

NAME

SoX – Sound eXchange, the Swiss Army knife of audio manipulation

EXAMPLES

Introduction

The core problem is that you need some experience in using effects in order to say ‘that any old sound file sounds with effects absolutely hip’. There isn’t any rule-based system which tell you the correct setting of all the parameters for every effect. But after some time you will become an expert in using effects.

Here are some examples which can be used with any music sample. (For a sample where only a single instrument is playing, extreme parameter setting may make well-known ‘typically’ or ‘classical’ sounds. Likewise, for drums, vocals or guitars.)

Single effects will be explained and some given parameter settings that can be used to understand the theory by listening to the sound file with the added effect.

Using multiple effects in parallel or in series can result either in a very nice sound or (mostly) in a dramatic overloading in variations of sounds such that your ear may follow the sound but you will feel unsatisfied. Hence, for the first time using effects try to compose them as minimally as possible. We don’t regard the composition of effects in the examples because too many combinations are possible and you really need a very fast machine and a lot of memory to play them in real-time.

However, real-time playing of sounds will greatly speed up learning and/or tuning the parameter settings for your sounds in order to get that ‘perfect’ effect.

Basically, we will use the ‘play’ front-end of SoX since it is easier to listen sounds coming out of the speaker or earphone instead of looking at cryptic data in sound files.

For easy listening of file.xxx (‘xxx’ is any sound format):

```
play file.xxx effect-name effect-parameters
```

Or more SoX-like (for ‘dsp’ output on a UNIX/Linux computer):

```
sox file.xxx -t oss -2 -s /dev/dsp effect-name effect-parameters
```

or (for ‘au’ output):

```
sox file.xxx -t sunau -2 -s /dev/audio effect-name effect-parameters
```

And for date freaks:

```
sox file.xxx file.yyy effect-name effect-parameters
```

Additional options can be used. However, in this case, for real-time playing you’ll need a very fast machine.

Notes:

I played all examples in real-time on a Pentium 100 with 32 MB and Linux 2.0.30 using a self-recorded sample (3:15 min long in ‘wav’ format with 44.1 kHz sample rate and stereo 16 bit). The sample should not contain any of the effects. However, if you take any recording of a sound track from radio or tape or CD, and it sounds like a live concert or ten people are playing the same rhythm with their drums or funky-grooves, then take any other sample. (Typically, less then four different instruments and no synthesizer in the sample is suitable. Likewise, the combination vocal, drums, bass and guitar.)

Echo

An echo effect can be naturally found in the mountains, standing somewhere on a mountain and shouting a single word will result in one or more repetitions of the word (if not, turn a bit around and try again, or climb to the next mountain).

However, the time difference between shouting and repeating is the delay (time), its loudness is the decay. Multiple echos can have different delays and decays.

It is very popular to use echos to play an instrument with itself together, like some guitar players (Queen’s

Brian May) or vocalists do. For music samples of more than one instrument, echo can be used to add a second sample shortly after the original one.

This will sound as if you are doubling the number of instruments playing in the same sample:

```
play file.xxx echo 0.8 0.88 60 0.4
```

If the delay is very short, then it sound like a (metallic) robot playing music:

```
play file.xxx echo 0.8 0.88 6 0.4
```

Longer delay will sound like an open air concert in the mountains:

```
play file.xxx echo 0.8 0.9 1000 0.3
```

One mountain more, and:

```
play file.xxx echo 0.8 0.9 1000 0.3 1800 0.25
```

Echos

Like the echo effect, echos stand for 'ECHO in Sequel', that is the first echos takes the input, the second the input and the first echos, the third the input and the first and the second echos, ... and so on. Care should be taken using many echos (see introduction); a single echos has the same effect as a single echo.

The sample will be bounced twice in symmetric echos:

```
play file.xxx echos 0.8 0.7 700 0.25 700 0.3
```

The sample will be bounced twice in asymmetric echos:

```
play file.xxx echos 0.8 0.7 700 0.25 900 0.3
```

The sample will sound as if played in a garage:

```
play file.xxx echos 0.8 0.7 40 0.25 63 0.3
```

Chorus

The chorus effect has its name because it will often be used to make a single vocal sound like a chorus. But it can be applied to other instrument samples too.

