EURASIP Journal on Applied Signal Processing 2003:11, 1147–1156

Multichannel Direction-Independent Speech Enhancement Using Spectral Amplitude Estimation

Thomas Lotter

Institute of Communication Systems and Data Processing, Aachen University (RWTH), Templergraben 55, D-52056 Aachen, Germany
Email: lotter@ind.rwth-aachen.de

Christian Benien

Philips Research Center, Aachen, Weißhausstraße 2, D-52066 Aachen, Germany
Email: christian.benien@philips.com

Peter Vary

Institute of Communication Systems and Data Processing, Aachen University (RWTH), Templergraben 55, D-52056 Aachen, Germany
Email: vary@ind.rwth-aachen.de

Received 25 November 2002 and in revised form 12 March 2003

This paper introduces two short-time spectral amplitude estimators for speech enhancement with multiple microphones. Based on joint Gaussian models of speech and noise Fourier coefficients, the clean speech amplitudes are estimated with respect to the MMSE or the MAP criterion. The estimators outperform single microphone minimum mean square amplitude estimators when the speech components are highly correlated and the noise components are sufficiently uncorrelated. Whereas the first MMSE estimator also requires knowledge of the direction of arrival, the second MAP estimator performs a direction-independent noise reduction. The estimators are generalizations of the well-known single channel MMSE estimator derived by Ephraim and Malah (1984) and the MAP estimator derived by Wolfe and Godsill (2001), respectively.

Keywords and phrases: speech enhancement, microphone arrays, spectral amplitude estimation.

1. INTRODUCTION

Speech communication appliances such as voice-controlled devices, hearing aids, and hands-free telephones often suffer from poor speech quality due to background noise and room reverberation. Multiple microphone techniques such as beamformers can improve the speech quality and intelligibility by exploiting the spatial diversity of speech and noise sources. Upon these techniques, one can differentiate between fixed and adaptive beamformers.

A fixed beamformer combines the noisy signals by a time-invariant filter-and-sum operation. The filters can be designed to achieve constructive superposition towards a desired direction (delay-and-sum beamformer) or in order to maximize the SNR improvement (superdirective beamformer) [1, 2, 3].

Adaptive beamformers commonly consist of a fixed beamformer towards a fixed desired direction and an adaptive null steering towards moving interfering sources [4, 5].

All beamformer techniques assume the target direction of arrival (DOA) to be known a priori or assume that it can be estimated sufficiently enough. Usually the performance of such a beamforming system decreases dramatically if the DOA knowledge is erroneous. To estimate the DOA during runtime, time difference of arrival (TDOA)-based locators evaluate the maximum of a weighted cross correlation [6, 7]. Subspace methods have the ability to detect multiple sources by decomposing the spatial covariance matrix into a signal and a noise subspace. However, the performance of all DOA estimation algorithms suffers severely from reverberation and directional or diffuse background noise.

Single microphone speech enhancement frequency domain algorithms are comparably robust against reverberation and multiple sources. However, they can achieve high noise reduction only at the expense of moderate speech distortion. Usually, such an algorithm consists of two parts. Firstly, a noise power spectral density estimator based on the assumption that the noise is stationary to a much higher

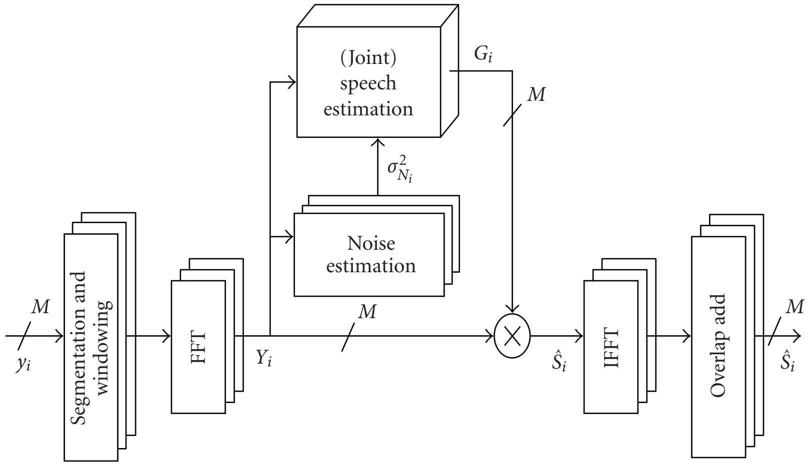


FIGURE 1: Multichannel noise reduction system.

degree than the speech. The noise power spectral density can be estimated by averaging discrete Fourier transform (DFT) periodograms in speech pauses using a voice activity detection or by tracking minima over a sliding time window [8]. Secondly, an estimator for the speech component of the noisy signal with respect to an error criterion. Commonly, a Wiener filter, the minimum mean square error (MMSE) estimator of the speech DFT amplitudes [9], or its logarithmic extension [10] are applied.

In this paper, we propose the extensions of two single channel speech spectral amplitude estimators for the use in microphone array noise reduction. Clearly, multiple noisy signals offer a higher-estimation accuracy possibility when the desired signals are highly correlated and the noise components are uncorrelated to a certain degree. The main contribution will be a joint speech estimator that exploits the benefits of multiple observations but achieves a DOA-independent speech enhancement.

Figure 1 shows an overview of the multichannel noise reduction system with the proposed speech estimators. The noisy time signals $y_i(k)$, $i \in \{1, \dots, M\}$, from M microphones are transformed into the frequency domain. This is done by applying a window $h(\mu)$, for example, a Hann window, to a frame of K consecutive samples and by computing the DFT on the windowed data. Before the next DFT computation, the window is shifted by Q samples. The resulting complex DFT values $Y_i(\lambda, j)$ are given by

$$Y_i(\lambda, k) = \sum_{\mu=0}^{K-1} y_i(\lambda Q + \mu) h(\mu) e^{-j2\pi k\mu/L}. \quad (1)$$

Here, k denotes the DFT bin and λ the subsampled time index. For the sake of brevity, k and λ are omitted in the following.

The noisy DFT coefficient Y_i consists of complex speech $S_i = A_i e^{j\alpha_i}$ and noise N_i components:

$$Y_i = R_i e^{j\vartheta_i} = A_i e^{j\alpha_i} + N_i, \quad i \in \{1, \dots, M\}. \quad (2)$$

The noise variances $\sigma^2_{N_i}$ are estimated separately for each channel and are fed into a speech estimator. If $M = 1$, the minimum mean square short-time spectral amplitude (MMS-STSA) estimator [9], its logarithmic extension [10], or less complex maximum a posteriori (MAP) estimators [11] can be applied to calculate real spectral weights G_1 for each frequency. If $M > 1$, a joint estimator can exploit information from all M channels using a joint statistical model of the DFT coefficients after IFFT and overlap-add M noise-reduced signals are synthesized. Since the phases are not modified, a beamformer could be applied additionally after synthesis.

The remainder of the paper is organized as follows. Section 2 introduces the underlying statistical model of multichannel Fourier coefficients. In Section 3, two new multichannel spectral amplitude estimators are derived. First, a minimum mean square estimator that evaluates the expectation of the speech spectral amplitude conditioned on all noisy complex DFT coefficients is described. Secondly, a MAP estimator conditioned on the joint observation of all noisy amplitudes is proposed. Finally, in Section 4, the performance of the proposed estimators in ideal and realistic conditions is discussed.

2. STATISTICAL MODELS

Motivated by the central limit theorem, real and imaginary parts of both speech and noise DFT coefficients are usually modelled as zero-mean independent Gaussian [9, 12, 13] with equal variance. Recently, MMSE estimators of the complex DFT spectrum S have been developed with Laplacian or Gamma modelling of the real and imaginary parts of the speech DFT coefficients [14]. However, for MMSE or MAP estimation of the speech spectral amplitude, the Gaussian model facilitates the derivation of the estimators. Due to the unimportance of the phase, estimation of the speech spectral amplitude instead of the complex spectrum is more suitable from a perceptual point of view [15].

The Gaussian model leads to Rayleigh distributed speech amplitudes A_i, that is,

$$p(A_i, \alpha_i) = \frac{A_i}{\pi \sigma_{S_i}^2} \exp\left(-\frac{A_i^2}{\sigma_{S_i}^2}\right). \tag{3}$$

Here, $\sigma_{S_i}^2$ describes the variance of the speech in channel i. Moreover, the pdfs of the noisy spectrum Y_i and noisy amplitude R_i conditioned on the speech amplitude and phase are Gaussian and Ricians, respectively,

$$p(Y_i \mid A_i, \alpha_i) = \frac{1}{\pi \sigma_{N_i}^2} \exp\left(-\frac{\left|Y_i - A_i e^{j\alpha_i}\right|^2}{\sigma_{N_i}^2}\right), \tag{4}$$

$$p(R_i \mid A_i) = \frac{2R_i}{\sigma_{N_i}^2} \exp\left\{-\frac{R_i^2 + A_i^2}{\sigma_{N_i}^2}\right\} I_0\left(\frac{2A_iR_i}{\sigma_{N_i}^2}\right). \tag{5}$$

Here, I_0 denotes the modified Bessel function of the first kind and zeroth order. To extend this statistical model for multiple noisy signals, we consider the typical noise reduction scenario of Figure 2, for example, inside a room or a car. A desired signal s arrives at a microphone array from angle θ. Multiple noise sources arrive from various angles. The resulting diffuse noise field can be characterized by its coherence function. The magnitude squared coherence (MSC) between two omnidirectional microphones i and j of a diffuse noise field is given by

$$\mathrm{MSC}_{ij}(f) = \frac{\left|\Phi_{ij}(f)\right|^2}{\Phi_{ii}(f)\Phi_{jj}(f)} = \mathrm{si}^2\left(\frac{2\pi f d_{ij}}{c}\right). \tag{6}$$

