Application Report
April 2001

GSM Interference Cancellation For Forensic Audio

Philip Harrison and Dr Boaz Rafaely (supervisor) *Institute of Sound and Vibration Research (ISVR)*
University of Southampton, England

ABSTRACT

An increasing problem in the field of forensic audio is the contamination of recordings with interference caused by radio transmissions from GSM mobile phones. Transmitting phones emit short duration radio-frequency pulses at a rate of 217 Hz. The interference pulses contain the fundamental frequency and a large number of harmonics overlapping the frequency range of speech, and therefore severely degrade speech intelligibility. Also, listener fatigue is increased due to the harsh sound of the interference, and overall such audio samples have significantly reduced forensic value. Furthermore, since the majority of recordings submitted for forensic analysis are still on analogue tapes, variations in tape speed produce variations in the interference spectrum, and fixed notch filters will usually not remove the interference for a range of recordings.

The aim of this project is to propose and investigate solutions to this problem through the use of digital signal processing. Single channel adaptive noise cancellation filters are studied, with the aim of removing as much of the interference as possible without adversely degrading the speech. The project studies various methods for cancellation of the GSM interference and presents a real-time implementation of the adaptive system on a TMS320C54, which can be used for instant evaluation of forensic audio recordings.

This document was an entry in the TI DSP Challenge 2000, an annual contest organized by TI to encourage students from around the world to find innovative ways to use DSPs. For more information on the TI DSP Challenge 2000, see TI's World Wide Web site at www.ti.com/sc/dsp_challenge.

Contents

Figures

Tables

Introduction

As the usage of GSM mobile phones increases the number of recordings submitted for forensic analysis that are contaminated with interference caused by transmitting phones has also increased. The transcription of such recordings is problematic because the frequencies present in the interference overlap the frequency range of speech causing masking of the speech. Also listener fatigue is increased due to the harsh sound of the interference.

The interference is caused by the changing electromagnetic field during transmission inducing a varying current in the electronic components of audio equipment. If the interference is induced in the signal path of equipment that is being used to record audio, the interference is added to the audio signal and recorded. The majority of material contaminated with GSM interference is either recordings of telephone conversations conducted on mobile phones or surveillance recordings. Since contaminated recordings submitted for forensic analysis are usually on analogue tape, the spectrum of the recorded interference does not remain constant either within a recording or between recordings, due to the variability of record and playback speeds.

The form of the induced interference is very regular as a result of the method of transmission used by the GSM standard. Time-Division Multiple Access (TDMA) is employed to allow several individual phones to transmit simultaneously to the same receiver on an individually assigned time slot [1]. A phone will transmit for the duration of one time slot, which is 15/26 ms (≈ 0.577 ms) and is called a burst period. The combination of eight burst periods forms a TDMA frame, with a duration of 120/26 ms (≈ 4.615 ms). The transmission of data is based on blocks of 26 TDMA frames. The phone transmits for the first 25 frames and the 26th frame is an idle frame during which the phone does not transmit. The repetitive nature of the interference and the idle frame can clearly been seen in Figure 1. The periodic form of the interference results in a very regular spectrum that contains many harmonics.

Figure 1. Time Domain Form Of GSM Interference Recorded in Absence Of Any Other Signal

