

MICROPHONE INTERFERENCE REDUCTION IN LIVE SOUND

Alice Clifford, Josh Reiss*

Centre for Digital Music
Queen Mary, University of London
London, UK

alice.clifford@eeecs.qmul.ac.uk

ABSTRACT

When multiple microphones are used to reproduce multiple sources microphone interference, or bleed, can occur due to each microphone picking up more than one source. This paper proposes combining the crosstalk resistant adaptive noise canceller (CTRANC) algorithm with centred adaptive filters using an estimation of delay to suppress the interference, while making little change to the target signal. The proposed method is compared with similar methods in both the anechoic and echoic cases. The method is shown to outperform the other methods in the anechoic case while in the echoic case it is shown to perform less well at reducing the level of the interference but still introduces the least artefacts. Extension to the proposed method to the N source and microphone case is also discussed.

1. INTRODUCTION

In a live sound performance it is common for multiple instruments or musicians to be performing at the same time. A common technique for setting microphones in this situation is to place a dedicated microphone to reproduce each sound source. Ideally, a given microphone signal will only contain the sound from a single source. In reality, a microphone may reproduce any number of sources surrounding it. This is similar to the concept of crosstalk in telecommunications and can be called bleed or leakage.

A microphone reproduces sound that enters the area surrounding it which is described by its polar pattern. When placing a microphone to reproduce a target sound source, it is placed to ensure the source is within this area. Sound from other sources may also enter this area and will also be reproduced, which can be referred to as interference. This interfering signal is assumed to consist of target signals of other microphones, as shown in Figure 1 and described in Equation (1)

$$\begin{aligned} x_1[n] &= \alpha_{11}s_1[n - \tau_{11}] + \alpha_{21}s_2[n - \tau_{21}] \\ x_2[n] &= \alpha_{12}s_1[n - \tau_{12}] + \alpha_{22}s_2[n - \tau_{22}], \end{aligned} \quad (1)$$

where x_1 and x_2 are microphone signals at timestep n , s_1 and s_2 are the sound sources, τ_{11} , τ_{12} , τ_{21} and τ_{22} are the delays of each source arriving at each microphone and α_{11} , α_{12} , α_{21} and α_{22} represent gain.

Interference of other sources causes a number of problems. An interfering signal can be a nuisance and can reduce the intelligibility of the target source. It can affect the overall gain of the microphone signal. It also means that if any processes are applied with

* The test audio in this research was excerpts of the raw audio of "Ana" by Vieux Farka Touré, available under a Creative Commons Attribution-NonCommercial license.

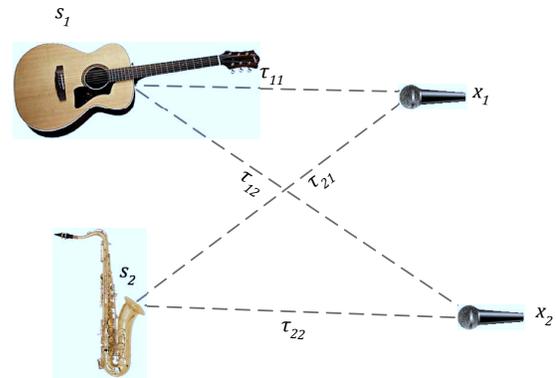


Figure 1: A configuration of 2 sources being reproduced by 2 microphones with the direct signal paths and equivalent delays shown.

the intention of being applied to the target signal, such as equalisation, it will also be applied to the interfering signal potentially causing problems in a mix.

Microphone interference can also cause comb filtering. Comb filtering occurs when a signal and a delayed version of the same signal are summed. A comb filter has defined peaks and troughs in the frequency response, caused by reinforcement and cancellation in the frequency domain. The comb filtering effect can be heard when the duplicated source is as much as 18dB quieter than the original [1].

If the microphone signals defined in Equation (1) are summed to the output y this becomes

$$y[n] = x_1[n] + x_2[n] \quad (2)$$

$$\begin{aligned} &= \alpha_{11}s_1[n - \tau_{11}] + \alpha_{12}s_1[n - \tau_{12}] + \\ &\quad \alpha_{21}s_2[n - \tau_{21}] + \alpha_{22}s_2[n - \tau_{22}] \end{aligned} \quad (3)$$

assuming

$$\tau_{11} < \tau_{21} \quad (4)$$

$$\tau_{22} < \tau_{12}. \quad (5)$$

Equation (3) shows that two versions of each source with different delays will be summed, thus causing comb filtering of both sources. The relative difference of the delay of each source arriving at each microphone is defined by

