From tone controls to audio forensics

Dr Christopher Hicks

CEDAR Audio Ltd

CSAR 16 June 2014

From tone controls to audio forensics

Sound, Noise and Hearing

- Physical fundamentals of sound
- What distinguishes noise from other sound?
- How is sound characterised and quantified?

Noise avoidance and reduction

- Avoidance or Reduction?
- The research landscape that enabled CEDAR to begin (pre-1980)
- Modern digital techniques – a taste of how they work (post-1990)
  - focus on forensic applications

The future

- Some current research

Sound, Noise and Hearing

What is sound?

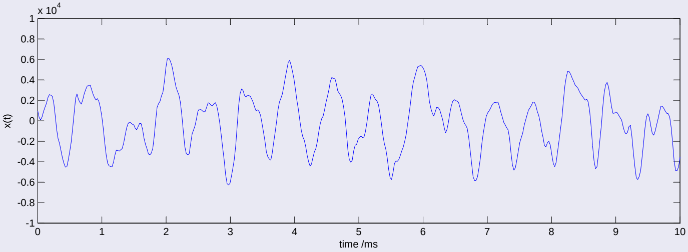
- Sound consists of pressure waves carried in air
- Detectable by human ear

What is noise?

- Sound that we don't want to hear!
- Technical defects of the "channel": hiss, clicks, hum
- Interference:
  - TV on during conversation
  - mobile phone "tic, tic-a-tic tic-a-tic"

Time or frequency?

Signal in the "time domain" $u(t)$



Time or frequency?

Sinusoidal vibrations

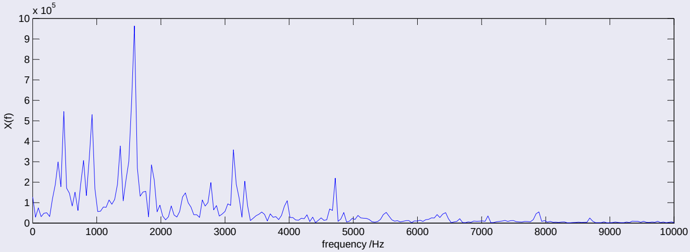
- Frequency: how many vibrations per second?
- Amplitude: how large are the oscillations?
- Phase: where is the time origin?

Limits of hearing

- Frequency: 20Hz to 20kHz
  - ten musical octaves
- Amplitude: 120dB (power ratio 1:10^{12})
  - total electrical power generated in UK to power required for LED
- Phase: of secondary importance only

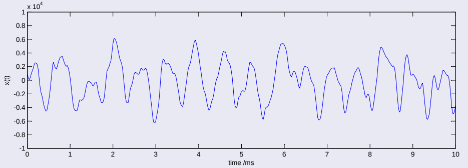
Time or frequency?

Signal in the 'frequency domain' $U(f)$

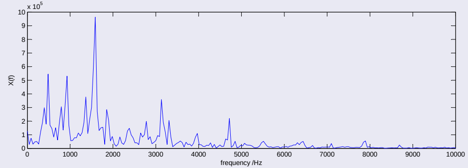


Time or frequency?

Fourier transform



$$u(t) = \int U(f) \exp(+i2\pi ft)\, df \qquad U(f) = \int u(t) \exp(-i2\pi ft)\, dt$$



Sound, Noise and Hearing

Sinusoidal vibrations

- Frequency: how many vibrations per second? pitch
- Amplitude: how large are the oscillations? volume

Complex vibrations

- Waveform: what "shape" are the vibrations? timbre
- Envelope: how does the waveform vary over time? attack/decay

Random vibrations

- Unpredictable, even in the short-term. hiss

Historical perspective: pre-1980

Noise avoidance

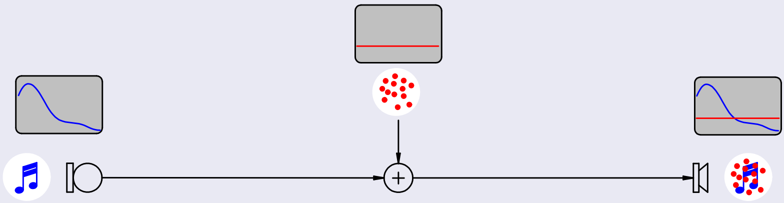
- Emphasis and de-emphasis
- Dynamic range compression and expansion

Noise reduction

- Low-pass filtering
- Downwards expansion
- "Packburn"

Analogue noise avoidance

Noisy Channel

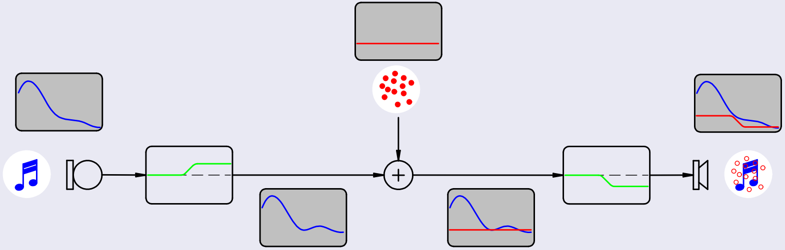


Demonstration

Original Noise The problem!

