High-Quality Inter-Channel Coherence Reduction For Stereo Acoustic Echo Cancellation

...or how to corrupt an audio signal and get away with it

Jean-Marc Valin, CSIRO ICT Centre

Introduction

Context: Acoustic echo cancellation with stereo signals
Problem: Inter-channel coherence makes the system ill-conditioned
Solution: Decorrelate the left and right channels using nonlinear processing while preserving the audio quality

Stereo Acoustic Echo Cancellation

Stereo acoustic echo cancellation can be improved by applying non-linear processing to both channels before playback

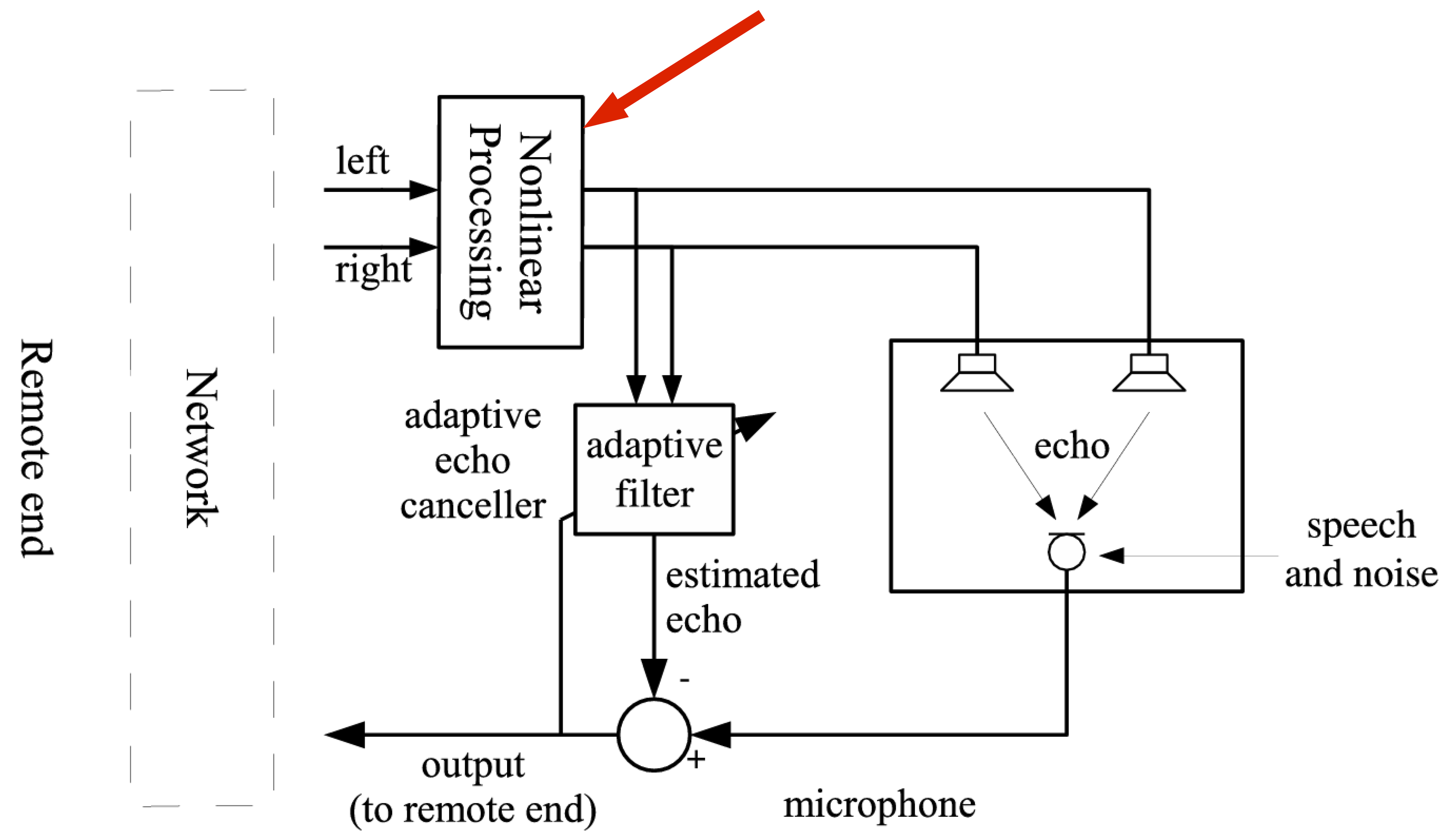

Figure 1: Stereo acoustic echo cancellation system.