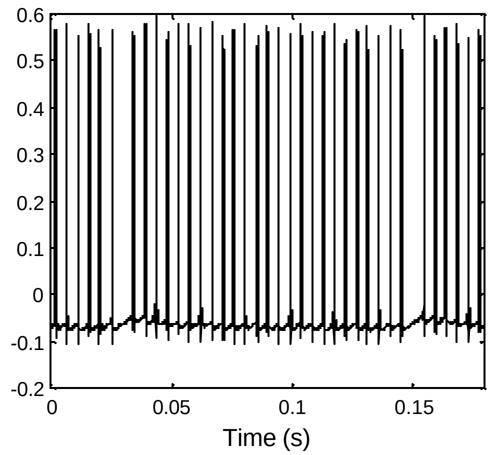


Figure 2. Spectrum Of GSM Interference Recorded in Absence Of Any Other Signal

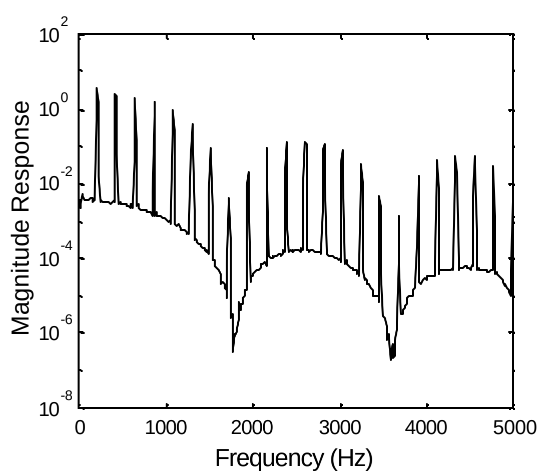


Figure 2 shows the first peak in the power spectral density occurs at 217 Hz. Since the pulses occur every 120/26 ms and frequency is the inverse of time period, the calculated natural frequency of the interference is 26000/120 = 216.667 Hz. The remaining peaks are all harmonics of the fundamental and occur at all integer multiples of 217 Hz.

The interference present in analogue recordings varies in frequency content due to the variation in tape speed. The differences are mainly between recordings made on different machines, but variation also exists within individual recordings. The time domain form of the interference and the relative frequency components are also varied between recordings because of the different internal electronic construction of the audio equipment used.

Partial removal of GSM interference is possible using a notch filter with troughs in the frequency response corresponding to the fundamental frequency of the GSM interference and its harmonics. This results in filtered speech that has a metallic sound due to the deep notches in the filter. The time domain form of the filter causes an inverted frame of interference to appear in the location of the idle frame. The filter also needs to be redesigned for each recording due to the variations in the spectrum of the interference.

A previous investigation into GSM interference in the signal paths of mobile phones [2] studied four methods of GSM interference removal and found that an orthogonal correlator produced the most favorable results. The filter works by correlating the contaminated signal with a set of predetermined functions to find the correct amplitude and phase to generate sinusoids that are subtracted from the contaminated signal to attenuate the interference. The methods evaluated in [2] are based on the induced interference having a constant spectrum and rely on information provided directly from the phone. This results in the methods being unsuitable for use with forensic recordings.

The use of adaptive filtering to remove the GSM noise is studied in this paper. The paper presents the algorithms investigated, Matlab simulations and a real-time implementation on the Texas Instruments TMS320C54.

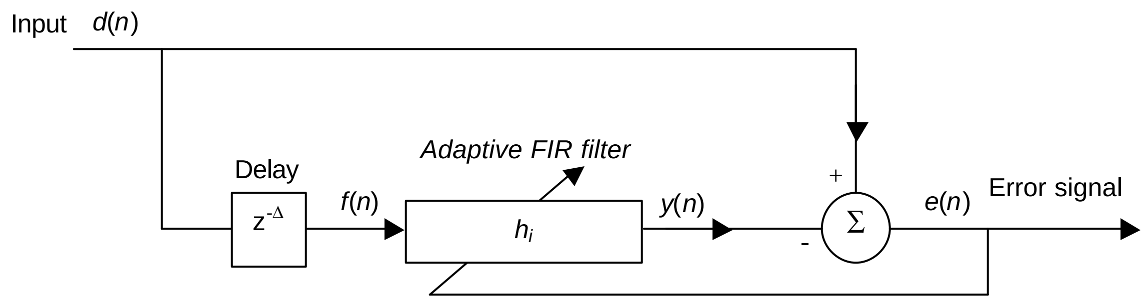
The Algorithm

The capability of adaptive filters to change their time and frequency domain characteristics in response to incoming data make them appear suited to deal with the differences present between recordings of interference. The types of adaptive filters suited to the task of interference removal are adaptive noise canceling filters. The basic architecture of adaptive noise canceling filters relies on the presence of a reference signal that only contains the noise that is going to be cancelled. The reference signal is filtered by an FIR filter and then subtracted from the contaminated input signal producing an error signal. For every signal there exists an optimal set of filter coefficients that will minimize the error signal and achieve the best possible cancellation of the interfering noise. Calculation of the optimal coefficients is a computationally expensive process. Alternatively the optimal coefficients can be approximated by adapting a set of coefficients using the ‘Least Mean Squares’ (LMS) algorithm [3].

Since there is no reference signal available from forensic recordings containing the GSM interference, the standard implementation is changed so that the reference signal is obtained by delaying the main input signal. This is called a single channel adaptive line enhancer [3] and is illustrated in

Figure 3. This filter arrangement is best suited to the task of separating periodic signals from broadband signals.

