EFFICIENT COMPUTATION OF MICROPHONE UTILITY IN A WIRELESS ACOUSTIC SENSOR NETWORK WITH MULTI-CHANNEL WIENER FILTER BASED NOISE REDUCTION

Joseph Szurley, Alexander Bertrand, Marc Moonen

Electrical Engineering Dept. (ESAT-SCD) Katholieke Universiteit Leuven
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
E-mail: joseph.szurley@esat.kuleuven.be, alexander.bertrand@esat.kuleuven.be, marc.moonen@esat.kuleuven.be

ABSTRACT

A wireless acoustic sensor network is considered with spatially distributed microphones which observe a desired speech signal that has been corrupted by noise. In order to reduce the noise the signals are sent to a fusion center where they are processed with a centralized rank-1 multi-channel Wiener filter (R1-MWF). The goal of this work is to efficiently compute an assessment of the contribution of each individual microphone with respect to either signal-to-noise ratio (SNR), signal-to-distortion ratio (SDR) or the minimized cost function referred to as the utility. These performance measures are derived by exploiting unique properties of the R1-MWF which can be computed efficiently from values that are known from the current signal estimation process. The performance measures may be used in unison or individually to determine the contributions of each microphone and help facilitate in selecting only a subset of the available signals in order to meet the bandwidth and power constraints of the system.

Index Terms— Wireless Acoustic Sensor Networks, Multi-Channel Wiener Filtering, Sensor Subset Selection

1. INTRODUCTION

Sensor networks are often deployed over large areas enabling greater information about the spatial properties of the sensing environment [1, 2]. Wireless sensor networks (WSN) take advantage of a collection of wireless devices that can be used to relay information between one another with some predefined task as an ultimate goal. In regards to audio applications these devices use available microphone signals on the devices to enhance an audio signal and form a wireless acoustic sensor network (WASN).

In WSNs there is often a desire to only use a fraction of the available signals in order to conserve network lifetime and adhere to bandwidth constraints while maintaining signal estimation accuracy. Finding the optimal *subset* of signals is often an intractable task and therefore a way to assess the signals in their order of the importance from the current estimation is essential.

This research work was carried out at the ESAT Laboratory of Katholieke Universiteit Leuven, in the frame of K.U.Leuven Research Council CoE EF/05/006 'Optimization in Engineering' (OPTEC) and PFV/10/002 (OPTEC), Concerted Research Action GOA-MaNet, the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 'Dynamical systems, control and optimization' (DYSCO) 2007-2011, Research Project FWO nr. G.0600.08 'Signal processing and network design for wireless acoustic sensor networks'. Alexander Bertrand is supported by a Postdoctoral Fellowship of the Research Foundation Flanders (FWO). The scientific responsibility is assumed by its authors.

In the WASN envisaged for this paper a desired speech signal which has been corrupted by noise is captured by a set of spatially distributed microphones. These microphone signals are sent to a fusion center where all of the data from the WASN is aggregated and processed. Using the available information an optimal filter in the linear minimum mean squared error (MMSE) sense is derived which in this paper takes the form of the rank-1 multi-channel Wiener filter (R1-MWF). The utility of each individual microphone is derived using the R1-MWF formulation which differs when compared to the derivation in [3, 4] which relied on the classical Speech Distortion Weighted Multi-channel Wiener Filter (SDW-MWF) formulation. The R1-MWF relies on the inversion of the noise correlation matrix which has been shown to be numerically more robust than the SDW-MWF [5], and due to its unique properties, allows for other pertinent information to be extracted in a computationally efficient manner.

In deriving the utility function from the R1-MWF other performance measures can be computed to assess the contribution of each microphone. In particular we show the contribution of each microphone to the output Signal-to-Noise Ratio (SNR) and Signal-to-Distortion Ratio (SDR) can be found concurrently with no addition to the computational complexity. These values may then be used in conjunction with a combination of thresholds or weights to determine an explicit trade-off between the full received signal and the optimal subset that is application dependent. They may also be applied in unison so that a psycho-acoustic model that mimics the human hearing spectrum can be used to facilitate signal subset selection but this is beyond the scope of this paper.

This paper is organized as follows. Section 2 introduces the problem formulation and notation used throughout the text. Section 3 defines three microphone specific performance measures, utility, SNR and SDR from the current known values of the R1-MWF. Section 4 discusses how to accurately monitor the individual signal contributions and their relationship to the performance measures. Section 5 employs a toy room scenario which gives the time averaged values of the performance measures in a simulated acoustic environment.

