Классификация музыки по жанрам

Кантор Виктор

Проблема

Захотели послушать джаз, открываем плейлист в vk или папку с музыкой, а там…

Imagine Dragons – Selene	4:04
Imagine Dragons – Radioctive	3:08
Imagine Dragons – Paradise	4:38
Imagine Dragons – Demons	3:35
Metallica – The Day That Never Comes	7:56
ESC 2014 - Austria – Conchita Wurst - Rise Like A Phoenix	3:05
Bon Jovi – Someday I'll Be Saturday Night	4:39
Бумбокс – Летний Дождь	3:19
Бумбокс – Сандали	4:42
Бумбокс – Полина	4:21
Бумбокс – Пошла вон	3:25
Бумбокс – Скажи как мне жить	3:24
Marty Robbins – Big iron	3:56
Jack Shaindlin – Let's go sunning	1:41
Frank Sinatra – Blue Moon	2:51
The Ink Spots – I Don't Want To Set The World On Fire	3:04
Peggy Lee – Why Don't You Do Right	2:27
The Ink Spots – Maybe	2:50
John Lee Hooker – Boom Boom Boom	2:44

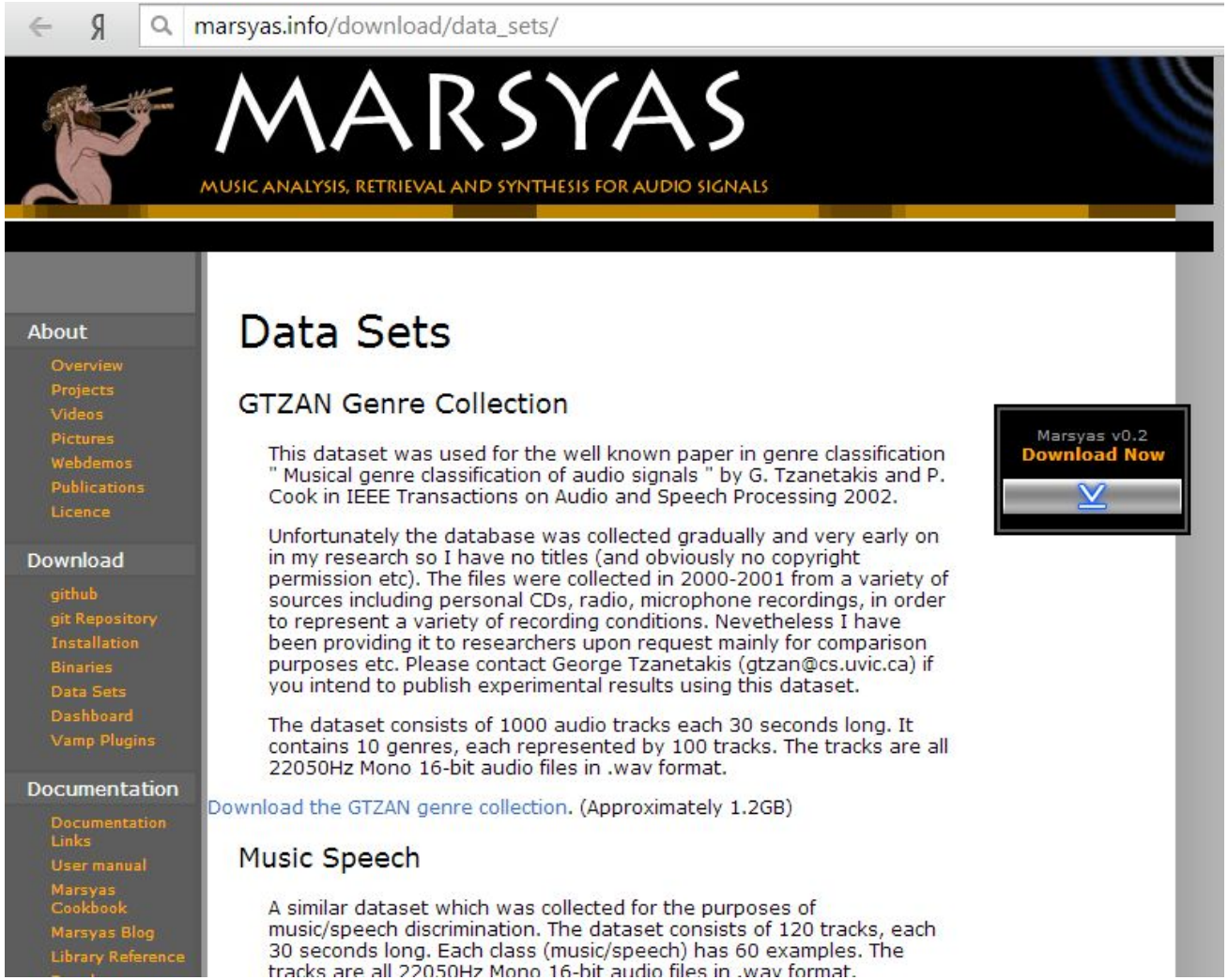
Имя	Дата изменения	Тип	Размер
The Piano Guys – Happy end (Piano C...	05.07.2013 12:05	Файл "MP3"	7 509 КБ
The Rolling Stones – Rain Fall Down	30.06.2013 11:49	Файл "MP3"	7 362 КБ
The White Stripes – Ball and Biscuit	22.06.2013 18:59	Файл "MP3"	10 247 КБ
This Will Destroy You – A Three-Legge...	21.12.2013 17:56	Файл "MP3"	17 228 КБ
This Will Destroy You – Quiet	21.12.2013 17:56	Файл "MP3"	7 604 КБ
Tiarah – Diane (Therapy! cover)	25.04.2014 13:02	Файл "MP3"	9 195 КБ
Toe – Goodbye	24.06.2013 10:31	Файл "MP3"	9 978 КБ
Toe – Kodoku no Hatsumei	24.06.2013 10:31	Файл "MP3"	5 556 КБ
Toe – Long Tomorrow	24.06.2013 10:31	Файл "MP3"	14 141 КБ
toe – Ordinary Days	24.06.2013 10:31	Файл "MP3"	7 100 КБ
Toe – Path	24.06.2013 10:31	Файл "MP3"	5 778 КБ
Toe – The Future Is Now	24.06.2013 10:31	Файл "MP3"	7 725 КБ
toe – You Go	24.06.2013 10:32	Файл "MP3"	6 221 КБ
Tony Hawk's Pro Skater 2 – Track 02	25.04.2014 13:08	Файл "MP3"	4 382 КБ
Tony Hawk's Pro Skater 2 – Track 10	25.04.2014 13:09	Файл "MP3"	7 049 КБ
Tony Hawk's Pro Skater 4 OST – 3	25.04.2014 13:07	Файл "MP3"	4 518 КБ
Trent Reznor and Atticus Ross – In Mo...	22.06.2013 18:59	Файл "MP3"	11 599 КБ
Trent Reznor and Atticus Ross – Intrig...	22.06.2013 18:59	Файл "MP3"	10 334 КБ
Uma2rman – А знаешь, всё еще буде...	11.03.2014 2:44	Файл "MP3"	3 204 КБ
Vorony_stereo	16.03.2014 22:31	Файл "MP3"	3 679 КБ
Woodkid – I Love You	07.11.2013 7:49	Файл "MP3"	9 064 КБ
Woodkid – Iron	26.06.2013 15:16	Файл "MP3"	6 139 КБ
Woodkid – Run Boy Run	26.06.2013 15:16	Файл "MP3"	8 384 КБ
Woodkid – The Great Escape	26.06.2013 15:16	Файл "MP3"	7 730 КБ
Woodkid – Wasteland	26.06.2013 15:17	Файл "MP3"	5 627 КБ
Zero 7 – In the Waiting Line (OST Hou...	19.11.2013 13:41	Файл "MP3"	7 382 КБ
Алина Орлова – Menulis (1)	09.08.2013 22:08	Файл "MP3"	3 793 КБ
Алина Орлова – Menulis	01.08.2013 15:47	Файл "MP3"	3 793 КБ
Би-2 – Падает снег (feat. Ю. Чичерин...	11.03.2014 2:46	Файл "MP3"	5 712 КБ