It works like the echo effect with a short delay, but the delay isn't constant. The delay is varied using a sinusoidal or triangular modulation. The modulation depth defines the range the modulated delay is played before or after the delay. Hence the delayed sound will sound slower or faster, that is the delayed sound tuned around the original one, like in a chorus where some vocals are a bit out of tune.

The typical delay is around 40ms to 60ms, the speed of the modulation is best near 0.25Hz and the modulation depth around 2ms.

A single delay will make the sample more overloaded:

```
play file.xxx chorus 0.7 0.9 55 0.4 0.25 2 -t
```

Two delays of the original samples sound like this:

```
play file.xxx chorus 0.6 0.9 50 0.4 0.25 2 -t 60 0.32 0.4 1.3 -s
```

A big chorus of the sample is (three additional samples):

```
play file.xxx chorus 0.5 0.9 50 0.4 0.25 2 -t 60 0.32 0.4 2.3 -t 40 0.3 0.3 1.3 -s
```

Flanger

The flanger effect is like the chorus effect, but the delay varies between 0ms and maximal 5ms. It sound like wind blowing, sometimes faster or slower including changes of the speed.

The flanger effect is widely used in funk and soul music, where the guitar sound varies frequently slow or a bit faster.

Now, let's groove the sample:

```
play file.xxx flanger
```

listen carefully between the difference of sinusoidal and triangular modulation:

```
play file.xxx flanger triangle
```

Reverb

A reverberation effect is sometimes needed in concert halls that are too small or contain so many people that the hall's natural reverberance is diminished.

Using the effect is easy:

```
play file.xxx reverb
```

gives the default reverberance (50%); or specify the desired reverberance as a percentage:

```
play file.xxx reverb 80
```

For fine tuning, see **sox(1)**.

If you run out of machine power or memory, then stop as many applications as possible.

Phaser

The phaser effect is like the flanger effect, but it uses a reverb instead of an echo and does phase shifting. You'll hear the difference in the examples comparing both effects. The delay modulation can be sinusoidal or triangular, preferable is the later for multiple instruments. For single instrument sounds, the sinusoidal phaser effect will give a sharper phasing effect. The decay shouldn't be too close to 1 which will cause dramatic feedback. A good range is about 0.5 to 0.1 for the decay.

We will take a parameter setting as before (gain-out is lower since feedback can raise the output dramatically):

```
play file.xxx phaser 0.8 0.74 3 0.4 0.5 -t
```

The drunken loudspeaker system (now less alcohol):

```
play file.xxx phaser 0.9 0.85 4 0.23 1.3 -s
```

A popular sound of the sample is as follows:

```
play file.xxx phaser 0.89 0.85 1 0.24 2 -t
```

The sample sounds if ten springs are in your ears:

```
play file.xxx phaser 0.6 0.66 3 0.6 2 -t
```

Compannder

The compander effect allows the dynamic range of a signal to be compressed or expanded. It works by calculating the input signal level averaged over time according to the given attack and decay parameters, and setting the output signal level according to the given transfer-function parameters.

For most situations, the attack time (response to the music getting louder) should be shorter than the decay time because our ears are more sensitive to suddenly loud music than to suddenly soft music.

For example, suppose you are listening to Strauss's 'Also Sprach Zarathustra' in a noisy environment such as a moving vehicle. If you turn up the volume enough to hear the soft passages over the road noise, the loud sections will be too loud. So you could try this:

```
sox asz.flac asz-car.flac compand 0.3,1 6:-70,-60,-20 -5 -90 0.2
```

The transfer function ('6:-70,...') says that very soft sounds (below -70dB) will remain unchanged. This will stop the compander from boosting the volume on 'silent' passages such as between movements. However, sounds in the range -60dB to 0dB (maximum volume) will be boosted so that the 60dB dynamic range of the original music will be compressed 3-to-1 into a 20dB range, which is wide enough to enjoy the music but narrow enough to get around the road noise. The '6:' selects 6dB soft-knee companding. The -5 (dB) output gain is needed to avoid clipping (the number is inexact, and was derived by experimentation).

The -90 (dB) for the initial volume will work fine for a clip that starts with near silence, and the delay of 0.2 (seconds) has the effect of causing the compander to react a bit more quickly to sudden volume changes.

