# The Role of Adaptive Filters and other audio processes in Audio Surveillance

**Gordon Reid**
Managing Director

**CEDAR Forensic**
20 Home End, Fulbourn, Cambridge CB1 5BS

# Abstract


- This tutorial discusses the increasing need for filtering systems that allow surveillance specialists to combat the effects of noise in non-ideal recording environments.

- Demonstrations of single-channel and cross-channel adaptive filters together with other noise reduction and signal enhancing techniques will be given using the CEDAR Cambridge Forensic System.

# The background to audio forensics

- The use of audio recordings of telephone conversations and interviews, as well as covert surveillance recordings, is an integral part of law enforcement.
- Such recordings are frequently of poor quality, but they are admissible if they are intelligible and meet the rules of evidence.
- It can be equally important that the recording is listenable, and that the information it contains is easily discerned by the jury.
- Written transcripts can be vital evidence, and these can only be made with confidence if the recording is intelligible. Listener fatigue and error rates will be significantly reduced if the recording can be made more listenable.

# Acquiring the recording

- Due to exposure of undercover techniques on TV, body recording has become a risky venture.
- Rub downs are common so reliance is now placed on smaller recorders, with a trade-off in recording quality.
- The use of probes is acceptable within strict guidelines. However, you normally have just one chance to install a probe and you are then at the mercy of whatever is happening in its vicinity.
- The average home, workplace or vehicle is a generator of acoustic and electrical noises, many of which are of a continuous and repetitive nature.

CEDAR

# Types of noise

- air conditioning noise
- excessive reverberation and echoes
- wind and rain
- road vehicles and aircraft
- engine noise
- domestic appliances
- other conversations
- radio, TV and live music
- induced hum and/or buzz from lighting and other sources
- interference from nearby transmitters such as mobile telephones
- faulty microphones and recording equipment
- noise inherent to the recording medium

# Why eliminate noise? (Transcription)

- Noise and unwanted sounds may lead to listener fatigue, reducing transcription accuracy. Reducing or eliminating the noises that lead to fatigue aids in obtaining faster, more accurate transcriptions.
- Listener fatigue generally takes three forms:
  - The ear becomes temporarily insensitive to the frequency bands containing continuous noise, so the listener may miss important signal components.
  - When presented with loud, unwanted sounds, the listener may become distracted and lose concentration, leading to signals being missed.
  - The listener becomes desensitised to the unwanted sounds and anything of similar nature. This may lead to transcription errors.

# Why eliminate noise? (Presentation)

- In general, court rooms suffer from unfavourable acoustics and lack sophisticated audio replay systems, so it is essential that audio prepared for presentation is as clean and unambiguous as possible.
- Untrained listeners may be confused by the presence of noise on recordings, limiting their ability to make correct judgements concerning the evidence that is being presented to them.
- Noise reduction allows you to prepare the evidence to minimise confusion without altering the nature of the wanted signal.

CEDAR

# Why ordinary filtering is inappropriate

- The cleaning of audio signals for forensic investigation imposes very different requirements from those encountered when restoring audio for CD, DVD and broadcast.
- In the former case, intelligibility is the overriding criterion. In the latter, listenability is paramount. This means that the tools are different for each.
- Some security forces and agencies attempt to improve recordings using filters and processors borrowed from the mainstream audio industry.
  - These can give satisfactory results if the audio is not severely degraded.
  - They will be inadequate for some of the types of noise and interference discussed within this presentation.

CEDAR

# What are adaptive filters?

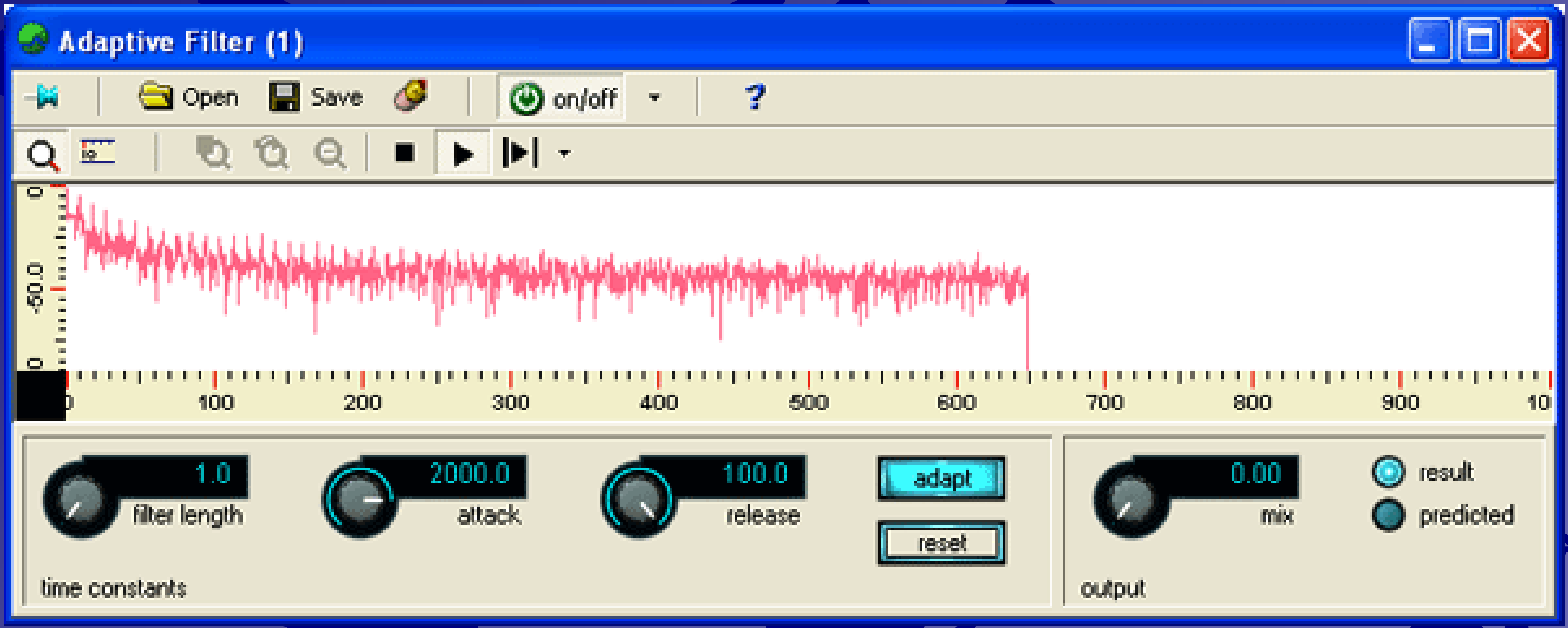
- If the characteristics of the noise are statistically constant it is possible to design a static filter that optimally separates the speech from the noise. Unfortunately, there are many circumstances when this is difficult or impossible. For example:
  - when the noise has a complicated spectrum
  - when the noise statistics are varying rapidly
- In these cases, we require a filter that can adjust itself in accordance with the varying signal characteristics. These are Adaptive Filters.

