

Sounds of Open Source

processing sound with sox



DMITRY DOLGOV

06-11-2021



- Why sound processing?
- Sound Exchange
- Recording and formats
- Cleaning up
- Merging layers
- Adding effects





@OpenClipart

Record audio

```
rec -r 44100 -b 16 -e signed-integer audio.mp3
```

Record audio

→ Self-describing formats



→ Raw or headerless formats



Cleaning up

```
# the noise is always different,  
# try to record it immediately  
# before the actual audio.
```

```
sox noise.mp3 -n noiseprof noise.prof  
sox audio.mp3 output.mp3 noisered noise.prof 0.21
```

Cleaning up

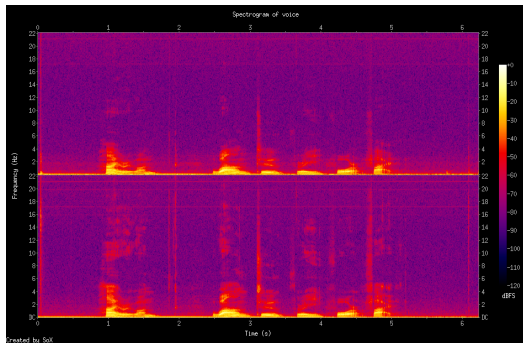


Figure 1: With noise

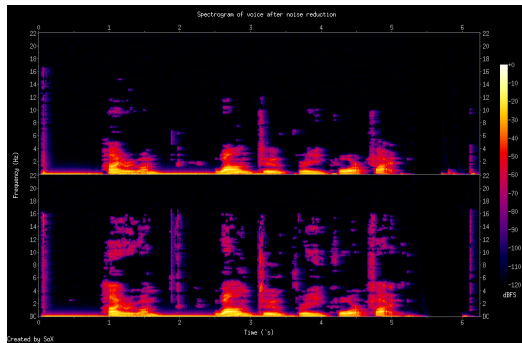


Figure 2: Without noise

Cleaning up

Fast Fourier transform

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-i2\pi kn}{N}}$$

Cleaning up

```
# Voice Activity Detector to trim  
# silence and quiet background noise  
# at the beginning and the end
```

```
sox audio.mp3 trimmed.mp3 vad
```

Cleaning up

-v, --volume to adjust volume by certain factor

```
sox audio.mp3 -n stat
```

```
Samples read:          441216
Length (seconds):      10.004898
Scaled by:             2147483647.0
Maximum amplitude:     0.019642
Minimum amplitude:     -0.019630
Mean    norm:          0.004128
Mean    amplitude:     -0.000000
Volume adjustment:     50.911
```

Automatic Gain Control

```
sox audio.mp3 agc.mp3 compand \  
  0.3,0.3 \ # attach, decay, time over which  
           \ # to measure for averaging  
  
6:         \ # transfer function, with 6dB soft knee  
  
    0, -40,      \ # do nothing below -40dB  
  -40,-35,      \ # compand -40dB to -35dB  
  -20,  0, -20 \ # everything else to -20dB  
  -10 -60 1
```

Automatic Gain Control

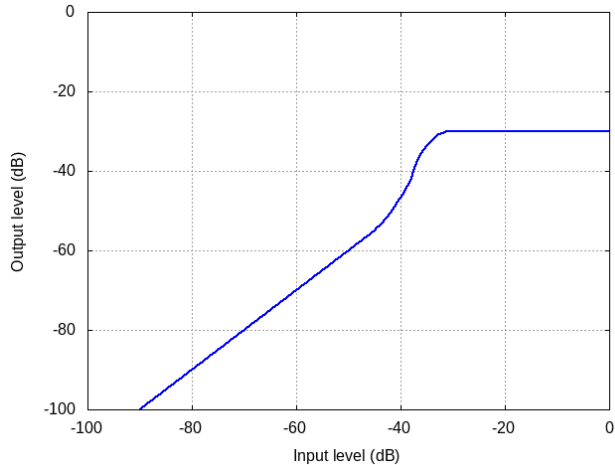


Figure 3: SoX compand (compress and expand) effect

Unite and lead

without the volume specified sox can try to
automatically adjust the volume to prevent
clipping, making everything unexpectedly quiet.

```
sox -v 1.0 part.1.mp3 part.2.mp3 concat.mp3
```

```
sox -m -v 1.0 part.1.mp3 part.2.mp3 mixture.mp3
```

