

ELM368 – DÖNEM PROJESİ

TEK KANALLI SES KAYNAĞI AYRIŞTIRICI

Mert Ayhan, Selin Ünlü, Beyzanur Karakulak

161024011, 161024016, 161024056

mert.ayhan2016@gtu.edu.tr, selin.unlu2016@gtu.edu.tr, b.karakulak2016@gtu.edu.tr

ABSTRACT (ÖZET)

Telefon konuşmalarında arkadan gelen çevre gürültüsünün azaltılması amacıyla bu proje yazılmıştır. İşaret ‘Tek Kanallı Ses Ayrıştırma’ yapılarak ardından ‘Alçak Geçiren Filtre’ ve ‘Yüksek Geçiren Filtre’ kullanılarak gürültünün azaltılmasına çalışılmıştır. Bu projede gürültü engelleme işleminin sonuçlarını daha iyi görmek için sesin ham hali ve işlenmiş hali yani gürültünün azaltılmış hali frekans uzayında incelenmiştir.

ANAHTAR KELİMELELER

IIR Filtre, Tek Kanallı Ses Ayrıştırma, Short-Time Frekans Dönüşümü, Spektrogram.

1. Giriş

Herhangi bir ortamda ses kaydı yaparken ortamda gürültü bulunması muhtemeldir[1]. Gürültü, birçok farklı etkiden kaynaklanabilecek rastgele değişen bir işarettir. Gürültü, halk arasında parazit olarak bilinir. Gürültüye sebep olan seslere saç kurutma makinesi sesi, su sesi, çatal bıçak sesi, rüzgâr sesi veya trafik sesleri örnek olarak verilebilir. Gürültünün varlığı ortamlarda insan sesinin anlaşılmayacak şekilde bozulmasına neden olabilir, iletimi tehlikeye atabilir ve iletilen bilgileri anlaşılmaz hale getirebilir. Gürültü azaltma için birçok yöntem geliştirilmesinin nedeni budur. Elbette bu yöntemlerin hiçbirisi gürültüyü tamamen ortadan kaldırmaz, ancak gürültü seviyesinde önemli bir azalma elde edilebilir. Böylece ses kalitesi artırılır. Bunu önlemek için bu proje geliştirilmiştir.

Bu projede giriş işaretleri ve filtrenin çıkışındaki işaretin zaman ve frekans uzayındaki grafikleri gösterilmiştir. Bunun için Short-Time Fourier Dönüşümü kullanılmıştır. Bu bir zaman-frekans analiz metodudur. Sürekli bir fonksiyonu, uygun bir pencere kullanmak suretiyle küçük parçalar halinde inceleyerek anlık frekans bulmakta kullanılır.

Bu projede ortamda gürültünün bulunduğu ses verileri kullanılarak ortamdaki gürültü bastırılmaya çalışıldı. Uzunluğu yaklaşık 30 saniye olan 5 adet ses verisi

kullanıldı. Bu sesler ogg formatında kaydedilmiştir. wav formatına çevrilerek Python’da işlenmiştir.

2. Deneyler ve Analiz

Bir ses sinyali olan $x(t)$ iki işaretin toplamı olarak kabul edilir. $x(t)$, insan sesi ve gürültünün birleşiminden meydana gelir[2].

$$x(t) = s_1(t) + s_2(t) \quad (1)$$

$s_1(t)$, insan sesi (foreground) ve $s_2(t)$, gürültü (background) ‘dır. Buradaki amaç $s_1(t)$ bileşenini elde etmektir. Bunu elde etmek için STFT kullanılır. Bu işlemden sonra (1)’deki denklem, n ve f e bağımlı fonksiyonlar haline gelir.

$$X(n, f) = S_1(n, f) + S_2(n, f) \quad (2)$$

Eşitlik (2) ‘nin mutlak değeri alınarak yaklaşık olarak elde edilir.

$$|X(n, f)| \approx |S_1(n, f)| + |S_2(n, f)| \quad (3)$$

Eşitlik (3)’te edilen işaret matris formunda yazılır. n değerleri sütunları ve f değerleri satırları oluşturur.

$$X(n, f) \approx S_1(n, f) + S_2(n, f) \quad (4)$$

$$X \approx S_1 + S_2 \quad (5)$$

Denklem matris formunda yazıldıktan böyle bir M seçilir ki X ile çarpılınca S_1 ifadesi elde edilir. M buna göre belirlenir.

$$\widehat{S}_1 = M \odot X \quad (6)$$

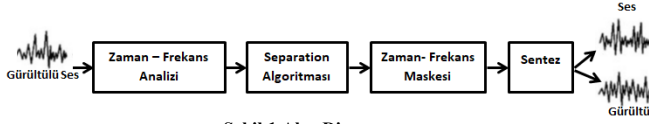
$$\widehat{S}_2 = (1 - M) \odot X \quad (7)$$

Eşitlik (7) ile elde edilen ifade \widehat{S}_2 ‘ye eşittir. Buradaki 1, birim matrisi ifade etmektedir.

