**İZMİR DEMOKRASİ ÜNİVERSİTESİ**

**MÜHENDİSLİK FAKÜLTESİ**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ BÖLÜMÜ**

**2020-2021 BAHAR DÖNEMİ**

**EEM 306 SAYISAL SİNYAL İŞLEME DERSİ PROJE RAPORU**

**VOKAL AYIRMA**

**Göktuğ GÖKMEN**

**1806102007**

1. Projenin Amacı

“Vokal Ayırma” projesindeki temel amaç, bir müzik parçasındaki vokalleri kendisini takip eden eşlikten ayırmak, daha net şekilde ifade etmek gerekirse parçayı vokalden arındırmaktır.

1. Hedef

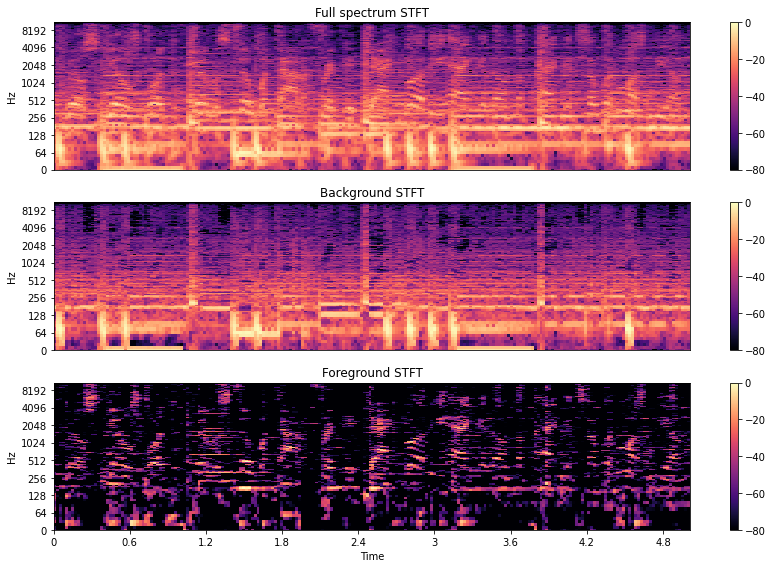
Proje sonunda elde edilmesini beklediğimiz çıktı, müzik parçalarını girdi olarak alan bu parçalara çeşitli yöntemler uygulayarak vokal ve eşlik olarak iki parçaya bölen bu parçaları hem görsel olarak (STFT (Short-time Fourier Transform), Spektrogram vb.) hem de işitsel olarak bir çıktı haline getiren sistem tasarlamak.