$$\tau_1 = \tau_{21} - \tau_{11} \quad (6)$$

$$\tau_2 = \tau_{12} - \tau_{22} \quad (7)$$

and the relative gain difference as

$$\alpha_1 = \alpha_{21} - \alpha_{11} \quad (8)$$

$$\alpha_2 = \alpha_{12} - \alpha_{22}. \quad (9)$$

Microphone and instrument placement plays an important role in the amount of microphone bleed that occurs. In a studio situation, for example, instruments can be isolated either in separate live rooms or by erecting baffles to provide some sound isolation. In a live sound situation this is not aesthetically appropriate. Microphone placement can be used to an advantage by using directional microphones and placing interfering signals in the null areas of a microphone's pick up area. This will not eliminate all interference and could cause other artefacts to occur, such as problems in the low frequencies due to the proximity effect.

1.1. Blind Source Separation

This problem can be looked at from a Blind Source Separation (BSS) point of view. BSS methods attempt to extract N sources from a mixture. The work in this paper is aimed at live sound where it is imperative that a method is able to run in real time. BSS methods generally are offline processes but a number of real-time implementations exist such as [2] and [3]. The method in [3] is taken from the DUET method of source separation, first presented in [4] and extended in [5]. Although stated to run in realtime, this method of source separation is aimed at the unmixing of N sources from 2 mixtures, i.e. from a stereo mix of panned sources. It is possible to use this method for 2 microphone recordings, but there are limitations to the distance between the microphones, which is reliant on sampling frequency for example at 16kHz the maximum distance allowed between the microphones for the method to run is when $d \leq 2.15\text{cm}$ [5]. The method in [2] is also used for stereo mixtures, assuming there is phase coherence between the mixtures and only intensity differences. This cannot be assumed in the multiple microphone case.

1.2. Noise Cancellation

Many of the problems that affect live sound are also present in telecommunications, for example noise and reverberation. Techniques exist in telecommunications for echo and noise cancellation, which share the same principles, and also run in real-time. The drawback is that most techniques are optimised for voice signals with lower bandwidths, for example a sampling rate of 8kHz is common [6] whereas in live sound we require a bandwidth to represent all the audible frequencies from 20Hz to 20kHz. For this reason, when an algorithm optimised for voice application is extended to incorporate wider bandwidth signals, the computational cost inherently increases.

In telecommunications, it is common that an external noise source will interfere with the direct source, for example a person speaking into a telephone may also have an interfering noise, such as air conditioning, in the same room. If an adequate estimation of the noise source is possible, this can be removed from the direct signal. This is where noise and echo cancellation can be used.

Common techniques for noise cancellation make use of an adaptive filter to estimate the impulse response of the interference of the noise signal to the main signal. These methods rely on a clean reference of the noise signal. In reality, this is not always the case. In a live sound scenario, a clean reference signal may not be

available as microphone bleed is assumed to be occurring on all signals.

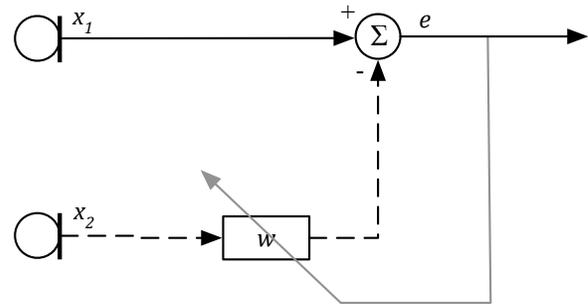


Figure 2: Block diagram of an adaptive filter

The Least Mean Squares (LMS) adaptive filter error is calculated by

$$e[n] = x_1[n] - \mathbf{W}^T[n]\mathbf{X}_2[n], \quad (10)$$

where the error e is also the clean output signal, where

$$\mathbf{X}_2[n] = [x_2[n], x_2[n-1], \dots, x_2[n-L+1]]^T \quad (11)$$

and the estimated update filter is

$$\mathbf{W}[n] = [w_0, w_1, \dots, w_{L-1}]^T. \quad (12)$$

The estimate of $\mathbf{W}[n]$ is then calculated by minimising $E\{e^2[n]\}$

$$\mathbf{W}[n+1] = \mathbf{W}[n] + \mu e[n]\mathbf{X}_2[n], \quad (13)$$

where μ is the adaptation step, which is generally a small value that effects convergence speed and accuracy, and the error signal e is the clean output.

Work in [7] addresses the same problem assuming close microphones and finding the Wiener filter solution by Power Spectral Density (PSD) estimation.

