Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 10-121

A Combined Multi-channel Wiener Filter based Noise Reduction and Dynamic Range Compression in Hearing Aids[1]

Kim Ngo[2], Ann Spriet[2,3], Marc Moonen[2],
Jan Wouters [3], Søren Holdt Jensen [4]

August 2011
Accepted for publication in Elsevier signal processing

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.be* in the directory *pub/sista/kngo/reports/10-121.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SCD (SISTA) Kasteelpark Arenberg 10, 3001 Leuven, Belgium, Tel. +32 16 321797, Fax +32 16 321970, WWW: *http://homes.esat.kuleuven.be/~kngo*. E-mail: *kim.ngo@esat.kuleuven.be*. 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), Concerted Research Action GOA-MaNet, the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011), Research Project FWO nr. G.0600.08 ('Signal processing and network design for wireless acoustic sensor networks'), IWT Project 'Signal processing and automatic fitting for next generation cochlear implants', EC-FP6 project SIGNAL: 'Core Signal Processing Training Program'. The scientific responsibility is assumed by its authors.

[3]Katholieke Universiteit Leuven, Department of Neurosciences, ExpORL, O. & N2, Herestraat 49/721, 3000 Leuven, Belgium, E-mail: Jan.Wouters@med.kuleuven.be

[4]Aalborg University, Department of Electronic Systems, MISP, Niels Jernes Vej 12 A6-3, 9220 Aalborg, Denmark, E-mail: shj@es.aau.dk

Abstract

Noise reduction (NR) and dynamic range compression (DRC) are basic components in hearing aids, but generally these components are developed and evaluated independently of each other. Hearing aids typically use a serial concatenation of NR and DRC. However, the DRC in such a concatenation negatively affects the performance of the NR stage: the residual noise after NR receives more amplification compared to the speech, resulting in a signal-to-noise-ratio (SNR) degradation. The integration of NR and DRC has not received a lot of attention so far. In this paper, a multi-channel Wiener filter (MWF) based approach is presented for speech and noise scenarios, where an MWF based NR algorithm is combined with DRC. The proposed solution is based on modifying the MWF and the DRC to incorporate the conditional speech presence probability in order to avoid residual noise amplification. The goal is then to analyse any undesired interaction effects by means of objective measures. Experimental results indeed confirm that a serial concatenation of NR and DRC degrades the SNR improvement provided by the NR, whereas the combined approach proposed here shows less degradation of the SNR improvement at a low increase in distortion compared to a serial concatenation.

A Combined Multi-channel Wiener Filter based Noise Reduction and Dynamic Range Compression in Hearing Aids

Kim Ngo[a,1], Ann Spriet[a], Marc Moonen[a], Jan Wouters[b], Søren Holdt Jensen[c]

[a] *Department of Electrical Engineering, Katholieke Universiteit Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium*
[b] *Division of Experimental Otorhinolaryngology, Katholieke Universiteit Leuven, ExpORL, O.& N2, Herestraat 49/721, B-3000 Leuven, Belgium*
[c] *Department of Electronic Systems, Aalborg University, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark*

Abstract

Noise reduction (NR) and dynamic range compression (DRC) are basic components in hearing aids, but generally these components are developed and evaluated independently of each other. Hearing aids typically use a serial concatenation of NR and DRC. However, the DRC in such a concatenation negatively affects the performance of the NR stage: the residual noise after NR receives more amplification compared to the speech, resulting in a signal-to-noise-ratio (SNR) degradation. The integration of NR and DRC has not received a lot of attention so far. In this paper, a multi-channel Wiener filter (MWF) based approach is presented for speech and noise scenarios, where an MWF based NR algorithm is combined with DRC. The proposed solution is based on modifying the MWF and the DRC to incorporate the conditional speech presence probability in order to avoid residual noise amplification. The goal is then to analyse any undesired interaction effects by means of objective measures. Experimental results indeed confirm that a serial concatenation of NR and DRC degrades the SNR improvement provided by the NR, whereas the combined approach proposed here shows less degradation of the SNR improvement at a low increase in

[1] Corresponding author. Tel.: +32 16 32 1797; fax: +32 16 321970. E-mail address: kim.ngo@esat.kuleuven.be (K. Ngo).

distortion compared to a serial concatenation.

Keywords: Hearing aids, noise reduction, dynamic range compression, multi-channel Wiener filter, speech presence probability.

---

1. Introduction

Reduced audibility and reduced dynamic range between threshold and discomfort level are some of the problems that hearing impaired people with a sensorineural hearing loss are dealing with. Furthermore, in scenarios where the target signal is a speech signal background noise (from competing speakers, traffic etc.) is a significant problem in that hearing impaired people indeed have more difficulty understanding speech in noise and so in general need a higher signal-to-noise-ratio (SNR) than people with normal hearing [1][2]. Therefore, noise reduction (NR) and dynamic range compression (DRC) are basic components in hearing aids nowadays [3], but generally these components are developed and evaluated independently of each other. The design and benefits of single-channel and multi-channel NR algorithms have been widely studied [4][5][6][7][8][9]. The same goes for the design and evaluation of different DRC algorithms [10][11][12][13][14]. Although sophisticated algorithms for NR and DRC exist there is still a question as to how these algorithms should be combined, which unfortunately, has not received a lot of attention so far. Combining hearing aid algorithms in general is indeed a challenging task since each algorithm can counteract and limit the functionality of other algorithms.

In this paper, the first aim is to analyse undesired interaction effects when the NR and DRC algorithms operate together. When NR and DRC are serially concatenated, undesired interaction effects typically occur, since each algorithm serves a different purpose. For instance, DRC can counteract NR by applying more amplification to the residual noise compared to the speech, which consequently degrades the SNR and defeats the purpose of using NR. In [15] [16] experiments have been conducted to evaluate different combinations of NR and DRC. One of the main conclusions was that a serial concatenation of NR and

DRC performs suboptimally due to the interaction effects between the NR and the DRC. In [17] it was shown that the NR algorithm does enhance the modulation depth of a noisy speech but when the DRC is activated the modulation depth of the speech envelope is greatly reduced. This indicates that the noise level is increased compared to the speech level, which is clearly undesirable. A combination of a single-channel NR and DRC was proposed in [18] where a minimum mean squared error and a maximum a posteriori optimal estimator are developed that incorporate DRC in the derivation of the NR algorithm. In [19] a SNR degradation was observed when NR and DRC are serially concatenated. Therefore, a dual-DRC concept was proposed to integrate NR and DRC. The basic idea behind this approach, is to identify speech dominant segments versus noise dominant segments such that less amplification is applied to the residual noise compared to the amplification in the speech dominant segments.