2. PROBLEM FORMULATION AND NOTATION

Consider a wireless acoustic sensor network with M spatially distributed microphones. The short-time Fourier transform (STFT) representation of the received signal at microphone k is given by

$$y_k[\omega, \mathrm{t}] = x_k[\omega, \mathrm{t}] + v_k[\omega, \mathrm{t}] \quad (1)$$

where x_k is the desired speech component of the received signal, v_k is the noise component, ω is the frequency bin, and t is the frame index. We will omit the ω and t variables, unless otherwise stated, bearing in mind that the following operations occur in the STFT domain.

All microphone signals are sent, un-processed, to a fusion center. The fusion center collects the received signals and places them in a stacked vector which takes the form

$$\mathbf{y} = [y_1 \ ... \ y_M]. \tag{2}$$

The speech vector $\mathbf{x}$ and noise vector $\mathbf{v}$ are constructed in a similar fashion.

If a single speech source is assumed the vector containing the speech component of each microphone signal is

$$\mathbf{x} = \mathbf{a}s \tag{3}$$

where s is the speech source signal and $\mathbf{a}$ is a steering vector that contains information pertaining to the room characteristics from the speech source to the microphones. The goal of the MWF is to minimize the MMSE between the desired speech signal and a linearly filtered version of the combined microphone signals. The linear MMSE cost function at the fusion center is

$$J(\mathbf{w}) = E\{|x_1 - \mathbf{w}^H\mathbf{y}|^2\} \tag{4}$$

where x_1 is the desired speech component of the reference microphone, $\mathbf{w}^H\mathbf{y}$ is the linearly filtered sensor signals and H denotes the conjugate transpose. For the ease of exposition and without loss of generality (w.l.o.g.) the first microphone signal x_1 is used as the reference microphone signal.

It is assumed that the source and the noise signals are statistically independent from one another so that the cost function may be written as

$$J(\mathbf{w}) = E\{|x_1 - \mathbf{w}^\mathbf{H}\mathbf{x}|^2\} + \mu E\{|\mathbf{w}^\mathbf{H}\mathbf{v}|^2\} \tag{5}$$

where a trade-off parameter $\mu > 0$ is added to place emphasis on either the speech distortion or noise reduction [6]. For the case where $\mu = 1$ (4) and (5) are equivalent. The optimal filter minimizing the cost function (5) is the SDW-MWF.

It has been shown in [7] that if only a single speech source is present the SDW-MWF is given by

$$\hat{\mathbf{w}} = \frac{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}\mathbf{e}_1}{\mu + \text{Tr}\{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}\}} \tag{6}$$

where Tr$\{\mathbf{A}\}$ is the trace of the matrix $\mathbf{A}$, $\mathbf{e}_1$ is a vector containing a one in the first entry (corresponding to the reference microphone) and zero otherwise, $\mathbf{R}_{\mathbf{vv}}^{-1}$ is the inverse of the noise correlation matrix $\mathbf{R}_{\mathbf{vv}} = E\{\mathbf{v}\mathbf{v}^H\}$ and $\mathbf{R}_{\mathbf{xx}} = E\{\mathbf{x}\mathbf{x}^H\}$ is the speech correlation matrix. This is referred to as the Rank-1 SDW-MWF (R1-MWF).

The so-called noise+speech correlation matrix $\mathbf{R}_{\mathbf{yy}} = E\{\mathbf{y}\mathbf{y}^H\}$ is often updated at discrete time intervals by means of a forgetting factor $0 < \lambda < 1$

$$\mathbf{R}_{\mathbf{yy}}[\omega, \text{t}] = \lambda\mathbf{R}_{\mathbf{yy}}[\omega, \text{t}-1] + (1-\lambda)\mathbf{y}[\omega, \text{t}]\mathbf{y}[\omega, \text{t}]^H \tag{7}$$

with the noise correlation matrix being updated in a similar fashion where it is assumed a voice activity detector (VAD) is able to distinguish between the noise+speech and noise only frames. This type of estimation allows for the combination of the current signal with older time-averaged statistics.

