

1- خواندن فایل mp3:

ابتدا با دستور audioread فایل viva.mp3 را در clip می ریزیم.

2- پیش پردازش:

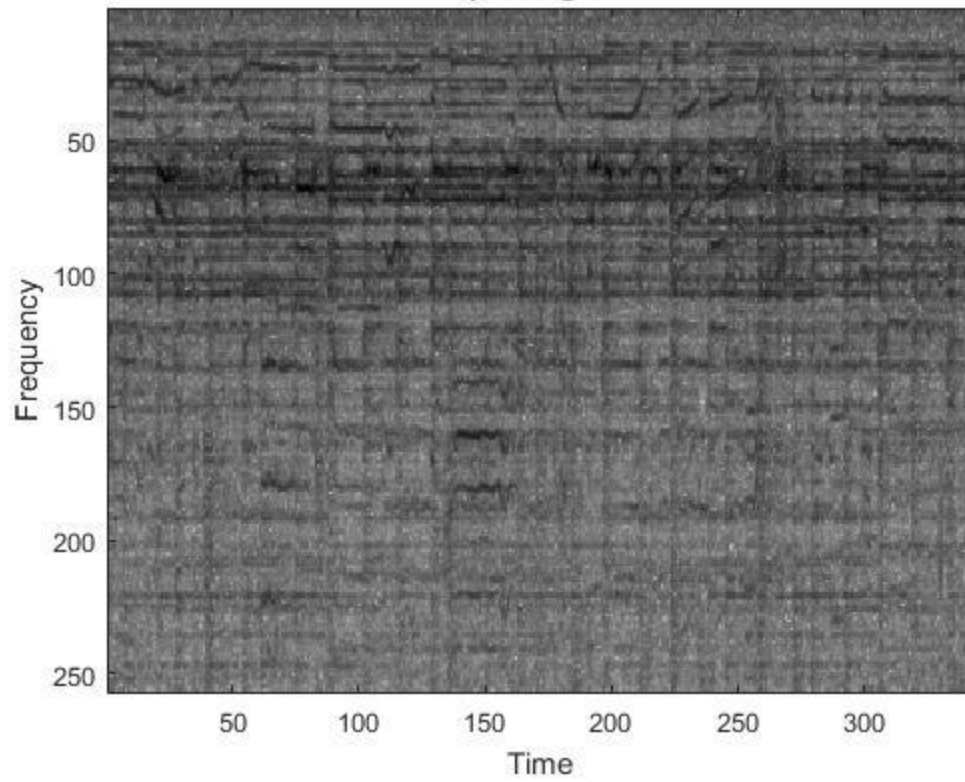
سپس به کمک if یک کانال از فایل صوتی در می آوریم و مقدار DC را از آن کم کرده و بعد از آن نرخ نمونه برداری را کاهش می دهیم.

3- اسپکتوگرام:

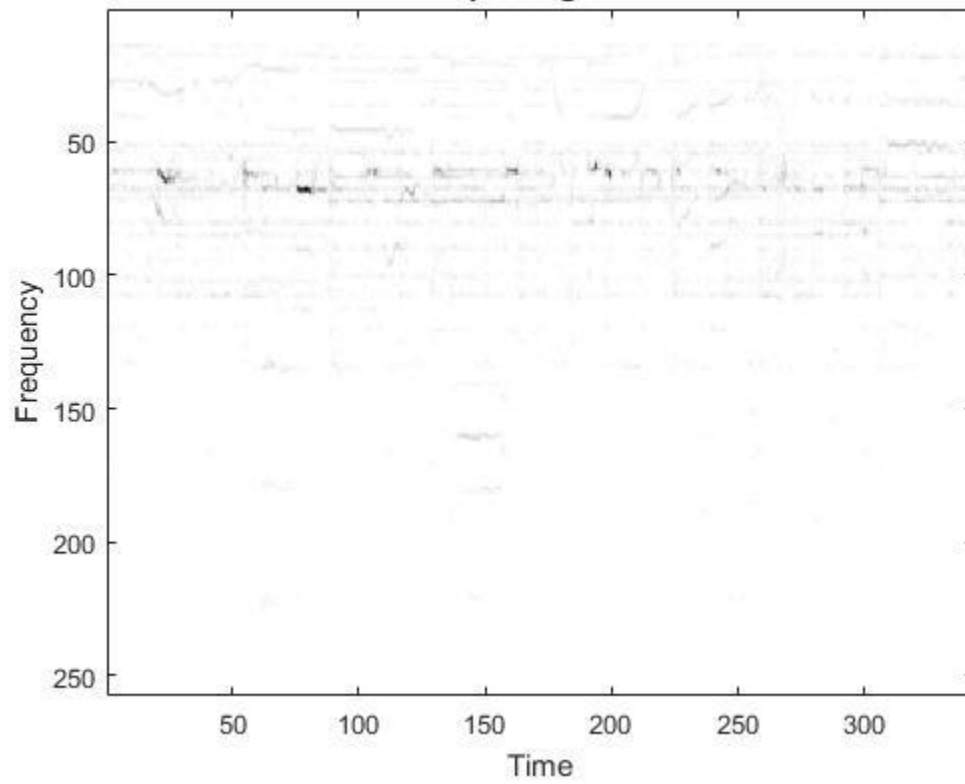
برای دستور spectrogram باید 3 ورودی window و noverlap و nfft را معین کنیم.