# Single channel adaptive filters

- Steady-state components are in general more predictable than changing content such as speech, so it is possible to identify these and filter them from the input.
- Single channel adaptive filters work best on low-frequency repetitive noises such as electrical and vehicle motors, wind noise, hums and low frequency background speech (babble) in public places. They are good for removing reverberation and echoes from recordings made in hard-walled environments such as prison cells.
- This filter may offer control of the filter length, and its coefficients may be displayed graphically. It will adapt quickly to the onset of new tones, but not as quickly when tones disappear from the signal. There may be Attack and Release controls to fine-tune this behaviour.

CEDAR

# The CEDAR 1-ch adaptive filter

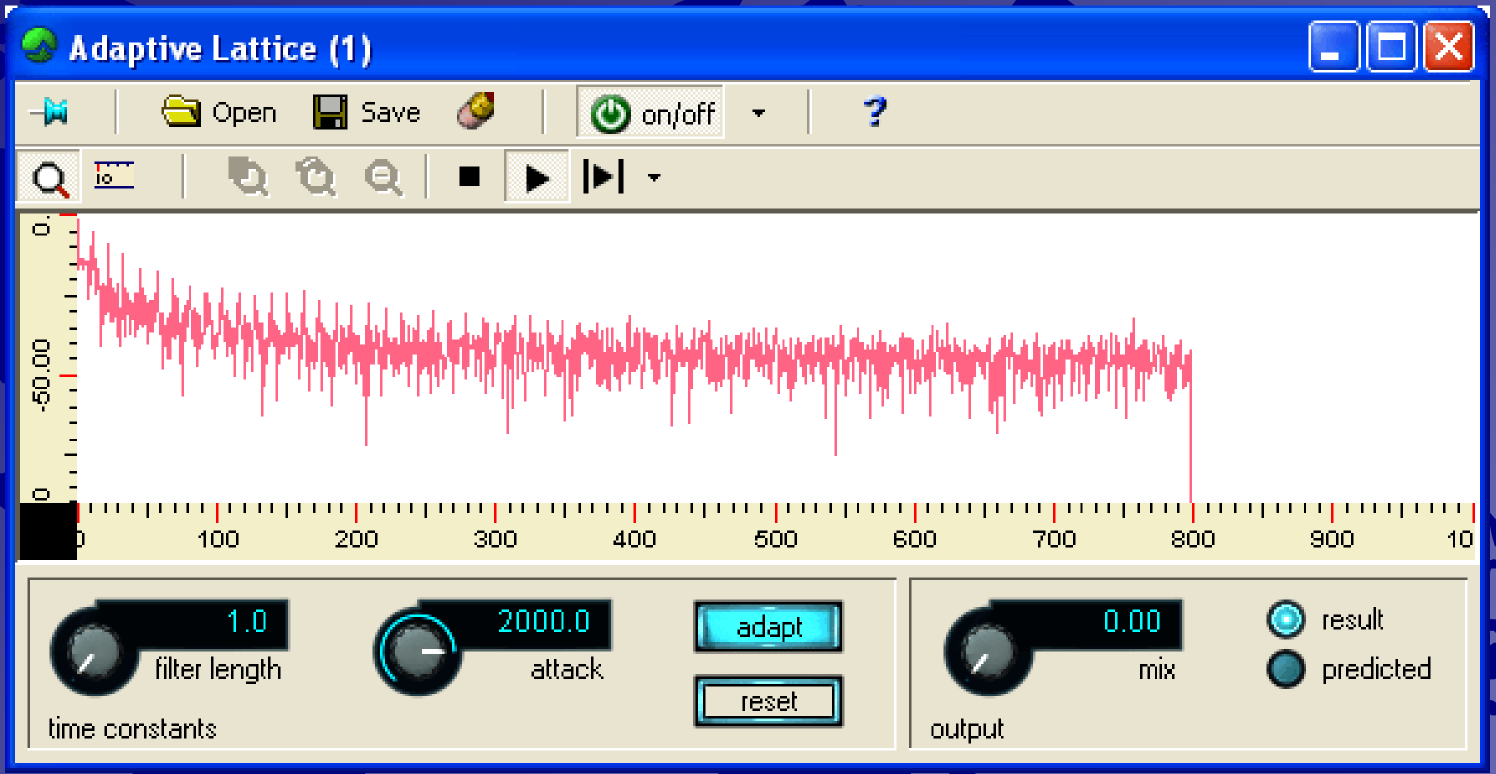


# The Lattice Filter

- A more powerful algorithm capable of digging deeper into a noisy recording, extracting more fricative sounds and consonants.
- It adapts more rapidly than the standard adaptive filter to new tones appearing in, and disappearing from, the input signal.
- With the exception of the Release control present on the standard adaptive filter, the operation of the CEDAR versions are all but identical.

CEDAR

# The CEDAR 1-ch lattice filter




CEDAR

# Using single channel adaptive filters

- Use for:
  - Single-channel recordings suffering from interference from signals such as engine noise and/or excessive broadband noise.
- Do not use for:
  - Impulsive noises such as clicks and crackle, for simple buzzes or for mild dehissing, for which other processes will be more suitable.
- Preparation:
  - It is useful to reduce the processing bandwidth to best match that of the wanted signal. This reduces the processing power required for any given filter length, and increases the flexibility and usefulness of the filter.