\hat{S}_2 'nın bu şekilde alınmasının nedeni \hat{S}_1 ve \hat{S}_2 'nin toplamından tekrar X işaretini elde edebilmektir. Bu nedenle birim matris kullanılır.

$$M \in [0,1] \quad (8)$$

(8) ile gösterilen ifade zaman-frekans maskesidir. Bu maskeyi bulamamızın sebebi iyi bir separation elde etmektir.



Şekil 1 Akış Diyagramı

Programın Algoritması:

1. Ses sinyalinin STFT'si alınarak Time-Frequency incelemesi yapıldı.
2. Soft mask Nearest Neighbour interpolasyonu ile yumuşatılarak maske elde edildi.
3. Ses sinyalinin STFT'si üzerine maske uygulanarak insan sesi sinyali tespit edildi.
4. İnsan sesi sinyali stop band'ı 3200 Hz olan IIR Butterworth alçak geçiren filtreden geçirildi.
5. Ardından stop band'ı 300 Hz olan IIR Butterworth yüksek geçiren filtreden geçirilerek gürültüsüz insan sesi elde edildi.

Projede önce gürültülü ortamda kayıt edilmiş insan sesinin zaman, frekans uzayındaki gösterimlerine ve spektrogramlarına bakılmıştır. Bunlar incelenerek ne tip bir filtre/filtreler kullanılacağı konusunda karar verilmiştir. IIR filtreler kullanılarak gürültü azaltılmaya çalışılmıştır.

Bu aşamada örnek ses verisi 2 için grafikler incelenecektir. Burada önemli olan daha çok spektrogram grafikleridir. Spektrogram grafiğindeki açık renkler (0 dB) gürültüleri ifade ederken koyu renkler ise gürültünün düştüğü yani daha az olduğu yerleri ifade eder. Buradaki asıl amaç açık renkli kısımları yok ederek insan sesini gürültüden ayıştırmaktır. Kullanılan ses verisi tek kanallı[3] bir sestir. Şekil 16'daki spektrogram insan sesi ve arka plan sesi yani gürültünün yoğun olarak bulunduğu bir grafikdir. Bu işaret filtreden geçirilince Şekil 19'da açık renklerin büyük ölçüde yok olduğu görülür. Renkler koyulaşmış ve gürültüler engellenmiştir. Şekil 14 ve Şekil 17 karşılaştırıldığında ise işaretin zaman uzayında bir hayli değiştiği görülür. İşaretin background ve foreground denilen iki sesin bileşiminden oluştuğu daha önceden belirtilmişti. Şekil 17'de ses background sesinden ayrıştırıldığı için ses işaretinin filtreden sonra bu şekilde değişmesi beklenen bir durumdur. Burada insan sesi elimine edilmiştir. Şekil 14'teki genlik değerlerinden daha düşük genlik değerlerine sahip bir grafik elde edilmek isteniyordu ve bu Şekil 17'de elde edildi. Şekil 15 ve Şekil 18'de işaretin frekans uzayındaki değişimi karşılaştırıldığında ise düzelmeler görüldü. Yan loblarda

daralma meydana geldi. Frekans uzayındaki bu farklılıklar da gürültünün azaltıldığını gösterdi.

Aslında işaret filtreden geçiriliyor denilse de burada interpolasyon yapılmaktadır. Nearest Neighbour filtresi/interpolasyonu[4] kullanılmaktadır. Bu filtre uzaklığı Cosinus Distance[5][6]'a göre hesaplanan en yakındaki komşuyu belirleyerek yeni bir sinyal oluşturur. Daha sonra oluşturulan bu yeni sinyal soft mask ile yumuşatılır.

Filtreden geçirilen sinyal alçak geçiren filtreden geçirildikten sonra Şekil 22'deki spektrograma bakıldığında 3000 Hz civarındaki değerlerin üzerindeki frekansların yok olması bekleniyordu ve bu gözlemlendi. Şekil 21'de 0.10 π civarındaki frekans bileşenlerinin kaybolduğu görüldü.

Son olarak yüksek geçiren filtreden geçirilen işaretin Şekil 25'deki spektrogramına bakıldığında 100 Hz değerine kadar bulunan bileşenlerinin yok olduğu görüldü. Şekil 24'e bakıldığında frekans uzayında 0 π civarındaki bileşenlerin kaybolduğu görüldü. İşaretin zaman uzayındaki şeklinde gözle görülür ölçüde değişim gözlemlenmedi.