Figure 3 plots the theoretical coherence of an ideal diffuse noise field and the measured coherence of the noise field inside a crowded cafeteria with a microphone distance of $d_{ij} = 12\,\mathrm{cm}$. For frequencies above $f_0 = c/2d_{ij}$, the MSC becomes very low and thus the noise components of the noisy spectra can be considered uncorrelated with

$$E\{N_i N_j^*\} = \begin{cases} \sigma_{N_i}^2, & i = j, \\ 0, & i \neq j. \end{cases} \tag{7}$$

Hence, (5) and (4) can be extended to

$$p(R_1, \dots, R_M \mid A_n) = \prod_{i=1}^{M} p(R_i \mid A_n), \tag{8}$$

$$p(Y_1, \dots, Y_M \mid A_n, \alpha_n) = \prod_{i=1}^{M} p(Y_i \mid A_n, \alpha_n), \tag{9}$$

for each $n \in \{1, \dots, M\}$. We assume the time delay of the speech signals between the microphones to be small compared to the short-time stationarity of speech and thus assume the speech spectral amplitudes A_i to be highly correlated. However, due to near-field effects and different microphone amplifications, we allow a deviation of the speech amplitudes by a constant channel-dependent factor c_i, that is,

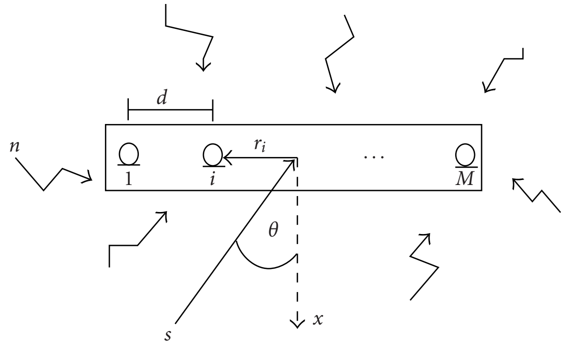


Figure 2: Speech and noise arriving at microphone array.

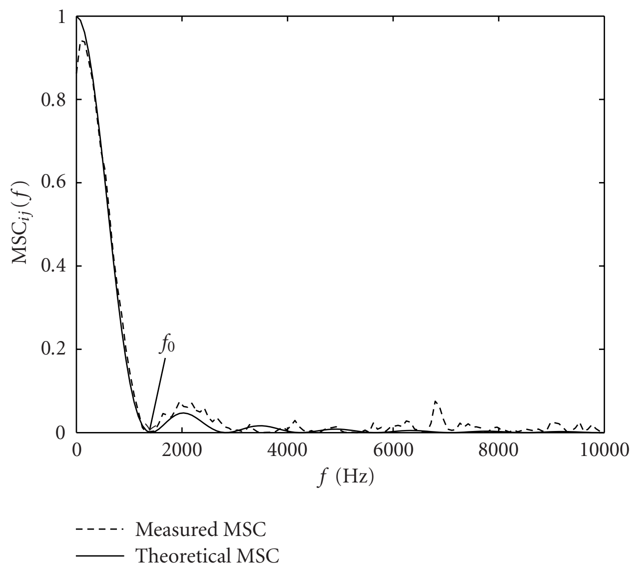


Figure 3: Theoretical MSC of a diffuse noise field and measured MSC inside a crowded cafeteria ($d_{ij} = 0.12\,\mathrm{m}$).

$A_i = c_i \cdot A$ and $\sigma_{S_i}^2 = c_i^2 \sigma_S^2$. Thus we can express $p(R_i \mid A_i = (c_i/c_n)A_n) = p(R_i \mid A_n)$. The joint pdf of all noisy amplitudes R_i given the speech amplitude of channel n can then be written as

$$\begin{aligned} p(R_1, \dots, R_M \mid A_n) = {} & \exp\left\{-\sum_{i=1}^{M} \frac{R_i^2 + (c_i/c_n)^2 A_n^2}{\sigma_{N_i}^2}\right\} \\ & \cdot \prod_{i=1}^{M} \left[\frac{2R_i}{\sigma_{N_i}^2} I_0\left(\frac{2(c_i/c_n)A_nR_i}{\sigma_{N_i}^2}\right)\right], \end{aligned} \tag{10}$$

where the c_i's are fixed parameters of the joint pdf. Similarly, the pdf of all noisy spectra Y_i conditioned on the clean speech amplitude and phase is

$$\begin{aligned} & p(Y_1, \dots, Y_M \mid A_n, \alpha_n) \\ & \quad = \prod_{i=1}^{M} \frac{1}{\pi \sigma_{N_i}^2} \cdot \exp\left(-\sum_{i=1}^{M} \frac{\left|Y_i - (c_i/c_n)A_n e^{j\alpha_i}\right|^2}{\sigma_{N_i}^2}\right). \end{aligned} \tag{11}$$

The unknown phases α_i can be expressed by α_n, the DOA, and the DFT frequency.

In analogy to the single channel MMSE estimator of the speech spectral amplitudes, the resulting joint estimators will be formulated in terms of a priori and a posteriori SNRs

$$\xi_i = \frac{\sigma_{S_i}^2}{\sigma_{N_i}^2}, \qquad \gamma_i = \frac{R_i^2}{\sigma_{N_i}^2}, \tag{12}$$

whereas the a posteriori SNRs γ_i can be directly computed, the a priori SNRs ξ_i are recursively estimated using the estimated speech amplitude $\hat{A}_i$ of the previous frame [9]:

$$\begin{aligned} \hat{\xi}_i(\lambda) &= \alpha \frac{\hat{A}_i^2(\lambda-1)}{\sigma_{N_i}^2} + (1-\alpha)P(\gamma_i(\lambda)-1) \\ &\text{with } P(x) = \begin{cases} x, & x > 0, \\ 0, & \text{else.} \end{cases} \end{aligned} \tag{13}$$

The smoothing factor α controls the trade-off between speech quality and noise reduction [16].

3. MULTICHANNEL SPECTRAL AMPLITUDE ESTIMATORS

We derive Bayesian estimators of the speech spectral amplitudes A_n, $n \in \{1, \dots, M\}$, using information from all M channels. First, a straightforward multichannel extension of the well-known MMSESTSA by Ephraim and Malah [9] is derived. Second, a practically more useful MAP estimator for DOA-independent noise reduction is introduced. All estimators output M spectral amplitudes A_n and thus M-enhanced signals are delivered by the noise reduction system.

3.1. *Estimation conditioned on complex spectra*

The single channel algorithm for channel number n derived by Ephraim and Malah calculates the expectation of the speech spectral amplitude A conditioned on the observed complex Fourier coefficient Y_n, that is, $E\{A_n | Y_n\}$. In the multichannel case, we can condition the expectation of each of the speech spectral amplitudes A_n on the joint observation of all M noisy spectra Y_i. To estimate the desired spectral amplitude of channel n, we have to calculate

$$\begin{aligned} \hat{A}_n &= E\{A_n | Y_1, \dots, Y_M\} \\ &= \int_0^\infty \int_0^{2\pi} A_n p(A_n, \alpha_n | Y_1, \dots, Y_M) d\alpha_n \, dA_n. \end{aligned} \tag{14}$$

This estimator can be expressed via Bayesian rule as

$$\hat{A}_n = \frac{\int_0^\infty A_n \int_0^{2\pi} p(A_n, \alpha_n) p(Y_1, \dots, Y_M | A_n, \alpha_n) d\alpha_n \, dA_n}{\int_0^\infty \int_0^{2\pi} p(A_n, \alpha_n) p(Y_1, \dots, Y_M | A_n, \alpha_n) d\alpha_n \, dA_n}. \tag{15}$$

To solve (15), we assume perfect DOA correction, that is, $\alpha_i := \alpha$ for all $i \in \{1, \dots, M\}$. Inserting $A_i = (c_i/c_n)A_n$ in (9) and (4), the integral over α in (15) becomes

$$\begin{aligned} I &= \int_0^{2\pi} \exp\left\{ -\sum_{i=1}^{M} \frac{\left| Y_i - (c_i/c_n) A_n e^{\alpha_n} \right|^2}{\sigma_{N_i}^2} \right\} d\alpha \\ &= \exp\left\{ -\sum_{i=1}^{M} \frac{\left| Y_i \right|^2 + \left((c_i/c_n) A_n \right)^2}{\sigma_{N_i}^2} \right\} \\ &\quad \times \int_0^{2\pi} \exp\{p \cos\alpha + q \sin\alpha\} d\alpha \end{aligned} \tag{16}$$

with

$$\begin{aligned} p &= \sum_{i=1}^{M} \frac{2c_i A_n}{c_n \sigma_{N_i}^2} \operatorname{Re}\{Y_i\}, \\ q &= \sum_{i=1}^{M} \frac{2c_i A_n}{c_n \sigma_{N_i}^2} \operatorname{Im}\{Y_i\}. \end{aligned} \tag{17}$$

The sum of sine and cosine is a cosine with different amplitude and phase:

$$p \cos\alpha + q \sin\alpha = \sqrt{p^2 + q^2} \cos\left(\alpha - \arctan\left(\frac{p}{q} \right) \right). \tag{18}$$