An important issue is the evaluation of such combined and integrated schemes, where the lack of an overall design criterion indeed makes the evaluation more difficult. In the evaluation the crucial question will be as to which effects are most damaging to speech intelligibility, e.g., the amount of background noise, distortion or the audibility. In this paper, a multi-channel Wiener filter (MWF) based approach is presented where an MWF based noise reduction algorithm is combined with DRC and compared to a serial concatenation. The work in [19] is based on a GSC based NR and the speech dominant segments and the noise dominant segments are estimated in a rather ad-hoc manner based on the power ratio between the output and the input of the NR algorithm. The solution here is based on using a modified MWF that incorporates the conditional speech presence probability (SPP) in the NR process [20][25][26] as well as a modified DRC that is also using this conditional SPP. Furthermore, the MWF based NR and DRC framework offers a way to analyze the undesired interaction effect by estimating the speech component and the noise component, which allows us to apply a speech DRC and a noise DRC. The combined approach is evaluated and compared to a serial concatenation based on objective measures such as the intelligibility-weighted signal-to-noise ratio and a frequency-weighted log-

spectral signal distortion measure. Experimental results indeed confirm that a serial concatenation of NR and DRC degrades the SNR improvement, whereas the combined approach proposed here shows less degradation of the SNR improvement at a low increase in distortion compared to a serial concatenation.

The paper is organized as follows. The problem statement and motivation are given in Section 2. In Section 3 the MWF based NR is described. In Section 4 the DRC algorithm is reviewed. The combined MWF based NR and DRC is given in Section 5. In Section 6 experimental results are presented. The work is summarized in Section 7.

2. Problem statement and motivation

When combining NR and DRC the main problem is that each algorithm serves a different purpose. The objective of the NR algorithm in speech and noise scenarios is to maximally reduce the noise while minimizing speech distortion, e.g., based on temporal, spectral and spatial signal information. The DRC on the other hand is designed to amplify sounds based on their intensity level and a compression characteristic. Fig. 1(a)-(b) shows the two ways to serially concatenate NR and DRC. The main issues can be stated as follows:

- When NR is performed before DRC, as in Fig. 1(a), the residual noise receives more amplification compared to the speech, which consequently defeats the purpose of using NR. From a DRC point of view there is no distinction between speech dominant segments and noise dominant segments, so all low intensity segments are amplified equally. This means that the reduced noise level, from the DRC point of view, is considered a low level signal which is then amplified while the speech is considered a high level signal, receives less amplification. This leads to the undesired SNR degradation.

- When DRC is performed before NR, as in Fig. 1(b), the DRC can negatively affect the NR especially so in a multi-channel NR where the correlation between the microphone signals can be affected by the independent

DRC on the microphone signals. Furthermore in this set-up the DRC is based on the speech+noise level rather than the speech+residual noise level and so the applied gain in this case may be too small to make the soft speech segments audible.

To avoid any undesired interaction effects it is desirable to combine NR and DRC in an integrated scheme, as in Fig. 1(c), which is the goal of this paper. In the sequel, the serial concatenation shown in Fig. 1(a) will serve as a reference system and the proposed solution will be referred to as the combined approach.

A combined NR and DRC system that could be viewed as the ideal system is shown in Fig. 2. The idea here is that if the clean speech and the noise-only contribution can be perfectly extracted then a speech DRC can be applied to the clean speech and a noise DRC to the noise-only contribution. The gain difference between the speech DRC and the noise DRC indicates a target noise suppression which means that the noise DRC gain can be set to zero, i.e., to suppress all noise, or it can be a scaled version of the speech DRC gain, i.e., $G^n_{\mathrm{dB}} < G^s_{\mathrm{dB}}$. The gain difference between the speech DRC curve and the noise DRC curve is defined as

$$\Delta G_{\mathrm{dB}} = G^s_{\mathrm{dB}} - G^n_{\mathrm{dB}}. \tag{1}$$

Finally, the overall output signal is the sum of the two compressed components. Since the ideal case does not contain residual noise then the SNR will improve when the noise DRC gain G^n_{dB} decreases compared to the speech DRC gain G^s_{dB}. The goal is then to compare the performance of the combined approach against this ideal performance, and any deviation from this will be considered as an undesired effect of having a NR and DRC combined.

3. Multi-channel Wiener filter based noise reduction

3.1. Preliminaries

Let $X_i(k,l), i = 1, ..., M$ denote the frequency-domain microphone signals

$$X_i(k,l) = X^s_i(k,l) + X^n_i(k,l) \tag{2}$$

where k is the frequency bin index, and l the frame index of a short-time Fourier transform (STFT), and the superscripts s and n are used to refer to the speech and the noise contribution in a signal, respectively. Let $\mathbf{X}(k,l) \in \mathbb{C}^{M\times 1}$ be defined as the stacked vector

$$\mathbf{X}(k,l) = [X_1(k,l)\ X_2(k,l)\ ...\ X_M(k,l)]^T \tag{3}$$

$$= \mathbf{X}^s(k,l) + \mathbf{X}^n(k,l) \tag{4}$$

where the superscript T denotes the transpose. In addition, we define the speech+noise, the speech and the noise correlation matrices as

$$\mathbf{R}^x(k,l) = \varepsilon\{\mathbf{X}(k,l)\mathbf{X}^H(k,l)\} \tag{5}$$

$$\mathbf{R}^s(k,l) = \varepsilon\{\mathbf{X}^s(k,l)\mathbf{X}^{s,H}(k,l)\} \tag{6}$$

$$\mathbf{R}^n(k,l) = \varepsilon\{\mathbf{X}^n(k,l)\mathbf{X}^{n,H}(k,l)\} \tag{7}$$

where $\varepsilon\{\}$ denotes the expectation operator, H denotes Hermitian transpose.