Analogue noise avoidance

Noisy Channel with Emphasis

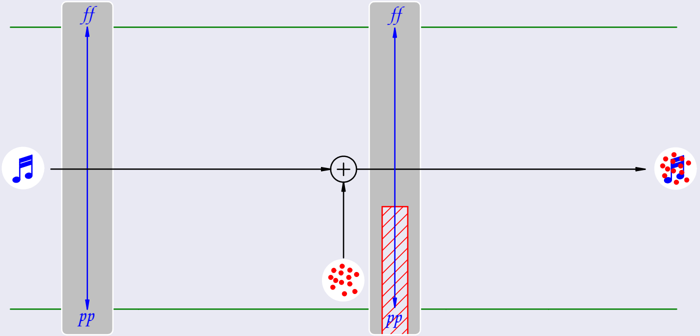


Demonstration

Original Emphasised Noise Emph+Noise Deemphasised

Analogue noise avoidance

Noisy Channel

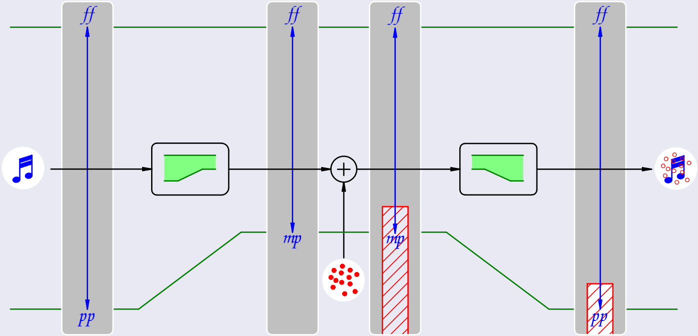


Demonstration

Original | Noise | The problem!

Analogue noise avoidance

Noisy Channel with Compression



Demonstration

Original Compressed Noise Comp+Noise Expanded

Analogue noise avoidance

A Cambridge success

- Take two simple ideas, emphasis and compression,...
- ...each basically sound, but neither without its problems,...
- ...and add a Research Student (Pembroke/Cavendish, 1958–63).

Ray Dolby

Noise reduction: pre-1980

Noise reduction: pre-1980

Digital techniques: post-1990

Noise avoidance

- Largely unnecessary

Noise reduction

- operate in frequency (or other) domain
- non-linear processes
- adaptive processes (process changes according to signal content)
- can use delay and lookahead
- statistical and model-based techniques

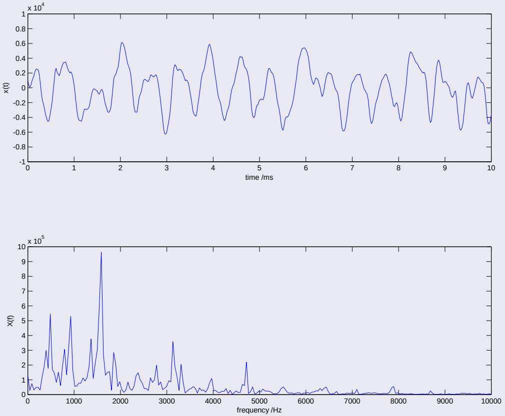
CEDAR20 – another Cambridge success

Interpolation of impulsive noise

Original | Declicked

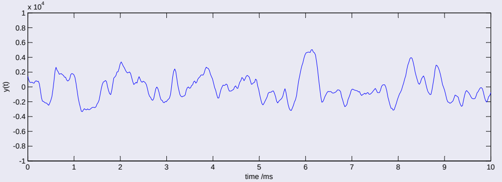
Digital techniques: post-1990

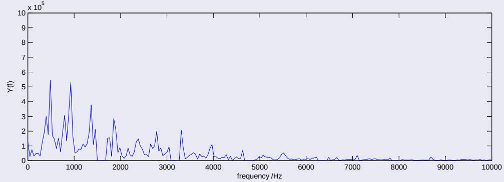
Frequency domain filters



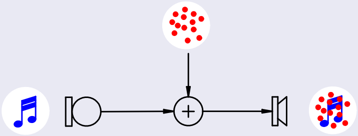
Digital techniques: post-1990

Frequency domain filters





The problem



Given a signal contaminated by noise $X(f) = U(f) + N(f)$

can we design a filter to recover the original? $H(f)\,X(f) = U(f)$?