Since we integrate from 0 to 2π, the phase shift is meaningless. With

$$\sqrt{p^2 + q^2} = \left| \sum_{i=1}^{M} \frac{(c_i/c_n) Y_i}{\sigma_{N_i}^2} \right| \tag{19}$$

and $\int_0^\pi \exp\{z \cos x\} dx = \pi I_0(z)$, the integral becomes

$$\begin{aligned} I &= 2\pi \exp\left\{ -\sum_{i=1}^{M} \frac{\left| Y_i \right|^2 + \left((c_i/c_n) A_n \right)^2}{\sigma_{N_i}^2} \right\} \\ &\quad \times I_0\left(2A_n \left| \sum_{i=1}^{M} \frac{(c_i/c_n) Y_i}{\sigma_{N_i}^2} \right| \right). \end{aligned} \tag{20}$$

The remaining integrals over A_n can be solved using [17, equation (6.631.1)]. After some straightforward calculations, the gain factor for channel n is expressed as

$$\begin{aligned} G_n &= \frac{\hat{A}_n}{|Y_n|} = 1.5\gamma \cdot \sqrt{\frac{\xi_n}{\gamma_n \left(1 + \sum_{i=1}^{M} \xi_i \right)}} \\ &\quad \cdot F_1\left(-0.5, 1, \frac{\left| \sum_{i=1}^{M} \sqrt{\gamma_i \xi_i} e^{j\vartheta_i} \right|^2}{1 + \sum_{i=1}^{M} \xi_i} \right), \end{aligned} \tag{21}$$

where F_1 denotes the confluent hypergeometric series and Γ the Gamma function. The argument of F_1 contains a sum of a priori and a posteriori SNRs with respect to the noisy phases ϑ_i, $i \in \{1, \dots, M\}$. The confluent hypergeometric series F_1 has to be evaluated only once since the argument is independent of n. Note that in case of $M = 1$, (21) is the single channel MMSE estimator derived by Ephraim and Malah. In a practical real-time implementation, the confluent hypergeometric series is stored in a table.

3.2. Estimation conditioned on spectral amplitudes

The assumption $\alpha_i := \alpha$, $i \in \{1,\dots,M\}$, introduces a DOA dependency since this is only given for speech from $\theta = 0°$ or after perfect DOA correction. For a DOA-independent speech enhancement, we condition the expectation of A_n on the joint observation of all noisy amplitudes R_i, that is, $\hat{A}_n = E\{A_n|R_1,\dots,R_M\}$.

When the time delay of the desired signal s in Figure 2 between the microphones is small compared to the short-time stationarity of speech, the noisy amplitudes R_i are independent of the DOA θ. Unfortunately, after using (10), we have to integrate over a product of Bessel functions, which leads to extremely complicated expressions even for the simple case $M = 2$.

Therefore, searching for a closed-form estimator, we investigate a MAP solution which has been characterized in [11] as a simple but effective alternative to the mean square estimator in the single channel application.

We search for the speech spectral amplitude $\hat{A}_n$ that maximizes the pdf of A_n conditioned on the joint observation of all R_i, $i \in \{1,\dots,M\}$:

$$\begin{aligned}\hat{A}_n &= \arg\max_{A_n} p(A_n|R_1,\dots,R_M)\\ &= \arg\max_{A_n} \frac{p(R_1,\dots,R_M|A_n)\,p(A_n)}{p(R_1,\dots,R_M)}.\end{aligned} \tag{22}$$

We need to maximize only $L = p(R_1,\dots,R_M|A_n)\cdot p(A_n)$ since $p(R_1,\dots,R_M)$ is independent of A_n. It is however easier to maximize $\log(L)$, without effecting the result, because the natural logarithm is a monotonically increasing function. Using (10) and (3), we get

$$\begin{aligned}\log L = \log\left(\frac{A_n}{\pi\sigma_{S_n}^2}\right) - \frac{A_n^2}{\sigma_{S_n}^2}\\ + \sum_{i=1}^{M}\Bigg[\log\left(\frac{2R_i}{\sigma_{N_i}^2}\right) - \frac{R_i^2+(c_i/c_n)^2A_n^2}{\sigma_{N_i}^2}\\ + \log\left(I_0\left(2\frac{(c_i/c_n)A_nR_i}{\sigma_{N_i}^2}\right)\right)\Bigg].\end{aligned} \tag{23}$$

A closed-form solution can be found if the modified Bessel function I_0 is considered asymptotically with

$$I_0(x) \approx \frac{1}{\sqrt{2\pi x}}e^x. \tag{24}$$

Figure 4 shows that the approximation is reasonable for larger arguments and becomes erroneous only for very low SNRs.

Thus the term in the likelihood function containing the Bessel function is simplified to

$$\begin{aligned}&\log\left(I_0\left(2\frac{(c_i/c_n)A_nR_i}{\sigma_{N_i}^2}\right)\right)\\ &\quad\approx \frac{2(c_i/c_n)A_nR_i}{\sigma_{N_i}^2} - \frac{1}{2}\log\left(4\pi\frac{(c_i/c_n)A_nR_i}{\sigma_{N_i}^2}\right).\end{aligned} \tag{25}$$

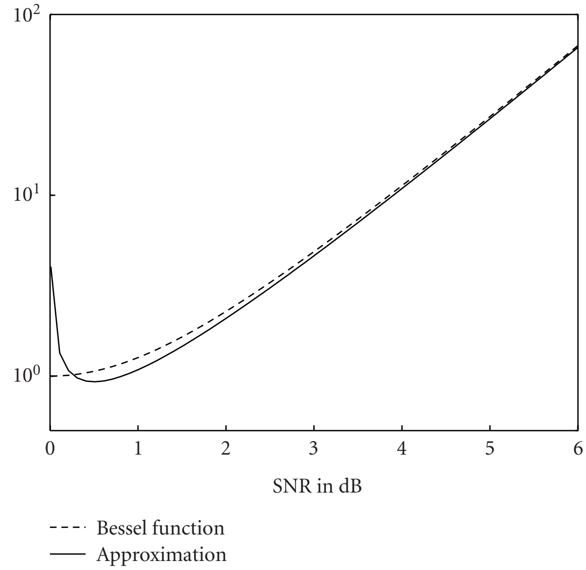


FIGURE 4: Bessel function and its approximation, $2(c_i/c_nA_nR_i)/\sigma_{N_i}^2 \approx 2\sqrt{\xi_i\gamma_i}$.

Differentiation of $\log L$ and multiplication with the amplitude A_n results in $A_n(\partial(\log L)/\partial A_n) = 0$:

$$A_n^2\left(-\frac{1}{\sigma_{S_n}^2} - \sum_{i=1}^{M}\frac{(c_i/c_n)^2}{\sigma_{N_i}^2}\right) + A_n\sum_{i=1}^{M}\frac{(c_i/c_n)R_i}{\sigma_{N_i}^2} + \frac{2-M}{4} = 0. \tag{26}$$

This quadratic expression can have two zeros; for $M > 2$, it is also possible that no zero is found. In this case, the apex of the parabolic curve in (26) is used as an approximation identical to the real part of the complex solution. The resulting gain factor of channel n is given as

$$\begin{aligned}G_n &= \frac{\hat{A}_n}{|Y_n|}\\ &= \frac{\sqrt{\xi_n/\gamma_n}}{2+2\sum_{i=1}^{M}\xi_i}\\ &\quad\cdot \mathrm{Re}\left\{\sum_{i=1}^{M}\sqrt{\gamma_i\xi_i}+\sqrt{\left(\sum_{i=1}^{M}\sqrt{\gamma_i\xi_i}\right)^2+(2-M)\left(1+\sum_{i=1}^{M}\xi_i\right)}\right\}.\end{aligned} \tag{27}$$

For the calculation of the gain factors, no exotic function needs to be evaluated any more. Also, $\mathrm{Re}\{\cdot\}$ has to be calculated only once since the argument is independent of n. Again, if $M = 1$, we have the single channel MAP estimator as given in [11].

4. EXPERIMENTAL RESULTS

In this section, we compare the performance of the joint speech spectral amplitude estimators with the well-known single channel Ephraim and Malah algorithm. Both M single channel estimators and the joint estimators output M-enhanced signals. In all experiments, we do not apply additional (commonly used) soft weighting techniques [9, 13] in order to isolate the benefits of the joint speech estimators compared to the single channel MMSE estimator.

All estimators were embedded in the DFT-based noise reduction system in Figure 1. The system operates at a sampling frequency of $f_s = 20$ kHz using half-overlapping Hann windowed frames. Both noise power spectral density $\sigma_{N_i}^2$ and variance of speech $\sigma_{S_i}^2$ were estimated separately for each channel. For the noise estimation task, we applied an elaborated version of *minimum statistics* [8] with adaptive recursive smoothing of the periodograms and adaptive bias compensation that is capable of tracking nonstationary noise even during speech activity.

To measure the performance, the noise reduction filter was applied to speech signals with added noise for different SNRs. The resulting filter was then utilized to process speech and noise separately [18]. Instead of only considering the segmental SNR improvement obtained by the noise reduction algorithm, this methods allows separate tracking of speech quality and noise reduction amount. The trade-off between speech quality and noise reduction amount can be regulated by, for example, changing the smoothing factor for the decision-directed speech power spectral density estimation (13). The speech quality of the noise-reduced signal was measured by averaging the segmental speech SNR between original and processed speech over all M channels. On the other hand, the amount of noise reduction was measured by averaging segmental input noise power divided by output noise power. Although the results presented here were produced with offline processing of generated or recorded signals, the system is well suited for real-time implementation.