Figure 3. Diagram Of Single Channel Adaptive Line Enhancer



The implementation consists of the input signal *d(n)*, which is delayed to form the reference input signal *f(n)*. The length of delay Δ is chosen so that the broadband signal (speech) components become decorrelated between the input and the adaptive filter. Due to their periodic nature the periodic signals (GSM interference) remain correlated. The delayed input signal is then convolved with the filter coefficients h_i to give $y(n)$ which should closely match the periodic signal. The filter length should be long enough to allow successful attenuation of the non-periodic signal components. The filter output is then subtracted from the input signal to give the error signal *e(n)*, that should contain only the broadband signal. If the output of the system were to be taken from the output of the adaptive filter the signal should contain only the periodic signal. The filter coefficients are then updated using the LMS algorithm [3], which for all i is:

$$h_i(n+1) = h_i(n) + m e(n) f(n-i) \qquad (1)$$

where m is the convergence coefficient. The convergence coefficient determines what size steps are taken towards the optimal filter coefficients. It effectively determines how fast the coefficients converge to the approximate optimal coefficients. If the coefficient is too large the coefficients will overshoot and diverge or go unstable.

Matlab Simulations

A simulation of the single channel adaptive line enhancer shown in

Figure 3 was written for Matlab [4] with an effective sampling frequency of 10 kHz, since the resulting 5 kHz bandwidth includes the majority of speech frequencies. An artificial form of interference was generated in Matlab to allow testing of the implementation without the variability present in recordings of real interference.

Testing of the implementation with artificial GSM interference revealed that the greatest attenuation of the interference is achieved with a delay just less than the period of the occurrence of the idle frame (i.e. 1190 samples) and a filter length of less than the period of the interference pulses (i.e. 40 samples), see Table 1, Configuration 1.

The delay and the FIR filter align the interference in the reference signal with the interference in the input signal so that the idle frames coincide. When the reference signal is subtracted from the input signal, the interference pulses cancel. The FIR filter operates as a delay and the coefficients converge to a spike in the time domain.

Table 1. Filter Configurations

Configuration	Delay (samples)	Filter Length (Samples)
1	1190	40
2	40	400

Figure 4. Configuration 1 - Resulting Error Signal For Artificial Interference

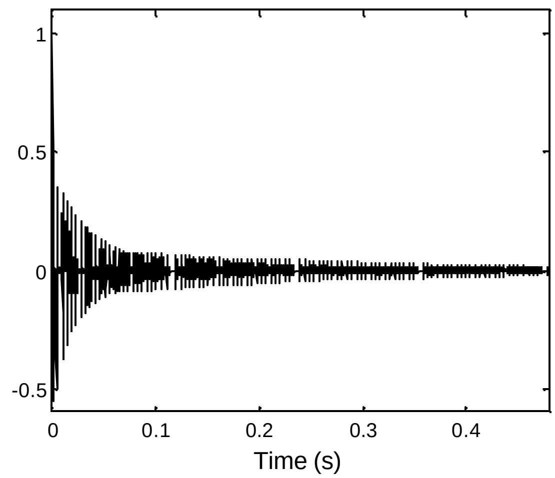
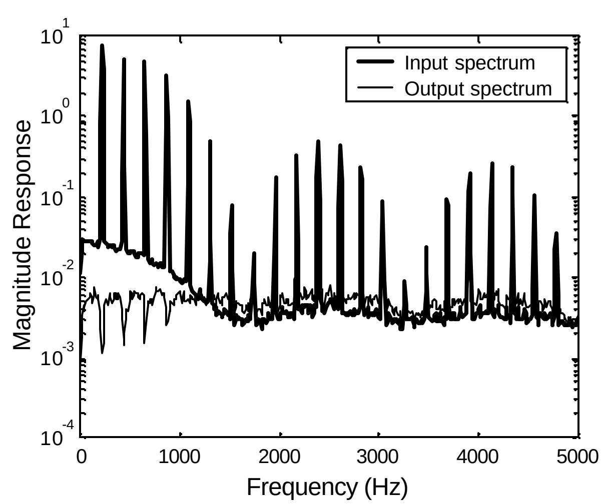


Figure 4 demonstrates the reduction in the error signal as the filter adapts to an input of artificial interference. When the coefficients have converged the interference signal is almost completely cancelled.

Figure 5 shows the performance of the configuration in the frequency domain when filtering artificial interference in white noise. The configuration achieves good attenuation of the interference in both the time and the frequency domain.

Figure 5. Configuration 1 - Input/Output Spectrum For Artificial Interference In White Noise



When filtering speech contaminated with artificial interference, the simulation successfully attenuates the artificial interference and improves the intelligibility of the speech. Due to the configuration of the filter an echo of speech with a similar amplitude to the original occurs 0.1196 seconds after the original in the output signal. This slightly reduces the level of intelligibility achieved by the cancellation of the interference.