2. CTRANC

A crosstalk resistant adaptive noise canceller (CTRANC) [8] assumes there is crosstalk or as it is referred to in this paper, microphone bleed, between the microphones. For this reason a clean reference is not assumed. Adaptive filters are then cascaded so the output of one becomes the input of the other, as shown in Figure 3. In this way, once one signal has the interference cancelled out it can be used as the reference for the interference cancellation of another source and vice versa [9], [6]. The LMS algorithm then becomes

$$e_1[n] = x_1[n] - \mathbf{W}_A^T \mathbf{E}_2[n] \quad (14)$$

$$e_2[n] = x_2[n] - \mathbf{W}_B^T \mathbf{E}_1[n], \quad (15)$$

where the FIR adaptive filters are

$$\mathbf{W}_A[n] = [w_1^A[n], w_2^A[n], \dots, w_N^A[n]] \quad (16)$$

$$\mathbf{W}_B[n] = [w_1^B[n], w_2^B[n], \dots, w_N^B[n]] \quad (17)$$

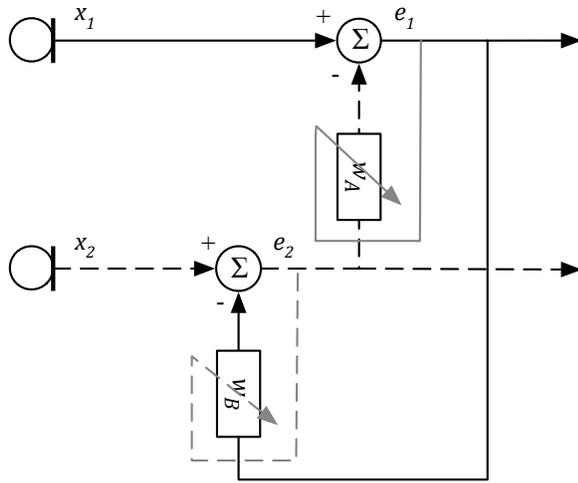


Figure 3: Block diagram of the CTRANC with 2 sources

and error vectors are

$$\mathbf{E}_1[n] = [e_1[n], e_1[n-1], \dots, e_1[n-N]] \quad (18)$$

$$\mathbf{E}_2[n] = [e_2[n], e_2[n-1], \dots, e_2[n-N]]. \quad (19)$$

Each filter is then updated by

$$\mathbf{W}_A[n+1] = \mathbf{W}_A[n] + \mu \mathbf{E}_2 e_1[n] \quad (20)$$

$$\mathbf{W}_B[n+1] = \mathbf{W}_B[n] + \mu \mathbf{E}_1 e_2[n]. \quad (21)$$

3. CENTRED ADAPTIVE FILTERS

In the purely anechoic case, the output of the adaptive filter in Equation (13) will simply be a single peak at a position representing delay and an amplitude representing gain and all other values are assumed to be 0. In reality, with the addition of reverberation and noise there will be a noise floor but there will still be a peak at the delay position. If the delay value is known, it is then possible to update fewer coefficients to get an accurate estimation of gain. Fewer coefficients means faster and more accurate convergence and less computational cost. Only a rough estimation of delay is required as a window of coefficients around the estimated delay value are updated. If the delay estimation is inaccurate by less than the window size then the method will still converge to the solution [10], [11], [12].

As in the LMS adaptive filter, the error is defined as

$$e[n] = x_1[n] - \mathbf{W}^T[n] \mathbf{X}_2[n] \quad (22)$$

and the filter coefficients updated using

$$\mathbf{W}[n+1] = \mathbf{W}[n] + \mu e[n] \mathbf{X}_2[n], \quad (23)$$

but where

$$\mathbf{W}[n] = [w_{\delta-D}[n], \dots, w_{\delta+D}[n]] \quad (24)$$

$$\mathbf{X}_2[n] = [x_2[n-\delta-D], \dots, x_2[n-\delta+D]], \quad (25)$$

and where δ is the estimation of the delay and D is a user-defined error distance around the delay to update the coefficients. A higher value of D will yield slower convergence but will encompass additional echoes or reverberation.

3.1. Delay estimation

There are many delay estimation methods [13] for estimating δ , equivalent to τ_1 and τ_2 in Equations (6) and (7). Adaptive filters themselves can be used to estimate delays [14], but this has the same computational cost that is trying to be avoided. A common method used is the Generalized Cross Correlation (GCC), first introduced in [15]. This method is computationally cheap and allows weightings to be applied to improve performance against noise and reverberation, such as the Phase Transform (PHAT).