The computational power needed is approximately M times that of one single channel Ephraim-Malah algorithm since for each microphone signal, an FFT, an IFFT, and an identical noise estimation algorithm are needed. The calculation of the a posteriori and a priori SNR (12) and (13) is also done independently for each channel. The joint estimators following (21) and (27) hardly increase the computational load, especially because $\mathrm{Re}(\cdot)$ and $F_1(\cdot)$ need to be calculated only once per frame and frequency bin.

4.1. Performance in artificial noise

To study the performance in ideal conditions, we first utilize the estimators on identical speech signals disturbed by spatially uncorrelated white noise. Figures 5 and 6 plot noise reduction and speech quality of the noise-reduced signal averaged over all M microphones for different number of microphones. While in Figure 5 the multichannel MMSE estimators according to (21) were applied, Figure 6 shows the performance of the multichannel MAP estimators according to (27). All joint estimators provide a significant higher speech quality and noise attenuation than the single channel MMSE estimator. The performance gain increases with the number of used microphones. The MAP estimators conditioned on the noisy amplitudes deliver a higher noise reduc-

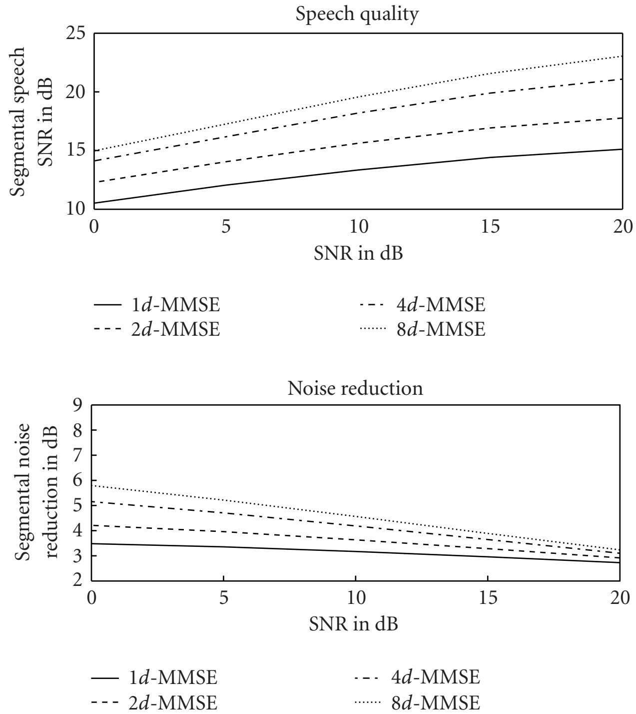


Figure 5: Speech quality and noise reduction of $1d$-MMSE estimators (reference) and Md-MMSE estimators with $M \in \{2, 4, 8\}$ for noisy signals containing identical speech and uncorrelated white noise.

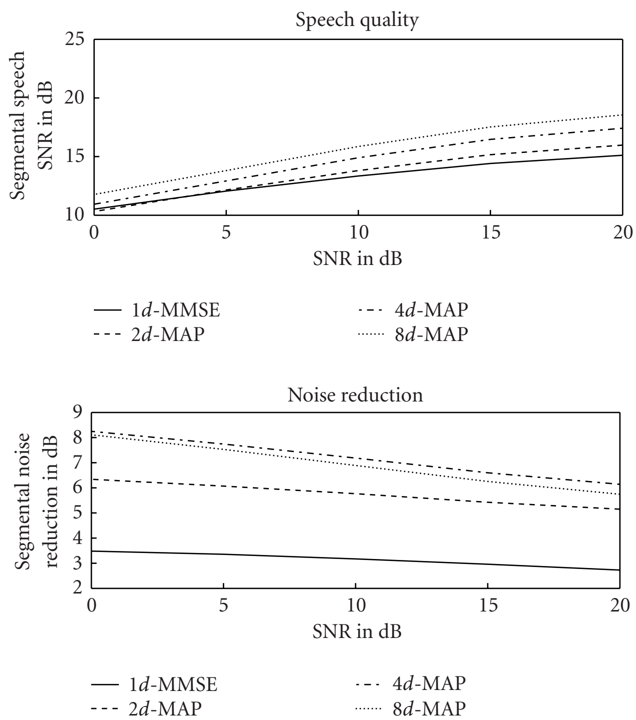


Figure 6: Speech quality and noise reduction of $1d$-MMSE estimators (reference) and Md-MAP with $M \in \{2, 4, 8\}$ for noisy signals containing identical speech and uncorrelated white noise.

tion than the multichannel MMSE estimator conditioned on the complex spectra at a lower speech quality. The gain in terms of noise reduction can be exchanged for a gain in terms of speech quality by different parameters.

4.2. *Performance in realistic noise*

Instead of uncorrelated white noise, we now mix the speech signal with noise recorded with a linear microphone array inside a crowded cafeteria. The coherence function of the approximately diffuse noise field is shown in Figure 3. Figure 7 plots the performance of the estimators using $M = 4$ microphones with an interelement spacing of $d = 12$ cm. Figure 8 shows the performance when using recordings with half the microphone distance, that is, $d = 6$ cm interelement spacing. The $4d$-MAP estimator provides both higher speech quality and higher noise reduction amount than the Ephraim-Malah estimator. In both cases, the multichannel MMSE estimator delivers a much higher speech quality at an equal or lower noise reduction. According to (6), the noise correlation increases with decreasing microphone distance. Thus, the performance gain of the multichannel estimators decreases. However, Figures 7 and 8 illustrate that significant performance gains are found at reasonable microphone distances.

Clearly, if the noise is spatially coherent, no performance gain can be expected by the multichannel spectral amplitude estimators. Compared to the $1d$-MMSE, the Md-MMSE and Md-MAP deliver a lower noise reduction amount at a higher speech quality when applied to speech disturbed by coherent noise.

4.3. *DOA dependency*

We examine the performance of the estimators when changing the DOA of the desired signal. We consider desired sources in both far and near field with respect to an array of $M = 4$ microphones with $d = 12$ cm.

4.3.1. *Desired signal in far field*

The far-field model assumes equal amplitudes and angle-dependent TDOAs:

$$s_i(t) = s(t - \tau_i(\theta)), \quad \tau_i = d \sin\left(\frac{\theta}{c}\right). \tag{28}$$

Figures 9 and 10 show the performance of the $4d$-estimators with cafeteria noise when the speech arrives from $\theta = 0°, 10°, 20°$, or $60°$ (see Figure 2). The performance of the MMSE estimator conditioned on the noisy spectra decreases with increasing angle of arrival. The speech quality decreases significantly, while the noise reduction amount is only slightly affected. This is because the phase assumption $\alpha_i = \alpha$, $i \in \{1, \ldots, M\}$ is not fulfilled.

On the other hand, the performance of the multichannel MAP estimator conditioned on the spectral amplitudes shows almost no dependency on the DOA.

4.3.2. *Desired signal in near field*

We investigate the performance when the source of the desired signal is located in the near field with distance ρ_i to

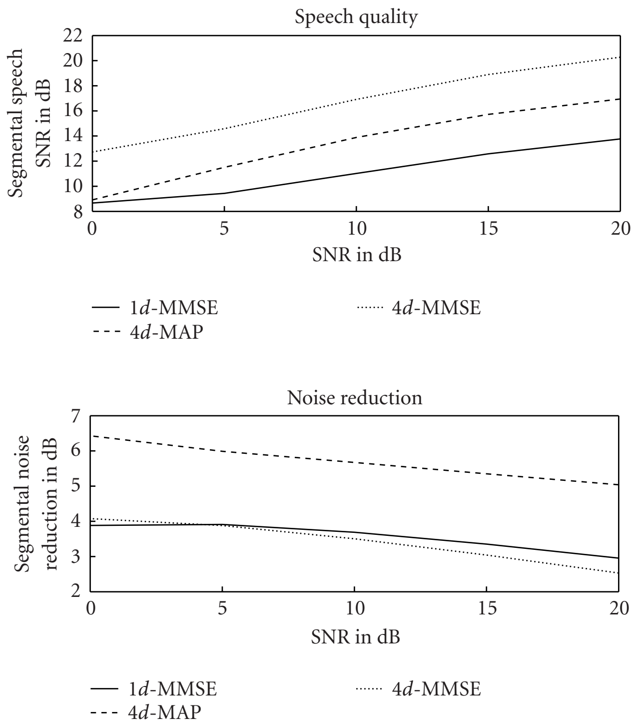


FIGURE 7: Speech quality and noise reduction of $1d/4d$-MMSE and $4d$-MAP for four signals containing identical speech and cafeteria noise (microphone distance $d = 12$ cm).

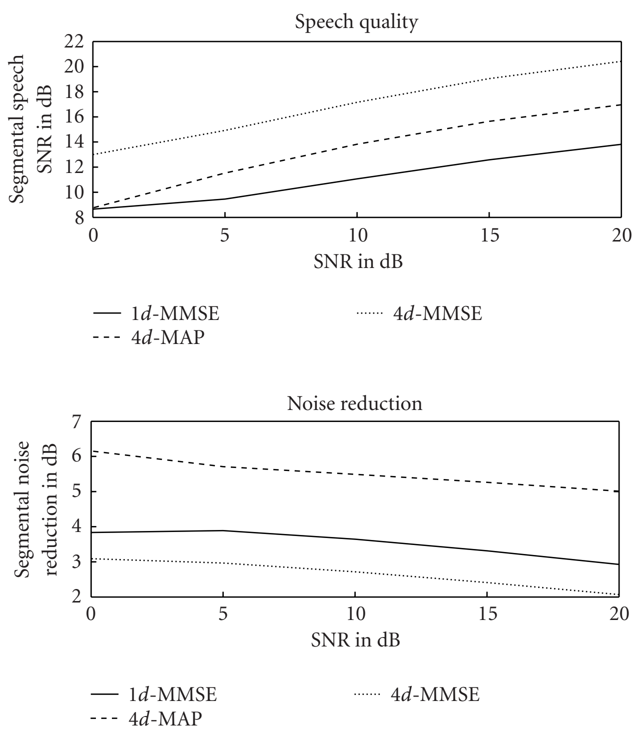


FIGURE 8: Speech quality and noise reduction of $1d/4d$-MMSE and $4d$-MAP for four signals containing identical speech and cafeteria noise (microphone distance $d = 6$ cm).