A two-state model for speech events can be expressed given two hypotheses $H_0(k,l)$ and $H_1(k,l)$ which represent speech absence and speech presence in frequency bin k and frame l, respectively, i.e.,

$$\begin{aligned} H_0(k,l) &: X_i(k,l) = X_i^n(k,l) \\ H_1(k,l) &: X_i(k,l) = X_i^n(k,l) + X_i^s(k,l), \end{aligned} \tag{8}$$

where the i-th microphone signal is used as a reference (in our case the first microphone signal $X_1(k,l)$ is used). Finally the conditional speech presence probability (SPP) $p(k,l) \triangleq P(H_1(k,l)|X_i(k,l))$ can be written as [27]

$$p(k,l) = \left\{1 + \frac{q(k,l)}{1-q(k,l)}(1+\xi(k,l))\exp(-\upsilon(k,l))\right\}^{-1} \tag{9}$$

where

$$\upsilon(k,l) \triangleq \frac{\gamma(k,l)\xi(k,l)}{(1+\xi(k,l))}. \tag{10}$$

and $q(k,l) \triangleq P(H_0(k,l))$ is the a priori speech absence probability (SAP), $\xi(k,l)$ and $\gamma(k,l)$ denote the a priori SNR and the a posteriori SNR, respectively.

Details on the estimation of the SAP, the a priori SNR and the a posteriori SNR can be found in [20][27]. Furthermore we introduce a detection of the H_0 and the H_1 state, which is a binary decision, obtained by averaging the conditional SPP $p(k,l)$ over all frequency bins k

$$P(l) = \begin{cases} 1 & \text{if} \quad \frac{1}{K}\sum_{k=1}^{K} p(k,l) \geq \alpha_{\text{frame}} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where $P(l) = 1$ means the H_1 state is detected and $P(l) = 0$ means the H_0 state is detected, and α_{frame} is a detection threshold.

3.2. Speech distortion weighted multi-channel Wiener filter (SDW-MWF$_\mu$)

The MWF optimally estimates the speech signal, based on a minimum mean squared error (MMSE) criterion, i.e.,

$$\mathbf{W}_{\text{MMSE}}(k,l) = \arg\min_{\mathbf{W}} \ \varepsilon\{|X_1^s(k,l) - \mathbf{W}^H\mathbf{X}(k,l)|^2\} \tag{12}$$

where the desired signal in this case is the (unknown) speech component $X_1^s(k,l)$ in the first microphone signal. The MWF has been extended to the SDW-MWF$_\mu$, which allows for a trade-off between noise reduction and speech distortion using a weighting factor μ [8][9]. Assuming that the speech and the noise signals are statistically independent the SDW-MWF$_\mu$ is defined by

$$\mathbf{W}_\mu(k,l) = \arg\min_{\mathbf{W}} \ \varepsilon\{|X_1^s(k,l) - \mathbf{W}^H\mathbf{X}^s(k,l)|^2\} + \mu\varepsilon\{|\mathbf{W}^H\mathbf{X}^n(k,l)|^2\}. \tag{13}$$

The SDW-MWF$_\mu$ is then given by

$$\mathbf{W}_\mu(k,l) = \left[\mathbf{R}^s(k,l) + \mu\mathbf{R}^n(k,l)\right]^{-1}\mathbf{R}^s(k,l)\mathbf{e}_1 \tag{14}$$

where the $M \times 1$ vector $\mathbf{e}_1$ equals the first canonical vector defined as $\mathbf{e}_1 = [1 \quad 0 \quad ... \quad 0]^T$. The second-order statistics of the noise are assumed to be (short-term) stationary which means that $\mathbf{R}^s(k,l)$ is estimated as $\mathbf{R}^s(k,l) = \mathbf{R}^x(k,l) - \mathbf{R}^n(k,\tilde{l})$ where $\mathbf{R}^x(k,l)$ and $\mathbf{R}^n(k,\tilde{l})$ are estimated (i.e. adapted) during periods of speech+noise (l) and periods of noise-only ($\tilde{l}$), respectively

(and "frozen" otherwise). For $\mu = 1$ the SDW-MWF$_\mu$ reduces to the MWF, while for $\mu > 1$ the residual noise level will be reduced at the cost of a higher speech distortion. The output $Z(k,l)$ of the SDW-MWF$_\mu$ can then be written as

$$Z(k,l) = \mathbf{W}_\mu^H(k,l)\mathbf{X}(k,l). \tag{15}$$

In a similiar manner the noise component $X_1^n(k,l)$ in the first microphone signal can be estimated with a SDW-MWF$_\mu$ given as

$$\begin{aligned}\mathbf{V}_\mu(k,l) &= \left(\mathbf{R}^s(k,l) + \mu\mathbf{R}^n(k,l)\right)^{-1}\mu\mathbf{R}^n(k,l)\mathbf{e}_1 \\ &= \mathbf{e}_1 - \mathbf{W}_\mu(k,l)\end{aligned} \tag{16}$$

which leads to the estimated noise component

$$\begin{aligned}\bar{Z}(k,l) &= \mathbf{V}_\mu^H(k,l)\mathbf{X}(k,l) \\ &= X_1(k,l) - Z(k,l).\end{aligned} \tag{17}$$

In addition to SDW-MWF$_\mu$ two other algorithms, namely SDW-MWF$_{\text{SPP}}$ and SDW-MWF$_{\text{Flex}}$, have been developed which will be shown to be valuable in the combined approach of NR and DRC. These algorithms will be reviewed in section 3.3 and section 3.4; further details can be found in [20][26].

3.3. SDW-MWF incorporating the conditional SPP (SDW-MWF$_{SPP}$)

The SDW-MWF$_{\text{SPP}}$ incorporates the conditional SPP in the SDW-MWF$_\mu$ defined in (9) to allow for a faster tracking of the spectral non-stationarity of the speech, as well as for exploiting the fact that speech may not be present at all time. The optimization criterion of the SDW-MWF$_{\text{SPP}}$ [20] can be written as

$$\begin{aligned}\mathbf{W}_{\text{SPP}}(k,l) = \arg\min_{\mathbf{W}}\ & p(k,l)\varepsilon\{|X_1^s - \mathbf{W}^H\mathbf{X}|^2|H_1\} \\ & + (1-p(k,l))\varepsilon\{|\mathbf{W}^H\mathbf{X}|^2|H_0\}\end{aligned} \tag{18}$$

where the first term corresponds to H_1 and is weighted by the conditional probability $p(k,l)$ that speech is present, while the second term corresponds to H_0 and

is weighted by the probability that speech is absent $(1 - p(k,l))$. The solution is then given by

$$\mathbf{W}_{\text{SPP}}(k,l) = \left[\mathbf{R}^s(k,l) + \left(\tfrac{1}{p(k,l)}\right)\mathbf{R}^n(k,l)\right]^{-1}\mathbf{R}^s(k,l)\mathbf{e}_1. \tag{19}$$

Compared to (14) the fixed weighting factor μ is replaced by $\frac{1}{p(k,l)}$, which is now adjusted for each frequency bin k and for each frame l, making the SDW-MWF$_{\text{SPP}}$ change with a faster dynamic. The SDW-MWF$_{\text{SPP}}$ offers more noise reduction when $p(k,l)$ is small, i.e., for noise dominant segments, and less noise reduction when $p(k,l)$ is large, i.e., for speech dominant segments.