# Demonstration of a 1-ch lattice filter

- Ensure that "Adapt" is ON, set the output mode to "result" and the Mix to 1.00 so that you hear the fully processed output from the filter.
- With the Attack set to 2,000ms, increase the filter length. In many cases, the graph will show a cusp between a gradient at low filter lengths and a horizontal profile at medium and high lengths. Adjust the length so that it lies on this cusp.
- Reduce the Attack. In many cases, you will find a setting where the noise is being significantly attenuated and the intelligibility is improved.
  - If you reduce the Attack too far, you will hear greater noise attenuation, but the intelligibility will be reduced
  - For many recordings, the optimum Attack lies around 100ms, which means that you are retaining sounds whose characteristics change on timescales of less than 100ms, and rejecting sounds that are stationary on scales greater than 100ms.

CEDAR

# Fine tuning the 1-ch filters

- Experiment with changes in the filter length and Attack until you are removing the most noise while retaining the most speech possible.
- Switch the filter on and off regularly. This will help you to judge the suitability of the values you have chosen.
- If there is a sudden change in the noise, pressing RESET will reinitialise the internal coefficients, and the filter will adapt to the new noise profile.
- If the nature of the noise is changing over timescales of a few seconds, reducing the Release time in the Adaptive Filter allows it to 'unlearn' its coefficients and re-adapt to the changing shape of the noise.
  - This will help when processing audio contaminated by, for example, vehicle engine noise, which can change markedly during the recording.

CEDAR

# Using the 1-ch output Modes

- In the CEDAR implementation, there are two output modes provided: the predicted signal, and the filtered signal. You can vary the output mix between the original audio and either of these to maximise intelligibility.
- You may find that optimum noise removal is obtained at the expense of the vowel sounds in the speech. In this case, you may improve matters by adding back some of the original signal. Reduce the Mix from 100% to a figure that improves the intelligibility without reintroducing excessive noise.
- On rare occasions when the noise is extremely non-stationary, you may find it more useful to retain the steady state components (vowel sounds) and reject the more rapidly varying signal components. To do this, switch the output mode from "Result" to "Predicted".
  - You may lose the plosives and fricatives in the speech, which will also have adverse effects on intelligibility.

CEDAR

# Notes for 1-ch adaptive filters

- The standard adaptive filter removes strong noise components first and then adapts to lower amplitude noise. The lattice filter tends to adapt to all noises at the same rate.
- The lattice filter removes more noise, particularly when the noise content is varying quickly or by a large amount. However, it is more prone to removing the tonal content of speech.
- Exceeding the optimum filter length can increase the amount of noise passed by the filter and may introduce echo artefacts.
- You must decide which of the filters offers the best result on a job-by-job basis.

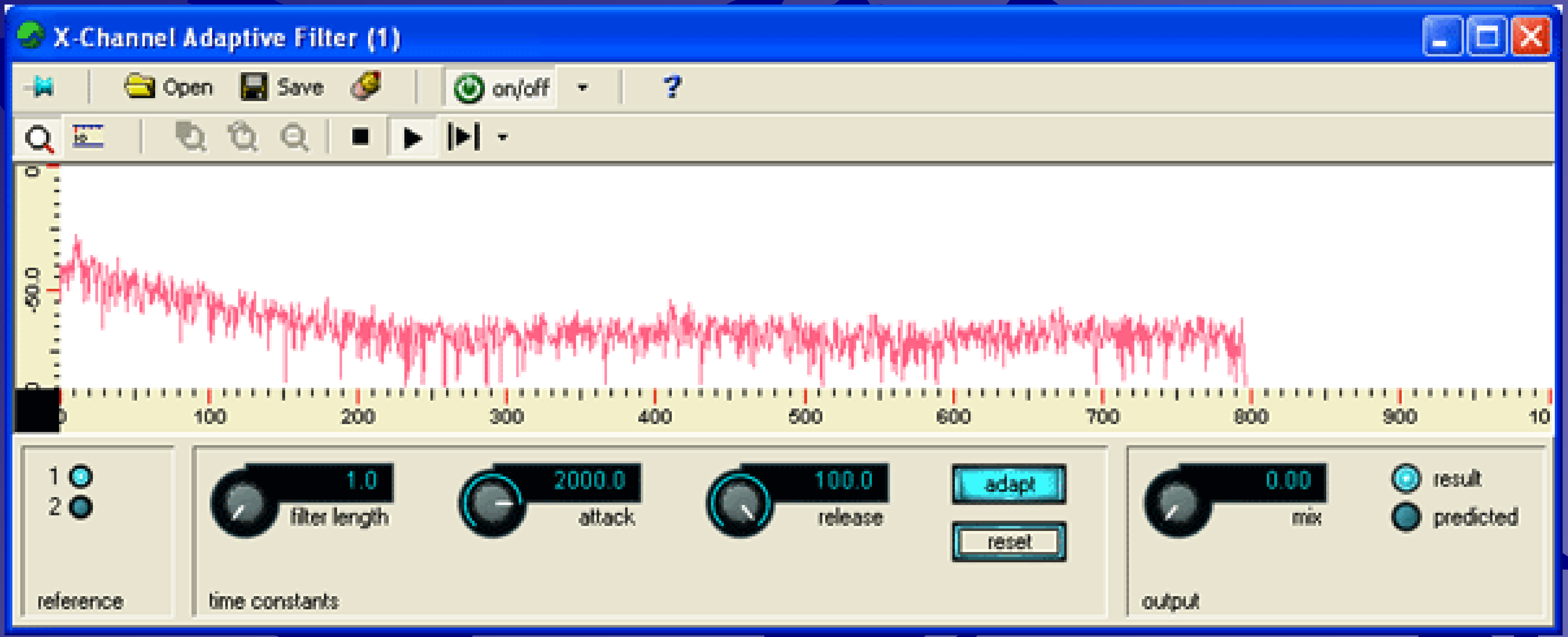
CEDAR

# Cross channel adaptive filters

- Villains are aware of surveillance and take steps to avoid it. A typical scenario involves a group of people turning on the TV to mask their speech.
- The Police can anticipate this and set up audio surveillance appropriately. This will entail putting a probe/bug in the room and recording its output to one channel of a recorder. A second recording will create a reference channel, using a receiver to record the interfering noise “off air”.
- It may seem that simple subtraction of the reference channel from the surveillance signal will eliminate the obscuring noise. This is not the case; the directly recorded reference signal is very different from the obscuring noise in the surveillance signal.
- Cross-channel adaptive filters determine which elements of the surveillance recording are due to the content within the reference, and which elements do not, and are therefore deemed to be the wanted signal.