Clipping

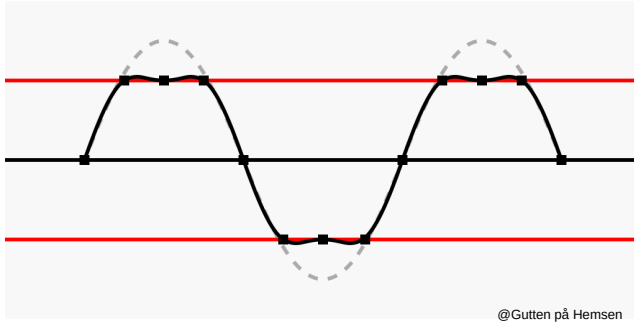


Figure 4: This PCM waveform is clipped between the red lines

Unite and lead

```
sox audio.mp3 output.mp3 \  
  trim 0 30 # cut out from the 0 to 30th second  
  fade 5    # fade in/out for 5 seconds  
  pad 7     # pad with 7 seconds of silence
```


Unite and lead

```
# splice two parts together at position 00:25 with  
# excess/leeway 1 second and half-cosine wave fading
```

```
sox part1.mp3 part2.mp3 mix.mp3 \  
    splice -h 00:25,1,1
```

Effects

add an echo with the delay 500 ms and loudness 0.3

sox voice.clean.mp3 echo.mp3 echo 0.8 0.9 500 0.3

Effects

```
# Kaiser-Bessel window band-pass filter  
# to generate "phone" effect  
  
sox audio.mp3 output.mp3 sinc 500-3000 vol +3  
  
# Add some clicking noise on top of it  
sox -n pinknoise.mp3 rate 44100 \  
    synth 10 pinknoise vol 0.01  
sox -m -v 1 output.mp3 pinknoise.mp3 phone.mp3
```

Kaiser-Bessel window

Effects

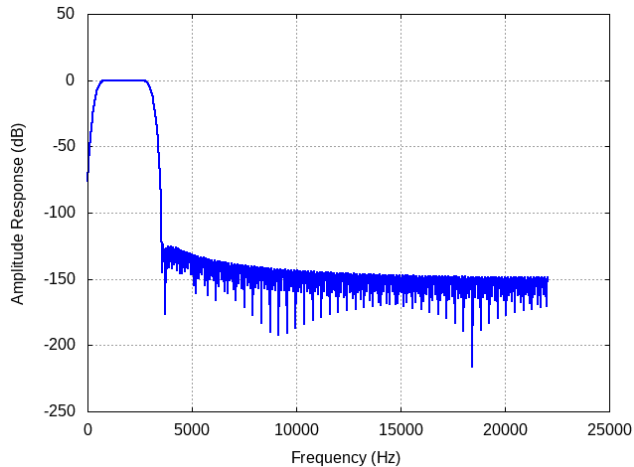


Figure 5: SoX sinc (band-pass filter) effect

Generating effects

```
# -n says the input is a special "null" file  
sox -n guitar-chortds.mp3 \
```

```
# sine waves for tones of specified note and octave  
synth pl G2 pl B2 pl D3 pl G3 pl D4 pl G4 \
```

```
# delay channels with each tone and mix with fading effect  
delay 0 .05 .1 .15 .2 .25 remix - fade 0 4 .1 norm -1
```

Triangular probability density function (TPDF) dither.





@Gerbrant

Figure 6: Dithering in image processing.

Questions?

 @erthalion

 dmitrii.dolgov at zalando dot de

 9erthalion6 at gmail dot com