In order to visualise the transfer function, SoX can be invoked with the **--plot** option, e.g.

```
sox -n -n --plot gnuplot compand 0,0 6:-70,-60,-20 -5 > my.plt
gnuplot my.plt
```

The following (one long) command shows how multi-band companding is typically used in FM radio:

```
play file.xxx vol -3dB filter 8000- 32 100 mcompand \
"0.005,0.1 -47,-40,-34,-34,-17,-33" 100 \
"0.003,0.05 -47,-40,-34,-34,-17,-33" 400 \
"0.000625,0.0125 -47,-40,-34,-34,-15,-33" 1600 \
"0.0001,0.025 -47,-40,-34,-34,-31,-31,-0,-30" 6400 \
"0,0.025 -38,-31,-28,-28,-0,-25" \
vol 15dB highpass 22 highpass 22 filter -17500 256 \
vol 9dB lowpass -1 17801
```

The audio file is played with a simulated FM radio sound (or broadcast signal condition if the lowpass filter at the end is skipped). Note that the pipeline is set up with US-style 75us preemphasis.

Changing the Rate of Playback

You can use stretch to change the rate of playback of an audio sample while preserving the pitch. For example to play at half the speed:

```
play file.wav stretch 2
```

To play a file at twice the speed:

```
play file.wav stretch 0.5
```

Other related options are 'speed' to change the speed of play (and changing the pitch accordingly), and pitch, to alter the pitch of a sample. For example to speed a sample so it plays in half the time (for those Mickey Mouse voices):

```
play file.wav speed 2
```

To raise the pitch of a sample 1 while note (100 cents):

```
play file.wav pitch 100
```

Reducing noise in a recording

First find a period of silence in your recording, such as the beginning or end of a piece. If the first 1.5 seconds of the recording are silent, do

```
sox file.wav -n trim 0 1.5 noiseprof /tmp/profile
```

Next, use the noisered effect to actually reduce the noise:

```
play file.wav noisered /tmp/profile
```

Making a recording

Thanks to Douglas Held for the following suggestion:

```
rec parameters filename other-effects silence 1 5 2%
```

This use of the **silence** effect allows you to start a recording session but only start writing to disk once non-silence is detected. For example, use this to start your favorite command line for recording and walk over to your record player and start the song. No periods of silence will be recorded.

Scripting with SoX

One of the benefits of a command-line tool is that it is easy to use it in scripts to perform more complex tasks. In marine radio, a Mayday emergency call is transmitted preceded by a 30-second alert sound. The

alert sound comprises two audio tones at 1300Hz and 2100Hz alternating at a rate of 4Hz. The following shows how SoX can be used in a script to construct an audio file containing the alert sound. The scripting language shown is 'Bourne shell' (sh) but it should be simple to translate this to another scripting language if you do not have access to sh.

```
# Make sure we append to a file that's initially empty:
rm -f 2tones.raw

for freq in 1300 2200; do
    sox -c1 -r8000 -n -t raw - synth 0.25 sine $freq vol 0.7 >> 2tones.raw
done

# We need 60 copies of 2tones.raw (0.5 sec) to get 30 secs of audio:
iterations=60

# Make sure we append to a file that's initially empty:
rm -f alert.raw

while [ $iterations -ge 1 ]; do
    cat 2tones.raw >> alert.raw
    iterations=`expr $iterations - 1`
done

# Add a file header and save some disc space:
sox -s2 -c1 -r8000 alert.raw alert.ogg

play alert.ogg
```

If you try out the above script, you may want to hit Ctrl-C fairly soon after the alert tone starts playing—it's not a pleasant sound! The **synth** effect is used to generate each of the tones; **-c1 -r8000** selects mono, 8kHz sampling-rate audio (i.e. relatively low fidelity, suitable for the marine radio transmission channel); each tone is generated at a length of 0.25 seconds to give the required 4Hz alternation. Note the use of 'raw' as the intermediary file format; a self-describing (header) format would just get in the way here. The self-describing header is added only at the final stage; in this case, **.ogg** is chosen, since lossy compression is appropriate for this application.

There are further practical examples of scripting with SoX available to download from the SoX web-site [1].

SEE ALSO

sox(1), **libsox(3)**

References

[1] *SoX—Sound eXchange / Scripts*, <http://sox.sourceforge.net/Docs/Scripts>

AUTHORS

This man page was written largely by Juergen Mueller (jmueller@uia.ua.ac.be). Other SoX authors and contributors are listed in the AUTHORS file that is distributed with the source code.