

A Distortion Reducing Subband Limiter Implementation for Conference Phones

Christian Schüldt^a, Fredric Lindstrom^b, Ingvar Claesson^a

^aBlekinge Institute of Technology, Department of Signal Processing, SE-37225, Ronneby, Sweden

^bKonftel AB, Research and Development, Box 268, SE-90106, Umeå, Sweden

Abstract- Distortion, most often caused by the loudspeaker, is a practical problem in acoustic echo cancellation-based conference phones. Typically, this distortion varies significantly with the frequency. This paper presents an implementation of a distortion reducing subband limiter to be used in conjunction with an acoustic echo canceller in a commercial conference phone, yielding improved cancellation performance. Verification of the performance was performed through evaluation of a fix-point real-time implementation.

I. INTRODUCTION

An acoustic echo cancellation-based conference phone typically comprises a loudspeaker, a microphone, a processing unit and suitable circuitry [1]. Whenever the loudspeaker is active, the microphone picks up an acoustic echo which is then more or less cancelled by a linear adaptive filter in the processing unit [2], see also figure 1. The problem is that there often are non-linear components, generally originating from loudspeaker driver distortion or resonance caused by the loudspeaker enclosure, embedded in the acoustic echo. Loudspeaker drivers used in low cost consumer electronic products often have a limited operating frequency range, most notably in the bass. A common approach to reduce the non-linearities caused by this is to simply filter the output, e.g. remove the bass, prior to feeding it to the loudspeaker. The downside with this approach is that even at low sound intensities which would not cause significant distortion, deterioration of the perceived audio quality is obvious due to lack of the removed frequency components.

This paper instead proposes a distortion reducing subband limiter. The advantage of this solution is that the signal fed to the speaker can easily be processed to keep distortion low as well as maintaining as good perceived audio quality as possible. A similar method has previously been used for increasing the perceived audio quality of inferior loudspeakers [3], while in this paper the aim is to reduce the distortion prior to acoustic echo cancellation while maintaining the perceived audio quality.

II. PROPOSED ALGORITHM

A number of I finite impulse response (FIR) filters, \mathbf{h}_i , where $i \in \{0, \dots, I-1\}$ denotes subband index, with N

Thanks to the Swedish Knowledge Foundation (KKS) for funding.

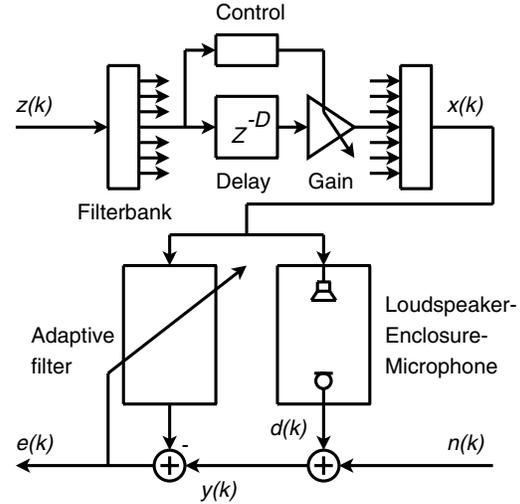


Fig. 1. A simplified schematic of the setup.

coefficients each are used in the proposed algorithm to construct a uniform filter bank, see figure 1. The filterbank is used to divide the input signal $z(k)$ into I subband signals $z_i(k)$. To maintain a low signal delay, no downsampling is used, also eliminating the need of synthesis filters. This particular filterbank structure is chosen to facilitate simple integration with existing conference phone software [4], although other filterbank structures are of course possible.

A maximum of each subband signal is calculated as

$$z_{i,\max}(k) = \max\{|z_i(k)|, \gamma_m z_{i,\max}(k-1)\}, \quad (1)$$

where γ_m is a parameter defining the decay of the function. The max values $z_{i,\max}(k)$ are then used as input to a limiter, with a threshold T_i in each subband. Attack- and release times are set to be the same for all subbands. Thus, the attenuation in each subband is

$$g_{i,\text{filt}}(k) = \gamma_i(k)g_{i,\text{filt}}(k-1) + (1 - \gamma_i(k))g_i(k), \quad (2)$$

where

$$\gamma_i(k) = \begin{cases} \gamma_a & \text{if } g_i(k) < g_{i,\text{filt}}(k-1) \\ \gamma_r & \text{otherwise} \end{cases} \quad (3)$$

and

$$g_i(k) = \begin{cases} \frac{T_i}{z_{i,\max}(k)} & \text{if } z_{i,\max}(k) > T_i \\ 1 & \text{otherwise} \end{cases} \quad (4)$$

The attenuation in each subband is applied to a delayed version of each subband signal and all attenuated subband signals are then summed to form the loudspeaker signal, see also figure 1,

$$x(k) = \sum_{i=0}^{I-1} z_i(k-D)g_{i,\text{filt}}(k), \quad (5)$$

where D is a delay to allow enough time between the calculation of the subband attenuation, equations (2), (3) and (4), and the application in equation (5).