می دانیم $nfft = window$ و بهترین مقدار برای noverlap نصف window است که در مقادیر داده شده نیز ۳۲ نصف ۶۴ است پس با تعیین window همه معین می شوند. در هر ثانیه 8000 نمونه داریم ضربدر 0.064 برابر است با 512 که window است. علت استفاده از لگاریتم این است که می توان مقادیر خیلی زیاد و خیلی کم را در کنارهم داشت و بررسی آن ها راحتتر خواهد بود. ابتدا با لگاریتم و سپس بون لگاریتم اسپکتوگرام می گیریم.

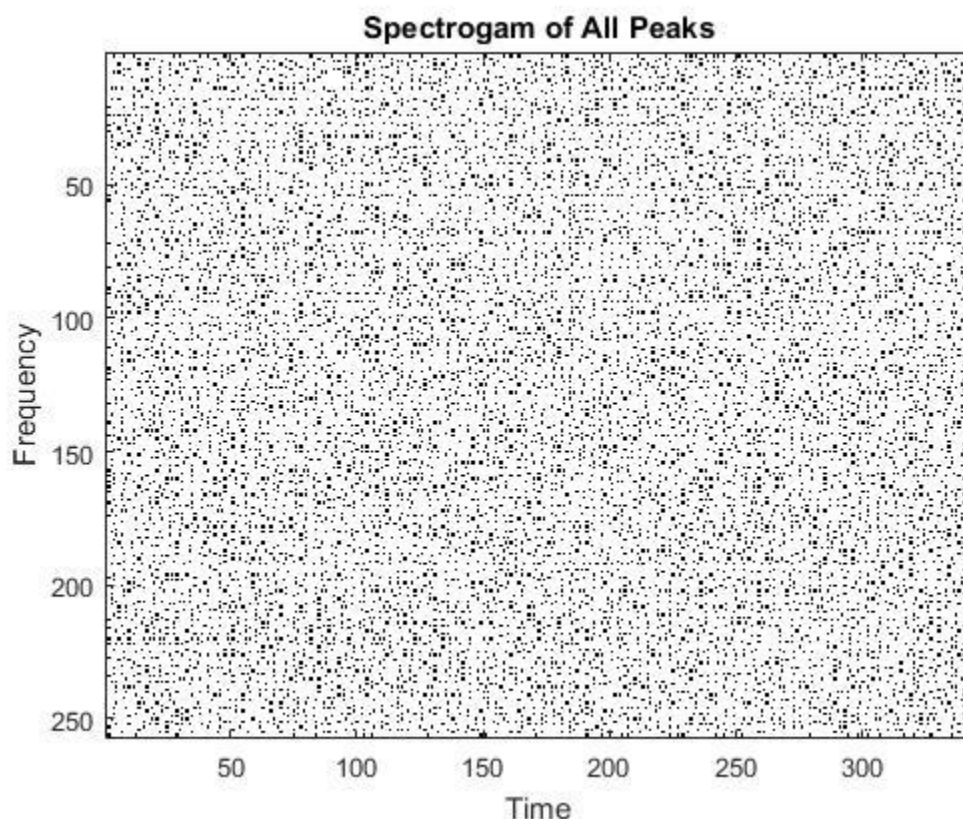
Spectrogram



Spectrogram

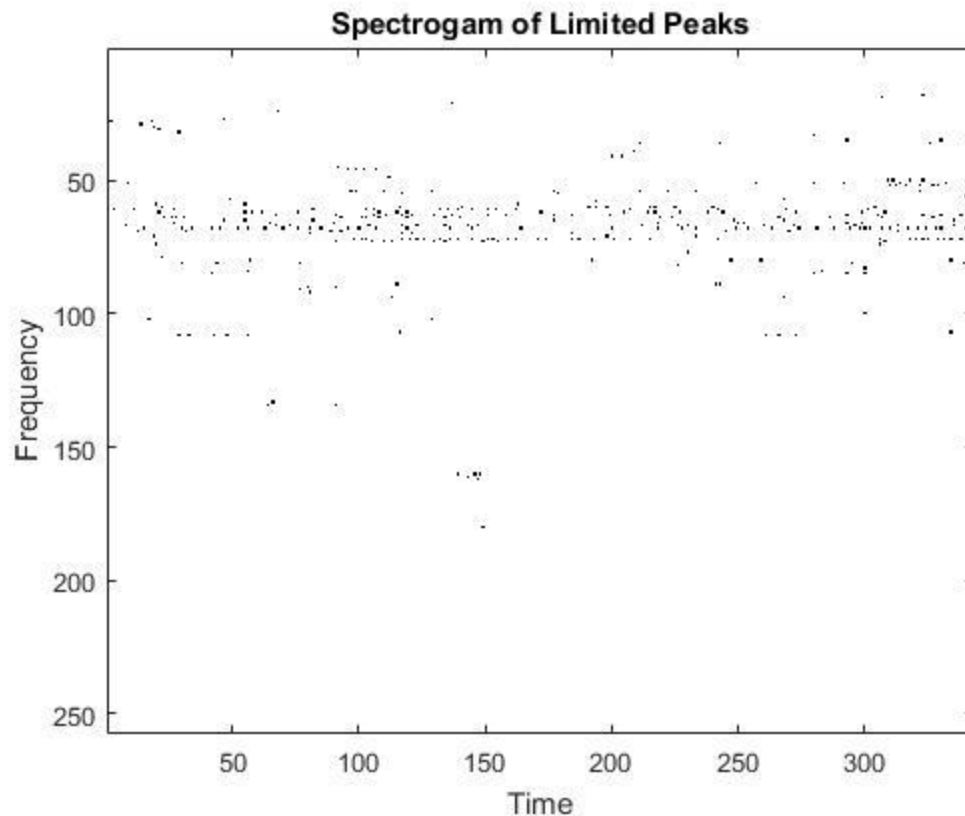


4- پیک های محلی اسپکتوگرام:



روش دیگر که به روش Philips مشهور است قسمتی از spectrogram را انتخاب کرده و با بررسی مثبت و منفی بودن تغییرات انرژی هر نقطه در صورت حرکت در زمان و فرکانس یک نقشه ی باینری کشیده که سیاه برای ۱ و سفید برای ۰ است. حال برای هر زمان یک کد باینری داریم که می توان نمونه ای که می خواهیم بررسی کنیم را روی زمان های مختلف جا به جا کنیم و مشابهت پیدا کنیم. روش دیگر استفاده از فیلترهای فراوان است که انتخاب آن ها به روش های پیشرفته machine-learning ربط دارد. راه دیگر پیدا کردن نت ها از روی اسپکتوگرام هر آهنگ است.

5- آستانه یابی:



پیک ها یکنواخت نیستند و در یک بازه ی فرکانسی فشرده شده اند. عموماً بهتر است ولی برای آهنگ هایی که شباهت زیادی در محل فشردگی پیک هایشان دارند مشکل ساز خواهد بود. اگر مطلوب نباشد می توان از در هر ثانیه و یا در هر فرکانش مقدار محدودی پیک نگه انتخاب کرد. می توان برای بهتر کردن عملیات نویز ها را حذف کنیم یا عدد 30 برای نسبت پیک ها به سائز سیگنال را عوض کنیم یا حتی معیار دیگری برای انتخاب تعداد پیک هی مهم در نظر بگیریم.

6- ساخت جدول:

Spectrogram local peaks and peak pairs