Digital techniques: post-1990

Wiener filter

If $X(f) = U(f) + N(f)$ and we choose

$$H(f) = \frac{S_{UX}(f)}{S_{XX}(f)} = \frac{S_{XX}(f) - S_{NN}(f)}{S_{XX}(f)} = \frac{S_{UU}(f)}{S_{UU}(f) + S_{NN}(f)}$$

then $H(f)\,X(f)$ is the "optimal estimate" of $U(f)$.

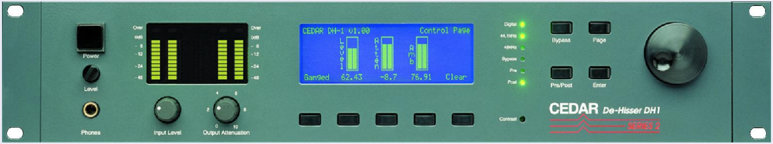
- You need to know the answer before you start.
- If the signal is dominant then $H(f) \to 1$. $(S_{UU}(f) \gg S_{NN}(f))$
- If the noise is dominant then $H(f) \to 0$. $(S_{NN}(f) \gg S_{UU}(f))$
- Pure implementations introduce unpleasant artefacts.
- How to find S_{NN} – 'fingerprint' or automatic.

Power Spectral Density

$S_{AB}(f) = \int R_{ab}(t) \exp(-i2\pi ft)\,dt$, where $R_{ab}(t) = E[a(\tau)b(\tau - t)]$.