CEDAR

# The CEDAR X-channel adaptive filter



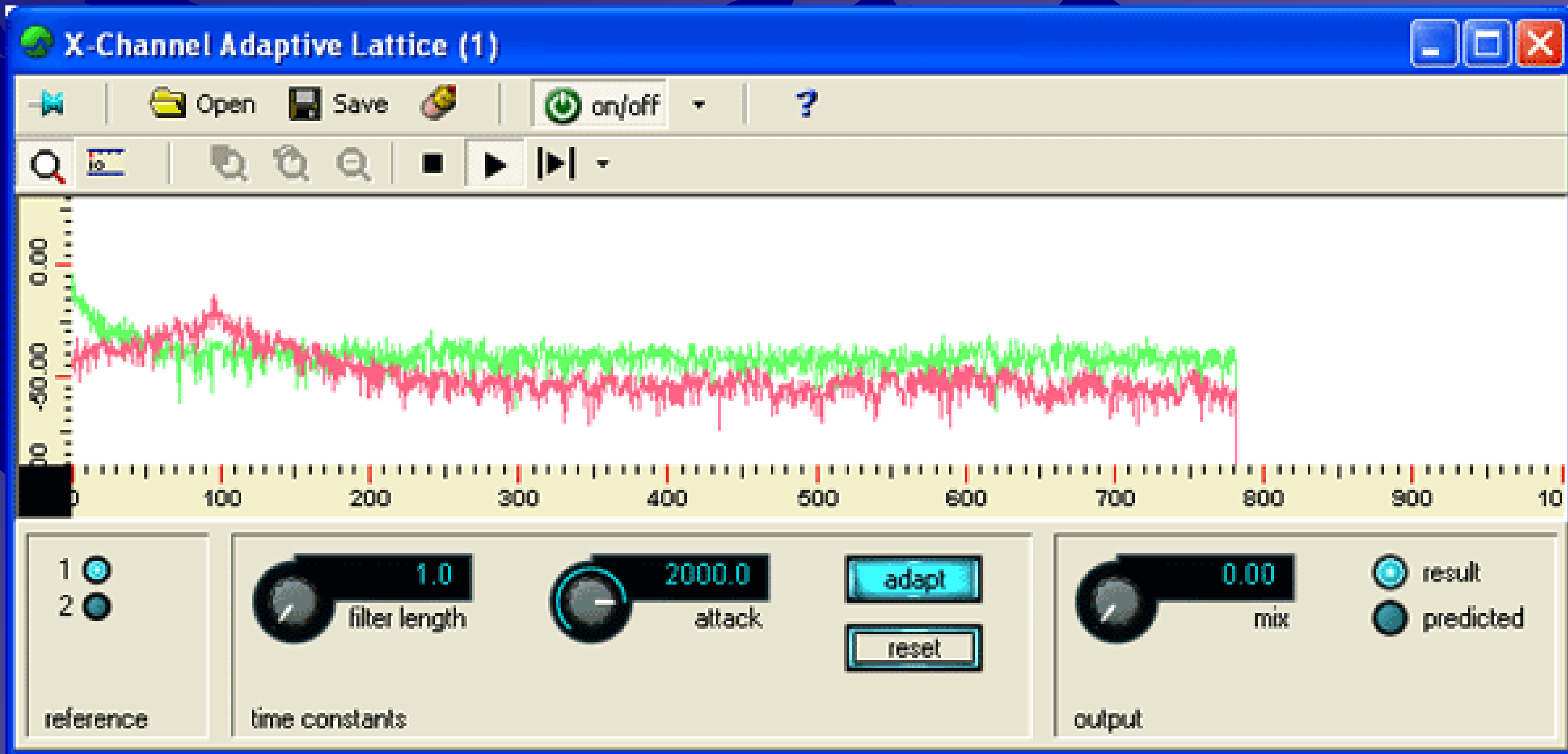

CEDAR

# The cross channel lattice filter

- This is a more powerful algorithm capable of digging deeper and extracting more of the reference channel from the surveillance recording.
- The lattice filter is unusual in that it shows coefficients for the reference channel AND the filtered channel, and the optimisation of both of these obtains the best results.
- Experience shows that the optimum filter length is slightly shorter when using the cross channel lattice filter than it is when using the cross channel adaptive filter.

CEDAR

# The CEDAR X-channel lattice filter

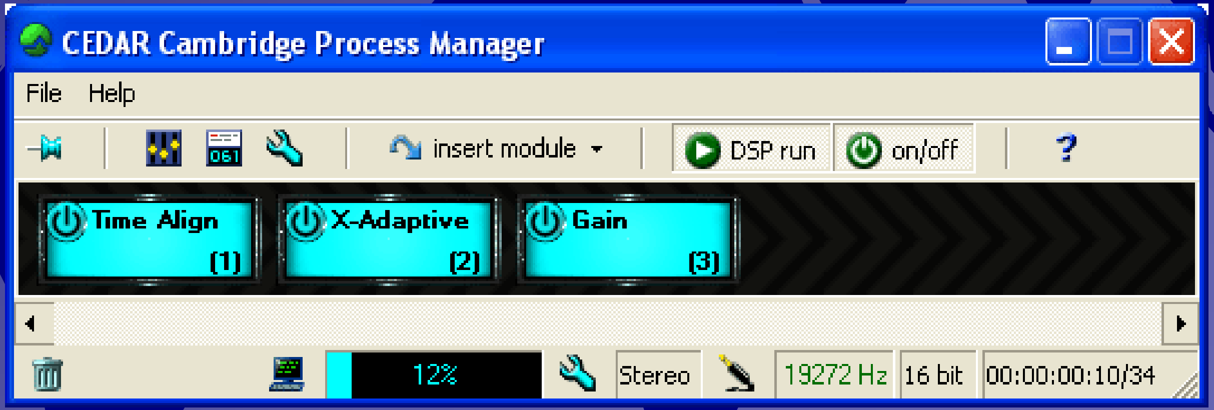


# Using cross channel adaptive filters

- Use for:
  - Recordings suffering from interference from complex signals such as recorded music, radio or TV broadcasts, and for which a suitable reference recording can be obtained.
- Do not use for:
  - Any other noises.
- Preparation:
  - Having obtained the reference recording, you must align it with the surveillance recording.

# Setting up a processing chain

- Optimise the processing bandwidth by using the lowest sample rate that retains the audio bandwidth of the wanted signal.
- In the Process Manager, insert a Time Align module, followed by the cross channel adaptive filter, followed by a Gain module. Use the "solo to all" facility to mute the reference channel and direct the wanted audio to all outputs.