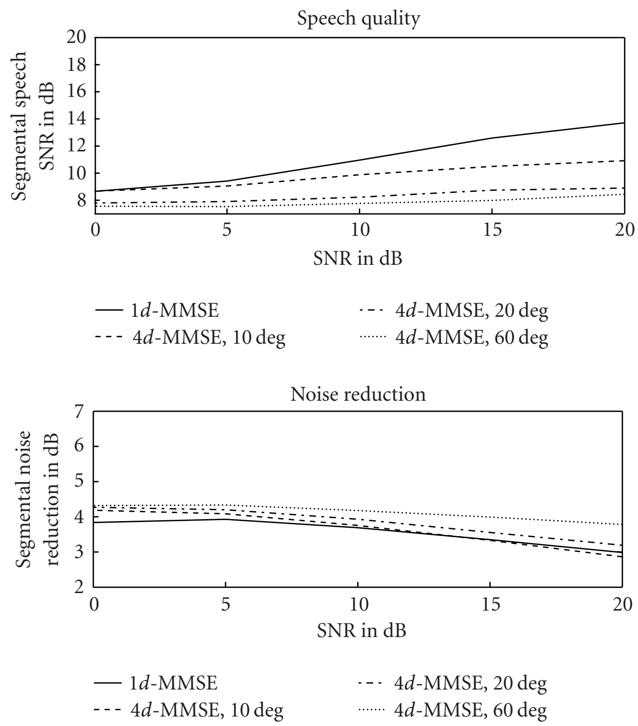


FIGURE 9: Speech quality and noise reduction of $4d$-MMSE compared to $1d$-MMSE for signals containing speech from $\theta = 10^\circ, 20^\circ$, and 60° and cafeteria noise (microphone distance $d = 12$ cm).

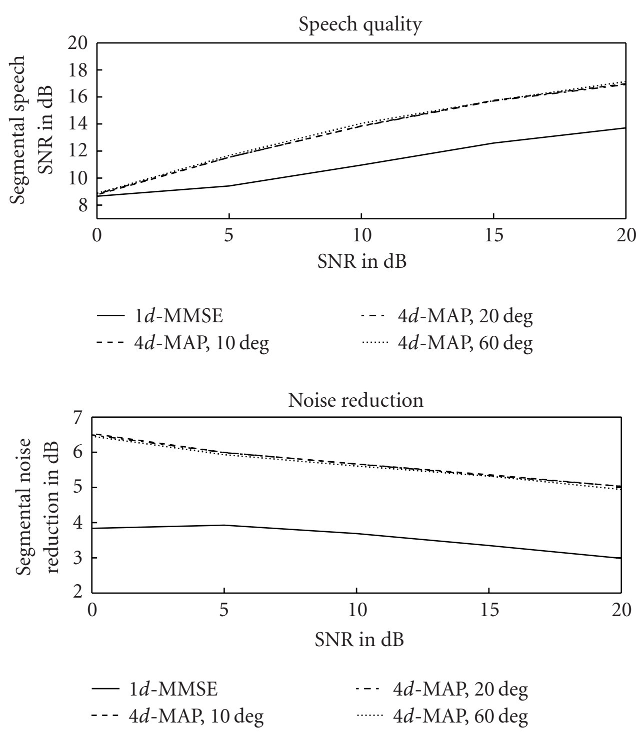


FIGURE 10: Speech quality and noise reduction of $4d$-MAP compared to $1d$-MMSE for signals containing speech from $\theta = 10^\circ, 20^\circ$ and 60° and cafeteria noise (microphone distance $d = 12$ cm).

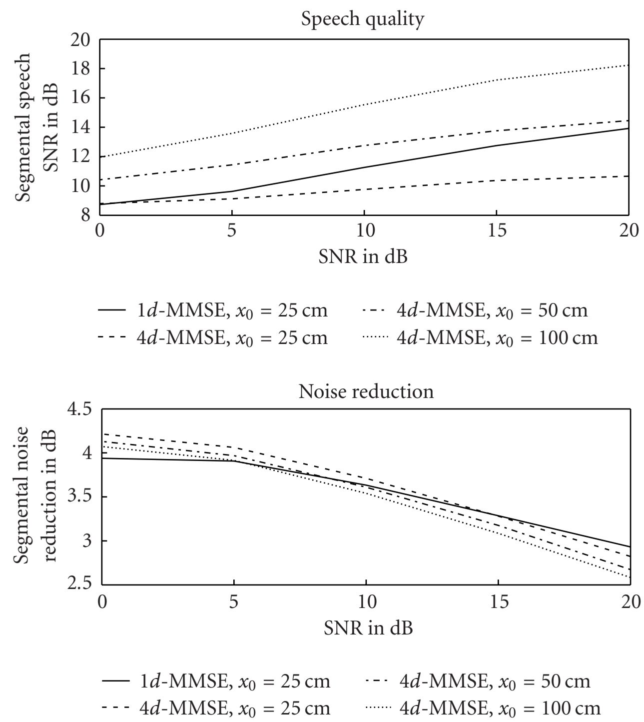


FIGURE 11: Speech quality and noise reduction of $4d$-MMSE compared to $1d$-MMSE for signals containing speech from $x_0 = 25$ cm, 50 cm, and 100 cm and cafeteria noise (microphone distance $d = 12$ cm).

microphone i. To simulate a near-field source, we use range-dependent amplifications and time differences:

$$s_i(t) = a_i s(t - \tau_i(\rho_i)), \tag{29}$$

where the amplitude factor for each channel decreases with the distance, $a_i \sim 1/\rho_i$. The source is located at different distances x_0 in front of the linear microphone array ($\theta = 0^\circ$) with $M = 4$ and $d = 12$ cm such that $\rho_i = \sqrt{x_0^2 + r_i^2}$, where r_i is defined in Figure 2.

Figures 11 and 12 show the performance of the $4d$-MMSE and $4d$-MAP estimators, respectively, when the source is located at $x_0 = 25$ cm, 50 cm, or 100 cm from the microphone array. The speech quality of the multichannel MMSE estimator decreases with decreasing distance. This is because at a higher distance from the microphone array, the time difference is smaller. Again, the multichannel MAP estimator conditioned on the noisy amplitudes shows nearly no dependency on the near-field position of the desired source.

4.4. *Reverberant desired signal*

Finally, we examine the performance of the estimators with a reverberant desired signal. Reverberation causes the spectral phases and amplitudes to become somewhat arbitrary, reducing the correlation of the desired signal. For the generation of reverberant speech signal, we simulate the acoustic situation depicted in Figure 13. The microphone array with

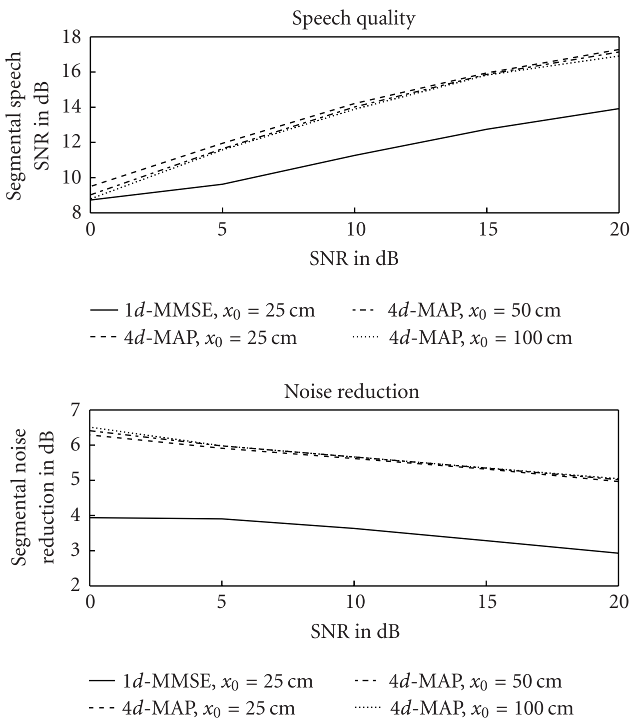


Figure 12: Speech quality and noise reduction of $4d$-MAP compared to $1d$-MMSE for signals containing speech from $x_0 = 25\,\text{cm}, 50\,\text{cm}$, and $100\,\text{cm}$ and cafeteria noise (microphone distance $d = 12\,\text{cm}$).

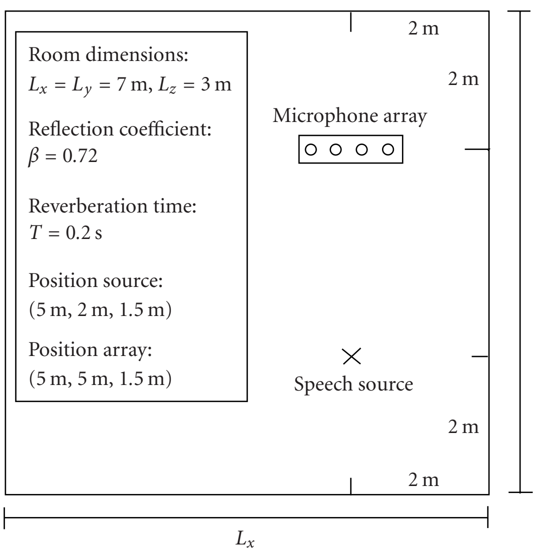


Figure 13: Speech source and microphone array inside a reverberant room.