In an attempt to increase the intelligibility of the filtered speech simulations were run with shorter filter lengths. Reducing the filter length causes the echo to occur closer to the original speech until the delay length is so short that the echo is perceived as part of the original speech signal. Increasing the filter length causes the FIR filter to act as a notch filter that attenuates the speech in the reference signal, resulting in a higher level of speech in the error signal. The best increase in intelligibility due to the parameter changes is achieved with a filter length of approximately 400 samples and a delay length less than the period of the interference pulses (i.e. 40 samples), see Table 1, Configuration 2.

Figure 6. Configuration 2 - Resulting Error Signal For Artificial Interference

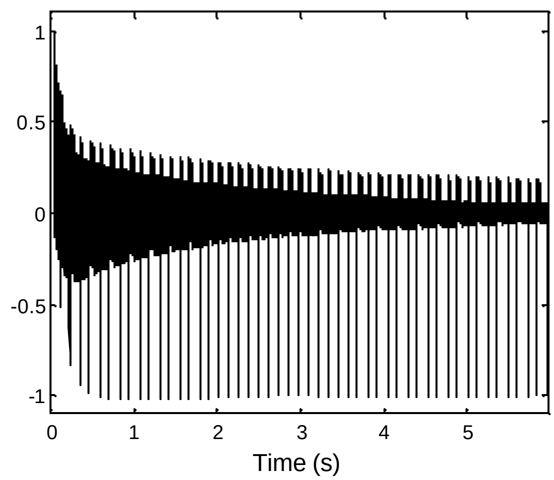


Figure 6 shows the reduction in the error signal as the filter adapts to an input of artificial interference. The filter adapts more slowly than configuration 1 because an inverted pulse of interference is generated in the error signal, which disrupts the convergence. The inverted pulse is a result of the idle frames in the reference signal and the input signal being unaligned. The filter attempts to remove a pulse of interference from the location of the idle frame in input signal and causes the inverted pulse. Even when converged, the filter is unable to adapt to the inverted pulse. However when filtering speech, configuration 2 provides greater intelligibility than configuration 1, even though the inverted pulse causes a regular clicking sound.

The testing of the two configurations with recordings of speech contaminated with real GSM interference generally produced good attenuation of the interference and an increase in intelligibility of the speech. Results of filtering digital recordings of contaminated speech produced results very similar to those achieved with the artificial interference. The configurations also produced good results when used when used on a recording made on a TDK SA-90 standard analogue compact audio cassette (CAC).

Figure 7. Configuration 2 - Input/Output Spectrum for CAC Recording

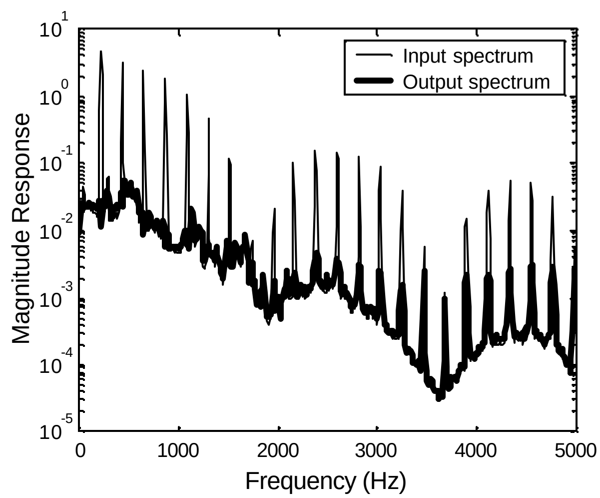


Figure 7 shows the attenuation produced by configuration 2 when filtering a CAC recording of contaminated speech. The slight speed variations present in the CAC recording resulted in a small reduction of performance in comparison with the digital recordings. The larger speed variations present in micro-cassette recordings resulted in a big reduction in performance when the configurations were tested with a recording on a TDK MC-60 micro-cassette. The filter is less able to adapt to the larger speed variations and the cancellation of the interference pulses is less accurate.

Real-Time Implementation

The implementation of the single channel adaptive filter on the fixed-point C5402 DSK [5] consists of two pieces of code. The main C code initializes the board and passes data between the processor, the codec and the assembler function. The assembler function performs the adaptive filtering and consists of two main sections. The first performs the adaptive filtering while the second section removes a DC component from the FIR filter coefficients. DC bias can become present in the FIR filter coefficients and cause the filter to saturate. The DC bias can originate from DC in the input or from finite precision errors. The DC offset is removed by calculating the mean value of the coefficients and subtracting it from each coefficient.