The GCC is calculated using

$$\Psi = \mathcal{F}^{-1} \{X_1^*[k] \cdot X_2[k]\}, \quad (26)$$

where Ψ is the GCC, \mathcal{F}^{-1} denotes the Inverse Fast Fourier Transform, $*$ denotes the complex conjugate and X_1 and X_2 are x_1 and x_2 in the frequency domain. By applying the PHAT, Ψ becomes

$$\Psi_P = \mathcal{F}^{-1} \left\{ \frac{X_1^*[k] \cdot X_2[k]}{|X_1^*[k] \cdot X_2[k]|} \right\}, \quad (27)$$

where $|\cdot|$ denotes the absolute magnitude. The estimate of delay δ is then calculated by

$$\delta = \arg \max_n \Psi_P[n]. \quad (28)$$

To reduce computational cost, it is also possible to calculate multiple delays from a single GCC-PHAT calculation [16] by extracting the position of N peaks rather than just 1.

4. CENTRED CTRANC

This paper proposes combining the CTRANC with the centred adaptive filters, known as centred CTRANC, to improve performance and convergence of the CTRANC method. As the CTRANC method the error signals are defined as

$$e_1[n] = d_1[n] - \mathbf{W}_A^T \mathbf{E}_2[n] \quad (29)$$

$$e_2[n] = d_2[n] - \mathbf{W}_B^T \mathbf{E}_1[n], \quad (30)$$

but where

$$\mathbf{W}_A[n] = [w_{\delta_1-D}^A[n], \dots, w_{\delta_1+D}^A[n]] \quad (31)$$

$$\mathbf{W}_B[n] = [w_{\delta_2-D}^B[n], \dots, w_{\delta_2+D}^B[n]] \quad (32)$$

and

$$\mathbf{E}_1[n] = [e_1[n-\delta_1-D], \dots, e_1[n-\delta_1+D]] \quad (33)$$

$$\mathbf{E}_2[n] = [e_2[n-\delta_2-D], \dots, e_2[n-\delta_2+D]] \quad (34)$$

and the filter coefficients are updated using

$$\mathbf{W}_A[n+1] = \mathbf{W}_A[n] + \mu \mathbf{E}_2 e_1[n] \quad (35)$$

$$\mathbf{W}_B[n+1] = \mathbf{W}_B[n] + \mu \mathbf{E}_1 e_2[n], \quad (36)$$

which requires estimation of both δ_1 and δ_2 .

5. DIRECT CALCULATION

The aim of using adaptive filters is to estimate the delay and gain changes of the sources arriving at each microphone, as defined by Equations (6) - (9). It is therefore possible to directly calculate the delay and gain difference for each frame of audio. This information is then used to scale and delay the interference, which is then subtracted from the direct signal, thus removing the interference.

This method is not commonly used as it requires averaging to simply provide a stable solution. For example the amplitude difference calculated for each frame will be slightly different, thus causing amplitude modulation of the interference. The adaptive filters have to converge to a solution which is then stable and will not modulate the signal. This method also does not take into account the crosstalk, relying on a clean interference signal.

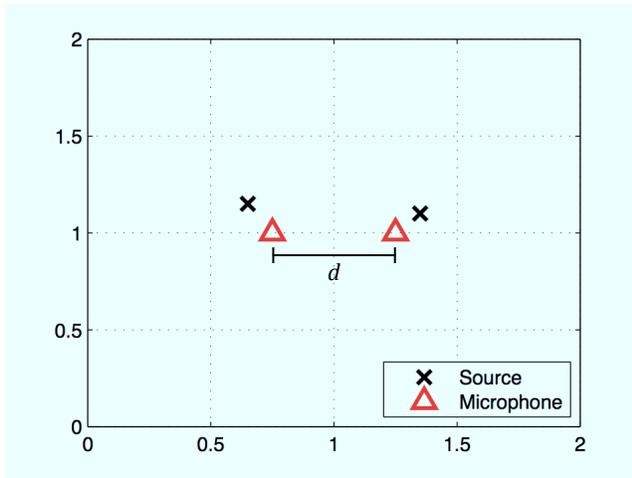


Figure 4: Simulation microphone and source layout where $d = 0.5m$

6. COMPARISON

The methods outlined in this paper were compared by measuring the amount of interference reduction and how artifacts and distortion effect the target sound. The CTRANC and centred CTRANC methods were optimised to produce the best results by selecting a suitable value for the adaption step, μ and the error distance D . The methods were compared in the 2 source, 2 microphone case.

6.1. Simulation Experimentation

The methods were first compared using simulated microphone signals. The sources and microphones were positioned virtually and the equivalent delay and gain calculated for each source to each microphone. The input sources were a guitar and vocal track. The sources were then combined to simulate each microphone signal with bleed. The microphones were assume to be omnidirectional in an anechoic environment. The sources were placed between 10cm and 12cm from the microphones, as shown in Figure 4. The distance d was increased from 10cm to 5m, producing different values for delay and gain. The relative position of each source to each microphone remained the same.