The goal is to minimise the inter-channel coherence at the output, while maximising the output quality (including stereo image)

Stereo localisation cues used by the human ear:
- Inter-aural intensity difference (high frequency)
- Inter-aural phase difference (low frequency)

We propose a two-part approach:
- Alter phase at high frequencies (where the ear is insensitive)
- Add noise at low frequencies

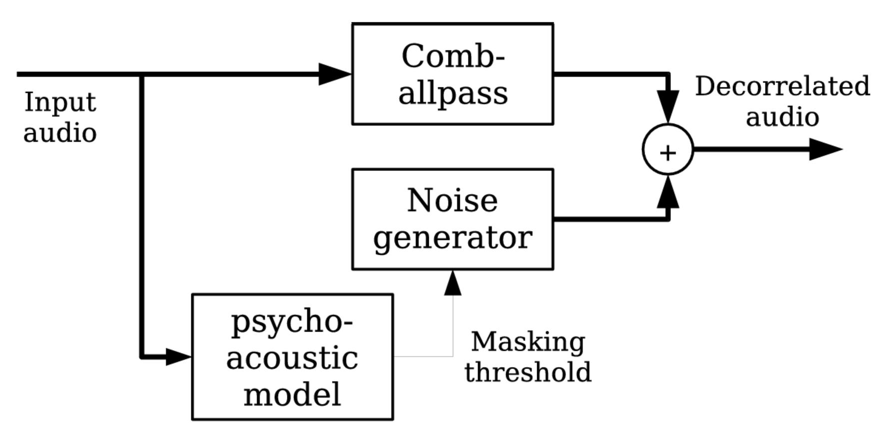

Figure 2: Overview of the proposed approach.

Psycho-acoustically masked noise

- Compute masking curve on current data
- Add in the frequency domain (WOLA)
- Only delay the noise, don't delay the signal (exploits temporal masking)

Shaped Comb-Allpass (SCAL) Filter

Comb all-pass filter

$$A(z) = \frac{\alpha + z^{-N}}{1 - \alpha z^{-N}}$$

Shaped Comb-Allpass (SCAL) adds a tilt to the phase response

$$A(z) = \frac{\alpha\left(1 - \beta z^{-1}\right) + z^{-N}}{1 - \alpha\left(-\beta z^{-N+1} + z^{-N}\right)}$$

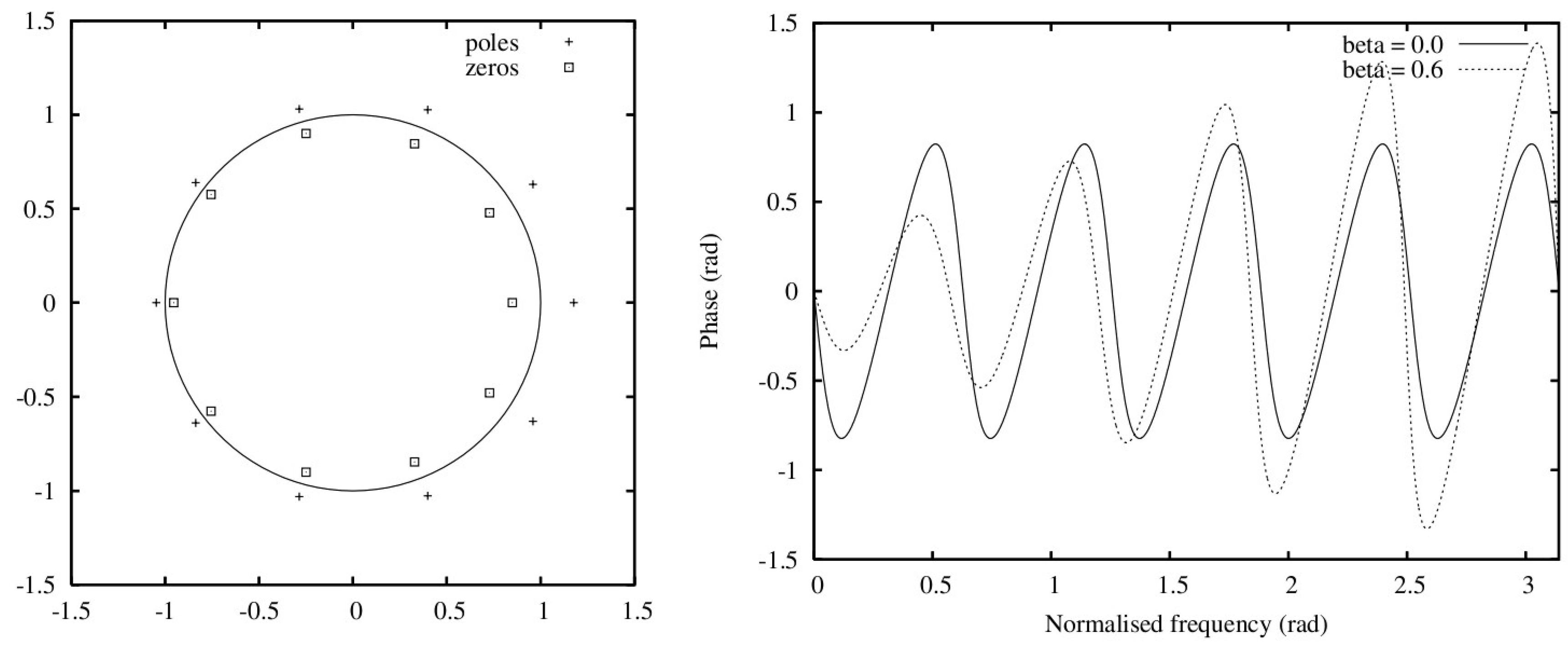

Figure 2: Left: pole-zero representation of the SCAL filter. Right: phase response of the comb-allpass and SCAL filters.

