

SOX

257522_PixelRule.tiff ↗

SOX(1)
SOX(1)

UNIX Programmer's Manual

NAME

sox - SOund eXchange - universal sound sample translator

SYNOPSIS

sox infile outfile

sox infile outfile [effect [effect options ...]]

sox infile -e effect [effect options ...]

sox [general options] [format options] ifile [format
options] ofile [effect [effect options ...]]

General options: [-V] [-v volume]

Format options: [-t filetype] [-r rate] [-s/-u/-U/-A]

[-b/-w/-l/-f/-d/-D] [-c channels] [-x]

Effects:

- copy

- rate

- avg

- stat

- echo delay volume [delay volume ...]

- vibro speed [depth]

- lowp center

- highp center

- band [-n] center [width]

DESCRIPTION

Sox translates sound files from one format to another, possibly doing a sound effect.

OPTIONS

The option syntax is a little grotty, but in essence:

sox file.au file.voc

translates a sound sample in SUN Sparc .AU format into a SoundBlaster .VOC file, while

`sox -v 0.5 file.au -rate 12000 file.voc rate`

does the same format translation but also lowers the amplitude by 1/2 and changes the sampling rate from 8000 hertz to 12000 hertz via the rate sound effect loop.

File type options:

`-t filetype`

gives the type of the sound sample file.

`-r rate` Give sample rate in Hertz of file.

`-s/-u/-U/-A`

The sample data is signed linear (2's complement), unsigned linear, U-law (logarithmic), or A-law (logarithmic). U-law and A-law are the U.S. and international standards for logarithmic telephone sound compression.

-b/-w/-l/-f/-d/-D

The sample data is in bytes, 16-bit words, 32-bit longwords, 32-bit floats, 64-bit double floats, or 80-bit IEEE floats. Floats and double floats are in native machine format.

-x

The sample data is in XINU format; that is, it comes from a machine with the opposite word order than yours and must be swapped according to the word-size given above. Only 16-bit and 32-bit integer data may be swapped. Machine-format floating-point data is not portable. IEEE floats are a fixed, portable format. ???

-c channels

The number of sound channels in the data file. This may be 1, 2, or 4; for mono, stereo, or quad sound data.

General options:

- e after the input file allows you to avoid giving an output file and just name an effect. This is only useful with the stat effect.

- v volume Change amplitude (floating point); less than 1.0 decreases, greater than 1.0 increases. Note: we perceive volume logarithmically, not linearly. Note: see the stat effect.

- V Print a description of processing phases. Useful for figuring out exactly how sox is mangling your sound samples.

The input and output files may be standard input and output. This is specified by '-'. The -t type option must be given in this case, else sox will not know the format of the given file. The -t, -r, -s/-u/-U/-A, -b/-w/-l/-f/-d/-D and -x options refer to the input data when given before the input file name. After, they refer to the output data.

If you don't give an output file name, sox will just read

the input file. This is useful for validating structured file formats; the stat effect may also be used via the -e option.

FILE TYPES

Sox needs to know the formats of the input and output files. File formats which have headers are checked, if that header doesn't seem right, the program exits with an appropriate message. Currently, raw (no header) binary and textual data, IRCAM Sound Files, Sound Blaster, SPARC .AU (w/header), Mac HCOM, PC/DOS .SOU, Sndtool, and Sounder, NeXT .SND, Windows 3.1 RIFF/WAV, Turtle Beach .SMP, CD-R, and Apple/SGI AIFF and 8SVX formats are supported.

.aiff AIFF files used on Apple IIc/IIgs and SGI. Note: the AIFF format supports only one SSND chunk. It does not support multiple sound chunks, or the 8SVX musical instrument description format. AIFF files are multimedia archives and can have multiple audio and picture chunks. You may need a separate archiver to work with them.

- .au** SUN Microsystems AU files. There are apparently many types of .au files; DEC has invented its own with a different magic number and word order. The .au handler can read these files but will not write them. Some .au files have valid AU headers and some do not. The latter are probably original SUN u-law 8000 hz samples. These can be dealt with using the .ul format (see below).
- .hcom** Macintosh HCOM files. These are (apparently) Mac FSSD files with some variant of Huffman compression. The Macintosh has wacky file formats and this format handler apparently doesn't handle all the ones it should. Mac users will need your usual arsenal of file converters to deal with an HCOM file under Unix or DOS.
- .raw** Raw files (no header).
The sample rate, size (byte, word, etc), and style (signed, unsigned, etc.) of the sample file must

be given. The number of channels defaults to 1.

.ub, .sb, .uw, .sw, .ul

These are several suffices which serve as a shorthand for raw files with a given size and style.

Thus, ub, sb, uw, sw, and ul correspond to "unsigned byte", "signed byte", "unsigned word", "signed word", and "ulaw" (byte). The sample rate defaults to 8000 hz if not explicitly set, and the number of channels (as always) defaults to 1.