6.2. Results

The simulated microphone outputs containing the direct source and lower amplitude bleed were then passed through each method. The results were analysed using the BSS_EVAL Matlab toolbox [17] to extract the signal-to-interference (SIR_{dB}), signal-to-artefact (SAR_{dB}) and signal-to-distortion (SDR_{dB}) ratios. The unprocessed microphone signals were also analysed for comparison. The results in this paper show the scenario where s_1 is the target signal and s_2 is the interfering signal.

Figure 5 shows the calculated SIR_{dB} for each method at each microphone distance of d . The centred CTRANC can be shown to have the highest values of SIR_{dB} for all but the $d = 0.1$ case, where DUET outperforms it. It is expected that the DUET method may perform well for small values of d as, although it is not aimed at microphones signals, it can perform source separation at small distances. The SIR_{dB} determines how much the interference has been reduced. As mentioned previously, studies have shown that comb filtering can be heard when the duplicate source is as much as 18dB lower in amplitude than the original [1]. It can be seen that the centred CTRANC reduces the level of the interference by more than 18dB for each value of d , therefore the possibility of comb filtering will be removed, even if the interference is not completely cancelled out.

The Wiener filter method [7] proved to outperform the proposed method in certain instances of d for SIR_{dB} but overall performed inconsistently over all values of d in the simulation experiment. The Wiener filter method assumes each microphone is an approximation of the ideal impulse response of the direct sound path and that is the interference is of a lower amplitude. If the interference is of a high enough amplitude, this assumption will no longer hold.

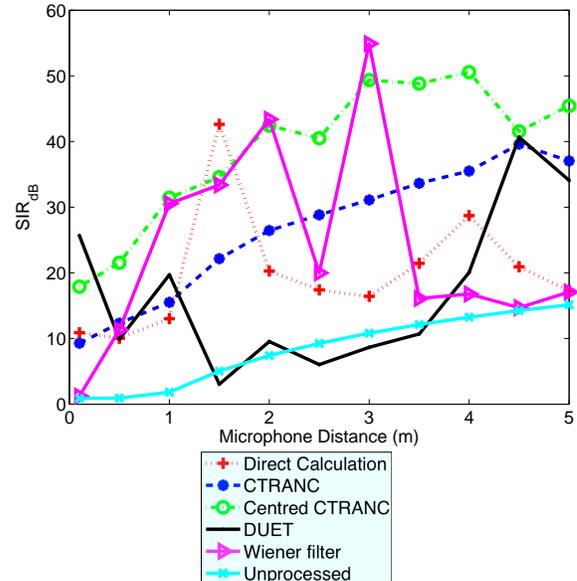


Figure 5: Signal-to-interference ratio of each method at each iteration of microphone distance for the simulated case.

Although DUET performs best on SIR_{dB} at $d = 0.1$ Figure 6 shows the centred CTRANC has a higher value of SAR_{dB} at the same distance. Signal to artefact ratio describes the amount of arte-

facts that have been introduced by a method. This shows that the DUET method introduces a lot of artefacts to the processed signal. Methods based on adaptive filters will generally not add additional artefacts to the target source as it is attempting to subtract the interfering source in the time domain. In live sound, this is desired as it would be preferable to remove some of the interference but leave the target signal intact rather than completely remove the interference but heavily distort the target signal. The results shown in Figure 7 also agree with this.

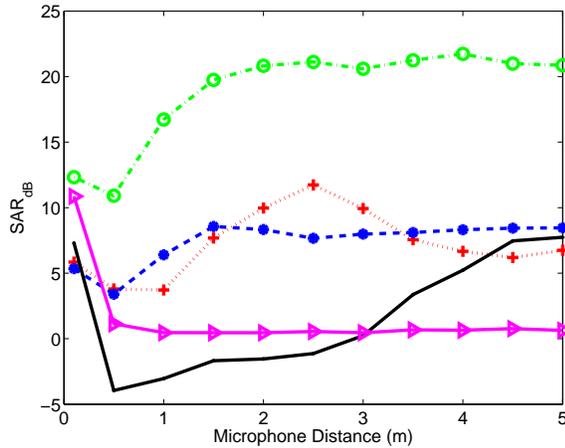


Figure 6: Signal-to-artefact ratio of each method at each iteration of microphone distance for the simulated case.

