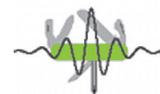




SoX – Sound eXchange

Benjamin.Barras@epfl.ch, EPFL - Domaine IT - responsable de la plate-forme Unix à l'EPFL



SoX – Sound eXchange, the Swiss Army knife of sound processing programs in command line.

SoX - Sound eXchange, le couteau suisse pour la manipulation des fichiers audio numériques en ligne de commande.

Fiche descriptive

SoX		
Domaine		
◆ Manipulation de fichiers audio numériques		
Licence	langue	version
◆ GPL	◆ anglais	◆ 14.4.0
Autres alternatives libres		
◆ ffmpeg (audio et vidéo)		
◆ Audacity (interface graphique)		
Sites Web		
◆ Projet et téléchargement: sox.sourceforge.net		
Plates-formes		
  		

Introduction

Nous connaissons bien **ImageMagick [1]** qui est l'outil magique pour la manipulation d'image. Je vous propose de découvrir **SoX [2]** qui est l'outil magique pour la manipulation des fichiers audio numériques en ligne de commande.

La commande `sox` permet de passer d'un format audio numérique à un autre avec une grande facilité. Par exemple, pour passer d'un fichier au format **WAV** à un autre fichier au format **FLAC**, la commande sera celle-ci:

```
sox monFichier.wav monFichier.flac
```

Difficile de faire plus simple et vous l'aurez compris, c'est l'extension du fichier qui fixe le format. Mais le format du fichier est fixé, dans l'ordre, de la manière suivante:

- 1 le format est donné à la ligne de commande à l'aide de l'option `-t`
- 2 le format est contenu dans l'en-tête du fichier.
- 3 le format est donné par l'extension du fichier.

Plus besoin de chercher le logiciel qui transforme tel ou tel fichier en tel ou tel format, SoX est là pour cela.

Installation

Avant d'aller plus loin, quelques informations concernant l'installation de ce logiciel:

Ubuntu: `sudo apt-get install sox`;

RedHat/Fedora: `yum install sox`;

Windows: on trouvera la version compressée `sox-<version>-win32.zip [3]` qu'il faut simplement déplier, ou prendre le binaire `sox-<version>-win32.exe` qu'il suffit d'exécuter;

MacOSX: télécharger et déplier la version compressée `sox-<version>-macosx.zip [3]`.

Formats

La question que l'on peut légitimement se poser est de savoir quels formats sont disponibles. Pour ceci, on utilisera la commande suivante:

```
sox --help
```

et en filtrant quelque peu le résultat, on trouvera la liste complète des formats supportés par `sox` sur votre machine:

```
AUDIO FILE FORMATS: 8svx aif aifc aiff aifc a1
amb au avr caf cdda cdr cvs cvsd dat dvms f4 f8
fap flac fssd gsm
hcom htk ima ircam la lpc lpc10 lu mat mat4 mat5
maud nist ogg paf prc pvf raw s1 s2 s3 s4 sb sd2
sds sf sl smp
snd sndfile sndr sndt sou sox sph sw txw u1 u2 u3
u4 ub ul uw vms voc vorbis vox w64 wav wavpcm wv
wve xa xi
```

Voici une commande qui permet de voir d'une autre manière tous les formats que l'on peut utiliser sur sa machine:

```
sox --help-format all
```

qui nous donnera une liste très complète du style:

```
...
Format: cdda
Description: Red Book Compact Disc Digital Audio
Also handles: cdr
Channels restricted to: stereo
Sample-rate restricted to: 44100
Reads: yes
Writes:
  16-bit Signed Integer PCM (16-bit precision)
...
Format: dat
Description: Textual representation of the
sampled audio
Reads: yes
Writes:
  Floating Point (text) PCM (53-bit precision)
```

SoX – Sound eXchange

```
...
Format: flac
Description: Free Lossless Audio CODEC compressed
audio
Reads: yes
Writes:
  8-bit FLAC (8-bit precision)
 16-bit FLAC (16-bit precision)
 24-bit FLAC (24-bit precision)
...
```

On remarquera la multitude de formats disponibles, mais un œil avisé aura remarqué l'absence du format *MP3* [4] en raison d'un brevet sur ce dernier. Mais on pourra y ajouter les formats supportés, dont la liste complète se trouve sur le site officiel de *SoX* [5], soit:

- 1 en compilant nous-mêmes une version personnalisée de *SoX*, en vérifiant préalablement que les bibliothèques du format souhaité se trouvent déjà sur notre machine;
- 2 en installant le paquet *libsox-fmt-all* (*SoX* format libraries) que l'on trouvera facilement dans son dépôt (Ubuntu) ou à l'aide d'un moteur de recherche.