3.4. SDW-MWF incorporating a flexible weighting factor (SDW-MWF$_{Flex}$)

The SDW-MWF$_{\text{Flex}}$ incorporates a flexible weighting factor based on $p(k,l)$ and $P(l)$ defined in (9) and (11), respectively. When such a detection is available the noise reduction, in the H_0 state and the H_1 state, can be applied with different weights leading to a more flexible noise reduction strategy. The optimization criterion for SDW-MWF$_{\text{Flex}}$ [26] is given by

$$\begin{aligned}\mathbf{W}_{\text{Flex}}(k,l) = \arg\min_{\mathbf{W}}\ P(l)\Big[&\mu_{H_1}\varepsilon\{|X_1^s - \mathbf{W}^H\mathbf{X}|^2|H_1\} \\ &+ (1-\mu_{H_1})\varepsilon\{|\mathbf{W}^H\mathbf{X}|^2|H_0\}\Big] + (1-P(l)) \\ &\Big[\tfrac{1}{\mu_{H_0}}\varepsilon\{|X_1^s - \mathbf{W}^H\mathbf{X}^s|^2\} + \varepsilon\{|\mathbf{W}^H\mathbf{X}^n|^2\}\Big]\end{aligned} \tag{20}$$

where $\mu_{H_1} = \max(p(k,l), \frac{1}{\alpha_{H_1}})$ which is a function of $p(k,l)$ and a lower threshold α_{H_1}, that defines the amount of noise reduction that can be applied in the H_1 state, whereas the term $\frac{1}{\mu_{H_0}}$ determines the noise reduction that can be applied in the H_0 state. The solution is given by

$$\mathbf{W}_{\text{Flex}}(k,l) = \left[\mathbf{R}^s(k,l) + \gamma_{\text{Flex}}(k,l)\mathbf{R}^n(k,l)\right]^{-1}\mathbf{R}^s(k,l)\mathbf{e}_1 \tag{21}$$

with the weighting factor defined as

$$\begin{aligned}\gamma_{\text{Flex}}(k,l) &= \left[P(l)\max(p(k,l), \tfrac{1}{\alpha_{H_1}}) + (1-P(l))\tfrac{1}{\mu_{H_0}}\right]^{-1} \\ &= \Big[P(l)\min(\tfrac{1}{p(k,l)}, \alpha_{H_1}) + (1-P(l))\mu_{H_0}\Big].\end{aligned} \tag{22}$$

In section 5 it is explained how the SDW-MWF$_{\text{SPP}}$ and SDW-MWF$_{\text{Flex}}$ are combined with DRC by exploiting the conditional SPP ($p(k,l)$) and the H_0 and H_1 detection $P(l)$.

3.5. Complexity and implementation

Several works have been published dealing with complexity reduction and the implementation aspects of the MWF. In [23][24] recursive implementations of the SDW-MWF have been proposed based on a generalized singular value decomposition (GSVD) or a QR decomposition. A subband implementation [21][22] has also been proposed which significantly reduces the complexity compared to a fullband approach. In [9] a cheap frequency domain approach has been introduced to implement the SDW-MWF based on a stochastic gradient algorithm where it has been shown that the stochastic gradient algorithm preserves the performance compared to the exact SDW-MWF. The complexity of the MWF has been further reduced in [8] where different implementations have been proposed based on an RLS or an LMS type update formula. The idea here is to exploit different structures in the estimation of the step size matrix, i.e., constrained vs. unconstrained or block-structured vs. diagonal.

4. Dynamic Range Compression

In this section, the basics of DRC are briefly reviewed. Further details can be found in [10][11][12][13][14]. The role of the DRC is to estimate a desirable gain to map the wide dynamic range of an input audio (e.g. speech) signal into the reduced dynamic range of a hearing impaired listener. The gain is then automatically adjusted based on the intensity level of the input signal. Segments with a high intensity level are amplified less compared to segments with a low intensity level. This makes weak sounds audible while loud sounds are not becoming uncomfortably loud. The DRC is typically defined by the following parameters:

- Compression threshold (CT).

- Compression ratio (CR).
- Attack (at) and release time (rt).
- DRC gain G^s_{dB}.

The CT is defined in dB and corresponds to the point where the DRC becomes active, i.e., where the gain is reduced. The CR determines the degree of compression. A CR of 2 (i.e. 2:1) means that for every 2dB SPL increase in the input signal, the output signal increases by 1dB SPL. The attack and release time are defined in milliseconds and specify how fast the gain is changed according to changes in the input signal. The attack time is defined as the time taken for the compressor to react to an increase in input signal level. The release time is the time taken for the compressor to react to a decrease in input SPL and G^s_{dB} is defined as the speech DRC gain. For the DRC the input level for each critical band in dB SPL is defined as

$$P^{\mathrm{in,s}}_{\mathrm{DRC,dB}}(k',l) = 20\log_{10}\left(\frac{|P^{\mathrm{in}}_{\mathrm{DRC}}(k',l)|}{P_{\mathrm{ref}}}\right) \tag{23}$$

where k' is used to indicate that the linear frequency is now mapped to the Bark scale and P_{ref} is the reference sound pressure (20 micro Pascal). The DRC curve is defined based on a linear curve and a compression curve defined in (24) and (25), respectively:

$$P_{\mathrm{lin,dB}}(k',l) = P^{\mathrm{in,s}}_{\mathrm{DRC,dB}}(k',l) + G^s_{\mathrm{dB}} \tag{24}$$

$$P_{\mathrm{cp,dB}}(k',l) = \mathrm{CT} + \frac{1}{\mathrm{CR}} \cdot (P^{\mathrm{in,s}}_{\mathrm{DRC,dB}}(k',l) - \mathrm{CT}) + G^s_{\mathrm{dB}} \tag{25}$$