For the process to be non-linear, we need to vary both α and *N* over time:
- *N* chosen randomly between [5,10]
- α chosen as a function of *N* with memory

$$\alpha(N) = \min\left(\left(\alpha(N-1) + r_0\right), \frac{1-\epsilon}{1+|\beta|} \right)$$

The SCAL filter preserves the stereo image because the low frequency phase is not heavily altered

Results

We compare to other algorithms:
- Non-shaped comb-allpass (no noise added)
- Smoothed absolute value non-linearity
- 1st order all-pass filter

Comparison on 41.1 kHz audio (speech and music)

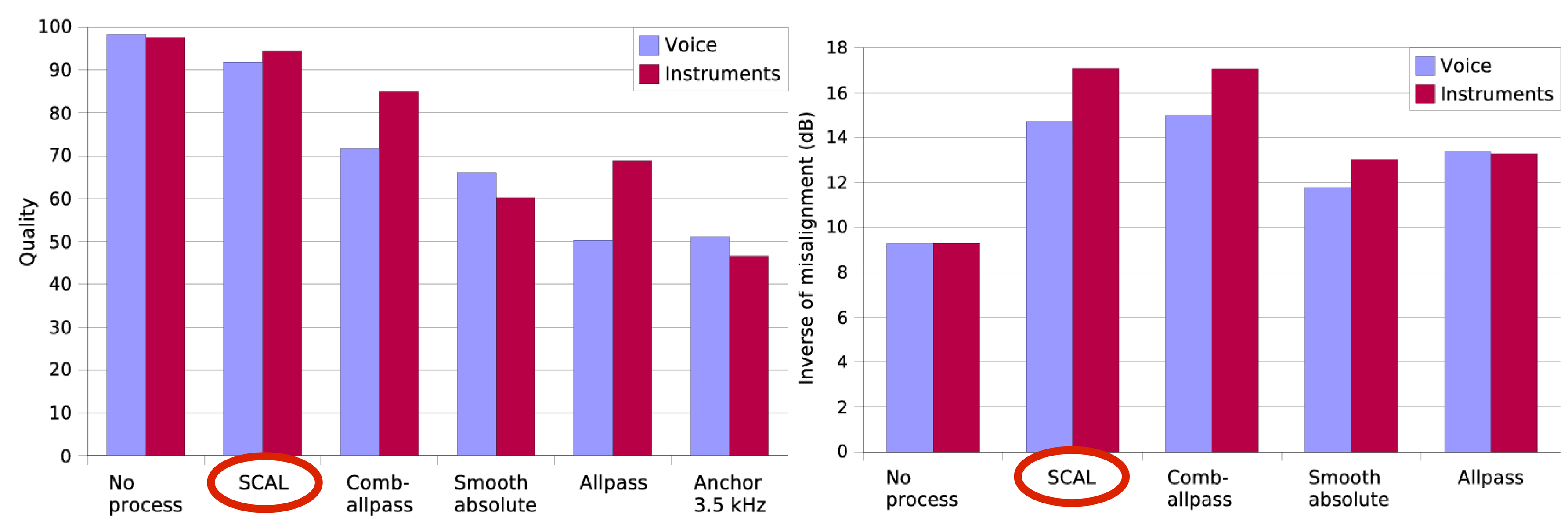

Figure 4: Left: MUSHRA quality of the signal from different algorithms (higher is better). Right: Inverse of the filter misalignment (lower is better).

Conclusion

The proposed solution achieves:
- Quality close to the unmodified signal
- Good coherence reduction (better convergence)
- Only 10 samples total algorithmic delay (< 1ms)

References

J. Benesty, D. R. Morgan, and M. M. Sondhi. A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation. IEEE Trans. SAP, 6(2):156–165, 1998.
T. Gänsler and J. Benesty. New insights to the stereophonic acoustic echo cancellation problem and an adaptive nonlinearity solution. IEEE Trans. Speech Audio Processing, 10(5), 2002.
M. M. Sondhi, D. R. Morgan, and J. L Hall. Stereophonic acoustic echo cancellation-an overview of the fundamental problem. IEEE Signal Processing Letters, 2(8):148–151, 1995.

References (cont.)

D. R. Morgan, J. L. Hall, and J. Benesty. Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation. IEEE Trans. SAP, 9(6):686–696, 2001.
M. Ali. Stereophonic acoustic echo cancellation system using time-varying all-pass filtering for signal decorrelation. In Proc. ICASSP, pages 3689–3692, 1998.
T. Gänsler and P. Eneroth. Influence of audio coding on stereophonic acoustic echo cancellation. In Proc. ICASSP, 1998.

Further information

contact: Jean-Marc Valin
phone: +61 (0)2 9889 3476
email: jean-marc.valin@ieee.org
web: http://www.ict.csiro.au/

www.csiro.au

CSIRO