If the speech and noise signals are assumed to be statistically independent, the speech correlation matrix is estimated by subtracting the noise+speech correlation matrix by the noise correlation matrix [6]

$$\mathbf{R}_{\mathbf{xx}} = \mathbf{R}_{\mathbf{yy}} - \mathbf{R}_{\mathbf{vv}}. \tag{8}$$

Since it is assumed that there is only a single speech source present $\mathbf{R}_{\mathbf{xx}}$ may be represented as

$$\mathbf{R}_{\mathbf{xx}} = P_s\mathbf{a}\mathbf{a}^H \tag{9}$$

where $P_s = E\{|s|^2\}$ is the power of the speech signal, $P_{x_1} = P_s|a_1|^2$ is the speech power in the reference microphone and a_1 is the first element of the steering vector.

Using the optimal filter value (6) the cost function takes the form

$$J(\hat{\mathbf{w}}) = P_{x_1} - \frac{\mathbf{e}_1^\mathbf{T}\mathbf{R}_{\mathbf{xx}}\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}\mathbf{e}_1}{\mu + \text{Tr}\{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}\}} \tag{10}$$

and using the fact that $\mathbf{R}_{\mathbf{xx}}$ is rank 1, the numerator in (10) can be reduced to $P_{x_1}(\text{Tr}\{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}\})$. This reduces the cost function to

$$J(\hat{\mathbf{w}}) = \frac{\mu P_{x_1}}{\mu + \text{Tr}\{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}\}}. \tag{11}$$

3. PERFORMANCE MEASURES

3.1. Utility

The signals in a WASN can be efficiently monitored to determine their *utility* or impact on the current cost function. The utility function U_k for monitoring one signal for deletion, as introduced in [4], is defined as the increase in the cost function by the removal of signal k,

$$U_k = J_{-k}(\hat{\mathbf{w}}_{-k}) - J(\hat{\mathbf{w}}) \tag{12}$$

where $\hat{\mathbf{w}}_{-k}$ is the optimal filter value missing the kth microphone signal.

By using the cost function given in (11) the utility for a given signal k is

$$U_k = \mu P_{x_1}\left[\frac{1}{\mu + \text{Tr}\{\mathbf{D}_{-k}\}} - \frac{1}{\mu + \text{Tr}\{\mathbf{D}\}}\right] \tag{13}$$

where $\mathbf{D} = \mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{xx}}$ and $\mathbf{D}_{-k} = \mathbf{R}_{\mathbf{vv}-k}^{-1}\mathbf{R}_{\mathbf{xx}-k}$. Note that the value of $\mathbf{D}_{-k}$ is not computed by simply removing the corresponding row and column from $\mathbf{D}$. The row and column must first be removed from the $\mathbf{R}_{\mathbf{xx}}$ and $\mathbf{R}_{\mathbf{vv}}$ matrices to give $\mathbf{R}_{\mathbf{xx}-k}$ and $\mathbf{R}_{\mathbf{vv}-k}$ and then an inverse needs to be performed on $\mathbf{R}_{\mathbf{vv}-k}$.

3.2. Signal-to-Noise Ratio based assessment

The output SNR at the fusion center is given by the ratio of power of the speech and noise components in the output signal

$$\begin{aligned} \text{SNR} &= \frac{E\{|\hat{\mathbf{w}}^H\mathbf{x}|^2\}}{E\{|\hat{\mathbf{w}}^H\mathbf{v}|^2\}} \\ &= \frac{\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{xx}}\hat{\mathbf{w}}}{\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{vv}}\hat{\mathbf{w}}}. \end{aligned} \tag{14}$$

It has been shown in [7] that (14) is equal to the Tr$\{\mathbf{D}\}$ using the rank-1 assumption. The decrease in the SNR from the removal of the kth microphone signal from the estimation can again be found by the difference in the trace value,

$$\begin{aligned}\Delta\text{SNR}_{-k} &\triangleq \text{SNR}_{-k} - \text{SNR} \\ &= \text{Tr}\{\mathbf{D}_{-k}\} - \text{Tr}\{\mathbf{D}\}\end{aligned} \tag{15}$$

which is independent of the speech distortion parameter μ and is already known from the calculation of the R1-MWF. The reader should note that the lack of dependence on μ only holds for the given single frequency bin solution as the full band solution takes the form

$$\text{SNR} = \frac{\sum_{\omega} E\{|\hat{\mathbf{w}}^H\mathbf{x}|^2\}}{\sum_{\omega} E\{|\hat{\mathbf{w}}^H\mathbf{v}|^2\}}. \tag{16}$$