Pour avoir à disposition le format Mp3

LAME (Lame Ain't an Mp3 Encoder)

C'est la librairie LAME [6] qui nous permet d'avoir le format MP3, librairie qui est une description d'un encodeur MP3 et pas un encodeur MP3, ceci afin de ne pas violer le brevet.

Important

Ce qu'il faut savoir et qu'il est important de bien comprendre, c'est la position des paramètres de cette commande ainsi que ses différentes options, cela vous évitera de mauvaises surprises. Le format de la commande est le suivant:

```
sox [global-options] [format-options] infile1
[[format-options] infile2] ...
[format-options] outfile [effect [effect-
options]]
```

La commande *sox* doit toujours avoir un fichier *infile* et un fichier *outfile*, si l'on ne désire pas avoir un tel fichier, on remplacera le nom du fichier par les options *-n* (*null file*) ou *-d* (*default audio device*).

Il y a trois types d'options:

global-options qui se placent juste après la commande.

format-options qui se placent devant le nom du fichier.

effect-options après une commande demandant un effet, effet qui se trouve par ailleurs toujours après le **outfile**.

On fera un **sox --help** ou **man sox** pour connaître tous les détails de ces différents paramètres.

Concernant les effets que l'on peut utiliser, une commande à connaître:

```
sox --help-effect all
```

et qui nous permettra de voir toutes les subtilités possibles avec cet outil, par exemple:

```
sox --help-effect synth
```

nous donnera ceci:

```
sox:      SoX v14.4.0

Effect usage:

synth [-j KEY] [-n] [length [offset [phase [p1
[p2 [p3]]]]]]]
      {type [combine] [[%]freq[k][:|+|/|- [%]
freq2[k]] [offset [phase [p1 [p2 [p3]]]]]]}
```

qui donne toutes les options, ainsi que leurs positions, pour l'effet **synth**. Cet effet consiste à synthétiser un son avec une multitude de paramètres et dont l'ordre, je me permets d'insister, est très important. Vous remarquerez qu'un effet, ainsi que ses paramètres, sera toujours disposé en fin de ligne de commande, soit après *outfile*. On peut cumuler les effets en les disposant, avec leurs propres paramètres, les uns après les autres. Par exemple, pour générer un son qui nous sera utile par la suite, nous utiliserons une commande du style:

```
sox -c 1 -r 44100 -n sin4k.flac synth 1 sine 4k \
vol -10dB
```

Nous voyons que cette commande contient deux options globales qui sont *-c 1* (1 canal) et *-r 44100* (fréquence d'échantillonnage de 44.1 kHz). Comme il n'y a pas de fichier en entrée puisque l'on génère un son, on mettra *-n* pour *infile*. Le nom du fichier *outfile* se place directement après le *infile* dont l'extension nous donne le format d'encodage. Deux effets sont présents, et leurs options associées, l'un *synth* (sinus d'une seconde et de 4 kHz) et l'autre *vol* (réduction du volume de 10dB).

Voir ses fichiers audionumériques

Il est vrai que les choses simples se font de plus en plus rares, et chose remarquable que l'on oublie peu à peu, est que l'on peut voir ses fichiers audionumériques, puisqu'en fin de compte, il ne s'agit que d'échantillons numériques. Si vous voulez voir à quoi ressemble un fichier audionumérique, vous pouvez utiliser le format *DAT* comme ceci:

```
sox Ratatouille.mp3 Ratatouille.dat
```

Ensuite, il ne reste plus qu'à éditer votre fichier, et vous y trouverez ceci:

```
; Sample Rate 44100
; Channels 2
      0          0          0
2.2675737e-05  0          0
4.5351474e-05  0          0
6.8027211e-05  0          0
9.0702948e-05  0          0
...
0.20011338    0.0010986328  0.0010681152
0.20013605    0          0
0.20015873    -0.0013122559  -0.0013427734
0.20018141    -0.0018920898  -0.001953125
0.20020408    -0.0014648438  -0.0014953613
0.20022676    -0.00057983398 -0.00061035156
0.20024943    3.0517578e-05  0
0.20027211    0.00036621094  0.00033569336
...
```

où l'on retrouve la fréquence d'échantillonnage de 44,1 kHz. La première colonne correspond au temps ($1/44100 = 2.2675737e-$

SoX – Sound eXchange

05), et les deux colonnes suivantes correspondent aux échantillons (normalisés à 1) des canaux gauche et droit. On utilise ce format pour construire des graphiques ou pour faire une transformée de Fourier rapide (*FFT, Fast Fourier Transform*). Cela nous permet de voir concrètement un fichier audionumérique de manière simple et élégante.