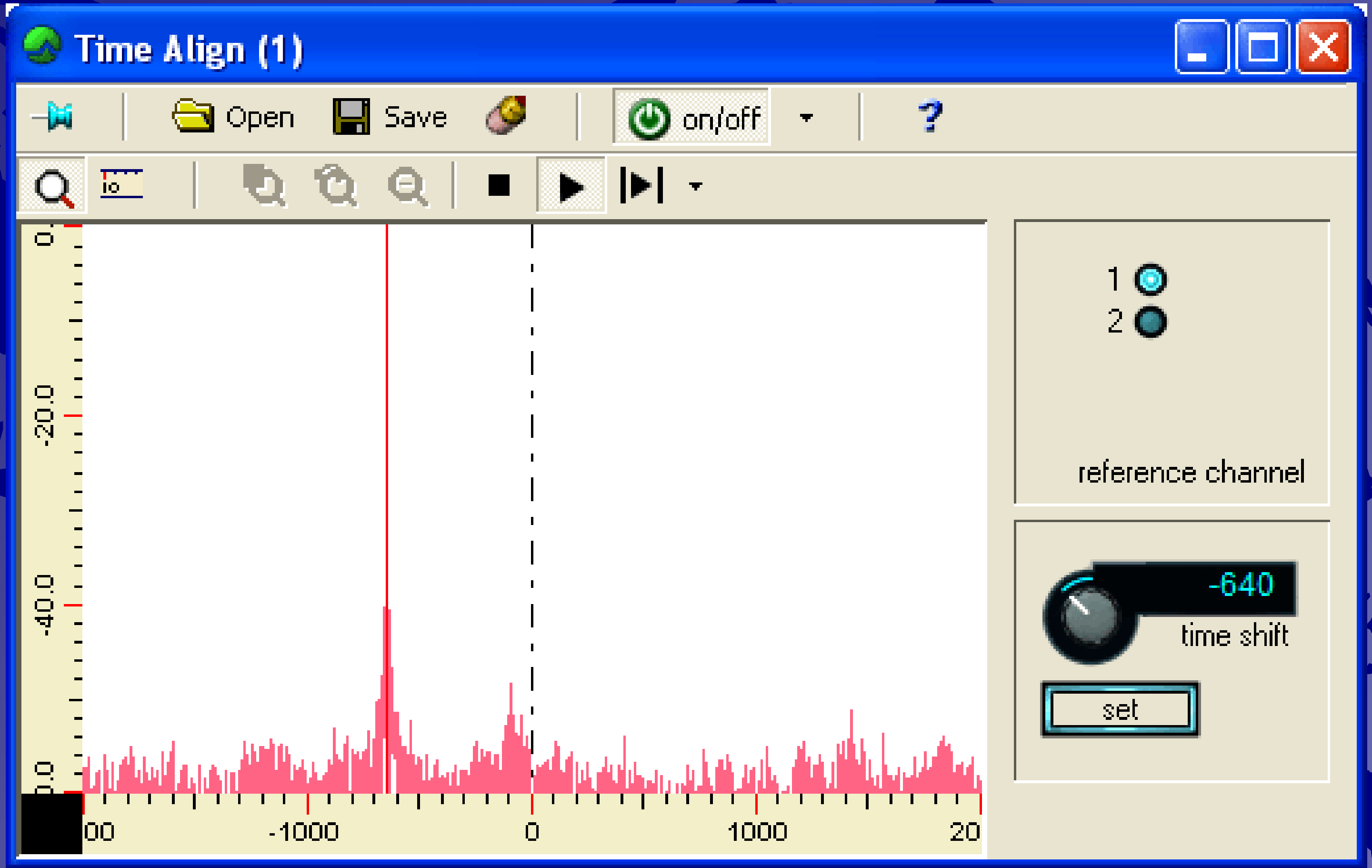


# Time Alignment

- It is important that the surveillance and reference channels are correctly aligned. If they are not, you will be unable to extract the speech with optimal intelligibility.
- A prominent peak in the time align display is a good indication that cross-channel adaptive filters can be used successfully, but if the highest peak is not particularly prominent, the filters are less likely to remove all of the unwanted sounds.
- If there is no correlation, cross-channel adaptive filters are unsuitable.

CEDAR

# The CEDAR time align module




CEDAR

# Demonstration of a X-ch adaptive filter

- Return to the cross-channel adaptive filter and select the reference channel.
- Set the output mode to "result”, the Mix to 1.00 and ensure that "Adapt" is ON.
- A good starting point is: Release set to 100s; Attack set to 2,000ms; filter length of 100.
- Play the audio and reduce the Attack to the point where you obtain maximum attenuation of the unwanted signal without unacceptable degradation.
  - If the filter length is too low, the amount of attenuation is reduced. If the filter length is too long, you may introduce echo artefacts.

CEDAR

# Fine tuning the X-ch adaptive filter

- If a static probe is being used, it is not beneficial to adjust the Release because the transfer function that relates the reference audio to the obscuring noise is largely constant.
- If a body wire is being worn, or if the relationship between the reference and the noise is otherwise changing, it may be helpful to reduce the Release so that the filter can adapt to the changing environment – for example, as the Officer moves around.

# Using the X-ch output modes

- It should not be necessary to change the output mode or move the Mix away from 100%.
- Unlike dual-channel adaptive filter systems, CEDAR Cambridge offers up to eight audio channels, allowing you to use the reference signal to clean up to seven other signals simultaneously.

# Using Adaptive Filters with other filters

- Clicks, crackle and thumps will cause the filter to readapt in unwanted ways.
  - For these problems, using audio restoration tools before the adaptive filter can yield excellent results.
- Adaptive filters are unsuitable for other noises including some forms of buzz, and broadband noise (hiss).
  - Users should also reduce or eliminate these prior to presenting the signal to the adaptive filter.

# Applying compression and limiting

- Placing a compressor or limiter before the adaptive filter can reduce the impact of sudden loud noises, helping to ensure that the filter’s internal parameters remain relevant to the audio.
- This also protects the hearing of transcription experts, who tend to listen at high volumes.