3.3. Signal-to-Distortion Ratio based assessment

The SDR is another important metric of a speech enhancement algorithm as it allows for the amount of speech-distortion to be measured. It was shown in [6] that the speech-distortion and SNR are closely related with one another. The SDR is given by

$$\text{SDR} = \frac{E\{x_1^2\}}{E\{|x_1 - \hat{\mathbf{w}}^H\mathbf{x}|^2\}} \tag{17}$$

and again using the rank-1 assumption, the SDR can be given as

$$\text{SDR} = \frac{(\mu + \text{Tr}\{\mathbf{D}\})^2}{\mu^2} \tag{18}$$

which is the inverse of the signal-to-distortion index described in [7].

Equation (18) also shows that there is a direct relationship between the SNR and SDR, i.e., an increase or decrease in SNR will have a similar effect on the SDR. Using (18) the decrease in the SDR due to the removal of a signal is then given by

$$\begin{aligned}\Delta\text{SDR}_{-k} &\triangleq \text{SDR}_{-k} - \text{SDR} \\ &= (\mu + \text{Tr}\{\mathbf{D}_{-k}\})^2 - (\mu + \text{Tr}\{\mathbf{D}\})^2\end{aligned} \tag{19}$$

which again relies on the calculation of the trace value when a signal is removed.

4. EFFICIENT COMPUTATION OF THE TRACE WHEN REMOVING A SIGNAL

We first describe an efficient manner in which to derive the trace value when a signal k is removed and then generalize this so that all signals can be monitored simultaneously. Before deleting the kth signal, the current value $\text{Tr}\{\mathbf{D}\}$ is known and therefore an efficient way to calculate $\text{Tr}\{\mathbf{D}_{-k}\}$ without taking a full matrix inverse of $\mathbf{R}_{\mathbf{vv}-k}$, which has a computationally complexity of $O(M-1)^3$, is desired.

For the ease of exposition we assume that the signal to be removed is the last element, i.e., $k = M$. This leads to the block partitioning of the inverse noise correlation matrix as

$$\mathbf{R}_{\mathbf{vv}}^{-1} = \left[\begin{array}{c|c} \mathbf{A}_k & \mathbf{b}_k \\ \hline \mathbf{b}_k^H & Q_k \end{array}\right] \tag{20}$$

the block partitioning of the speech correlation matrix as

$$\mathbf{R}_{\mathbf{xx}} = \left[\begin{array}{c|c} \mathbf{R}_{\mathbf{xx}-k} & \mathbf{d}_k \\ \hline \mathbf{d}_k^H & V_k \end{array}\right] \tag{21}$$

and the block partitioning of the steering vector as

$$\mathbf{a} = \left[\begin{array}{c} \mathbf{a}_{-k} \\ \hline a_k \end{array}\right]. \tag{22}$$

Based on (22) the vector quantity $\mathbf{d}_k$ is defined as

$$\mathbf{d}_k = P_s|\mathbf{a}_{-k}a_k^*| \tag{23}$$

where * represents the complex conjugate and the scalar quantity V_k is defined as

$$V_k = P_s|a_k|^2. \tag{24}$$

We define a diagonal matrix containing the current diagonal elements of $\mathbf{D}$ as

$$\mathbf{\Lambda}_D = \mathbf{I}_M \circ \mathbf{D} \tag{25}$$

where $A \circ B$ is the Hadamard or element-wise product of two matrices and $\mathbf{I}_M$ is the identity matrix. The diagonal elements for the correlation matrices can be constructed in a similar fashion as $\mathbf{\Lambda}_V = \mathbf{I}_M \circ \mathbf{R}_{\mathbf{vv}}^{-1}$ and $\mathbf{\Lambda}_X = \mathbf{I}_M \circ \mathbf{R}_{\mathbf{xx}}$ and the product of the two diagonal matrices $\mathbf{\Lambda}_V$ and $\mathbf{\Lambda}_X$ is given as $\mathbf{\Lambda}_{VX}$. Using the matrices defined in (20) and (21) the current trace is given by

$$\text{Tr}\{\mathbf{D}\} = \text{Tr}\{\mathbf{A}_k\mathbf{R}_{\mathbf{xx}-k}\} + 2\mathcal{R}\{\mathbf{b}_k^H\mathbf{d}_\mathbf{k}\} + Q_kV_k \tag{26}$$

where $\mathcal{R}\{.\}$ extracts the real component of its argument. It was shown in [4] that the inverse correlation matrix with the deletion of row and column k can be found by

$$\mathbf{R}_{\mathbf{vv}-k}^{-1} = \mathbf{A}_k - \frac{1}{Q_k}\mathbf{b}_k\mathbf{b}_k^H. \tag{27}$$