…одна большая свалка :)

Несколько светлых идей

- Навести порядок
- Не делать это вручную для 1000+ аудиозаписей
- Определение жанра – в чистом виде задача классификации
- Остается придумать признаки
- По запросам в гугле про классификацию музыки узнаем, что область имеет название Music Information Retrieval
- Находим много интересных статей :)

Сначала найдем датасет



http://marsyas.info/download/data_sets/

Что в датасете

- Десять жанров
- По 100 wav-файлов с первыми 30 секундами аудиозаписи для каждого жанра
- Один канал

Как посмотреть на датасет

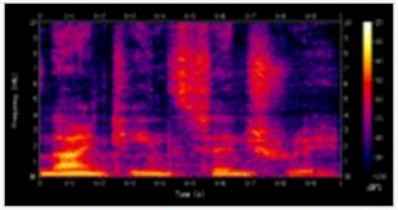
Гуглим, как люди визуализируют звук и сразу узнаем про спектрограммы. Вот, что пишет про это Вики:

Спектрограмма

Материал из Википедии — свободной энциклопедии

[править | править исходный текст]

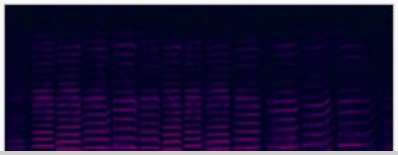
Спектрограмма (**сонограмма**) — изображение, показывающее зависимость спектральной плотности мощности сигнала от времени. Спектрограммы применяются для идентификации речи, анализа звуков животных, в различных областях музыки, радио- и гидролокации, обработке речи, сейсмологии и в других областях.