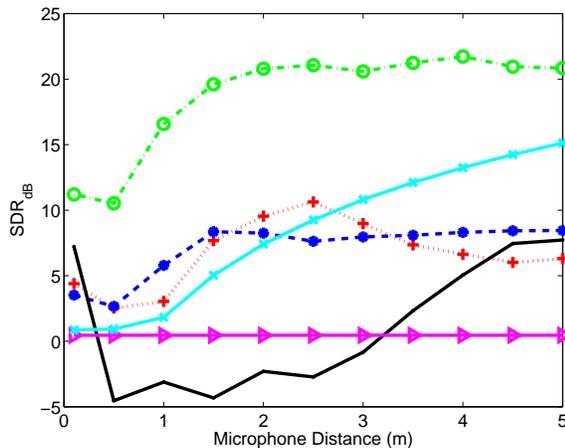


Figure 7: Signal-to-distortion ratio of each method at each iteration of microphone distance for the simulated case.

6.3. Real Recordings

To test each method’s effectiveness in a real space, a test was setup using 2 speakers and 2 microphones. The speakers were spaced from 10cm to 100cm at 10cm intervals while the microphones were always placed 10cm from each speaker, with an error of ± 1 cm as in Figure 8. This distance was chosen to simulate a close microphone configuration. It is not assumed the layout is symmetric.

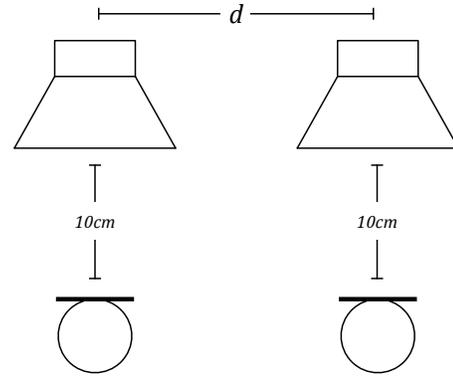


Figure 8: Layout of speakers and microphones in the test recordings

6.4. Results

As with the simulation, the SIR_{dB} , SAD_{dB} and SDR_{dB} for each method and value of d was calculated and can be seen in Figures 9, 10 and 11. Figure 9 shows the method with the highest performance is the original CTRANC method. The reason for this is that due to the excess noise and reverberation, the centred CTRANC would produce errors in the impulse response at the edges of the window but would estimate the amplitude and delay so would improve the SIR_{dB} , as it has higher SIR_{dB} than the unprocessed microphone signals. Using a higher value of D may improve this, but by increasing D the computational cost increases. The Wiener filter method performed only slightly lower than the traditional CTRANC method. Unlike in the simulation experiments, the Wiener filter method performs more consistently with real recordings. The DUET method proved to be more successful at some lengths of d but is not consistent over all the distances tested.

Figures 10 and 11 show the SAR_{dB} and SDR_{dB} for the real test. As shown in the simulations, the DUET method adds additional delay and artefacts. The centred CTRANC overall performs best when measuring SAR_{dB} , agreeing with the simulations that the centred CTRANC does not add artefacts or distortion and is consistent over all values tested of d . In the real recordings the CTRANC has shown to perform consistently well, particularly by the SIR_{dB} measure, but performs worse than the centred CTRANC in measures of SAR_{dB} and SDR_{dB} , thus using the centred CTRANC also reduces the amount of artefacts and distortion over the CTRANC. The Wiener filter method performed worse than the centred CTRANC method but with slightly higher values of SAR_{dB} and SDR_{dB} than the traditional CTRANC. It can therefore be said that for SAR_{dB} and SDR_{dB} these methods had similar performance.

7. CONCLUSIONS AND FUTURE WORK

A centred CTRANC method has been proposed that combines centred adaptive filters with the CTRANC system of noise cancellation. The proposed method outperformed other methods for interference reduction in the simulated anechoic case with little additional artefacts compared to the other methods under test. The

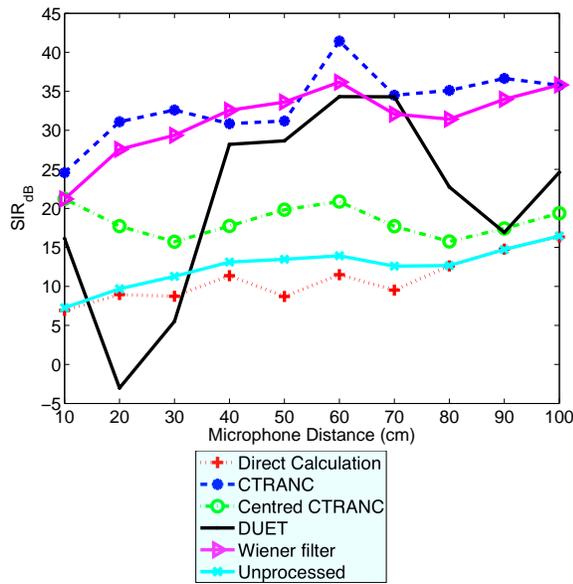


Figure 9: Signal-to-interference ratio of each method at each iteration of microphone distance for the real case.