The trace with the removal of signal k can therefore be calculated as

$$\text{Tr}\{\mathbf{D}_{-k}\} = \text{Tr}\{\mathbf{A}_k\mathbf{R}_{\mathbf{xx}-k}\} - \frac{1}{Q_k}\text{Tr}\{\mathbf{b}_k\mathbf{b}_k^H\mathbf{R}_{\mathbf{xx}-k}\}. \tag{28}$$

Using (9) along with (22), (23), and (24) produces

$$\text{Tr}\{\mathbf{b}_k\mathbf{b}_k^H\mathbf{R}_{\mathbf{xx}-k}\} = \frac{|\mathbf{b}_k^H\mathbf{d}_k|^2}{V_k} \tag{29}$$

which leads to an alternative representation of (28) given by

$$\text{Tr}\{\mathbf{D}_{-k}\} = \text{Tr}\{\mathbf{A}_k\mathbf{R}_{\mathbf{xx}-k}\} - \frac{1}{Q_kV_k}|\mathbf{b}_k^H\mathbf{d}_\mathbf{k}|^2. \tag{30}$$

The vector product $\mathbf{b}_k^H\mathbf{d}_\mathbf{k}$ in (30) may be represented as the kth diagonal element of $\mathbf{\Lambda}_D$ subtracted by the product of the kth diagonal elements of $\mathbf{R}_{\mathbf{vv}}^{-1}$ and $\mathbf{R}_{\mathbf{xx}}$, i.e., $\mathbf{\Lambda}_D(k) - Q_kV_k$. Using this fact and rearranging (26), the trace with element k removed becomes

$$\text{Tr}\{\mathbf{A}_k\mathbf{R}_{\mathbf{xx}-k}\} = \text{Tr}\{\mathbf{D}\} - 2\mathcal{R}\{\mathbf{\Lambda}_D(k)\} + Q_kV_k. \tag{31}$$

Finally plugging (31) into (28) gives the trace with the signal k removed as

$$\text{Tr}\{\mathbf{D}_{-k}\}=\text{Tr}\{\mathbf{D}\}-2\mathcal{R}\{\mathbf{\Lambda}_D(k)\}+Q_kV_k-\frac{1}{Q_kV_k}|\mathbf{\Lambda}_D(k)-Q_kV_k|^2. \tag{32}$$

Suppose now we wish to monitor all M signals in the WASN. This would entail taking an inverse at again an $O((M-1)^3)$ computationally complexity for all M signals yielding an $O(M^4)$ operation. Using the notation above, the trace with each element missing can be given in vector form where $\mathbf{v} = [\text{Tr}\{\mathbf{D}_{-1}\}\ldots\text{Tr}\{\mathbf{D}_{-M}\}]^T$ is

$$\mathbf{v} = \text{Tr}\{\mathbf{D}\}\mathbb{1} - (2\mathcal{R}\{\mathbf{\Lambda}_D\} - \mathbf{\Lambda}_{VX} + \mathbf{\Lambda}_{VX}^{-1}|\mathbf{\Lambda}_D - \mathbf{\Lambda}_{VX}|^2)\mathbb{1} \tag{33}$$

