$, the decimation factor is $D = 2$ and the prototype filter lengths are set to $L_\mathbf{h} = L_\mathbf{g} = 4M = 16$. The desired system delay is set to $\tau_T = 8$ in all cases. The delay of the prototype analysis filter τ_H is varied from 0 to τ_T, which means that the group-delay of the prototype synthesis filter τ_G varies from τ_T to 0 since $\tau_T = \tau_H + \tau_G$.*

V. Evaluation in Subband Beamforming

A. Least Squares Subband Beamforming

We evaluate the performance of the filter banks in the case of a subband LS optimal beamformer with real data recorded in a hands-free car situation, [26], [27]. In this situation we have a target signal, an interference signal causing echo at the far end of the communication link, and background noise, see Fig. 12. The microphone array has I input signals $x_i(n)$, where i is the microphone index, $i = 1, \ldots, I$, and the output signal is denoted by $y(n)$.

A linear array of six microphones was mounted in a car where $w = 2\pi f$, and f is the normalized frequency. The constant, C_d, is defined as

$$C_d = \frac{\int_{-\pi}^{\pi} \hat{P}_{x_s}(w)dw}{\int_{-\pi}^{\pi} \hat{P}_{y_s}(w)dw} \quad (55)$$

where $\hat{P}_{x_s}(w)$, is a spectral power estimate of a single sensor observation and, $\hat{P}_{y_s}(w)$ is the spectral power estimate of the beamformer output, when the source signal is active alone. The constant C_d normalizes the mean output spectral power to that of the single sensor spectral power. The measure of distortion in Eq. (54), is the mean output spectral power deviation from the observed single sensor spectral power. Ideally, the distortion is zero.

To measure the performance of the microphone array, the normalized noise suppression quantity, S_N, is introduced as

$$\mathcal{S_N} = C_s \frac{\int_{-\pi}^{\pi} \hat{P}_{y_N}(w)dw}{\int_{-\pi}^{\pi} \hat{P}_{x_N}(w)dw} \quad (56)$$

and the normalized interference suppression quantity, S_I, is introduced as

$$\mathcal{S_I} = C_s \frac{\int_{-\pi}^{\pi} \hat{P}_{y_I}(w)dw}{\int_{-\pi}^{\pi} \hat{P}_{x_I}(w)dw} \quad (57)$$

where

$$C_s = \frac{1}{C_d} \quad (58)$$

and where, $\hat{P}_{y_N}(w)$ and $\hat{P}_{x_N}(w)$ are spectral power estimates of the beamformer output and the reference sensor observation, respectively, when the surrounding noise is active alone. In the same way $\hat{P}_{y_I}(w)$ and $\hat{P}_{x_I}(w)$ are spectral power estimates when the interference signal/signals are active alone.

The beamformer performance measures for the different filter banks are shown in Table II for 1-tap filters in the subband beamformers and in table III for 6-tap filters in the subband beamformers. Fig. 13 shows short-time power estimates of the reference microphone signal and the microphone array output signal for the first case shown in Table III, with $\tau_T = 256$. It can be observed that the background noise is suppressed by about 15 dB and that the interference signal (the male speaker) is suppressed by about 17 dB.

C. Aliasing Distortion Measures

To determine the level of aliasing present in the subband signals $x_i^{(m)}(n)$, $m = 0, \ldots, M-1$, and in the output enhanced speech, $y(n)$, Signal-to-Aliasing Ratio (SAR) measures are introduced. The average SAR of the subband signals $x_i^{(m)}(n)$ is defined as

$$\overline{\mathcal{SAR}}_{x_i^{(m)}} = \frac{1}{MI} \sum_{i=1}^{I} \sum_{m=0}^{M-1} \frac{\int_{-\pi}^{\pi} \hat{P}_{x_i^{(m)},S}(\omega)d\omega}{\int_{-\pi}^{\pi} \hat{P}_{x_i^{(m)},A}(\omega)d\omega}, \quad (59)$$

where $P_{x_i^{(m)},S}(\omega)$ denotes the PSD of the desired signal component and $P_{x_i^{(m)},A}(\omega)$ denotes the PSD estimate of the aliasing component in $x_i^{(m)}(n)$. The average subband SAR measures from the evaluation are shown in Table IV.

