

Dejavu: audio fingerprinting and recognition in Python

Audio fingerprinting and recognition algorithm implemented in Python, see the explanation here:

[How it works](#)

Dejavu can memorize audio by listening to it once and fingerprinting it. Then by playing a song and recording microphone input, Dejavu attempts to match the audio against the fingerprints held in the database, returning the song being played.

Note that for voice recognition, Dejavu is not the right tool! Dejavu excels at recognition of exact signals with reasonable amounts of noise.

Installation and Dependencies:

Read [INSTALLATION.md](#)

Setup

First, install the above dependencies.

Second, you'll need to create a MySQL database where Dejavu can store fingerprints. For example, on your local setup:

```
[crayon-6a67d0f188d23676065505/]
```

Now you're ready to start fingerprinting your audio collection!

Quickstart

```
[crayon-6a67d0f188d29211605727/]
```

Fingerprinting

Let's say we want to fingerprint all of July 2013's VA US Top 40 hits.

Start by creating a Dejavu object with your configurations settings (Dejavu takes an ordinary Python dictionary for the settings).

```
[crayon-6a67d0f188d2b023787791/]
```

Next, give the `fingerprint_directory` method three arguments:

- input directory to look for audio files
- audio extensions to look for in the input directory
- number of processes (optional)

```
[crayon-6a67d0f188d2c547195885/]
```

For a large amount of files, this will take a while. However, Dejavu is robust enough you can kill and restart without affecting progress: Dejavu remembers which songs it fingerprinted and converted and which it didn't, and so won't repeat itself.

You'll have a lot of fingerprints once it completes a large folder of mp3s:

```
[crayon-6a67d0f188d2e313678415/]
```

Also, any subsequent calls to `fingerprint_file` or `fingerprint_directory` will fingerprint and add those songs to the database as well. It's meant to simulate a system where as new songs are released, they are fingerprinted and added to the database seamlessly without stopping the system.

Example

Fingerprint files of extension mp3 in the `./mp3` folder

```
[crayon-6a67d0f188d2f304112309/]
```

Run a test suite on the `./mp3` folder by extracting 1, 2, 3, 4, and 5 second clips sampled randomly

[crayon-6a67d0f188d31319607467/]

Results

```
results — -bash — 165x38
DEBUG:root:Running Dejavu fingerprinter on files in ./mp3...
DEBUG:root:-----
DEBUG:root:file: The-Lights-Galaxia--While-She-Sleeps_132_4sec.mp3
DEBUG:root:song: The-Lights-Galaxia--While-She-Sleeps
DEBUG:root:song_result: The-Lights-Galaxia--While-She-Sleeps
DEBUG:root:correct match
DEBUG:root:query duration: 0.921
DEBUG:root:confidence: 33
DEBUG:root:song start_time: 132
DEBUG:root:result start time: 148.0
DEBUG:root:inaccurate match
DEBUG:root:-----
DEBUG:root:file: Sean-Fournier--Falling-For-You_49_1sec.mp3
DEBUG:root:song: Sean-Fournier--Falling-For-You
DEBUG:root:song_result: Sean-Fournier--Falling-For-You
DEBUG:root:correct match
DEBUG:root:query duration: 0.296
DEBUG:root:confidence: 34
DEBUG:root:song start_time: 49
DEBUG:root:result start time: 49.0
DEBUG:root:accurate match
DEBUG:root:-----
DEBUG:root:file: The-Lights-Galaxia--While-She-Sleeps_156_1sec.mp3
DEBUG:root:song: The-Lights-Galaxia--While-She-Sleeps
DEBUG:root:song_result: The-Lights-Galaxia--While-She-Sleeps
DEBUG:root:correct match
DEBUG:root:query duration: 0.253
DEBUG:root:confidence: 60
DEBUG:root:song start_time: 156
DEBUG:root:result start time: 156.0
DEBUG:root:accurate match
DEBUG:root:-----
DEBUG:root:-----
```

More information

<https://github.com/worldveil/dejavu>