$M = 4$ and an interelement spacing of $d = 12\,\text{cm}$ are positioned inside a reverberant room of size $L_x = 7\,\text{m}$, $L_y = 7\,\text{m}$, and $L_z = 3\,\text{m}$. A speech source is located three meters in front of the array.

The acoustical transfer functions from the source to each

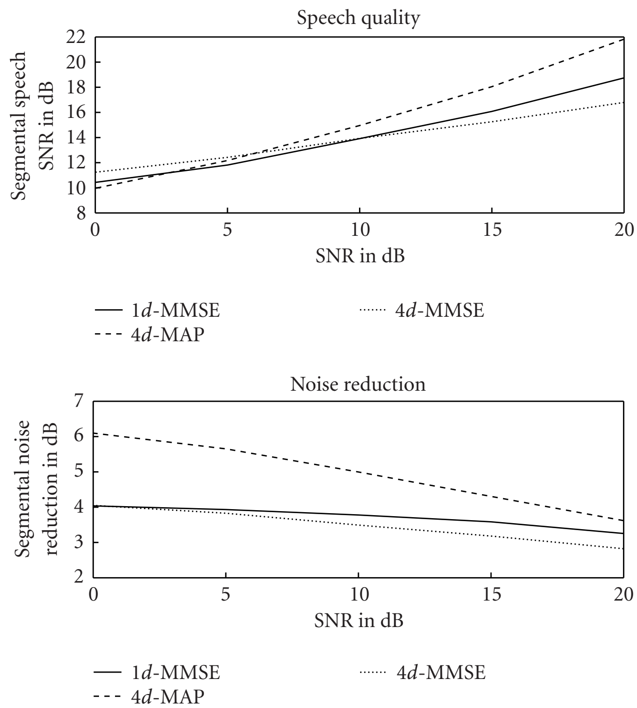


Figure 14: Speech quality and noise reduction of $1d/4d$-MMSE and $4d$-MAP for reverberant speech. (Figure 13) and cafeteria noise (microphone distance $d = 12\,\text{cm}$).

microphone were simulated with the image method [19], which models the reflecting walls by several image sources. The intensity of the sound from an image source at the microphone array is determined by a frequency-independent reflection coefficient β and by the distance to the array.

In our experiment, the reverberation time was set to $T = 0.2$ second, which corresponds to a defection coefficient $\beta = 0.72$ according to Eyring's formula

$$\beta = \exp\left\{-\frac{13.82}{\left(c\left(\frac{1}{L_x}+\frac{1}{L_y}+\frac{1}{L_z}\right)T\right)}\right\}. \tag{30}$$

Figure 14 shows the performance of the estimators when the reverberant speech signal is mixed with cafeteria noise. As expected, the overall performance gain obtained by the multichannel estimators decreases. However, there is still a significant improvement by the multichannel MAP estimator conditioned on the spectral amplitudes left. The multichannel MMSE estimator conditioned on the complex spectra performs worse due to its sensitivity to phase errors caused by reverberation.

5. CONCLUSION

We have derived analytically a multichannel MMSE and a MAP estimator of the speech spectral amplitudes, which can be considered as generalizations of [9, 11] to the multichannel case. Both estimators provide a significant gain compared to the well-known Ephraim-Malah estimator when the highly correlated speech components are in phase and the noise components are sufficiently uncorrelated.

The MAP estimator conditioned on the noisy spectral amplitudes performs multichannel speech enhancement independent of the position of the desired source in the near or the far field and is only moderately susceptible to reverberation. The multichannel noise reduction system is well suited for real-time implementation. It outputs multiple enhanced signals which can be combined by a beamformer for additional speech enhancement.

ACKNOWLEDGMENT

The authors would like to thank Rainer Martin for many inspiring discussions.

REFERENCES

[1] E. Gilbert and S. Morgan, "Optimum design of directive antenna arrays subject to random variations," *Bell System Technical Journal*, vol. 34, pp. 637–663, May 1955.

[2] M. Dörbecker, *Multi-channel algorithms for the enhancement of noisy speech for hearing aids*, Ph.D. thesis, Aachener Beiträge zu digitalen Nachrichtensystemen, vol. 10, Wissenschaftsverlag Mainz, Aachen, Germany, 1998, Aachen University (RWTH), P. Vary, Ed.

[3] J. Bitzer and K. Simmer, "Superdirective microphone arrays," in *Microphone Arrays: Signal Processing Techniques and Applications*, M. Brandstein and D. Ward, Eds., pp. 19–38, Springer-Verlag, Berlin, Germany, May 2001.

[4] L. Griffith and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagation*, vol. 30, no. 1, pp. 27–34, 1982.

[5] O. Hoshuyama and A. Sugiyama, "Robust adaptive beamforming," in *Microphone Arrays: Signal Processing Techniques and Applications*, M. Brandstein and D. Ward, Eds., pp. 87–109, Springer-Verlag, Berlin, Germany, 2001.

[6] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[7] J. DiBiase, H. Silverman, and M. Brandstein, "Robust localization in reverberant rooms," in *Microphone Arrays: Signal Processing Techniques and Applications*, M. Brandstein and D. Ward, Eds., pp. 157–180, Springer-Verlag, Berlin, Germany, 2001.

[8] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. Speech, and Audio Processing*, vol. 9, no. 5, pp. 504–512, 2001.

[9] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, 1984.

[10] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 443–445, 1985.

[11] P. Wolfe and S. Godsill, "Simple alternatives to the Ephraim and Malah suppression rule for speech enhancement," in *Proc. 11th IEEE Workshop on Statistical Signal Processing (SSP '01)*, pp. 496–499, Orchid Country Club, Singapore, August 2001.

[12] D. Brillinger, *Time Series, Data Analysis and Theory*, McGraw-Hill, New York, NY, USA, 1981.

[13] R. McAulay and M. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 28, no. 2, pp. 137–145, 1980.

[14] R. Martin, "Speech enhancement using MMSE short time spectral estimation with Gamma distributed speech priors," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '02)*, Orlando, Fla, USA, May 2002.

[15] P. Vary, "Noise suppression by spectral magnitude estimation - mechanism and theoretical limits," *Signal Processing*, vol. 8, no. 4, pp. 387–400, 1985.

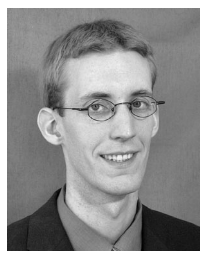
[16] O. Cappe, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Trans. Speech, and Audio Processing*, vol. 2, no. 2, pp. 345–349, 1994.

[17] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, San Diego, Calif, USA, 1994.

[18] S. Gustafsson, R. Martin, and P. Vary, "On the optimization of speech enhancement systems using instrumental measures," in *Proc. Workshop on Quality Assessment in Speech, Audio, and Image Communication*, pp. 36–40, Darmstadt, Germany, March 1996.

[19] J. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal Acoustic Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

Thomas Lotter received the Diploma of Engineering degree in electrical engineering from Aachen University of Technology (RWTH), Germany, in 2000. He is now with the Institute of Communication Systems and Data Processing (IND), Aachen University of Technology, where he is currently pursuing the Ph.D. degree. His main research interests are in the areas of speech and audio processing, particularly in speech enhancement with single and multimicrophone techniques.



Christian Benien received the Diploma of Engineering degree in electrical engineering from Aachen University of Technology (RWTH), Germany, in 2002. He is now with Philips Research in Aachen. His main research interests are in the areas of speech enhancement, speech recognition, and the development of interactive dialogue systems.

Peter Vary received the Diploma of Engineering degree in electrical engineering in 1972 from the University of Darmstadt, Darmstadt, Germany. In 1978, he received the Ph.D. degree from the University of Erlangen-Nuremberg, Germany. In 1980, he joined Philips Communication Industries (PKI), Nuremberg, where he became Head of the Digital Signal Processing Group. Since 1988, he has been Professor at Aachen University of Technology, Aachen, Germany, and Head of the Institute of Communication Systems and Data Processing. His main research interests are speech coding, channel coding, error concealment, adaptive filtering for acoustic echo cancellation and noise reduction, and concepts of mobile radio transmission.

Special Issue on Inverse Synthetic Aperture Radar

Call for Papers

Inverse synthetic aperture radar (ISAR) is a powerful signal processing technique that can provide an electromagnetic image of a target. ISAR images may be obtained by coherently processing the received target echoes of wide bandwidth transmitted pulses. This technique is typically applied to a stationary monostatic radar configuration observing a moving target, and relies on the target's rotation to form the synthetic aperture. ISAR imaging techniques have been extensively employed over the last few years in improving target classification algorithms, specifically those attempting to identify ship, airborne, and orbiting targets. These improvements have been made possible through the advances that have occurred in signal processing techniques, such as those made in ISAR blind motion compensation or autofocussing, polarimetry-based classification, super resolution, and the suppression or exploitation of multipath effects.

Furthermore, the recent resurgence of bistatic and multistatic radars has resulted in an awareness of extra parameters in the fundamental ISAR imaging processes that provide, for example, 3D and/or interferometric capabilities. The significant advances made in computing technology also impacts on the ability of the user to employ these new signal processing techniques in applications that require rapid target identification.

The goal of this special issue is to discuss the state of the art in ISAR imaging and signal processing techniques.