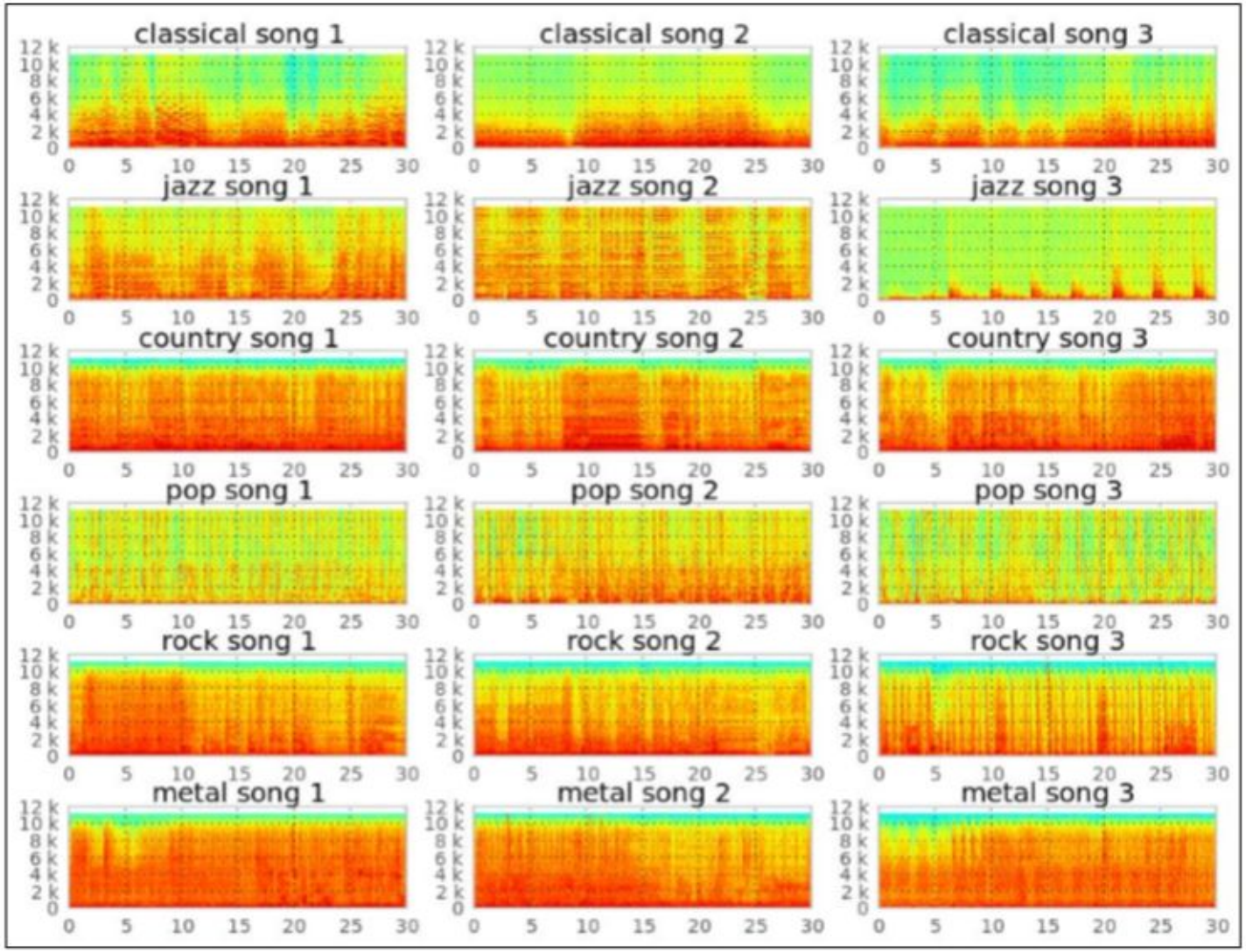
Спектрограмма мужского голоса

Представление [править | править исходный текст]

Наиболее распространенным представлением спектрограммы является двумерная диаграмма: на горизонтальной оси представлено время, по вертикальной оси — частота; третье измерение с указанием амплитуды на определенной частоте в конкретный момент времени представлено интенсивностью или цветом каждой точки изображения.