If the specifications of the loudspeaker at hand is known, setting of T_i is performed manually. If not, the thresholds are set adaptively, i.e. $T_i = T_i(k)$ is raised as long as sufficient echo cancellation is achieved, otherwise it is reduced,

$$T_i(k) = T_i(k-1) + \Delta T_i \quad \text{if } \frac{P_e(k)}{P_y(k)} < c_1 \hat{J}(\infty), \quad (6)$$

$$T_i(k) = T_i(k-1) - \Delta T_i \quad \text{if } \frac{P_e(k)}{P_y(k)} > c_2 \hat{J}(\infty), \quad (7)$$

where ΔT_i , c_1 , and c_2 are constants, $\hat{J}(\infty)$ is the estimated steady-state mean square error [2] and $P_e(k)$ and $P_y(k)$ are the short-term average power of the echo cancelled signal $e(k)$ and the microphone signal $y(k)$, respectively. Care should be taken to ensure that $\hat{J}(\infty)$ is evaluated in a single-talk state and when the filter is well converged. For an adaptive filter $\hat{\mathbf{w}}(k)$ this can be achieved by verifying that

$$\sum_{j=0}^J \|\hat{\mathbf{w}}(k) - \hat{\mathbf{w}}(k - jL)\| < T_{\text{conv}}, \quad (8)$$

where J , L and T_{conv} are constants, and using a state of the art doubletalk-detector [1].

III. REAL-TIME IMPLEMENTATION

The proposed algorithm was implemented in real time on a fix-point digital signal processor [5] together with a normalized least mean square adaptive filter [2]. The sampling rate was 8khz and the parameter settings of the key parameters were $I = 32$, $N = 64$, $D = 100$, $\gamma_m = 0.9$, $\gamma_a = 0.9$ and $\gamma_r = 0.96$.

In figure 2, plot (a) and (b) the short term average power of the microphone signal $y(k)$ and the echo cancelled signal $e(k)$, respectively, for an input signal $z(k)$ consisting of a 12 seconds long speech signal is shown. In plot (a), the proposed algorithm is used and in plot (b), the proposed algorithm is bypassed, i.e. $z(k)$ is fed to the loudspeaker without processing ($x(k) = z(k)$). In plot (c), the ratio $\frac{P_y(k)}{P_e(k)}$ for both the situation with the proposed subband limiter and without is shown. The proposed subband limiter manages to reduce a significant amount of the distortion, resulting in much better echo cancellation, as can be seen in the figure. This allows better conference phone duplex, which in turn enhances the perceived audio quality. Subjective tests

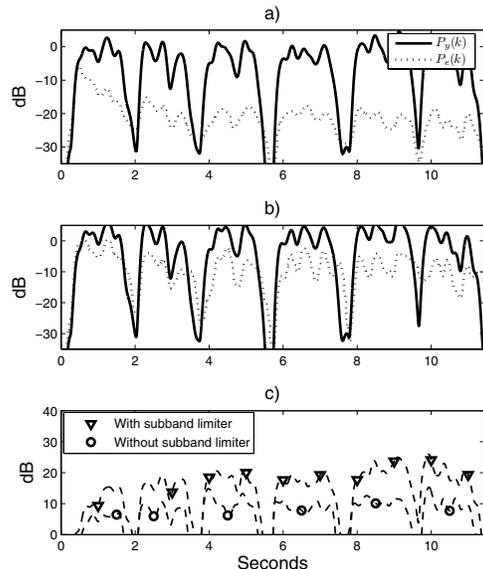


Fig. 2. Plots (a) and (b) show $P_e(k)$ and $P_y(k)$ with the proposed method and without the proposed method, respectively. Plot (c) shows the ratio $\frac{P_y(k)}{P_e(k)}$.

showed that for bass intensive signals at high volumes, effects of the proposed algorithm was audible but not annoying. For low volumes, effects of the proposed algorithm was not audible. Furthermore, duplex performance was improved, resulting in improved overall user comfort.

IV. CONCLUSION

This paper presented an implementation of a subband limiter for reducing distortion in conjunction with an acoustic echo canceller. Effects of the proposed subband limiter in terms of increased echo cancellation performance in a real-time fix-point environment were shown. Subjective tests showed an increase of the perceived sound quality.

V. REFERENCES

- [1] E. Hänsler and G. Schmidt, *Acoustic Echo and Noise Control: A Practical Approach*, Wiley, 2004.
- [2] S. Haykin, *Adaptive Filter Theory*, 4th ed., Prentice, 2002.
- [3] R. M. Aarts, "Methods to Improve the Sound Reproduction of Limited Size Systems," *IEEE Int. Conf. on Consumer Electronics (ICCE)*, 2007.
- [4] C. Schüldt, F. Lindstrom and I. Claesson, "A Combined Implementation of Echo Suppression, Noise Reduction and Comfort Noise in a Speaker Phone Application," *IEEE Int. Conf. on Consumer Electronics (ICCE)*, 2007.
- [5] *ADSP-BF533 hardware reference*, Analog Devices, 2005.