This special issue will focus on such seamless integration of visual analysis methods in, or joint design with, robust compression and transmission solutions.

Topics of interest include (but are not limited to):

- Bistatic/multistatic ISAR
- 3D / interferometric ISAR
- ISAR-in the presence of multipath signals
- ISAR-based target classification
- ISAR autofocussing / blind motion compensation algorithms
- Polarimetric ISAR
- Super-resolution techniques for ISAR
- ISAR imaging for maneuvering targets

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	June 1, 2005
Acceptance Notification	September 1, 2005
Final Manuscript Due	December 1, 2005
Publication Date	1st Quarter, 2006

GUEST EDITORS:

John Homer, School of Information Technology and Electrical Engineering, Cooperative Research Centre for Sensor Signal and Information Processing, The University of Queensland, Brisbane, Australia; homerj@itee.uq.edu.au

James Palmer, School of Information Technology and Electrical Engineering, Wedgetail TRDC, The University of Queensland, Brisbane, Australia; palmer@itee.uq.edu.au

Marco Martorella, Department of Information Engineering, University of Pisa, Via Caruso, Pisa, Italy; m.martorella@iet.unipi.it

Brad Littleton, School of Physics, The University of Queensland, Brisbane, Australia; littleton@physics.uq.edu.au

Fabrizio Berizzi, Department of Information Engineering, University of Pisa, Via Caruso, Pisa, Italy; f.berizzi@iet.unipi.it

Victor Chen, US Naval Research Laboratory, Washington DC, USA; vchen@radar.nrl.navy.mil

Dennis Longstaff, School of Information Technology and Electrical Engineering, Cooperative Research Centre for Sensor Signal and Information Processing, The University of Queensland, Brisbane, Australia; idl@itee.uq.edu.au

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on Reliable Communications over Rapidly Time-Varying Channels

Call for Papers

Wireless communications have become an important part of everyday life. Think for instance about mobile telephone applications, wireless local area networks (WLANs), wireless ad hoc networks, and so forth. Most of these systems have been designed assuming that the channel can be regarded as constant over a block of data. Nonetheless, market studies predict a rapid growth of high data rate mobile applications such as watching TV on mobile phones. In such mobile applications, Doppler shifts introduce temporal channel variations, which become more pronounced as the carrier frequency increases, and basically violate the time-invariance assumption. As a result, many existing wireless systems can only provide low data rates at high mobility (e.g., UMTS) or even break down completely at high speeds (e.g., DVB-T and IEEE802.16).

This special issue therefore focuses on communications over rapidly time-varying channels, which can not be viewed as time invariant over a frame. Different time-varying channel models have recently been proposed, such as the basis expansion model and the Gauss-Markov model. Results are welcomed on how to estimate the channel parameters for such models, and, related to that, what is the optimal training strategy. In addition, low-complexity equalization schemes for time-varying channels should receive some attention, as well as joint precoder-decoder designs to boost the performance. Also, the behavior of existing multiple-access schemes in rapidly time-varying channels, such as the well-known code-division multiple-access (CDMA) scheme, as well as the development of novel multiple-access schemes for rapidly time-varying channels are important research topics that require further investigation. Finally, multiple-input multiple-output (MIMO) communications and space-time coding (STC) over time-varying channels are very new areas that urgently need to be covered.

This special issue is intended to gather new and insightful results on wireless communications over rapidly time-varying channels, a challenging research topic that gains increasing attention due to its importance in future wireless applications. The results might for instance be useful in the frame of the mobile extensions of DVB-T and IEEE802.16, e.g., DVB-M and IEEE802.20.

Topics of interest include (but are not limited to):

- Channel modeling, estimation, and equalization
- Optimal training
- Joint precoder/decoder design
- Multiple-access schemes
- MIMO communications
- Space-time coding
- Information theoretic results

Authors should follow the EURASIP JASP manuscript format described at the journal site http://asp.hindawi.com/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	June 1, 2005
Acceptance Notification	October 1, 2005
Final Manuscript Due	January 1, 2006
Publication Date	2nd Quarter, 2006

GUEST EDITORS:

Geert Leus, Delft University of Technology, 2611 WE Delft, The Netherlands; leus@cas.et.tudelft.nl

Georgios Giannakis, University of Minnesota, Minneapolis, MN 55455, USA; georgios@ece.umn.edu

Jean-Paul Linnartz, Philips Research Laboratories, 5656AA Eindhoven, The Netherlands; j.p.linnartz@philips.com

Xiaoli Ma, Auburn University, Auburn, AL 36849, USA; xiaoli@eng.auburn.edu

Ananthram Swami, Army Research Laboratory, Adelphi, MD 20783, USA; a.swami@ieee.org

Cihan Tepedelenlioglu, Arizona State University, Tempe, AZ 85287-5706, USA; cihan@asu.edu

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on
Wireless Location Technologies and Applications

Call for Papers

The development of communications systems that include location and tracking capabilities has generated great interest in cellular and wireless local/personal area networks. A host of potential services can be enabled by suitably accurate location and tracking facilities in conjunction with appropriate communications and data transfer platforms.

From established radio techniques, such as WLAN (e.g. IEEE 802.11 a/b/g), to emerging WPAN networks (e.g. Bluetooth and IEEE 802.15.4) to newer Ultra-Wideband (UWB) systems (e.g. IEEE 802.15.3a and IEEE 802.15.4a), a common denominator to drive adoption and growth is implementing innovative services in addition to data transfer. Whether the positioning techniques are based on signal strength, time of flight, or on fingerprinting techniques, they offer the potential for new applications which rely on the knowledge of the location of the wireless nodes.

A large number of issues must be addressed to move from coarse delay measurement to useful range estimation for tracking purposes. Issues include generation of accurate delay/ranging estimates, proper operation in dense multipath environments, delay/ranging information sharing between nodes, computationally efficient algorithms, algorithms for low infrastructure environments, dealing with NLOS as well as integration with usable applications. This special issue will address the state of the art in wireless location technologies and applications with particular emphasis on accurate results in low infrastructure environments.

Topics of interest include (but are not limited to):

- Delay/ranging estimation techniques for wireless systems
- Location techniques for ad hoc networks
- Tracking algorithm performance in indoor systems
- Estimation of processing power requirements of competing techniques
- System architectures
- Multiantenna systems for positioning
- Fundamental accuracy limits
- Delay/ranging estimation techniques for NLOS conditions
- Field trials of ranging or positioning devices

Authors should follow the EURASIP JASP manuscript format described at the journal site http://asp.hindawi.com/. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	June 1, 2005
Acceptance Notification	October 1, 2005
Final Manuscript Due	January 1, 2006
Publication Date	2nd Quarter, 2006

GUEST EDITORS:

Ian Oppermann, CWC, University of Oulu, PO Box 45000, 90014 Oulun yliopisto, Finland; ian@ee.oulu.fi

Domenico Porcino, Philips Research Laboratories, Wireless Group, Cross Oak Lane, Redhill, Surrey, UK, RH15HA; domenico.porcino@philips.com

Thomas Kaiser, Department of Communication Systems, University of Duisburg-Essen, Bismarckstr. 81, 47057 Duisburg, Germany; thomas.kaiser@uni-duisburg.de

Special Issue on

Advanced Video Technologies and Applications for H.264/AVC and Beyond

Call for Papers

The recently developed video coding standard H.264/MPEG-4 AVC significantly outperforms previous standards in terms of coding efficiency at reasonable implementation complexity and costs in VLSI realization. Real-time H.264 coders will be available very soon. Many applications, such as surveillance systems with multiple video channel recording, multiple channel video services for mobile devices, will benefit from the H.264 coder due to its excellent coding efficiency. The new video coding technology introduces new opportunities for video services and applications. However, advanced video coding is only one aspect for successful video services and applications. To enable successful new applications, additional technologies to cope with time-varying channel behaviors and diverse usage characteristics are needed. For serving multiple videos, some extended designs such as joint rate-distortion optimization and scheduling of multiple parallel video sessions are also required to achieve fair and robust video storage and delivery. For video surveillance systems, intelligent video content analysis and scalabilities in video quality, resolution, and display area, coupled with wireless transmission, can offer new features for the application. Finally, computational complexity reduction and low-power design of video codecs as well as content protection of video streams are particularly important for mobile devices.

The goal of this special issue is to discuss state-of-the-art techniques to enable various video services and applications on H.264/AVC technologies and their new developments.

Topics of interest include (but are not limited to):

- Video over DVB-H
- Error resilience of video over mobile networks
- Video delivery in multiuser environments
- Rate-distortion optimization for multiple video sources
- Multipath delivery of video streams
- Optimization of video codecs for quality improvement and power reduction
- Security and content protection of video streams
- Transcoding techniques
- Scalable video
- Other advanced video coding technologies
- Video quality measures

Authors should follow the EURASIP JASP manuscript format described at the journal site http://asp.hindawi.com/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP's manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	August 1, 2005
Acceptance Notification	November 1, 2005
Final Manuscript Due	March 1, 2006
Publication Date	3rd Quarter, 2006

GUEST EDITORS:

Jar-Ferr Yang, National Cheng Kung University, Tainan, Taiwan; jfyang@ee.ncku.edu.tw

Hsueh-Ming Hang, National Chiao Tung University, Hsinchu 300, Taiwan; hmhang@mail.ncku.edu.tw

Eckehard Steinbach, Munich University of Technology, Munich, Germany; eckehard.steinbach@tum.de

Ming-Ting Sun, University of Washington, Seattle, Washington, USA; sun@ee.washington.edu

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on Advances in Blind Source Separation

Call for Papers

Almost every multichannel measurement includes mixtures of signals from several underlying sources. While the structure of the mixing process may be known to some degree, other unknown parameters are necessary to demix the measured sensor data. The time courses of the source signals and/or their locations in the source space are often unknown a priori and can only be estimated by statistical means. In the analysis of such measurements, it is essential to separate the mixed signals before beginning postprocessing.