# Equalisation

- Adaptive filters can make material sound more hissy than before processing. They can also make speech sound thin and reedy. Therefore, it is often desirable to apply some equalisation to the revealed speech.
- Typical settings include low-pass filtering at a few kHz to reduce the hiss left after the adaptive process, and gain in the region 700Hz - 1.6 kHz to accentuate the speech frequencies.
- If the material is very hissy after processing, following the adaptive filter with a high-quality single-ended noise reduction algorithm (perhaps before further equalisation) is suggested.

CEDAR

# Some examples of speech enhancement

- Mobile phone interference
- Problems with body wires
- Problems with distortion
- Telephones: volume and noise problems
- Hum on interview tapes
- Excessive reverberation
- Black box recordings

# Example 1. Mobile phone interference

- If mobile telephones are brought into proximity with unshielded recording equipment, they can generate bursts of high amplitude buzz on the recording. This noise can be of sufficient energy to mask the wanted information in the recording.
- You require two processes to remove the interference:
  - Remove the larger impulses using a declicker.
  - This will leave a smoother sounding residual, with a well identified fundamental frequency (217Hz for GSM interference) and regularly spaced harmonics. You may attenuate - or in many cases, eliminate - these using a debuzzer.

CEDAR

# Example 2. Problems with body wires

- Body wires suffer from poor signal to noise ratios and limited bandwidth. In addition, the recordings made using them are highly susceptible to unwanted background sounds and clothing rustle.
- If clothing rustle has not swamped the wanted signal, you might approach these problem in three stages:
  - Reduce the background noise using a single channel adaptive filter.
  - Use equalisation to increase the high frequency content and improve intelligibility.
  - Suppress residual background sounds using dialogue noise suppression.

CEDAR

# Example 3. Problems with distortion

- When microphones and recorders are used outside their specifications, and also when batteries begin to fail, various types of distortion will result. The resulting harsh sound has increased high frequency content that is correlated with the sound but which does not contain wanted information. This can reduce intelligibility, and will increase listening fatigue when monitoring or transcribing the audio.

- You require two processes to remove the interference:
  - Remove the clipping using a dedicated declip algorithm.
  - If necessary, remove the residual buzz and crackle using a decrackler.

CEDAR

# Example 4. Telephones: Volume & noise

- Many telephone recordings are made by holding a small recorder to the telephone earpiece. As a result, the near speaker is recorded clearly, while the far signal is of lower signal/noise ratio, limited bandwidth, and significantly lower volume, leading to it being obscured by local noise.
- Three processes are required:
  - Reduce the background noise using a single channel adaptive filter or dehisser (as appropriate) in order to reveal the lower level voice obtained from the telephone earpiece.
  - Balance and normalise the volumes of the two voices using a compressor/limiter.
  - Use equalisation to improve the presence and intelligibility of the voice recorded from the telephone earpiece.

CEDAR

# Example 5. Hum on interview tapes

- Many interview tapes suffer from electrical hum loud enough to obscure the speakers. This 50Hz or 60Hz hum has a strong fundamental and an extended harmonic series. Due to the nature of the recorders, the frequency may drift considerably, making conventional filtering unsuitable.
- You can completely eliminate most of these hums using an auto-tracking debuzzer.

# Example 6. Excessive reverberation

- Covert recordings made in reverberant spaces such as custody cells, corridors and stairwells can suffer from excessive reverberation that reduces intelligibility.
- Use a single channel adaptive filter or dialogue noise suppression unit to attenuate the extended reverberation tails. This will ‘dry’ the recording and limit the adverse effects of the recording environment without unduly reducing intelligibility.

# Example 7. Black box recordings

- These are limited bandwidth recordings, frequently distorted, made in high (electrical and acoustic) noise environments. They suffer from loud buzz, usually with a fundamental of 200Hz or 400Hz, loud broadband noise caused by airflow over the cockpit, distortion, and electrical crackle.
- You need to restore these recordings in many stages, reducing each problem individually.
  - Start by using a decrackle algorithm to eliminate the electrical crackle and to reduce distortion.
  - Eliminate the buzz induced by the aircraft's electrical system.
  - Use a single channel adaptive filter to reduce the broadband noise.
  - Use equalisation to improve the bandwidth and increase intelligibility.

CEDAR

# Conclusions about adaptive filtering

- Adaptive filters are important tools for forensic audio processing, capable of achieving results that cannot be obtained by other methods.
- There are at least four forms of adaptive filter, and it is the responsibility of the user to select the correct one for the task at hand.
- Adaptive filters are not universal panaceae, and they often work best in combination with other audio restoration processes.
- It is important to be able to hear and adjust the output from all the subtractive processes simultaneously. Performing multiple processes individually will often produce less than optimum results.
- CEDAR Cambridge supports multiple, simultaneous real-time algorithms, allowing users to construct appropriate processing chains on a job-by-job basis to perform complex noise reduction tasks.

CEDAR

# Retouch™

- Until recently, filtering systems were limited as to the types of noises that they could remove: clicks and scratches, crackle and buzz, pops, thumps and broadband noise.
- CEDAR Retouch, introduced in February 2002, offered a huge leap forward in processing technology, allowing users to identify individual sounds and words within a recording.
- Once identified, these sounds can be extracted, accentuated, or even eliminated - often seamlessly.