The real-time implementation was tested with the two preferred filter configurations from the Matlab simulations. At the chosen sampling rate of the codec, 9142 Hz, configuration 1 requires an FIR filter length of 32 samples (0.0035 seconds) and a delay of 1090 samples (0.1192 seconds). Configuration 2 requires an FIR filter length of 366 samples. Since the real-time implementation is restricted by the operating speed of the processor a filter length of 64 samples (0.007 seconds) was used with a delay of 35 samples (0.0038 seconds).

Figure 8. Configuration 2 - Input/Output Spectrum for CAC Recording

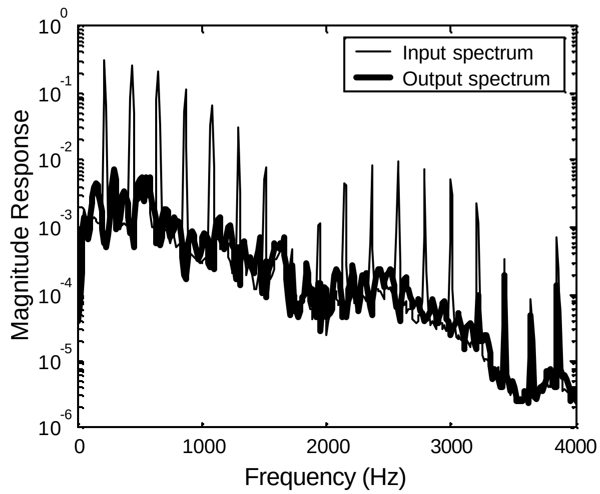


Figure 8 shows the performance achieved in the frequency domain by the real-time implementation of configuration 2 when filtering the TDK SA-90 CAC recording of contaminated speech. Even though the filter length has been reduced the configuration still provides good attenuation of the interference.

The real-time implementations of the two configurations produced results similar to those in the Matlab simulations for all the recordings tested. Overall the quality of the real-time results was less then for the simulations. One of the reasons for this is that Matlab uses floating point arithmetic while the DSP used in the real-time implementation uses fixed point arithmetic. This will cause the real-time implementation to adapt to coefficients that are not as accurate as those which could be achieved in the floating-point arithmetic of Matlab. The real-time implementation of configuration 2 resulted in a reduction in performance due to the necessary shortening of the filter length.

Conclusions

A study was conducted in this paper into GSM interference and the current methods available for removing the interference. A single channel adaptive line enhancer was simulated in Matlab and two configurations were found that satisfactorily attenuated the interference and increased the intelligibility of the contaminated speech. The adaptive filter was then implemented in real-time and achieved similar results to the simulations.

Filter configuration 1 produced excellent attenuation of the interference, achieving almost perfect cancellation and results in an increase in the intelligibility of the contaminated speech. As a consequence of the configuration a speech echo is introduced in the error signal, which reduces the potential intelligibility. The second configuration provides greater speech intelligibility than the first configuration, by not producing an echo, whilst achieving less attenuation of the interference.

The filter configurations achieved successful attenuation of the interference for a range of recordings made on both digital and analogue recording equipment. The inherent speed instability with micro-cassette recordings severely reduced performance of the configurations when filtering such material. The real-time implementation of the adaptive filter provides an instant evaluation of the two filter configurations and achieves levels of performance similar to the simulations. Even greater real-time performance could be achieved with highly optimized code.

Further work that could be carried out would be the investigation of algorithms with faster convergence than the LMS. Such algorithms may be able to adapt to the speed variations of microcassette recordings and expand the range of applicability of the filter. This could cause a more accurate cancellation of the interference and improve the results obtained from recordings of interference on CACs. Algorithms with a convergence faster than the LMS that could be used are the frequency domain LMS and the recursive least squares (RLS).

The development of an efficient and robust method of removing the inverted interference pulse in configuration 2 could increase the performance of the filter without a significant increase in computation.

References

[1] Scourias, J, A Brief Overview of GSM, http://kbs.cs.tu-berlin.de/~jutta/gsm/js-intro.html

[2] Rosengren, P and Nilsson, A, Bumblebee Killer (Thesis: MME 99-07), University of Karlskrona/Ronneby, 1999

[3] Widrow, B and Stearns, S, Adaptive Signal Processing, Prentice-Hall, 1985

[4] The Mathworks, Inc, *Matlab Version 5.3*, 1999

[5] Texas Instruments, *Code Composer Studio DSK Version 1.21*, 2000