Note historique

D'où provient ce fameux 44,1 KHz, tout bon ingénieur connaît bien le critère de Shanon qui dit que la fréquence d'échantillonnage doit être, au moins le double de la fréquence maximale, qui pour nos oreilles sera d'environ 20 KHz.

Historiquement [7] cela provient de l'idée d'utiliser des bandes magnétiques vidéos pour y stocker, dans les trames vidéos, des échantillons de son numérique. Pour les systèmes à 50 Hz, nous avons 625 lignes dont 37 lignes cachées. Il faut tenir compte des lignes impaires et paires, ce qui nous fait: $(625 - 37) / 2 = 294$ lignes utiles. Pour les systèmes à 60 Hz, nous avons 525 lignes dont 35 lignes cachées, cela nous donne $(525 - 35) / 2 = 245$ lignes utiles. Donc, si l'on met 3 échantillons numériques (le binaire est représenté avec le blanc et le noir) par ligne utile, cela nous donne

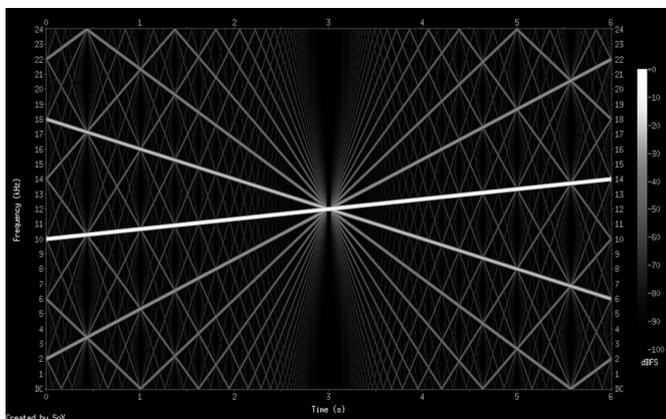
$$50 * 294 * 3 = 44100$$

$$60 * 245 * 3 = 44100$$

Une autre manière de voir ses fichiers audionumériques est d'en faire un spectrogramme afin de visualiser le spectre de fréquence. Une commande amusante pour voir un fichier sous forme spectrale sera:

```
sox -n -n synth 6 tri 10k:14k spectrogram \
-z 100 -w kaiser
```

qui consiste à générer un signal triangulaire qui commence à 10kHz et qui finit à 14kHz, le tout en 6 secondes. L'effet *spectrogramme* va nous générer une image (*spectrogram.png*) avec 100 dB de contraste (axe z) et utilise une fenêtre de type *kaiser* pour le filtrage. Cela nous donne cette très belle image:



Filtrer un fichier audionumérique

Ici aussi, on dispose de plusieurs filtres dont les plus utiles sont le passe-bas, le passe-bande et le passe-haut. Mais avant de faire du

filtrage, on va générer quelques fichiers que l'on va mixer par la suite, afin d'apprendre à utiliser notre outil. Voyons cela:

```
# Sinus 2kHz
sox -c 1 -r 44100 -n sin2k.flac synth 1 sine 2k

# Sinus 8kHz
sox -c 1 -r 44100 -n sin8k.flac synth 1 sine 8k

# Sinus 18kHz
sox -c 1 -r 44100 -n sin18k.flac synth 1 sine 18k
```

On va mixer le tout pour n'avoir qu'un seul fichier, cela d'une manière très élégante:

```
sox -m sin2k.flac sin8k.flac sin18k.flac \
result.flac
```

Il faut savoir que l'amplitude est la somme des amplitudes prises individuellement divisée par trois. Le résultat est un fichier audio-numérique qui contient trois fréquences (2, 8 et 18 kHz). Vérifions-le à l'aide d'un spectrogramme:

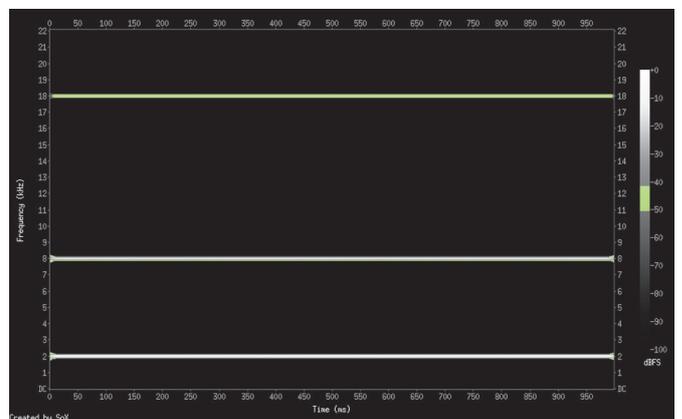
```
sox result.flac -n spectrogramme -z 100 -w kaiser
```

ce qui nous donne l'image suivante:



Nous allons effectuer un filtrage passe-bas afin de couper les fréquences dépassant 6 kHz et refaire un spectrogramme pour vérifier le résultat, soit:

```
sox result.flac filterResult.flac lowpass 6000
sox filterResult.flac -n spectrogramme -z 100 \
-w kaiser
```



où nous voyons clairement que la fréquence de 18kHz est nettement affaiblie.