1. Yöntem

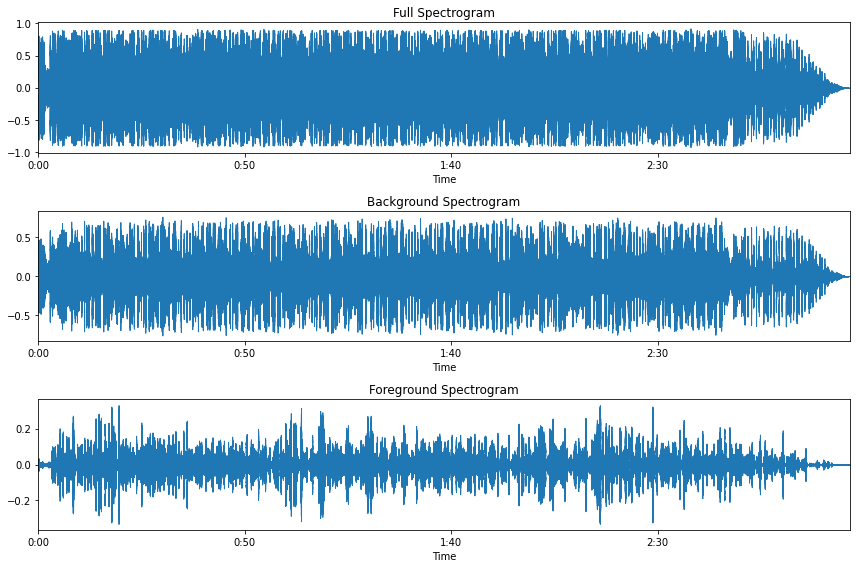
Projemizi Google Colab’de gerçekleyeceğiz ve proje sonunda çıktıları elde edeceğimiz yer de Colab defterimiz olacak. Projeye başlarken ilk olarak kütüphaneleri tanımlayarak başlıyoruz, projede yoğun olarak kullanılan kütüphaneler : Numpy, Matplotlib, Librosa, IPython ve SoundFile. Kütüphaneleri tanımlamamızın ardından kullanıcıdan girdi olarak bir parça seçmesini ve bunun yolunu (drive) girmesini istedik. Ardından girilen parçayı büyüklük ve faz olarak iki parçaya ayırdık (D (Spektrogram) = S (Magnitude) \* P (Phase) ve sinyalimize fourier transform uygulayarak frekans düzlemine dönüştürdük (uygulamanın sonunda inverse fourier transform uygulayarak tekrar , zaman düzlemine geçeceğiz.) bundan sonrasında tamamen bu bileşenler üzerinden ilerleyeceğiz. Maskeleri uyguladığımızda oluşacak parçalar için birer çıktı fonksiyonu oluşturduk. Filtre oluşturma aşamasında Librosa kütüphanesinin nn\_filter fonksiyonu aracılığıyla cosine benzerliğini kullanarak kareleri karşılaştıracağımız ve benzer kareleri toplayacağımız bir filtre geliştirdik. Daha sonra marjinleri ve üsleri belirleyip maske oluşturma aşamasına geçtik formülünü kullanan Librosa kütüphanesi softmask fonksiyonunu kullandık ve iki adet maske oluşturduk.Elde ettiğimiz maskeleri (matematiksel çarpım yoluyla) giriş spektrumumuza uyguladık. Ardından uygulamanın başlarında iki bilşene ayırdığımız spektrogramı maske uygulanmış sinyallerde tekrar birleştireceğiz bunun için D (Spektrogram) = S (Magnitude) \* P (Phase) formülünü kullanacağız ve maske uygulayarak elde ettiğimiz sinyallerle, en başta spektrogram’ın bileşeni olan phase’a matematiksel çarpım uygulayacağız son olarak elde ettiğimiz sinyallere Librosa kütüphanesinin istft fonksiyonu aracılığıyla Inverse Short-time Fourier Transform uygulayacağız ve çıktıları oluşturduğumuz fonksiyon aracılığıyla oluşturduğumuz dosyaları yazdıracağız.

1. Sonuç

Sonuç olarak sisteme girdi olarak verdiğimiz vokali ve eşliği bütün olan bir parçadan iki parça çıktısı aldık. Birinci parça eşlik ve ikinci parça vokal olarak sistem çıktısı elde ettik ve sistemin başarısını gözlemlemek ve zihinde daha net bir görüntü oluşabilmesi açısından parçanın bütünün ve elde ettiğimiz çıktıların Short-time Fourier Transform (şekil.a) ve Spektrogramlarını (şekil b.) çizdirdik.



Şekil a : Short-time Fourier Transform



Şekil.b : Spektrogram

1. Başarım

* Sistemin genel olarak enstrümental sesleri ayırmada daha başarılı olduğu gözlendi.
* Vokal sesleri ayırırken seste ciddi bir azalma mevcut (neredeyse duyulmayacak düzeyde). Bu da vokal seslerde ciddi bir başarısızlığa neden oluyor.
* Sistem, genel olarak backgroundu daha soft ve vokal seslerin baskın olduğu parçalarda daha başarılı.

1. Kaynakça
2. <https://medium.com/datarunner/librosa-9729c09ecf7a>
3. <https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1086&context=argcon>
4. <https://towardsdatascience.com/audio-ai-isolating-instruments-from-stereo-music-using-convolutional-neural-networks-584ababf69de>
5. <https://towardsdatascience.com/audio-ai-isolating-vocals-from-stereo-music-using-convolutional-neural-networks-210532383785>
6. <https://www.kaggle.com/ashishpatel26/feature-extraction-from-audio>
7. <https://www.audiolabs-erlangen.de/resources/MIR/FMP/C2/C2_STFT-Inverse.html>
8. <http://librosa.org/doc/main/generated/librosa.magphase.html>