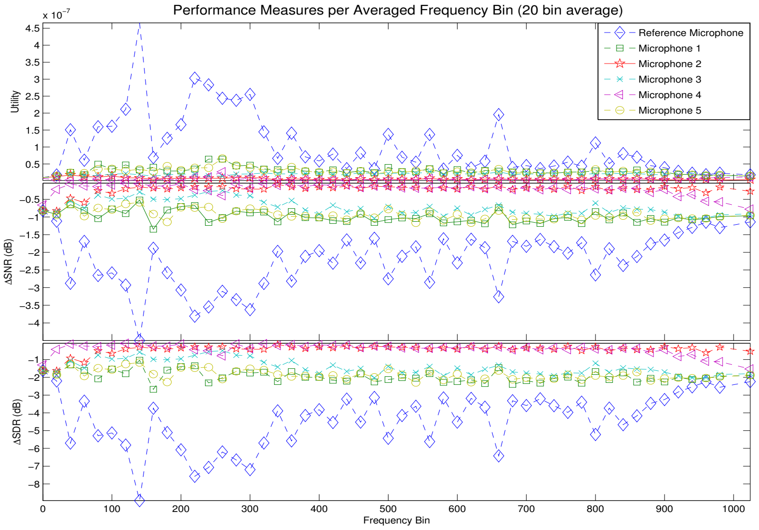

Fig. 2. Subband Utility, SNR, and SDR

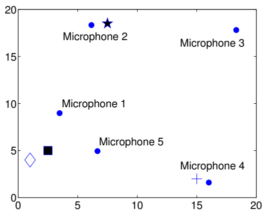

Fig. 1. Simulated Room Environment

and $\mathbb{1}$ is a vector with all entries equal to one. Expression (33) may then be further reduced to

$$\mathbf{v} = \mathrm{Tr}\{\mathbf{D}\}\mathbb{1} - \mathbf{\Lambda}_{VX}^{-1}|\mathbf{\Lambda}_D|^2\mathbb{1} \tag{34}$$

which has terms that are only composed of diagonal matrices making it an $O(M)$ operation. The utility, SNR and SDR can now be calculated simultaneously with the values from (34).

5. SIMULATIONS

Figure 1 depicts a simulated room environment (20x20x5m) where there is a single speech source ■, a babble noise source $+$, a white noise source ★, a reference microphone $\diamond$, and 5 other microphones •. There is also white additive noise on each microphone equal to 10% of the speech source power representative of thermal noise. The microphones, speech, and noise sources are positioned at a height of 1.5 m from the ground. A reflection coefficient of 0.4 was used for the room and a sampling frequency of 8 kHz was used for the signals.

A weighted overlap-add technique, as introduced in [8], was used with a DFT block size of 2048. The utility, SNR and SDR values were averaged over the entire collection time so that the individual microphones could be analyzed in regards to the performance measures. In real-time applications, an updating similar to the one used to update the correlation matrices in (7) could be used enabling the performance measures to be analyzed in varying environments.

Figure 2 shows the corresponding utility, SNR and SDR. The performance measures mimic one another due to the dependence on the trace elements of the current estimation. The performance measures are highly effected by the input SNR, where the reference microphone has the largest impact due to having the largest input SNR. Microphones with low input SNRs do not significantly contribute to the output SNR and SDR which indicate that these signals could be removed without severely impacting the noise reduction or signal distortion.

6. CONCLUSION

The utility function derived shows the signal components that contribute the most to the noise reduction. By using unique properties of the R1-MWF formulation other information such as the output SNR and SDR were extracted efficiently from the utility calculation compared to previous utility formulations where only the difference in the cost was observed. This allows for the direct impact of the removal of signal components to be viewed in terms that can be custom tailored to the specific application of the WASN.

7. REFERENCES

[1] Liljana Gavrilovska and Ramjee Prasad, *Ad-Hoc Networking Towards Seamless Communications (Signals and Communication Technology)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[2] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, 2001, vol. 4, pp. 2033 –2036 vol.4.

[3] J. Szurley, A. Bertrand, M. Moonen, P. Ruckebusch, and I. Moerman, "Utility based cross-layer collaboration for speech enhancement in wireless acoustic sensor networks," in *Proc. of the European signal processing conference (EUSIPCO)*, Barcelona - Spain, August 2011.

[4] A. Bertrand and M. Moonen, "Efficient sensor subset selection and link failure response for linear MMSE signal estimation in wireless sensor networks," in *Proc. of the European signal processing conference (EUSIPCO)*, Aalborg - Denmark, August 2010, pp. 1092–1096.

[5] B. Cornelis, M. Moonen, and J. Wouters, "Performance analysis of multichannel wiener filter-based noise reduction in hearing aids under second order statistics estimation errors," july 2011, vol. 19, pp. 1368 –1381.

[6] Benesty J., Makino S., and Chen J., *Speech Enhancement*, Springer-Verlag, 2005.

[7] M. Souden, J. Benesty, and S. Affes, "On optimal frequency-domain multichannel linear filtering for noise reduction," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 2, pp. 260 –276, feb. 2010.

[8] A. Bertrand, J. Callebaut, and M. Moonen, "Adaptive distributed noise reduction for speech enhancement in wireless acoustic sensor networks," in *Proc. of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Tel Aviv, Israel, August 2010.