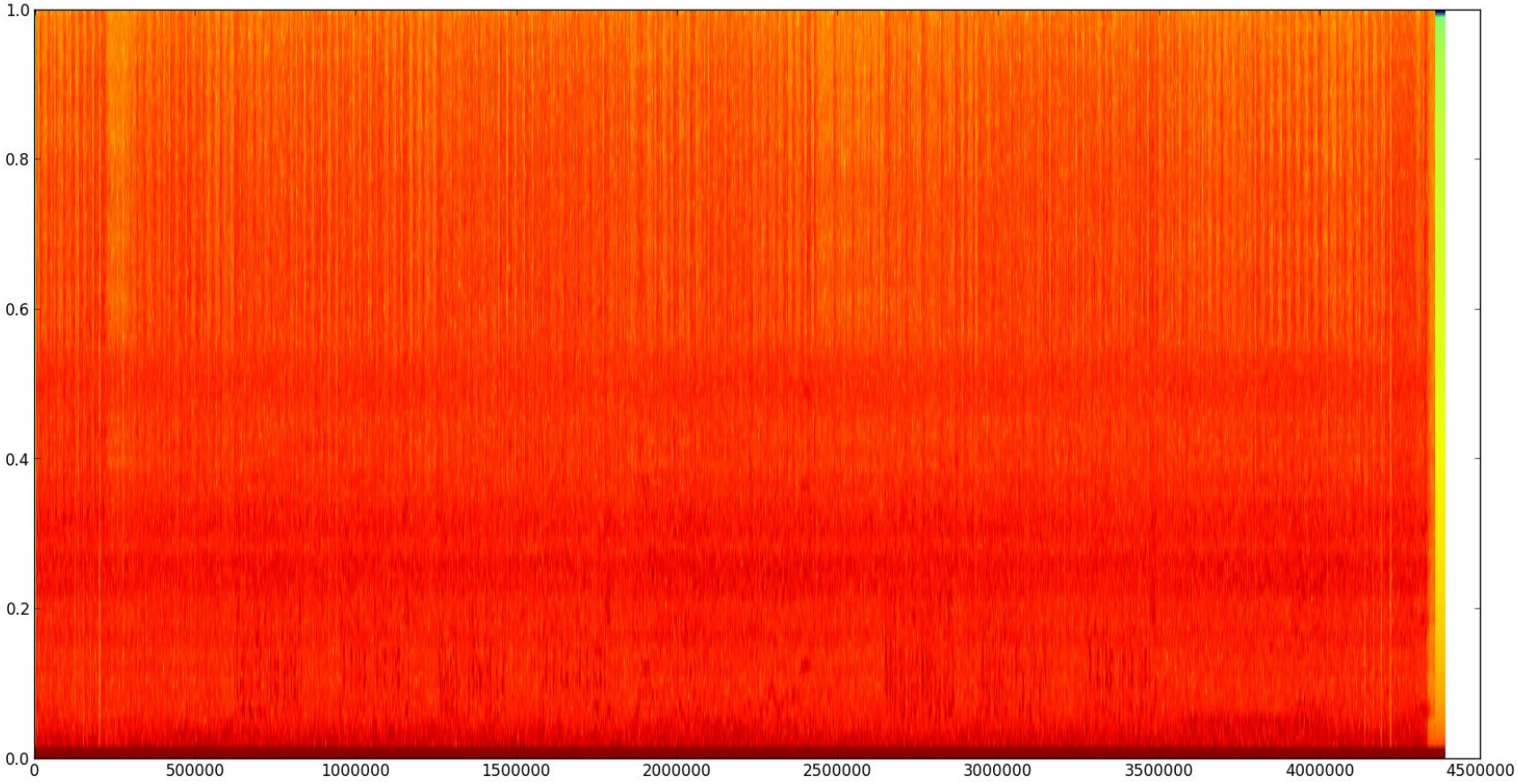
Спектрограммы для разных жанров



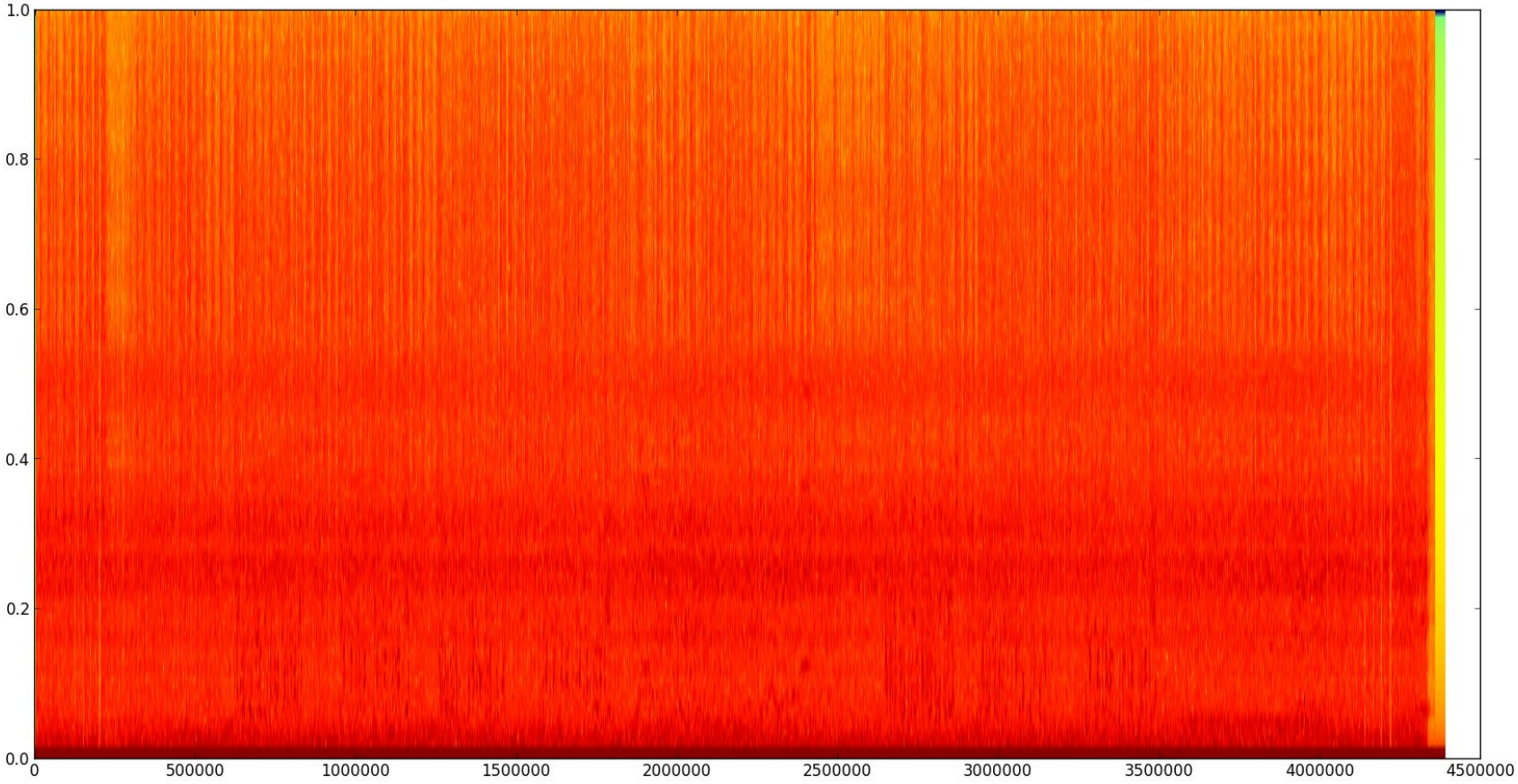
Как построить?

```
>>> import scipy
>>> from matplotlib.pyplot import specgram
>>> sample_rate, X = scipy.io.wavfile.read(wave_filename)
>>> print sample_rate, X.shape
22050, (661794,)
>>> specgram(X, Fs=sample_rate, xextent=(0,30))
```

Какой жанр?


1.0
0.8
0.6
0.4
0.2
0.0
0
500000
1000000
1500000
2000000
2500000
3000000
3500000
4000000
4500000

Какой жанр?