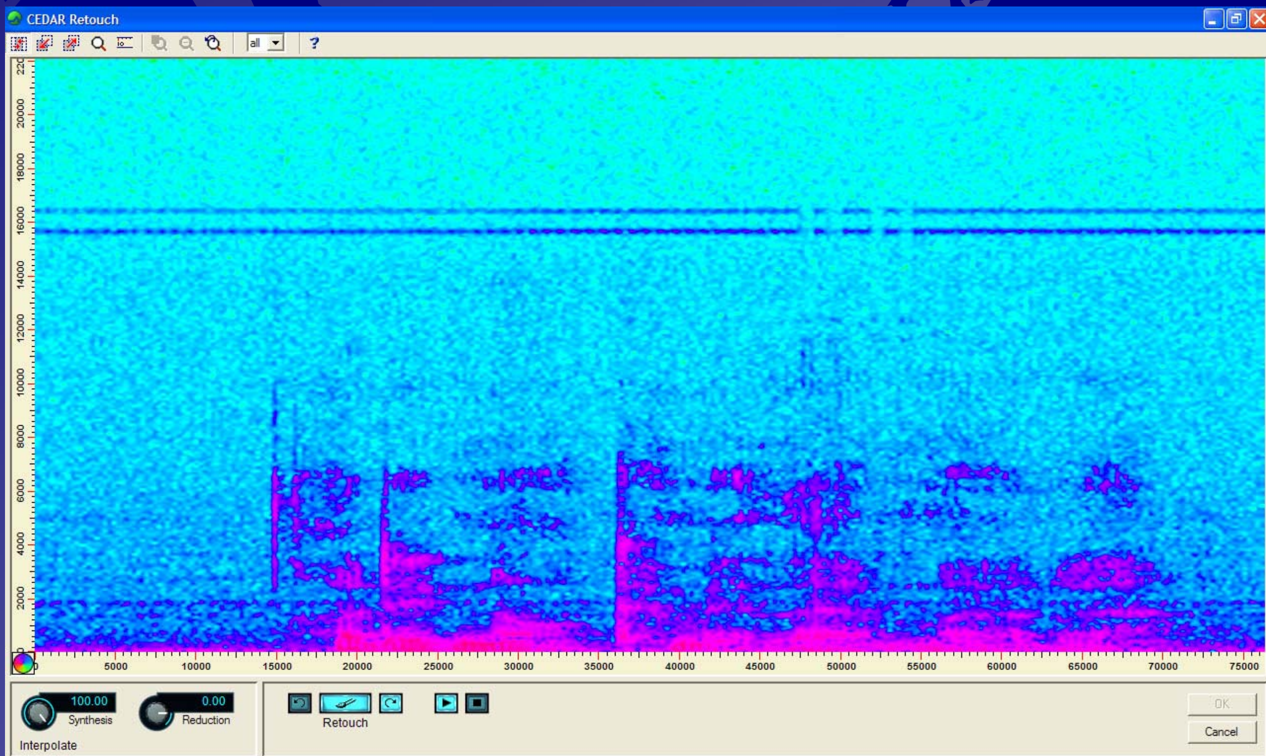
# Retouch™ - the dangers

- In the wrong hands, Retouch™ can be used to manipulate recordings to falsify evidence.
- The following four slides demonstrate this…

CEDAR

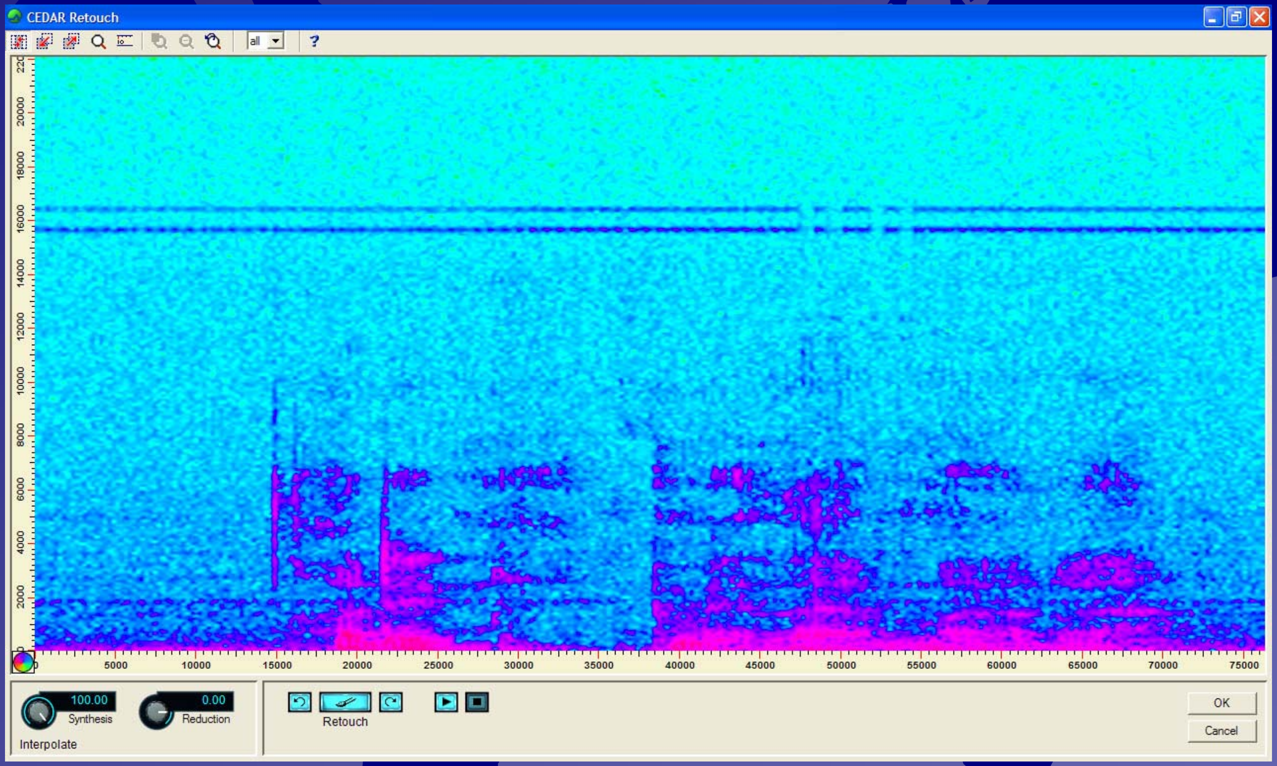
# Original recording

- “I did not commit that murder”



