Tuto / commandes pour entrainer un wake-word

I - Trouver plein de bruits

Extraire des bandes-sons de film:

```
ffmpeg -i /data/films/La\ grande\ vadrouille.avi -acodec pcm_s16le -ar 16000 -ac 1 -f wav la-grande-vadrouille.wav
ffmpeg -i /data/films/Kill\ Bill\ Vol\ 1\ \[1080p\]\ MULTi\ 2003\ BluRay\ x264-Pop\ .mkv -acodec pcm_s16le -ar 16000 -ac 1 -f wav kill-bill.wav
ffmpeg -i /data/films/Harry\ Potter\ 4\ Et\ La\ Coupe\ De\ Feu\ 2005\ \[1080p\]\ MULTI\ 2004\ Bluray\ X264\ lorraines.mkv -acodec pcm_s16le -ar 16000 -ac 1 -f wav hp-4.wav
ffmpeg -i /data/films/Harry\ Potter\ 1\ A\ L\'Ecole\ Des\ Sorciers\ \[1080p\]\ MULTI\ 2001\ Bluray\ X264\ lorraines.mkv -acodec pcm_s16le -ar 16000 -ac 1 -f wav hp-1.wav
ffmpeg -i /data/films/Hunger\ Games\ 1.mkv -acodec pcm_s16le -ar 16000 -ac 1 -f wav hunger-games-1.wav
ffmpeg -i /data/films/Pulp\ Fiction.avi -acodec pcm_s16le -ar 16000 -ac 1 -f wav pulp-fiction.wav
ffmpeg -i /data/films/DRAGONS\ \(2010\).avi -acodec pcm_s16le -ar 16000 -ac 1 -f wav dragon.wav
ffmpeg -i /data/films/DRAGONS\ \(2010\).avi -acodec pcm_s16le -ar 16000 -ac 1 -f wav galdiator.wav
ffmpeg -i /data/films/Gladiator\ 2000\ 10th\ Anniversary\ Edition\ MULTi\ TrueFrench\ 1080p\ HDLight\ x265\ HEVC.GHT.mkv -acodec pcm_s16le -ar 16000 -ac 1 -f wav galdiator.wav
ffmpeg -i /data/films/Alice\ au\ pays\ des\ merveilles.avi -acodec pcm_s16le -ar 16000 -ac 1 -f wav alice.wav
```

Télécharger des sons de pluie, et de bruit exterieur depuis Youtube:

```
youtube-dl -x "https://www.youtube.com/watch?v=mPZkdNFkNps"
youtube-dl -x "https://www.youtube.com/watch?v=xNN7iTA57jM"
youtube-dl -x "https://www.youtube.com/watch?v=BOdLmxy06H0"
youtube-dl -x "https://www.youtube.com/watch?v=BOdLmxy06H0"
youtube-dl -x "https://www.youtube.com/watch?v=cO_IFJaWmhA"
```

Télécharger des sons de pubs et de voiture depuis soundarchive

Télécharger des podcasts

J'ai utilisé l'appli de podcast de mon telephone: AntennaPod.

- telecharger le podcast
- Appui long sur la piste -> partager -> envoi avec kde connect

J'ai 20 fichiers qui font chacun entre 30 minutes et 1h30 de long.

Récuperer de la musique:

J'ai copié 140 pistes de differents style d'albums (j'ai mis une grosse part de chanson francaise).

Récuperer des wake-word opensource:

https://github.com/MycroftAI/precise-community-data

j'ai pris ceux qui ressemblaient au miens, à savoir:

- athena
- amelia

II - Enregistrer plein de wake word

- j'en ai enregistré une centaine via le micro de la reconnaissance vocal en me mettant à plein d'endroit dans ma chambre, avec `precise-collect`
- j'en ai enregistré ensuite 10 par personnes volontaire via le micro de mon pc:
  - Ma famille (6 personnes)
  - Des amis (3 personnes)
- J'en ai enregistré une bonne 50aine, avec les micros de differents pcs

III - Nettoyer les enregistrements de wake word

J'ai utilisé audacity, et je l'ai fait à la main (ptètre qu'avec des script ca se fait aussi).

En gros, j'ai verifié qu'il n'y ai que le mot-clé dans l'enregistrement, et j'ai coupé si il y avait trop de silence au début ou à la fin.

Ca me fait des enregistrement qui font entre 0.8s et 2s, avec un peu de blanc au début et à la fin.

IV - Convertir et découper les source de bruit

Convertir en wav:

```
SOURCE_DIR=music
DEST_DIR=music-wav/
# selon les extensions de fichier, changez *.mp3 par *.<extension> (youtube-dl télécharge en m4a)
for i in $SOURCE_DIR/*.mp3; do echo "Converting $i..."; fn=${i##*/}; ffmpeg -i "$i" -acodec pcm_s16le -ar 16000 -ac 1 -f wav "$DEST_DIR/${fn%.*}.wav"; done
for i in $SOURCE_DIR/*.flac; do echo "Converting $i..."; fn=${i##*/}; ffmpeg -i "$i" -acodec pcm_s16le -ar 16000 -ac 1 -f wav "$DEST_DIR/${fn%.*}.wav"; done
```

Decouper par tranche de 30s:

```
for f in *.wav; do sox "$f" "split.$f" trim 0 30 : newfile : restart ; done
```

V - Augmenter artificiellement vos mots-clés:

Il faut d'abord choisir certaines partie du bruit précedents qu'on va ajouter par dessus nos enretgistrement de mots-clé J'ai pris:

- Les videos youtubes
- Le son de soundarchive
- Et des morceaux de classique

Pour être tranquille, on les normalise avec sox:

```
SOURCE_DIR=noise-short/
DEST_DIR=noise-norm/
```

```
for i in $SOURCE_DIR/*.wav; do echo "Converting $i..."; fn=${i##*/}; sox --norm "$i"
"$DEST_DIR/${fn%.*}.wav"; done
```

Pis pour chaque groupe, on va les appliquer avec ce script sur nos wake-word (vous pouvez changer le coeff 0.1 par 0.05 ou 0.2 ou 0.3 selon vos enregistrement):

```
#!/bin/bash

NOISEDIR=$1

DIR="$( cd "$( dirname "${BASH_SOURCE[0]}" )" >/dev/null 2>&1 && pwd )"

for f in *.wav
do
    NOISEFILE=$(find ${NOISEDIR} -type f | shuf -n 1)

    sox -m -v 0.9 $f -v 0.1 ${NOISEFILE} noise.$f trim 0 `soxi -D $f`
done
```

Normalement après ca on a triplé nos wake-words.

VI - Installer mycroft precise:

https://github.com/MycroftAI/mycroft-precise#source-install

1. il faut python 3.6 (pour pas se prendre la tète installer ubuntu 18.04)
2. il y a un bug dans le requirements.txt, après l'installation: `pip install h5py<3.0.0`

VII - Entrainement:

- Mettre 80/90% de vos wake-word dans le dossier wake-word et le reste dans test/wake-word
- `precise-train yoda1.net maitre_yoda/ -e 100 -s 0.5`
- Faire un dossier avec tout vos noise puis `precise-train-incremental yoda1.net maitre_yoda/ -r path/to/noise-folder -e 50 -th 0.4 -s 0.5`
- Attendre une bonne heure avec un pc de gamer de l'an 2019

IIX - Profitez !

Révision #3
Créé 2020-12-20 17:33:04 UTC par tjiho
Mis à jour 2021-03-05 23:41:56 UTC par tjiho