The output level in dB SPL is then given by

$$P^{\mathrm{out,s}}_{\mathrm{DRC,dB}}(k',l) = \begin{cases} P_{\mathrm{lin,dB}}(k',l) & \text{if } P^{\mathrm{in,s}}_{\mathrm{DRC,dB}}(k',l) < \mathrm{CT} \\ P_{\mathrm{cp,dB}}(k',l) & \text{if } P^{\mathrm{in,s}}_{\mathrm{DRC,dB}}(k',l) \geq \mathrm{CT} \end{cases} \tag{26}$$

A DRC curve that shows the output SPL as a function of the input SPL with CR=2, CT=30dB and G^s_{dB}=30dB is shown in Figure 3. Finally the DRC gain in dB is calculated as the output level minus the input level, i.e.,

$$G_{\mathrm{DRC,dB}}(k',l) = P^{\mathrm{out,s}}_{\mathrm{DRC,dB}}(k',l) - P^{\mathrm{in,s}}_{\mathrm{DRC,dB}}(k',l). \tag{27}$$

The attack and release time are then applied to the DRC gain $G_{\mathrm{DRC,dB}}(k', l)$ typically using a first-order recursive averaging filter, before the DRC gain is applied to the input signal.

5. Combined MWF based NR and DRC

This section presents three different approaches to combine a MWF based NR and DRC. A SDW-MWF$_\mu$ serially concatenated with a DRC is described first and is considered to be the baseline system. The SDW-MWF$_\mu$ is then replaced by the SDW-MWF$_{\mathrm{SPP}}$ and SDW-MWF$_{\mathrm{Flex}}$ to obtain combined schemes with improved performance.

5.1. SDW-MWF$_\mu$ based NR and DRC

First the perfect extraction of the clean speech and the noise-only contribution in Fig. 2 is replaced with a SDW-MWF$_\mu$ based NR. The speech component and the corresponding noise component in the reference microphone signal are then estimated using (15) and (17), respectively, as shown in Fig. 4. At this point it is important to emphasize that the main challenge is the estimation of the speech component, which is shown with the solid box in Fig. 4. The estimated speech component can be written as

$$\begin{aligned} Z(k,l) &= \mathbf{W}^H(k,l)\big(\mathbf{X}^s(k,l) + \mathbf{X}^n(k,l)\big) \\ &= Z^s(k,l) + Z^n(k,l) \end{aligned} \tag{28}$$

where $Z^s(k,l)$ is the speech component in $Z(k,l)$ and $Z^n(k,l)$ is residual noise. This is where the usual problem with a cascade of NR and DRC appears since the estimated speech component $Z(k,l)$ is indeed bound to have residual noise, which then could be amplified by the DRC. i.e.,

$$\hat{Z}(k,l) = Z(k,l)G_{\mathrm{DRC,dB}}(k,l). \tag{29}$$

Any such residual noise, from the speech DRC point of view, is now considered a low level signal which is then amplified, while the actual speech component

is considered a high level signal which is then compressed. This leads to the undesired SNR degradation. An example of this is shown in Fig. 5 where the speech and the noise input SPL are 50dB and 30dB, respectively. This shows that with the given DRC curve the output SPL between the speech and the noise is reduced by 10dB which is obviously undesired. On the other hand, the estimated noise component $\bar{Z}(k,l)$ in (17) is better controlled since the noise DRC can be set to zero, i.e., to suppress all noise, or it can be a scaled version of the speech DRC as explained in section 2.

5.2. SDW-MWF$_{SPP}$ based NR and dual-DRC

The DRC described in section 4 amplifies signals based on their intensity level and makes no distinction between speech dominant segments and noise dominant segments. The aim could then be to identify the speech dominant segments and the noise dominant segments such that the residual noise amplification can be avoided. By reusing the conditional SPP $p(k,l)$ estimated in the SDW-MWF$_{\text{SPP}}$ a dual-DRC approach is introduced [19] such that a different DRC curve is applied to the speech dominant segments and to the noise dominant segments. The two DRC curves are defined similarly as in (24)-(26) and the overall DRC output power is then defined as

$$P^{out}_{\text{dual-DRC,dB}}(k,l) = p(k,l) \cdot P^{out,s}_{\text{DRC,dB}}(k,l) + (1 - p(k,l)) \cdot P^{out,n}_{\text{DRC,dB}}(k,l). \quad (30)$$

where $P^{out,s}_{\text{DRC,dB}}(k,l)$ and $P^{out,n}_{\text{DRC,dB}}(k,l)$ are defined by the speech DRC curve and the noise DRC curve, respectively. The dual-DRC gain is then defined as

$$G_{\text{dual-DRC,dB}}(k,l) = P^{out}_{\text{dual-DRC,dB}}(k,l) - P^{in,s}_{\text{DRC,dB}}(k,l). \quad (31)$$

The dual-DRC approach is illustrated in Fig. 6 with an example where the input SPL is 60dB and the output SPL now depends on the conditional SPP $p(k,l)$. The procedure is as follows:

- If speech is present ($p(k,l)$=1) the speech DRC curve is applied.

- If speech is absent ($p(k,l)$=0) it is undesirable to amplify the residual noise compared to the speech and therefore a lower gain is applied, i.e., the noise DRC curve is applied.

- For $0 < p(k,l) < 1$ a weighted sum of the two DRC curves is used.

The rationale behind the noise DRC curve is that it results in a lower gain compared to the speech DRC curve, as the goal indeed is to apply a lower gain to the noise dominant segments compared to the speech dominant segments.

The proposed MWF based NR and dual-DRC using SDW-MWF$_{\text{SPP}}$ is shown in Fig. 7. The main difference between this approach and the MWF based NR and DRC using SDW-MWF$_{\mu}$ is that the speech DRC in Fig. 4 implicitly assumes that the estimated speech component does not contain residual noise. The gain difference between the noise DRC curve and the speech DRC curve in the dual-DRC is given by

$$\Delta G_{\text{dual,dB}} = G^{s}_{\text{dB}} - G^{n}_{\text{H}_1,\text{dB}} \tag{32}$$

where $G^{n}_{\text{H}_1,\text{dB}}$ is the noise DRC curve in the dual-DRC approach. Based on the example given in Fig. 5 it is shown in Fig. 8 that the noise DRC gain $G^{n}_{H_1,\text{dB}}$ need to be 10dB lower than G^{s}_{dB} to compensate for the 10dB reduction between the speech and the noise output SPL. The properties of $G^{n}_{H_1,\text{dB}}$ can be summarized as follows:

- If $G^{n}_{H_1,\text{dB}}$ is set too low the desired hearing aid gain G^{s}_{dB} may be compromised.