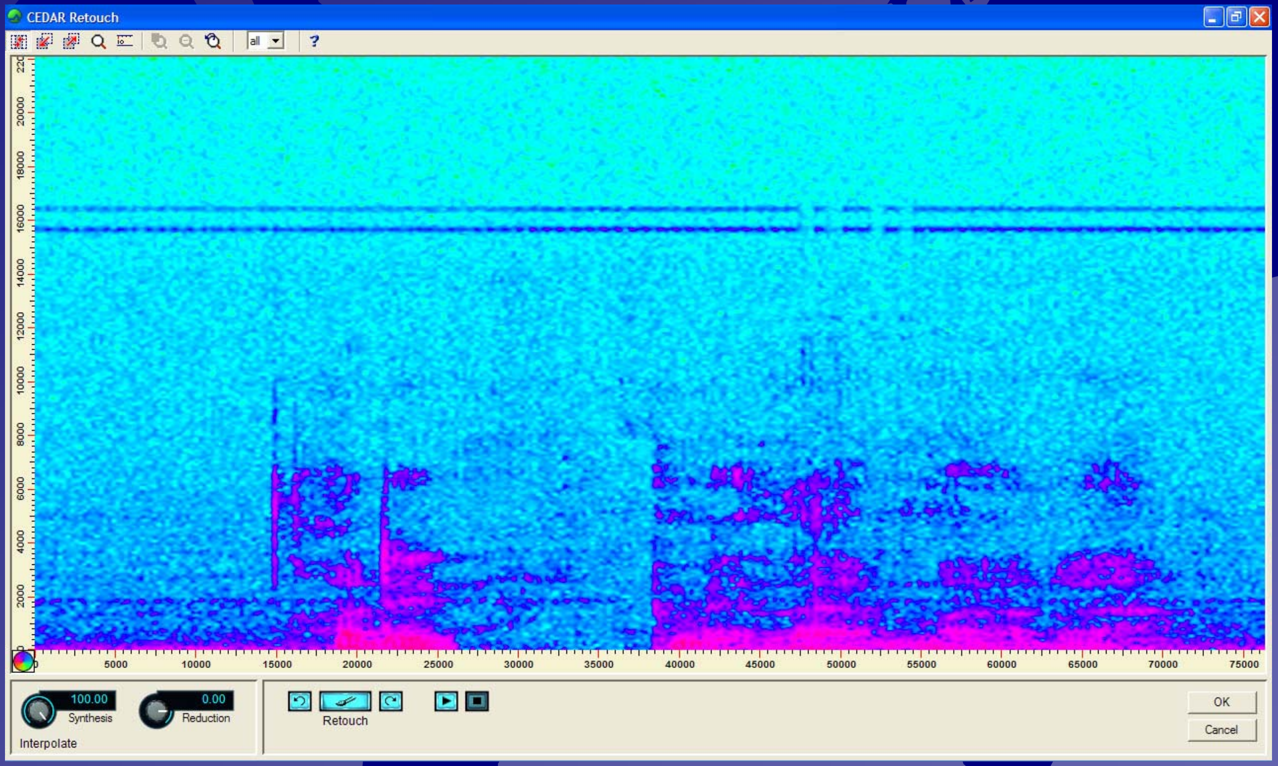
# First application of Retouch™

- “I did n__ commit that murder”



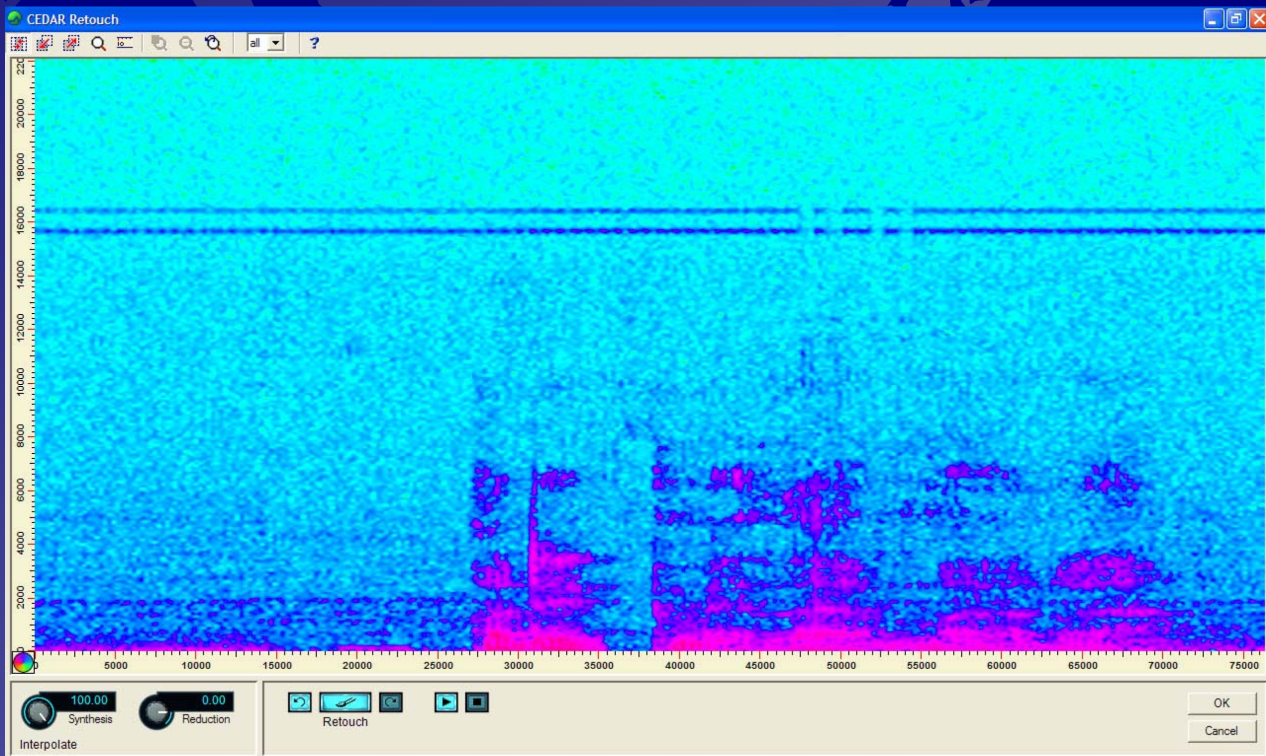
# Second application of Retouch™

- “I did ___ commit that murder”



# Final recording…

- λ “I did commit that murder”




CEDAR

# Legitimate uses of Retouch™

- To lift individual words out of the background noise to aid in the recognition of names, places, dates and so on.
- To eliminate loud noises such as sirens and claxons that are obscuring conversations while criminal activities are taking place.
- Existing users find increasing uses in aiding transcription.

# Contacts

- CEDAR Forensic

  *20 Home End, Fulbourn, Cambridge, CB1 5BS, UK*
  *Telephone: (+44) 1223 881771 · Fax: (+44) 1223 881778*
  *Email: forensic@cedaraudio.com*
  *Web: www.cedaraudio.com*

- CEDAR Audio USA

  *Telephone: (+1) 207 828 0024 · Fax: (+1) 207 773 2422*
  *Email: cedarusa@cedaraudio.com*

- CEDAR Asia

  *Telephone: (+66) 1 822 9227 · Fax: (+66) 2 619 1277*
  *Email: asia@cedaraudio.com*