Un autre exemple, sur un système *UNIX* on peut générer un signal aléatoire de manière très simple, comme ceci:

SoX – Sound eXchange

```
cat /dev/urandom | sox -t raw -r 44100 -b 16 \  
-e unsigned-integer - urandom.dat vol -10dB \  
trim 0 00:03
```

où nous avons le format raw (-t raw), l'échantillonnage à 44.1 kHz (-r 44100) et des entiers non signés (-e unsigned-integer) de 16 bits (-b 16). Le signal arrive en entrée standard (- pour infile), le tout rangé dans un fichier (urandom.dat) dont le volume est réduit de 10 dB (vol -10dB) et coupé au bout d'une durée de 3 secondes (trim 0 00:03). Ce type de signal est très utile pour générer ce que l'on appelle un *bruit blanc*, c'est un signal qui contient toutes les fréquences avec la même amplitude. Vous pouvez le vérifier à l'aide de la commande suivante:

```
sox urandom.dat -n spectrogramme -z 100 \  
-w kaiser -X 200 \  
display spectrogram.png
```

Écouter ses fichiers audio numériques

SoX possède une commande play pour écouter ses morceaux de musique préférés. Cela se fait simplement comme ceci:

```
play monMorceauDeMusique.flac  
ou  
sox monMorceauDeMusique.flac -d
```

Il n'y a pas besoin de préciser le format du fichier, il est donné par l'extension du fichier. Si l'on souhaite plus d'information sur un morceau de musique, on pourra utiliser la commande suivante:

```
sox --i Ratatouille.mp3
```

dont le contenu est le suivant:

```
Input File:      'Ratatouille.mp3'  
Channels:       2  
Sample Rate:    44100  
Precision:      16-bit  
Duration:       00:03:34.91 = 9477487 samples =  
                16118.2 CDDA sectors  
File Size:      3.44M  
Bit Rate:       128k  
Sample Encoding: MPEG audio (layer I, II or III)  
Comments:  
Title=Ratatouille  
Artist=Sonalp  
Album=Le Son Des Alpes  
Tracknumber=8/8  
Year=2002  
Genre=80  
Discnumber=1/1
```

Enregistrer des fichiers audio numériques

SoX possède également une commande rec pour enregistrer son entrée numérique. Cela se fait simplement comme ceci:

```
rec -c 1 monEnregistrement.flac trim 0 00:03  
ou  
sox -c 1 -d monEnregistrement.flac trim 0 00:03
```

dont l'effet *trim* consiste à couper l'enregistrement, ici ce sera 3 secondes, et le format est donné par l'extension du fichier.

Conclusion

J'espère vous avoir fait découvrir quelques facettes intéressantes de SoX qui mérite bien son appellation de couteau suisse. Reste que les options non explorées dans cet article ne sont utiles qu'à des spécialistes qui savent manipuler du son. Je pense que la chose la plus intéressante à retenir est la conversion de fichiers audio numériques.

Références

- [1] ImageMagick, www.imagemagick.org/script/index.php & Flash informatique no 8/2011, flashinformatique.epfl.ch/spip.php?article2184
- [2] SoX, sox.sourceforge.net
- [3] sourceforge.net/projects/sox/files/sox
- [4] MP3, fr.wikipedia.org/wiki/MPEG-1/2_Audio_Layer_3
- [5] Formats supportés par SoX, sox.sourceforge.net/soxformat.html
- [6] LAME Ain't an MP3 Encoder, fr.wikipedia.org/wiki/LAME
- [7] WATKINSON, John. *An Introduction to Digital Audio*. Focal Press. ISBN 0-24-051643-5.



Article du FI-EPFL 2012 sous licence CC BY-SA 3.0 / B. Barras

Le DIT vous forme

Toute l'équipe des cours du DIT vous souhaite de joyeuses fêtes de fin d'année ainsi qu'une **année 2013** riche en formation. Notre programme complet pour les 6 prochains mois se trouve dans notre site [Web dit.epfl.ch/formation](http://Web.dit.epfl.ch/formation).

cours.dit@epfl.ch