Digital techniques: post-1990

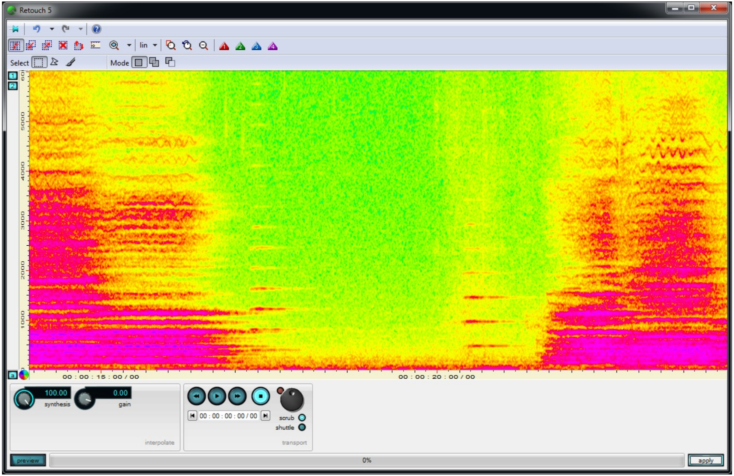
Spectral subtraction



Demonstration

Carpenters	Noisy Wiener NR4
Howells	Noisy Wiener NR4

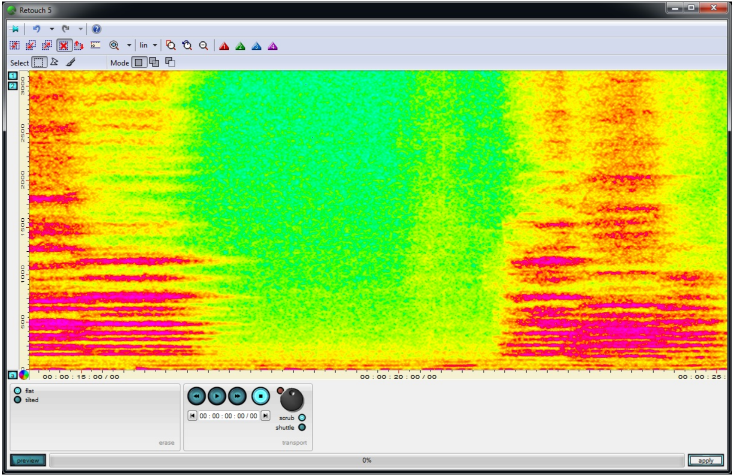
Retouch



Demonstration

Original

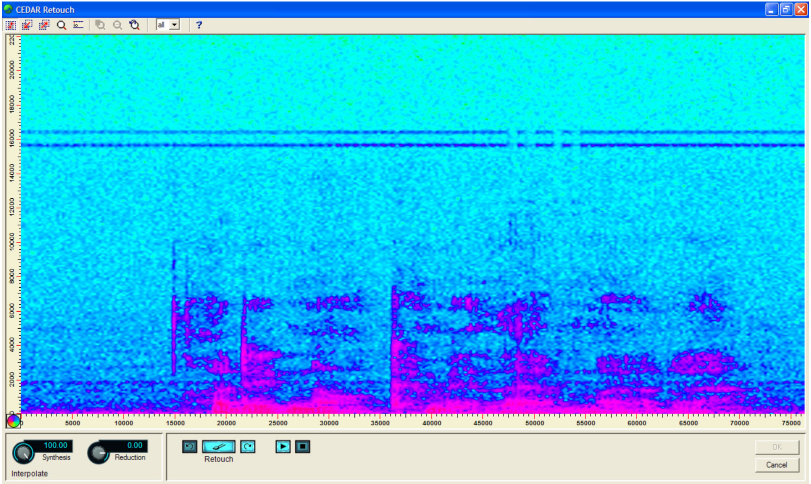
Retouch



Demonstration

Retouched

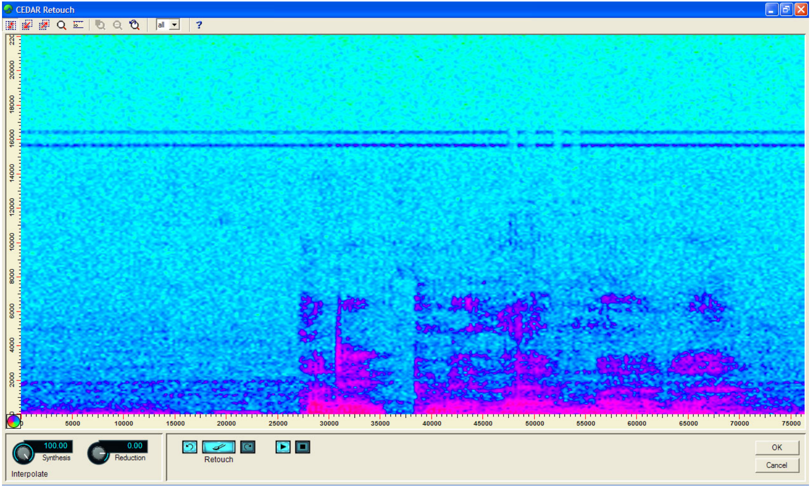
Retouch



Demonstration

Denial...

Retouch



Demonstration

...or confession?

Surveillance

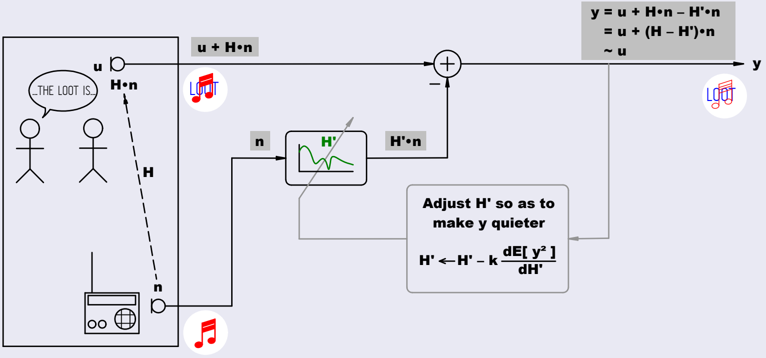
The scenario

James Bond has bugged the hotel room where the baddies are due to meet. The baddies suspect that they might be being bugged, so they turn on the radio to mask their conversation.

Q has a plan...

Naturally, Q has anticipated the situation, and has equipped Bond with a 'cross-channel adaptive lattice filter,' disguised as an expensive watch.

Cross-channel adaptive filter



Demonstration

Surveillance Reference Subtract Adaptive

Mobile phone interference

Problem

- Mobile telephones are radio transmitters.
- Radio transmitters interfere with other electronic apparatus, such as microphones, radio receivers etc.
- Isolated clicks, and bursts of harsh-sounding buzz.

A solution

A combination of click-removal processes (derivatives of those used to remove clicks from records) plus a buzz-removal process tuned to 217Hz (for GSM).

Demonstration

Original Processed

Air accident investigation

Problem

- Cockpit communications are made in a very harsh, noisy environment.
- Aircraft electrical systems operate at 400Hz, and interference at this frequency and its harmonics is a particular hindrance to speech intelligibility.
- In the event of an incident, recordings made of the communications with ATC and recordings retrieved from the "black box" form vital clues as to what happened.

A solution

A single-channel adaptive filter removes the "stationary" part of the signal.

Demonstration

Original Processed

Speed errors

Problem

- mechanical faults
- stretched tape
- failing batteries
- warped discs

A solution

Use the fact that musical notes are largely constant to estimate a speed variation function, and then correct by variable-rate interpolation.

Demonstration

Original Processed

The future

Some current research

- noise reduction and speech enhancement
- sound classification (speech, breaking glass, gunshots)
- speaker authentication and identification (‘voice-print’)
- signal separation
  - at what point does noise reduction become signal separation?
- packet loss concealment
- acoustics, localisation, room identification (‘room-print’)
- forensic musicology
- electrical network frequency analysis, and similar

Questions?

Thank you!