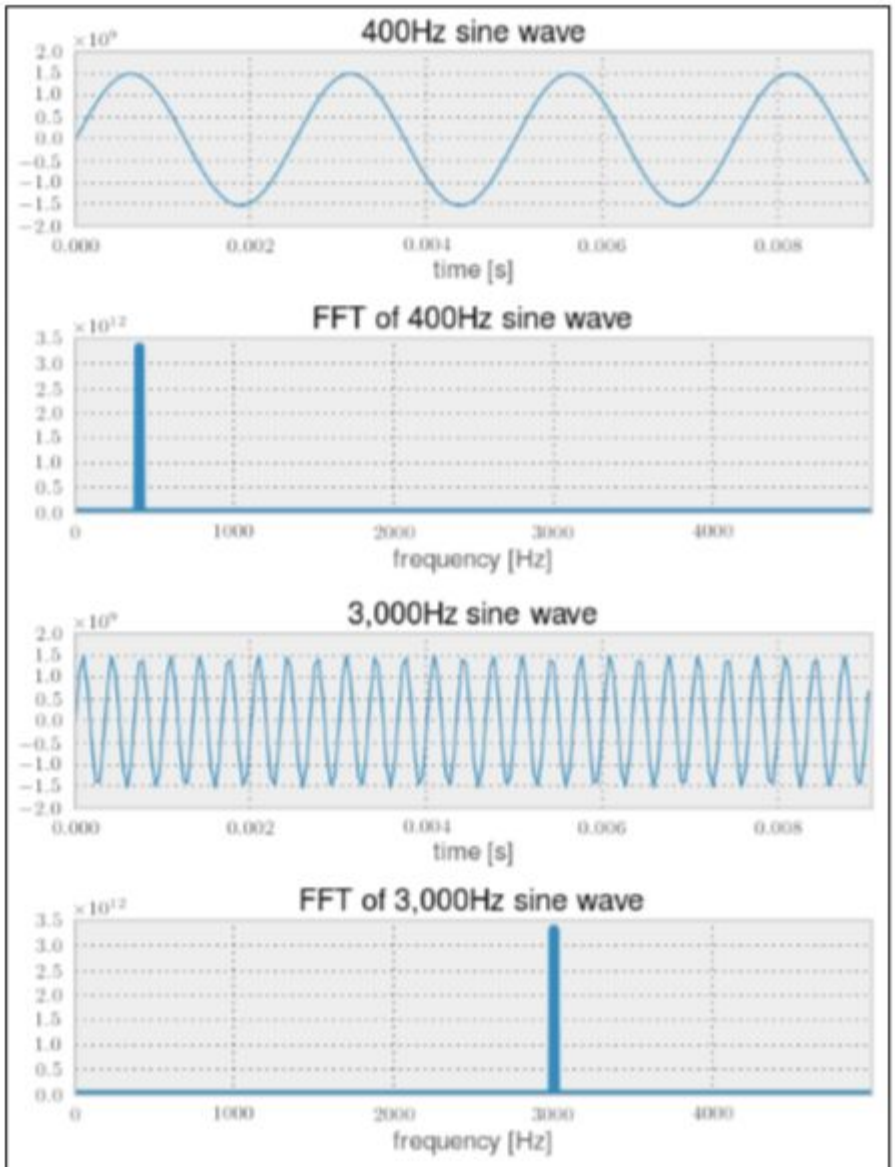


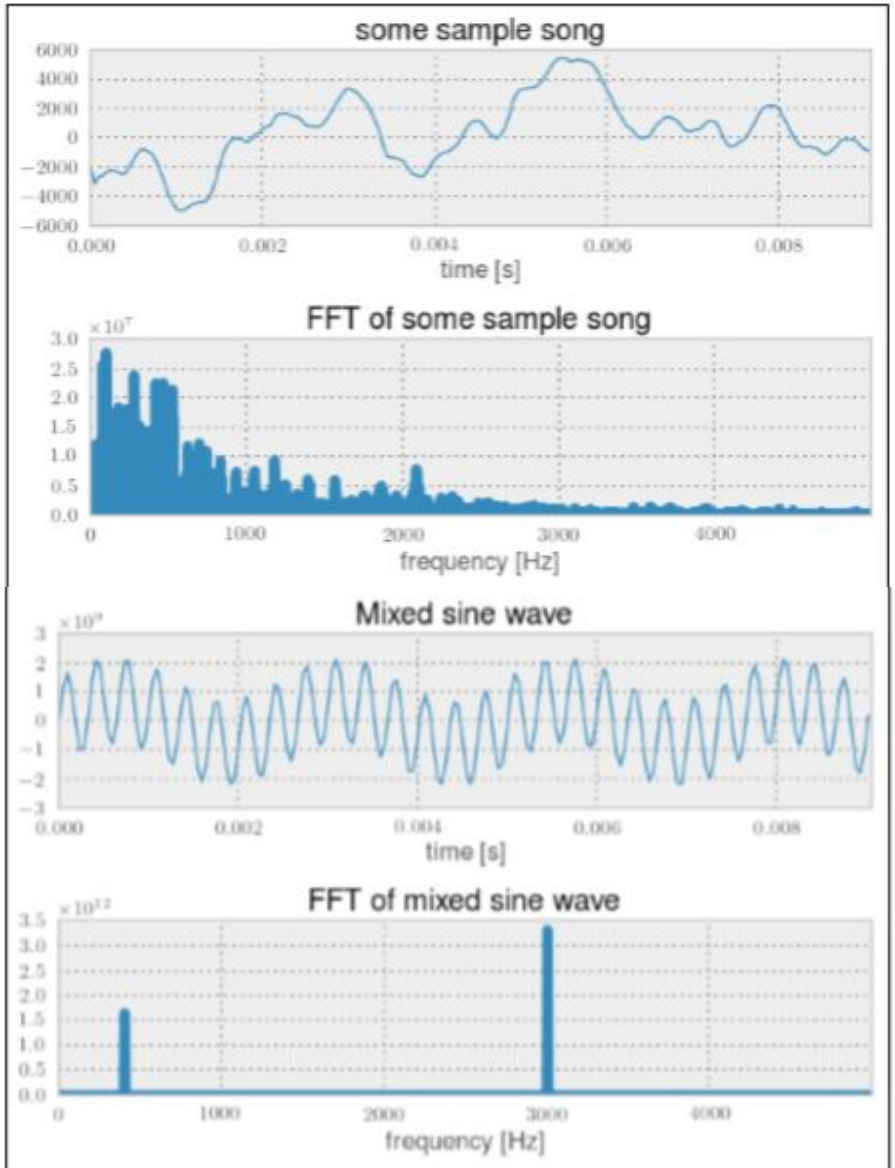
Метал (Metallica – St.Anger)

Как бы построить один набор чисел?