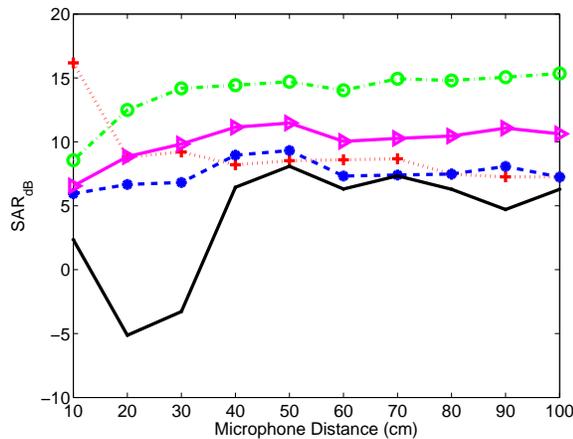


Figure 10: Signal-to-artefact ratio of each method at each iteration of microphone distance for the real case.

proposed method was shown to be outperformed by the simple CTRANC system in real recordings but was shown to also introduce less artefacts.

The efficacy of the centred CTRANC is not effected by the level of the interference but by the environment within which the sources and microphones are placed and the reverberation and noise present therefore it is currently best suited to close microphone applications. Work continues into modifying the centred CTRANC to improve the robustness in real environments, such as manipulating the filter output at each iteration using the estimate of delay by, for example, applying a weighting to the coefficient update range, assuming the correct delay lies close to the centre of the range. The time domain filter can also be assumed to only have positive coefficients therefore the negative components can be set to 0. The problem with this approach is it may effect the convergence prop-

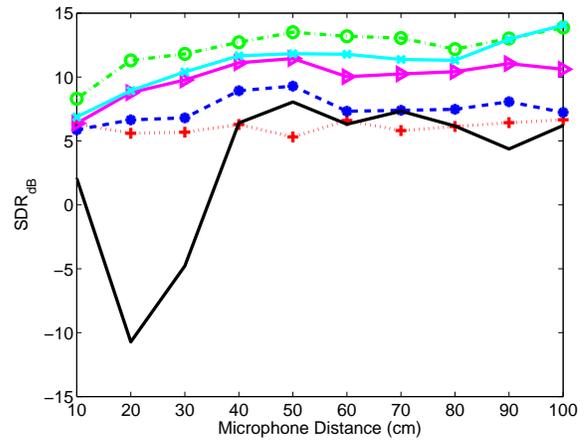


Figure 11: Signal-to-distortion ratio of each method at each iteration of microphone distance for the real case.

erties as the error is now no longer the calculated error.

7.1. CTRANC - N source and N microphones

Theoretically, the CTRANC method can be scaled to N sources and microphones, assuming the number of sources is equal to the number of microphones and therefore the system is homogenous. A diagram of this can be seen in Figure 12. As the number of source and microphones increases, the number of adaptive filters F required increases, with the relation $F = N(N - 1)$. For this reason computational complexity increases as N increases. Work will continue looking at whether centred filters can be implemented in the N source case to improve bleed reduction.

8. REFERENCES

- [1] S. Brunner, H.-J. Maempel, and S. Weinzierl, "On the audibility of comb-filter distortions," in *Proceedings of the 122nd Audio Engineering Society Convention*, (Vienna, Austria), 2007.
- [2] D. Barry, E. Coyle, and B. Lawlor, "Real-time sound source separation: Azimuth discrimination and resynthesis," in *Audio Engineering Society Convention 117*, 10 2004.
- [3] M. Baeck and U. Zölzer, "Real-time implementation of a source separation algorithm," in *Proceedings of the 6th International Conference on Digital Audio Effects (DAFx-03)*, (London, UK), 2003.
- [4] A. Jourjine, S. Rickard, and Ö. Yilmaz, "Blind separation of disjoint orthogonal signals: Demixing n sources from 2 mixtures," in *ICASSP '00: Proceedings of the Acoustics, Speech, and Signal Processing, 2000. on IEEE International Conference*, vol. 5, (Istanbul, Turkey), 2000.
- [5] Ö. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Transactions on Signal Processing*, vol. 52, July 2004.
- [6] G. Mirchandani, J. Zinser, R.L., and J. Evans, "A new adaptive noise cancellation scheme in the presence of crosstalk