There are lots of Sparc samples floating around in u-law format with no header and fixed at a sample rate of 8000 hz. (Certain sound management software cheerfully ignores the headers.) Similarly, most Mac sound files are in unsigned byte format with a sample rate of 11025 or 22050 hz.

.sf

IRCAM Sound Files.

SoundFiles are used by academic music software such as the CSound package, and the MixView sound sample editor.

.voc Sound Blaster VOC files.
VOC files are multi-part and contain silence parts, looping, and different sample rates for different chunks. On input, the silence parts are filled out, loops are rejected, and sample data with a new sample rate is rejected. Silence with a different sample rate is generated appropriately. On output, silence is not detected, nor are impossible sample rates.

.auto This is a ``meta-type": specifying this type for an input file triggers some code that tries to guess the real type by looking for magic words in the header. If the type can't be guessed, the program exits with an error message. The input must be a plain file, not a pipe. This type can't be used for output files.

.cdr CD-R
CD-R files are used in mastering music Compact

Disks. The file format is, as you might expect, raw stereo raw unsigned samples at 44khz. But, there's some blocking/padding oddity in the format, so it needs its own handler.

.dat

Text Data files

These files contain a textual representation of the sample data. There is one line at the beginning that contains the sample rate. Subsequent lines contain two numeric data items: the time since the beginning of the sample and the sample value. Values are normalized so that the maximum and minimum are 1.00 and -1.00. This file format can be used to create data files for external programs such as FFT analyzers or graph routines. SOX can also convert a file in this format back into one of the other file formats.

.smp

Turtle Beach SampleVision files.

SMP files are for use with the PC-DOS package SampleVision by Turtle Beach Softworks. This package

is for communication to several MIDI samplers. All sample rates are supported by the package, although not all are supported by the samplers themselves. Currently loop points are ignored.

.wav

Windows 3.1 .WAV RIFF files.

These appear to be very similar to IFF files, but not the same. They are the native sound file format of Windows 3.1. Obviously, Windows 3.1 is of such incredible importance to the computer industry that it just had to have its own sound file format.

EFFECTS

Only one effect from the palette may be applied to a sound sample. To do multiple effects you'll need to run sox in a pipeline.

copy

Copy the input file to the output file. This is the default effect if both files

have the same sampling rate, or the rates are "close".

rate

Translate input sampling rate to output sampling rate via linear interpolation to the Least Common Multiple of the two sampling rates. This is the default effect if the two files have different sampling rates. This is fast but noisy: the spectrum of the original sound will be shifted upwards and duplicated faintly when up-translating by a multiple.

avg

Mix 4- or 2-channel sound file into 2- or 1-channel file by averaging the samples for different speakers.

stat

Do a statistical check on the input file, and print results on the standard error file. stat may copy the file untouched from input to output, if you select an output file. The "Volume Adjustment:" field in the statistics gives you the argument to the -v number which will make the sample as loud as possible.

echo [delay volume ...]

Add echoing to a sound sample. Each delay/volume pair gives the delay in seconds and the volume (relative to 1.0) of that echo. If the volumes add up to more than 1.0, the sound will melt down instead of fading away.

vibro speed [depth]

Add the world-famous Fender Vibro-Champ sound effect to a sound sample by using a sine wave as the volume knob. Speed gives the Hertz value of the wave. This must be under 30. Depth gives the amount the volume is cut into by the sine wave, ranging 0.0 to 1.0 and defaulting to 0.5.

lowp center

Apply a low-pass filter. The frequency response drops logarithmically with center frequency in the middle of the drop. The slope of the filter is quite gentle.

highp center

Apply a high-pass filter. The frequency response drops loga-

rithmically with center frequency in the middle of the drop. The slope of the filter is quite gentle.

band [-n] center [width] Apply a band-pass filter. The frequency response drops logarithmically around the center frequency. The width gives the slope of the drop. The frequencies at center + width and center - width will be half of their original amplitudes. Band defaults to a mode oriented to pitched signals, i.e. voice, singing, or instrumental music. The -n (for noise) option uses the alternate mode for un-pitched signals. Band introduces noise in the shape of the

filter, i.e. peaking at the center frequency and settling around it.

Sox enforces certain effects. If the two files have different sampling rates, the requested effect must be one of copy, or rate, If the two files have different numbers of channels, the avg effect must be requested.

reverse

Reverse the sound sample completely. Included for finding Satanic subliminals.

BUGS

The syntax is horrific. It's very tempting to include a default system that allows an effect name as the program name and just pipes a sound sample from standard input to standard output, but the problem of inputting the sample rates makes this unworkable.

FILES

SEE ALSO
NOTICES

The echoplex effect is:

Copyright (C) 1989 by Jef Poskanzer.

Permission to use, copy, modify, and distribute this software and its documentation for any purpose and without fee is hereby granted, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation. This software is provided "as is" without express or implied warranty.