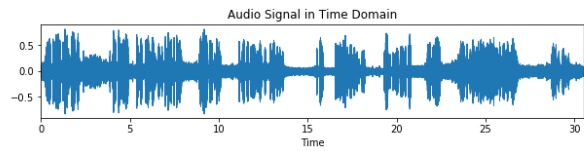
Spektrogram[7] özel bir dalga formu içinde mevcut çeşitli frekanslarda zaman bir sinyalin gücü, ya da "ses şiddeti" temsil eden görsel bir yöntemdir.

Diğer bilimlerde spektrogramlar, mikrofonlar tarafından kaydedildiği gibi insanlar, makineler, hayvanlar, balinalar, jetler vb. Tarafından üretilen ses dalgalarının frekanslarını görüntülemek için yaygın olarak kullanılır.

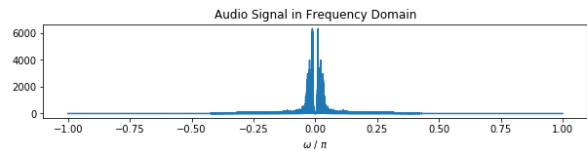
Spektrogramlar temel olarak iki boyutlu grafiklerdir ve üçüncü bir boyut renklerle temsil edilir. Renklerin tonunun açıktan koyuya doğru gitmesi genlik değerinin azalması anlamına gelir.

2.1 Grafik, Tablo ve Şekiller

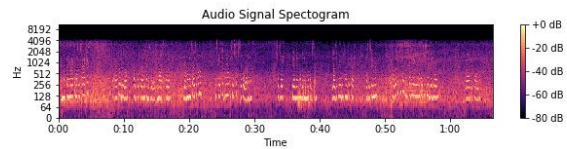
Örnek Ses 1'in Grafikleri:



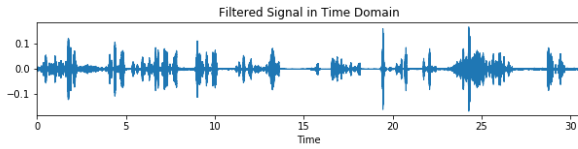
Şekil 2 Zaman Uzayında Ses İşareti



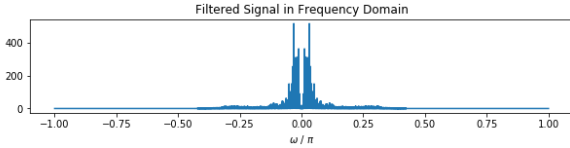
Şekil 3 Frekans Uzayında Ses İşareti



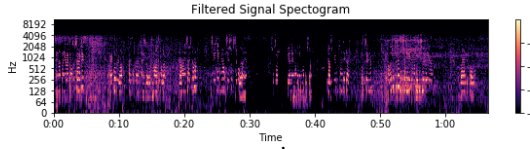
Şekil 4 Ses İşaretinin Spektrogramı



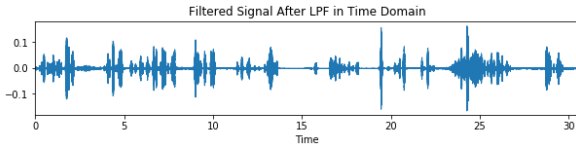
Şekil 5 Zaman Uzayında Filtrenmiş İşaret



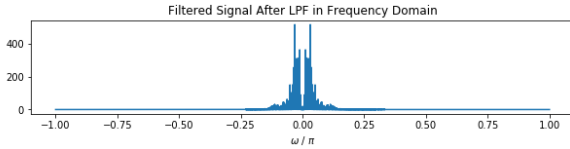
Şekil 6 Frekans Uzayında Filtrenmiş İşaret



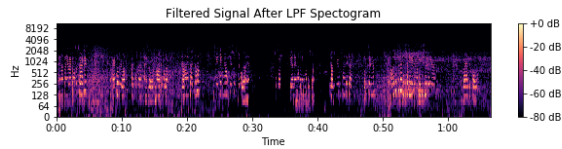
Şekil 7 Filtrenmiş İşaretin Spektrogramı



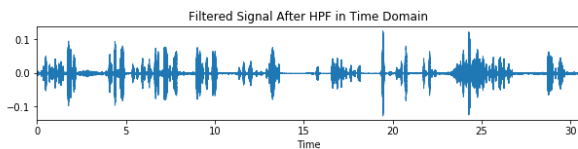
Şekil 8 Zaman Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrenmiş İşaret



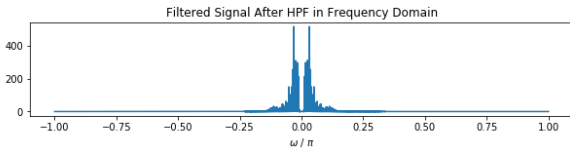
Şekil 9 Frekans Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrenmiş İşaret