- If $G^{n}_{H_1,\text{dB}}$ is set too high the impact of $p(k,l)$ may be too small to compensate for the residual noise amplification.

The goal of the dual-DRC is thus to find a trade-off between NR and DRC, i.e., SNR improvement and the desired DRC gain by adjusting $G^{n}_{H_1,\text{dB}}$ as discussed above.

5.3. SDW-MWF$_{Flex}$ based NR and flexible dual-DRC

Following the above discussion it is desirable to minimize the term in (32) without sacrificing the SNR improvement. This can be achieved by not only using the conditional SPP $p(k,l)$ introduced in the SDW-MWF$_{\text{SPP}}$ but also the H_0 and H_1 state detection $P(l)$ introduced in the SDW-MWF$_{\text{Flex}}$. A flexible dual-DRC can then be written as

$$\begin{aligned} P^{out}_{\text{flex-DRC,dB}}(k,l) &= P(l)\Big[p(k,l)P^{out,s}_{\text{DRC,dB}}(k,l) + (1-p(k,l))P^{out,n}_{\text{DRC,dB}}(k,l)\Big] \\ &\quad + (1-P(l))P^{out,n}_{\text{DRC,dB}}(k,l) \\ &= \begin{cases} H_1 : p(k,l)P^{out,s}_{\text{DRC,dB}}(k,l) + (1-p(k,l))P^{out,n}_{\text{DRC,dB}}(k,l) \\ H_0 : P^{out,n}_{\text{DRC,dB}}(k,l) \end{cases} \end{aligned} \tag{33}$$

where the noise DRC curve $P^{out,n}_{\text{DRC,dB}}(k,l)$ in the H_1 and H_0 can be similiar or in the H_0 state the gain can be set lower. The flexible dual-DRC gain is given by

$$G_{\text{flex-DRC,dB}}(k,l) = P^{out}_{\text{flex-DRC,dB}}(k,l) - P^{in,s}_{\text{DRC,dB}}(k,l). \tag{34}$$

The rationale behind the flexible dual-DRC is:

- When a H_1 state is detected, i.e., $P(l)$=1, a dual-DRC is applied using G^s_{dB} and $G^n_{H_1,\text{dB}}$.
- When a H_0 state is detected, i.e., $P(l)$=0, a DRC is applied with $G^n_{\text{H}_0,\text{dB}}$ $\leq G^n_{\text{H}_1,\text{dB}}$ since in the H_0 state it is not required to get close the desired DRC gain.

The DRC gain difference between the noise DRC curve and the speech DRC curve in the flexible dual-DRC is then given by

$$\begin{aligned} \Delta G_{\text{flex,dB}} &= P(l)\Big[G^s_{\text{dB}} - G^n_{\text{H}_1\text{dB}}\Big] + (1-P(l))G^n_{\text{H}_0\text{dB}} \\ &= \begin{cases} H_1 : G^s_{\text{dB}} - G^n_{\text{H}_1\text{dB}} \\ H_0 : G^n_{\text{H}_0\text{dB}} \end{cases} \end{aligned} \tag{35}$$

The proposed MWF based NR and the flexible dual-DRC using SDW-MWF$_{\text{Flex}}$ is shown in Fig. 9.

6. Experimental Results

In this section, experimental results for the combined approaches are presented. The simulations aim at showing the undesired interaction effects when a MWF based NR and DRC are serially concatenated, and to compare this approach to the proposed combined approaches using the introduced dual-DRC.

6.1. Performance measures

To assess the SNR performance the intelligibility-weighted signal-to-noise ratio (SNR) improvement [28] is used which is defined as

$$\Delta\text{SNR}_{\text{intellig}} = \sum_i I_i(SNR_{i,out} - SNR_{i,in}) \tag{36}$$

where I_i is the band importance function defined in [29] and where $\text{SNR}_{i,out}$ and $\text{SNR}_{i,in}$ represent the output SNR and the input SNR (in dB) for the i^{th} weighted band, respectively. To measure the signal distortion a frequency-weighted log-spectral signal distortion (SD) is used, i.e.,

$$\text{SD} = \frac{1}{K}\sum_{k=1}^{K}\sqrt{\int_{f_l}^{f_u} w_{ERB}(f)\Big(10log_{10}\frac{P^s_{out,k}(f)}{P^s_{in,k}(f)}\Big)^2 df} \tag{37}$$

where K is the number of frames, $P^s_{out,k}(f)$ is the output power spectrum of the kth frame, $P^s_{in,k}(f)$ is the input power spectrum of the kth frame and f is the frequency index. The SD measure is calculated with a frequency-weighting $w_{ERB}(f)$ giving equal weight for each auditory critical band, as defined by the equivalent rectangular bandwidth (ERB) of the auditory filter [30].

6.2. Experimental set-up

Simulations have been performed with a 2-microphone behind-the-ear hearing aid mounted on a CORTEX MK2 manikin. The loudspeakers (FOSTEX 6301B) are positioned at 1 meter from the center of the head. The resulting

Method	SDW-MWF$_\mu$	SDW-MWF$_{\text{SPP}}$	SDW-MWF$_{\text{Flex}}$
Input SNR	0dB	0dB	0dB
ΔSNR	13.1dB	13.2dB	13.9dB
SD	4.2dB	4.3dB	4.2dB

Table 1: SNR improvement and SD of the different SDW-MWF based NR.

room reverberation time $T_{60} = 0.21s$. The speech is located at 0° and the two multi-talker babble noise sources are located at 120° and 180°. The speech signals consist of male sentences from the HINT-database [31] and the noise signals consist of a multi-talker babble from Auditec [32]. The signals are sampled at 16kHz. Both the MWF based NR and the DRC are implemented using an FFT length of 128 with half overlapping frames. The DRC is implemented based on critical bands [33] which is realized by using individual FFT bins at low frequencies and by combining FFT bins at higher frequencies [10]. The following parameters are fixed during all simulations:

- Input level is set to 65dB SPL at the microphone.
- Attack and release time are set to at=10ms and rt=20ms.
- Compression ratio CR=2.
- Compression threshold is set to CT=30dB.