The SAR of $y(n)$ is defined as

$$\mathcal{SAR}_y = \frac{\int_{-\pi}^{\pi} \hat{P}_{y,S}(\omega)d\omega}{\int_{-\pi}^{\pi} \hat{P}_{y,A}(\omega)d\omega}. \quad (60)$$

where $\hat{P}_{y,S}(\omega)$ denotes the PSD estimate of the desired signal component in $y(n)$ and $\hat{P}_{y,A}(\omega)$ denotes the PSD estimate of the aliasing component in $y(n)$. The output SAR measures are also shown in Table II and Table III. In order for the aliasing components in the output signal to be masked by the speech, the $\mathcal{SAR}_y$ measures should be at least 30 dB. If the measure is lower, the modulated character of the aliasing will become audible in the signal. Table II and III show that the processing delay can be reduced with the proposed method, without audible aliasing in the output signal.

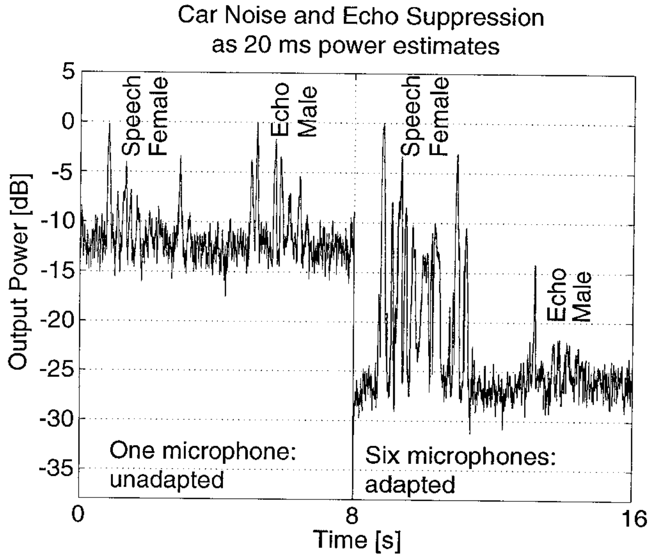

Fig. 13. *Short time (20 ms) power estimates of the reference microphone input signal (left) and the beamformer output signal (right). The figure corresponds to the first filter bank in Table III in the evaluation.*

VI. Conclusion

A design method for uniform DFT-filter banks with reduced delay has been proposed for the application of oversampled subband adaptive microphone arrays. The design aims at minimizing aliasing components individually, in order to bound the aliasing effects when phase alterations are applied by adaptive beamformers. The evaluation of a subband beamformer in a microphone array in a automobile environment shows that noise and interference suppression levels are improved by 1-2 dB with the proposed method, while the distortion caused by the subband beamformer is reduced substantially. Further, the evaluation shows that a filter bank with significantly reduced delay can be used, without severe loss of performance.

[19] T. Karp, A. Mertins, G. Schuller, "Recent Trends in the Design of Biorthogonal Modulated Filter Banks," TICSP Workshop on Transforms and Filter Banks, 1998.
[20] H. Kobayashi, H. Kiya, "Design of Oversampled DFT Filter Banks with Minimum Delay," Electronics and Communications in Japan, Part 3, Vol. 81, No. 1, 1998.
[21] T. W. Parks, C. S. Burrus, *Digital Filter Design* , John Wiley & Sons, Inc. 1987.
[22] D. R. Xiangkun Chen, T. W. Parks, "Design of FIR filters in the Complex Domain," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. ASSP-35, no. 2, Feb. 1979.
[23] N. Grbić, J. M. de Haan, S. Nordholm, I. Claesson, "Design of Oversampled Uniform DFT-Filter Banks with Reduced Inband Aliasing and Delay Constraints," Proc. of ISSPA'01, Kuala Lumpur, Malaysia, pp. 104-107, Aug. 2001.
[24] I. Sodagar, K. Nayebi, T.P. Barnwell, "Time-varying filter banks and wavelets," *IEEE transactions on Signal Processing*, vol. 42, no. 11, pp. 2983 -2996, Nov. 1994.
[25] J. M. de Haan, *Filter Bank Design for Subband Adaptive Filtering - Methods and Applications*, Licentiate Thesis, Blekinge Institute of Technology, pp. 45-52, May 2001, ISBN 91-7295-003-X.
[26] M. Dahl, I. Claesson, S. Nordholm, "Acoustic Noise and Echo Cancelling with Microphone Array," *IEEE transactions on Vehicular Technology*, vol. 48, no. 5, pp. 1518-1526, Sept. 1999.
[27] N. Grbić, S. Nordholm, I. Claesson, "Optimal and Adaptive Beamforming for Speech Signals in a mixture of Spatially Coherent and Incoherent Noise Fields," submitted to *IEEE transactions on Speech and Audio Processing.*