Blind source separation (BSS) techniques then allow separation of the source signals from the measured mixtures. Many BSS problems may be solved using independent component analysis (ICA) or alternative approaches such as sparse component analysis (SCA) or nonnegative matrix factorization (NMF), evolving from information theoretical assumptions that the underlying sources are mutually statistically independent, sparse, smooth, and/or nonnegative.

The aim of this special issue is to focus on recent developments in this expanding research area.

The special issue will focus on one hand on theoretical approaches for single- and multichannel BSS, evolving from information theory, and especially on nonlinear blind source separation methods, and on the other hand or their currently ever-widening range of applications such as brain imaging, image coding and processing, dereverberation in noisy environments, and so forth.

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	October 1, 2005
Acceptance Notification	February 1, 2006
Final Manuscript Due	May 1, 2006
Publication Date	3rd Quarter, 2006

GUEST EDITORS:

Scott Makeig, Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California, San Diego, La Jolla, CA 92093-0961, USA; smakeig@ucsd.edu

Andrzej Cichocki, Laboratory for Advanced Brain Signal Processing, Brain Science Institute, The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan; cia@brain.riken.go.jp

Frank Ehlers, Federal Armed Forces Underwater Acoustics and Marine Geophysics Research Institute, Klausdorfer Weg 2-24, 24148 Kiel, Germany; frankehlers@ieee.org

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on Tracking in Video Sequences of Crowded Scenes

Call for Papers

Object tracking in live video is an enabling technology that is in strong demand by large application sectors, such as video surveillance for security and behavior analysis, traffic monitoring, sports analysis for enhanced TV broadcasting and coaching, and human body tracking for human-computer interaction and movie special effects.

Many techniques and systems have been developed and demonstrated for tracking objects in video sequences. The specific goal of this special issue is to provide a status report regarding the state of the art in object tracking in crowded scenes based on the video stream(s) of one or more cameras. The objects can be people, animals, cars, and so forth. The cameras can be fixed or moving. Moving cameras may pan, tilt, and zoom in ways that may or may not be communicated to the tracking system.

All papers submitted must address at least the following two issues:

- Processing of live video feeds

For many applications in surveillance/security and TV sports broadcasting, the results of processing have value only if they can be provided to the end user within an application-defined delay. The submitted papers should present algorithms that are plausibly applicable to such incremental ("causal") processing of live video feeds, given suitable hardware.

- Handling of crowded scenes

Crowded-scene situations range from relatively simple (e.g., players on a planar field in a soccer match) to very difficult (e.g., crowds on stairs in an airport or a train station). The central difficulties in crowded scenes arise from the constantly changing occlusions of any number of objects by any number of other objects.

Occlusions can be resolved to some degree using a single video stream. However, many situations of occlusion are more readily resolved by the simultaneous use of several cameras separated by wide baselines. In addition to resolving ambiguities, multiple cameras also ease the exploitation of 3D structure, which can be important for trajectory estimation or event detection.

Topics of interest include principles and evaluation of relevant end-to-end systems or important components thereof, including (but not limited to):

- Handling of occlusions in the image plane in single-camera scenarios
- Handling of occlusions in a world coordinate system (3D, possibly degenerated to 2D) in single- or multi-camera scenarios
- Fusion of information from multiple cameras and construction of integrated spatiotemporal models of dynamic scenes
- 3D trajectory estimation
- Tracking of multiple rigid, articulated, or nonrigid objects
- Automatic recovery of camera pose from track data
- Detection and recognition of events involving multiple objects (e.g., offside in soccer)

Papers must present a thorough evaluation of the performance of the system or method(s) proposed in one or more application areas such as video surveillance, security, sports analysis, behavior analysis, or traffic monitoring.

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	October 1, 2005
Acceptance Notification	February 1, 2006
Final Manuscript Due	May 1, 2006
Publication Date	3rd Quarter, 2006

GUEST EDITORS:

Jacques G. Verly, Department of Electrical Engineering and Computer Science, University of Liège (ULg), Sart Tilman, Building B28, 4000 Liège, Belgium; jacques.verly@ulg.ac.be

John MacCormick, Microsoft Research, Silicon Valley, 1065 La Avenida Mountain View, CA 94043, USA; jmacc@microsoft.com

Stephen McKenna, Division of Applied Computing, University of Dundee, Dundee DD1 4HN, Scotland, UK; stephen@computing.dundee.ac.uk

Justus H. Piater, Department of Electrical Engineering and Computer Science, University of Liège (ULg), Sart Tilman, Building B28, 4000 Liège, Belgium; justus.piater@ulg.ac.be

Hindawi Publishing Corporation
http://www.hindawi.com

EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING

Special Issue on

Advances in Subspace-Based Techniques for Signal Processing and Communications

Call for Papers

Subspace-based techniques have been studied extensively over the past two decades and have proven to be very powerful for estimation and detection tasks in many signal processing and communications applications. Such techniques were initially investigated in the context of super-resolution parametric spectral analysis and the related problem of direction finding. During the past decade or so, new potential applications have emerged, and subspace methods have been proposed in several diverse fields such as smart antennas, sensor arrays, system identification, time delay estimation, blind channel estimation, image segmentation, speech enhancement, learning systems, and so forth.

Subspace-based methods not only provide new insight into the problem under investigation but they also offer a good trade-off between achieved performance and computational complexity. In most cases they can be considered as low cost alternatives to computationally intensive maximum likelihood approaches.

The interest of the signal processing community in subspace-based schemes remains strong as is evident from the numerous articles and reports published in this area each year. Research efforts are currently focusing on the development of low-complexity adaptive implementations and their efficient use in applications, numerical stability, convergence analysis, and so forth.

The goal of this special issue is to present state-of-the-art subspace techniques for modern applications and to address theoretical and implementation issues concerning this useful methodology.

Topics of interest include (but are not limited to):

- Efficient and stable subspace estimation and tracking methods
- Subspace-based detection techniques
- Sensor array signal processing
- Smart antennas
- Space-time, multiuser, multicarrier communications
- System identification and blind channel estimation
- State-space model estimation and change detection
- Learning and classification
- Speech processing (enhancement, recognition)
- Biomedical signal processing
- Image processing (face recognition, compression, restoration)

Authors should follow the EURASIP JASP manuscript format described at http://www.hindawi.com/journals/asp/. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at http://www.mstracking.com/asp/, according to the following timetable:

Manuscript Due	October 1, 2005
Acceptance Notification	February 1, 2006
Final Manuscript Due	May 1, 2006
Publication Date	3rd Quarter, 2006

GUEST EDITORS:

Kostas Berberidis, University of Patras, 26500 Patras, Greece; berberid@ceid.upatras.gr

Benoit Champagne, McGill University, Québec, Canada H3A 2T5; champagne@ece.mcgill.ca

George V. Moustakides, University of Thessaly, 38221 Volos, Greece; moustaki@uth.gr

H. Vincent Poor, Princeton University, Princeton, NJ 08544, USA; poor@princeton.edu

Peter Stoica, Uppsala University, 75105 Uppsala, Sweden; peter.stoica@it.uu.se

EURASIP Journal on

Embedded Systems

:: Open Access ::

Website: http://www.hindawi.com/journals/es/

Aims and Scope

"EURASIP Journal on Embedded Systems" is an international journal that serves the large community of researchers and professional engineers who deal with the theory and practice of embedded systems, particularly encompassing all practical aspects of theory and methods used in designing homogeneous as well as heterogeneous embedded systems that combine data-driven and control-driven behaviors.

Original full and short papers, correspondence, and reviews on design and development of embedded systems, methodologies applied for their specification, modeling and design, and adaptation of algorithms for real-time execution are encouraged for submission.

The coverage includes complex homogeneous and heterogeneous embedded systems, specification languages and tools for embedded systems, modeling and verification techniques, hardware/software trade-offs and codesign, new design flows, design methodologies and synthesis methods, platform-based design, component-based design, adaptation of signal processing algorithms to limited implementation resources, rapid prototyping, computing structures and architectures for complex embedded systems, real-time operating systems, methods and techniques for the design of low-power systems, interfacing with the real world, and novel application case studies and experiences. The coverage, however, does not exclude other interesting related and emerging topics like software-defined radio. Example applications include wireless and data communication systems, speech processing, image and video processing, digital signal processing applications, as well as control and instrumentation.

There are two types of issues: regular issues and special issues. The regular issues publish collections of papers without special solicitation. The special issues have specifically aimed and targeted topics of interest contributed by authors responding to a particular Call-for-Papers or by invitation, edited by invited guest editor(s). Regular papers can be submitted at any time, while special issue papers can be submitted only based on planned schedules and submission guidelines of the Call-for-Papers.

Special Issue Proposals

Proposals for special issues can be submitted directly to the Editor-in-Chief or to es.si@hindawi.com.

Open Access

The EURASIP JES, as an open access journal, enables immediate, worldwide, barrier-free online access to the full text of research articles for the best interests of the scientific community. All interested readers can read, download, and/or print any article published in the EURASIP JES.

Hindawi Publishing Corporation
410 Park Avenue, 15th Floor, #287 pmb, New York, NY 10022, USA

HINDAWI

A publication of the European Association for Speech, Signal, and Image Processing