In order to evaluate the effect from the DRC on the different SDW-MWF based NR and to make a fair comparison each SDW-MWF algorithms are adjusted such that the SNR improvement and SD are as similar as possible, see table 1.

6.3. SNR improvement

The gain settings in the first experiment are shown in Table 2. Notice that G^n_{dB} is set to zero since since the aim is to show the effect of the DRC on the SNR improvement for the NR shown in Table 1. The results for these experiments are shown in Fig. 10 and 11. This shows that the DRC degrades

	G^s_{dB} (1)	G^n_{dB} (1)	$G^n_{H_1,\text{dB}}$ (32)	$G^n_{H_0,\text{dB}}$ (35)
SDW-MWF$_\mu$	30dB	0	N/A	N/A
SDW-MWF$_{\text{SPP}}$	30dB	0	10dB-30dB	N/A
SDW-MWF$_{\text{Flex}}$	30dB	0	27.5dB	20dB-25dB

Table 2: Gain settings for the first experiment.

	G^s_{dB} (1)	G^n_{dB} (1)	$G^n_{H_1,\text{dB}}$ (32)	$G^n_{H_0,\text{dB}}$ (35)
SDW-MWF$_\mu$	30dB	0dB-30dB	N/A	N/A
SDW-MWF$_{\text{SPP}}$	30dB	0dB-30dB	20dB-27.5dB	N/A
SDW-MWF$_{\text{Flex}}$	30dB	0dB-30dB	27.5dB	20dB-25dB

Table 3: Gain settings for the second experiment.

the SNR improvement of the SDW-MWF$_\mu$ and the SDW-MWF$_{\text{SPP}}$ by 6dB which is illustrated at $\Delta G_{\text{dual,dB}}$=0dB compared to Table 1. The dotted line shows the SNR improvement for SDW-MWF$_{\text{SPP}}$ and dual-DRC as a function of $\Delta G_{\text{dual,dB}}$. Better performance is achieved when $\Delta G_{\text{dual,dB}}$ increases as this increases the impact of the dual-DRC. The SDW-MWF$_{\text{Flex}}$ based NR and the flexible dual-DRC is seen to achieve larger SNR improvement at a small increase in SD as low as 1dB.

6.4. Output SNR

The gain settings in the second experiment are shown in Table 3. In this experiment, the performance of the different schemes is compared to the ideal performance, i.e., when the speech DRC is applied to the clean speech and the noise DRC is applied to the noise-only signal, see section 2. The results for these experiments are shown in Fig. 12 and 13. The dashed line shows the ideal output SNR which as expected improves when ΔG_{dB} is increased. For the combined schemes the output SNR also improves but at the same time it is clear that the

SNR improvement is smaller than in the ideal case. This happens because the signal filtered by $\mathbf{W}_{\mu}(k,l)$, $\mathbf{W}_{\text{SPP}}(k,l)$ and $\mathbf{W}_{\text{Flex}}(k,l)$ contains residual noise which is subsequently receives more amplification compared to the speech. It is also worth noting that when $\Delta G_{\text{dB}} < 20\text{dB}$ the output SNR is higher for SDW-MWF$_{\mu}$ and DRC which is due to the fact that the overall gain with the DRC is higher than with the dual-DRC gain. Using the flexible dual-DRC improves the output SNR but it is still far from the ideal performance.

7. Conclusion

In this paper, the undesired interaction effects in a serial concatenation of a MWF based NR and DRC are analysed. First of all it is shown that having a traditional SDW-MWF$_{\mu}$ based NR and DRC leads to a SNR degradation. The reason for this is that a traditional DRC only uses the intensity level of a signal segment to estimate the gain independently of whether a speech dominant segment or a noise dominant segment is considered. This is highly undesirable since this consequently defeats the purpose of having a NR algorithm, as the residual noise recieves more amplification compared to the speech after the NR stage.

The combined solutions proposed here are based on two modifications both in the MWF based NR process and in the DRC. The first modification is to incorporate the conditional SPP in the NR process, which is referred to as SDW-MWF$_{\text{SPP}}$. Using the conditional SPP serves the purpose of identifying the speech dominant segments and the noise dominant segments. The second modification is based on reusing the conditional SPP estimated in the SDW-MWF$_{\text{SPP}}$ to change the DRC into a dual-DRC that incorporates the conditional SPP. The dual-DRC uses two compression curves instead of one compression curve in a traditional DRC. The two compression curves allow a switchable compression characteristic based on the conditional SPP, i.e., smaller gain is applied to the noise dominant segments whereas in the speech dominant segments the aim is to apply a gain similar to a traditional DRC. Experimental

results indeed confirm that a serial concatenation of NR and DRC degrades the SNR improvement provided by the NR, whereas the combined approach proposed here shows less degradation of the SNR improvement at a low increase in distortion compared to a serial concatenation.

Acknoledgements

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), Concerted Research Action GOA-MaNet, the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011), Research Project FWO nr. G.0600.08 ('Signal processing and network design for wireless acoustic sensor networks'), IWT Project 'Signal processing and automatic fitting for next generation cochlear implants', EC-FP6 project SIGNAL: 'Core Signal Processing Training Program'. The scientific responsibility is assumed by its authors.

References

[1] H. Dillon, *Hearing Aids*. Boomerang Press, Thieme, 2001.

[2] J. M. Kates, *Digital Hearing Aids*. Plural Publishing, 2008.

[3] V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, and U.Rass, "Signal processing in high-end hearing aids: State of the art, challenges, and future trends," *EURASIP Journal on Applied Signal Processing*, vol. 18, pp. 2915–2929, 2005.

[4] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, no. 2, pp. 113–120, Apr. 1979.

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

[6] I. Frost, O.L., "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[7] L. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 1, pp. 27–34, Jan. 1982.

[8] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for the speech distortion weighted multichannel wiener filter for robust noise reduction," *Speech Communication*, vol. 7-8, pp. 636–656, Jul. 2007.

[9] A. Spriet, M. Moonen, and J. Wouters, "Stochastic gradient based implementation of spatially pre-processed speech distortion weighted multichannel wiener filtering for noise reduction in hearing aids," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 911–625, Mar. 2005.

[10] J. M. Kates and K. H. Arehart, "Multichannel dynamic-range compression using digital frequency warping," *EURASIP Journal on Applied Signal Processing*, vol. 18, pp. 3003–3014, 2005.