- [speech signals],” *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 39, pp. 681–694, Oct. 1992.
- [7] E. K. Kokkinis and J. Mourjopoulos, “Unmixing acoustic sources in real reverberant environments for close-microphone applications,” *J. Audio Eng. Soc.*, vol. 58, no. 11, pp. 907–922, 2010.
- [8] J. Zinser, R., G. Mirchandani, and J. Evans, “Some experimental and theoretical results using a new adaptive filter structure for noise cancellation in the presence of crosstalk,” in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP ’85.*, vol. 10, pp. 1253 – 1256, Apr. 1985.
- [9] L. Lepaulox, P. Scarlart, and C. Marro, “An efficient low-complexity algorithm for crosstalk-resistant adaptive noise canceller,” in *17th European Signal Processing Conference (EUSIPCO 2009)*, 2009.
- [10] V. Margo, D. Etter, N. Carlson, and J. Gross, “Multiple short-length adaptive filters for time-varying echo cancellations,” in *Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on*, vol. 1, pp. 161–164 vol.1, Apr. 1993.
- [11] Y. Lu, R. Fowler, W. Tian, and L. Thompson, “Enhancing echo cancellation via estimation of delay,” *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4159–4168, 2005.
- [12] J. Gross, D. Etter, V. Margo, and N. Carlson, “A block selection adaptive delay filter algorithm for echo cancellation,” in *Circuits and Systems, 1992., Proceedings of the 35th Midwest Symposium on*, pp. 895–898 vol.2, Aug. 1992.
- [13] J. Chen, J. Benesty, and Y. A. Huang, “Time delay estimation in room acoustic environments: An overview,” *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1 – 19, 2006.
- [14] F. Reed, P. Feintuch, and N. Bershard, “Time delay estimation using the lms adaptive filter - static behaviour,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 3, p. 561, 1981.
- [15] C. H. Knapp and G. C. Carter, “Generalized correlation method for estimation of time delay,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.
- [16] A. Clifford and J. Reiss, “Calculating time delays of multiple active sources in live sound,” in *Proceedings of the 129th Audio Engineering Society Convention*, November 2010.
- [17] C. Févotte, R. Gribonval, and E. Vincent, “Bss eval toolbox user guide,” tech. rep., IRISA Technical Report 1706, Rennes, France, April 2005.

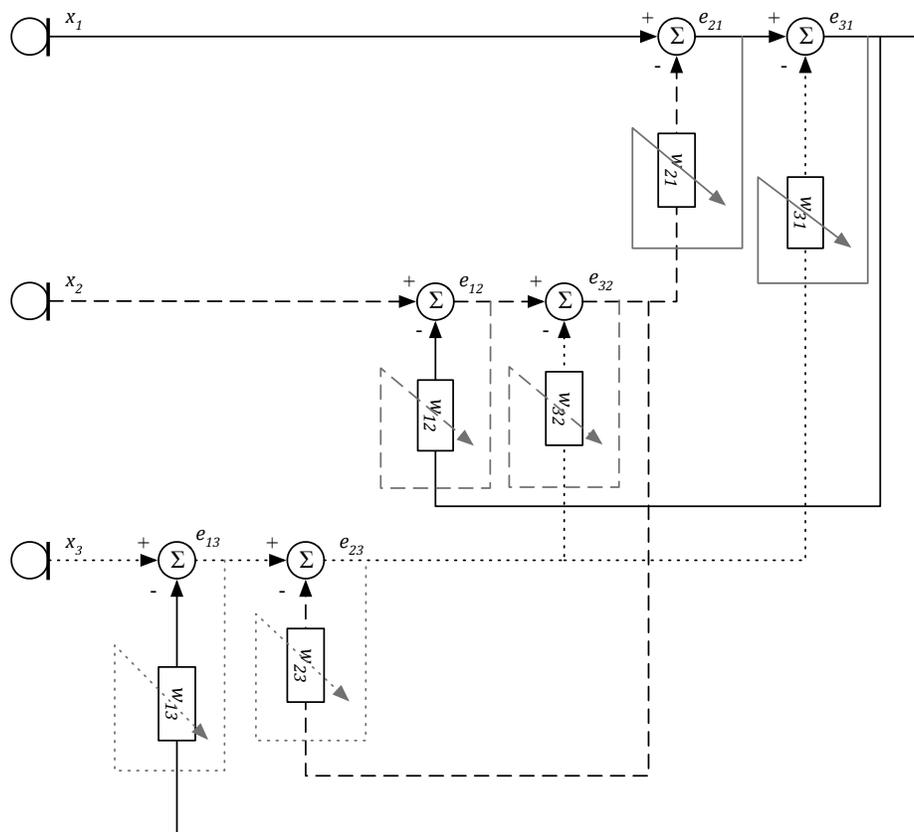


Figure 12: Block diagram of the CTRANC method with 3 microphone inputs and 3 sources.