- В обработке сигналов часто используют разложение в ряд Фурье
- Несколько основных гармоник уже неплохо описывают сигнал
- Почему бы и здесь не сделать так?

Вспоминаем второй курс (если он уже позади :)





Как сделать?

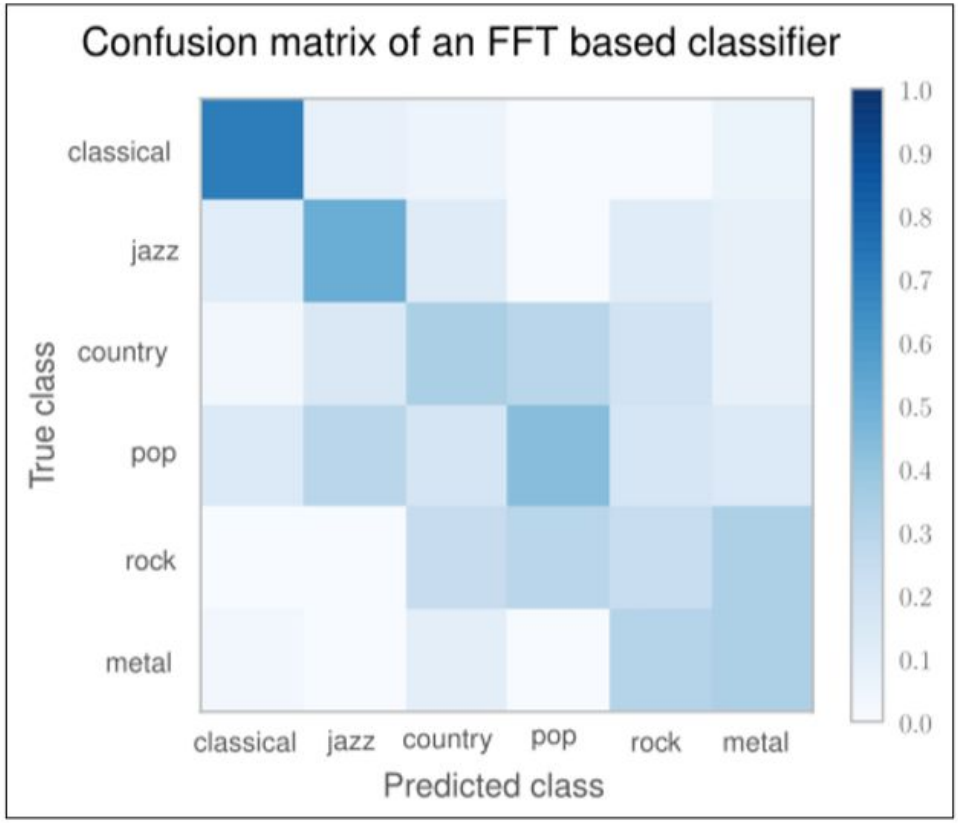
```
def create_fft(fn):
    sample_rate, X = scipy.io.wavfile.read(fn)
    fft_features = abs(scipy.fft(X)[:1000])
    base_fn, ext = os.path.splitext(fn)
    data_fn = base_fn + ".fft"
    np.save(data_fn, fft_features)

def read_fft(genre_list, base_dir=GENRE_DIR):
    X = []
    y = []
    for label, genre in enumerate(genre_list):
        genre_dir = os.path.join(base_dir, genre, "*.fft.npy")
        file_list = glob.glob(genre_dir)
        for fn in file_list:
            fft_features = np.load(fn)

            X.append(fft_features[:1000])
            y.append(label)

    return np.array(X), np.array(y)
```

Результат 1


Confusion matrix of an FFT based classifier
True class
classical
jazz
country
pop
rock
metal
Predicted class
classical
jazz
country
pop
rock
metal
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Повышаем качество

- Круто? Не очень. Читаем статьи.
- Узнаем, например, про Mel Frequency Cepstral Coefficients (MFCC) – преобразование Фурье логарифма спектра
- Используем в качестве признаков

Как обычно, есть готовый пакет

```
>>>from scikits.talkbox.features import mfcc
>>>sample_rate, X = scipy.io.wavfile.read(fn)
>>>ceps, mspec, spec = mfcc(X)
>>> print(ceps.shape)
(4135, 13)
```