[11] T. Herzke and V. Hohmann, "Effects of instantaneous multiband dynamic compression on speech intelligibility," *EURASIP Journal on Applied Signal Processing*, vol. 18, pp. 3034–3043, 2005.

[12] P. J. Blamey, D. S. Macfarlane, and B. R. Steele, "An intrinsically digital amplification scheme for hearing aids," *EURASIP Journal on Applied Signal Processing*, vol. 18, pp. 3026–3033, 2005.

[13] T. Schneider and R. Brennan, "A multichannel compression strategy for a digital hearing aid," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 1, pp. 411–414 vol.1, Apr. 1997.

[14] M. Li, H. McAllister, N. Black, and T. De Perez, "Wavelet-based nonlinear AGC method for hearing aid loudness compensation," *IEE Proceedings - Vision, Image and Signal Processing* , vol. 147, no. 6, pp. 502–507, Dec. 2000.

[15] K. Chung, "Effective compression and noise reduction configurations for hearing protectors," *Journal of the Acoustical Society of America*, vol. 121, no. 2, pp. 1090–1101, Feb. 2007.

[16] A. K. H. Anderson, M. C. and J. M. Kates, "The acoustic and peceptual effects of series and parallel processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 619805, 20 pages, 2009, doi:10.1155/2009/619805.

[17] K. Chung, "Challenges and recent developments in hearing aids part i: Speech understanding in noise, microphone technologies and noise reduction algorithms," *Trends in Amplification*, vol. 8, no. 3, pp. 83–124, 2004.

[18] D. Mauler, A. M. Nagathil, and R. Martin, "On optimal estimation of compressed speech for hearing aids," *Interspeech*, pp. 826–829, Aug. 27-31, 2007.

[19] K. Ngo, S. Doclo, A. Spriet, M. Moonen, J. wouters, and S. H. Jensen, "An integrated approach for noise reduction and dynamic range compression in hearing aids," *European Signal Processing Conference (EUSIPCO), Lausanne, Switzerland*, Aug. 2008.

[20] K. Ngo, A. Spriet, M. Moonen, J. Wouters, and S. H. Jensen, "Incorporating the conditional speech presence probability in multi-channel wiener filter based noise reduction in hearing aids," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 930625, 11 pages, 2009, doi:10.1155/2009/930625.

[21] A. Spriet, M. Moonen, and J. Wouters, "A multichannel subband GSVD based approach for speech enhancement in hearing aids," *IEEE Interna-*

tional Workshop on Acoustic Echo and Noise Control), (IWAENC), Darmstadt, Germany, pp. 187-190, Sept. 2001.

[22] A. Spriet, M. Moonen, and J. Wouters, "A multichannel subband GSVD approach to speech enhancement," *European Transactions on Telecommunications, Special Issue on Acoustic Echo and Noise Control*, vol. 13, no. 2, pp. 149-158, Mar. 2002.

[23] S. Doclo and M. Moonen, "GSVD-Based Optimal Filtering for Single and Multi-Microphone Speech Enhancement," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2230-2244, Sept. 2002.

[24] G. Rombouts and M. Moonen, "QRD-based unconstrained optimal filtering for acoustic noise reduction," *Signal Processing*, vol. 83, no. 9, pp. 1889-1904, Sept. 2003.

[25] K. Ngo, A. Spriet, M. Moonen, J. Wouters, and S. H. Jensen, "Variable speech distortion weighted multichannel wiener filter based on soft output voice activity detection for noise reduction in hearing aids," *IEEE International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Seattle, USA, Sept. 2008.

[26] K. Ngo, M. Moonen, J. Wouters, and S. H. Jensen, "A Flexible Speech Distortion Weighted Multi-Channel Wiener Filter for Noise Reduction in Hearing Aids," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2008.

[27] I. Cohen, "Optimal speech enhancement under signal presence uncertainty using log-spectral amplitude estimator," *IEEE Signal Processing Letters*, vol. 9, no. 4, pp. 113–116, Apr. 2002.

[28] J. E. Greenberg, P. M. Peterson, and P. M. Zurek, "Intelligibility-weighted measures of speech-to-interference ratio and speech system performance," *Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 3009–3010, Nov. 1993.

[29] Acoustical Society of America, "ANSI S3.5-1997 American National Standard Methods for calculation of the speech intelligibility index," Jun. 1997.

[30] B. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. Academic Press, 2003.

[31] M. Nilsson, S. D. Soli, and A. Sullivan, "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," *Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 1085–1099, Feb. 1994.

[32] Auditec, "Auditory Tests (Revised), Compact Disc, Auditec, St. Louis," St. Louis, 1997.

[33] E. Zwicker, "Subdivision of the audible frequency range into critical bands," *Journal of the Acoustical Society of America*, vol. 33, no. 2, pp. 248–249, Feb. 1961.

Figure 1: NR and DRC in a serial concatenation compared to an integrated scheme.

Figure 2: Ideal system where the speech and the noise-only contributions are extracted which then are compressed separately.

Figure 3: DRC curve (CR defines how the slope is changed and CT is the point at which the slope changes).

Figure 4: A serial concatenation of a SDW-MWF$_\mu$ based NR and DRC.

Figure 5: Illustration of the output SPL after the DRC with the noise located at 30dB input SPL and the speech at 50dB SPL.

Figure 6: Dual-DRC with the conditional speech presence probability $p(k,l)$ to provide a weighting between the two DRC curves.

Figure 7: A combined approach of a SDW-MWF$_{\mathrm{SPP}}$ based NR and dual-DRC.

Figure 8: Illustration of the output SPL after the dual-DRC with the noise located at 30dB input SPL and speech at 50dB SPL.

Figure 9: A combined approach of a SDW-MWF$_{\mathrm{Flex}}$ based NR and a flexible dual-DRC.

Figure 10: The SNR improvement for the different SDW-MWF based NR and DRC.

Figure 11: The distortion for the different SDW-MWF based NR and DRC.

Figure 12: The effect of $\Delta \mathrm{G}_{\mathrm{dB}}$ on the output SNR for SDW-MWF$_\mu$ and SDW-MWF$_{\mathrm{SPP}}$ based NR and DRC.

Figure 13: The effect of $\Delta \mathrm{G}_{\mathrm{dB}}$ on the output SNR for SDW-MWF$_\mu$ and SDW-MWF$_{\mathrm{Flex}}$ based NR and DRC.