Hint: `x = np.mean(ceps[int(num_ceps*1/10):int(num_ceps*9/10)], axis=0)`

Итоговый код построения фич

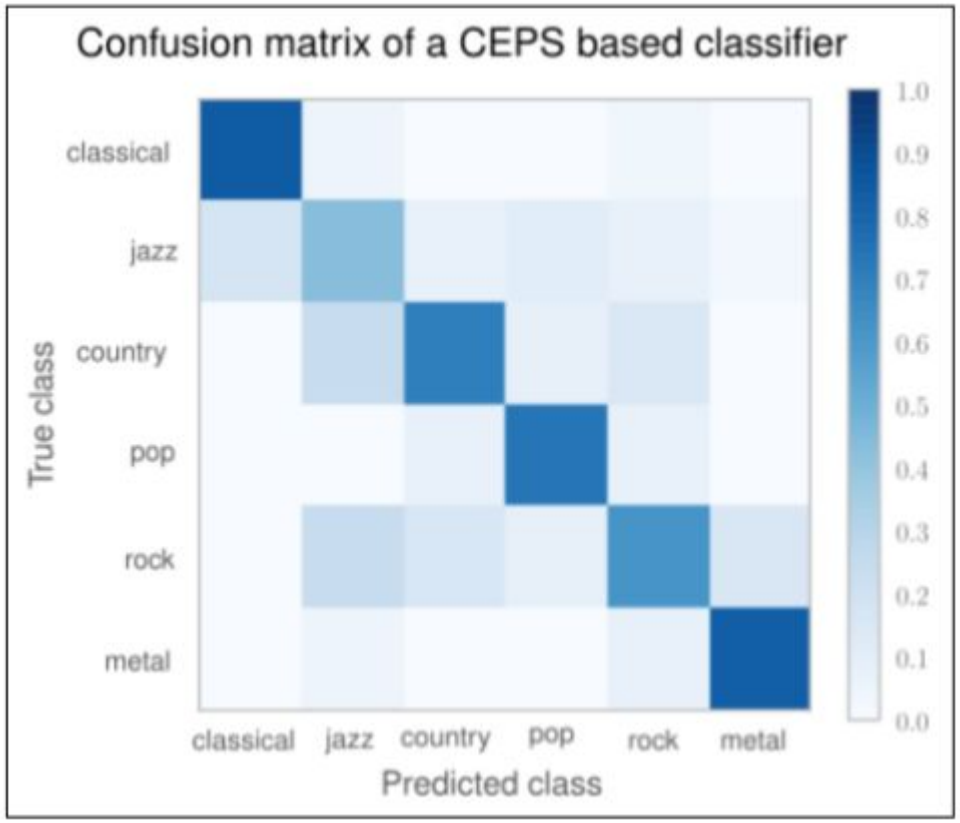
```
def write_ceps(ceps, fn):
    base_fn, ext = os.path.splitext(fn)
    data_fn = base_fn + ".ceps"
    np.save(data_fn, ceps)
    print("Written %s" % data_fn)

def create_ceps(fn):
    sample_rate, X = scipy.io.wavfile.read(fn)
    ceps, mspec, spec = mfcc(X)
    write_ceps(ceps, fn)

def read_ceps(genre_list, base_dir=GENRE_DIR):
    X, Y = [], []
    for label, genre in enumerate(genre_list):
        for fn in glob.glob(os.path.join(
base_dir, genre, "*.ceps.npy")):
            ceps = np.load(fn)
            num_ceps = len(ceps)
X.append(np.mean(
ceps[int(num_ceps*1/10):int(num_ceps*9/10)], axis=0))
y.append(label)

    return np.array(X), np.array(y)
```

Результат 2


Confusion matrix of a CEPS based classifier
True class
classical
jazz
country
pop
rock
metal
Predicted class
classical
jazz
country
pop
rock
metal
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Как лучше расшарить тему:

1. Воспроизвести все описанные действия и построить классификатор музыки по жанрам

2. Разобрать статью с описанием признаков, которых не было на семинаре, реализовать их и сравнить результаты на кросс-валидации

3. * Реализовать скачивание плейлиста из vk, конвертацию mp3 в wav и классификацию по жанрам. При необходимости, сделать свою обучающую выборку.

4. ** Сделать классификатор жанров по словам песен (где они есть), сравнить качество с MFCC.

Что еще можно придумать?

- Можно сделать подбор плейлиста из музыки одного жанра или из разных
- Можно сделать поиск похожих мелодий и рекомендации песен по звучанию, советуя те песни, которые ближе в пространстве признаков
- Можно сделать автоматическое составление плейлиста с плавным переходом от одного жанра к другому (например, чтобы постепенно исправить плохое настроение :)

Бонус-трек

- Конвертация из mp3 в wav из консоли:
  ffmpeg, sox
- Выкачивание музыки из vk на Python:
  http://habrahabr.ru/post/157925/
- Приятная и интересная книжка по машинке на Python:

Есть вопросы? Пишите :)

e-mail: xead.wl@gmail.com

vk: http://vk.com/id6356042