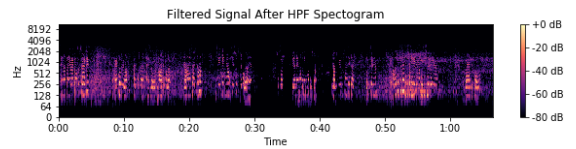
Şekil 10 Alçak Geçiren Filtre Çıkışındaki Filtrenmiş İşaretin Spektrogramı



Şekil 11 Zaman Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrenmiş İşaret

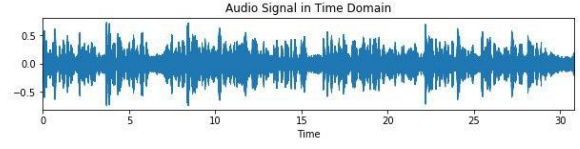


Şekil 12 Frekans Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrenmiş İşaret

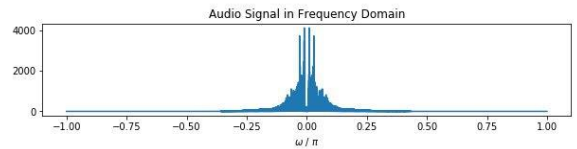


Şekil 13 Yüksek Geçiren Filtre Çıkışındaki Filtrenmiş İşaretin Spektrogramı

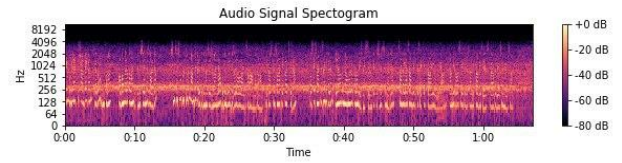
Örnek Ses 2'nin Grafikleri:



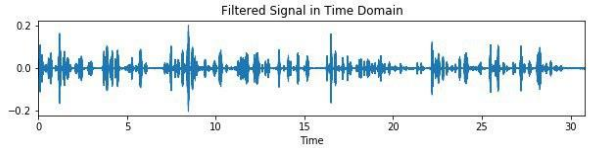
Şekil 14 Zaman Uzayında Ses İşareti



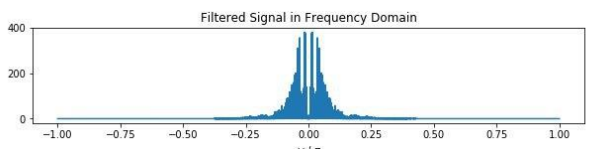
Şekil 15 Frekans Uzayında Ses İşareti



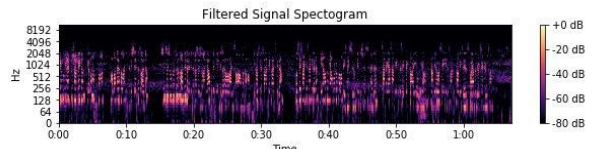
Şekil 16 Ses İşaretinin Spektrogramı



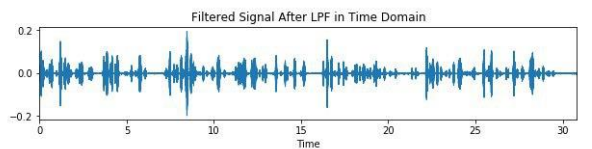
Şekil 17 Zaman Uzayında Filtrenmiş İşaret



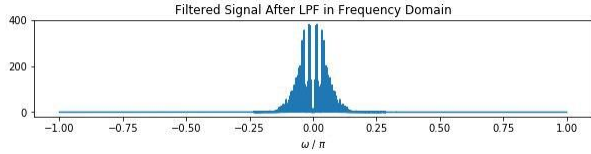
Şekil 18 Frekans Uzayında Filtrenmiş İşaret



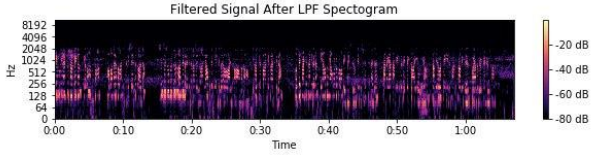
Şekil 19 Filtrenmiş İşaretin Spektrogramı



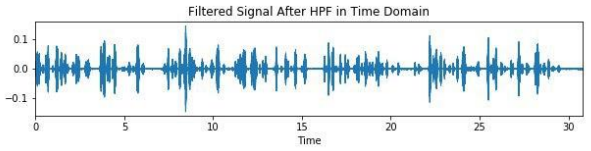
Şekil 20 Zaman Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrenmiş İşaret



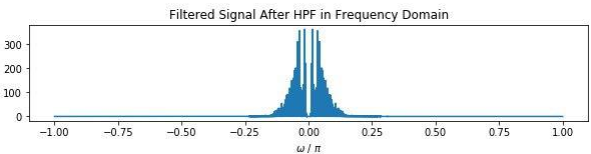
Şekil 21 Frekans Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



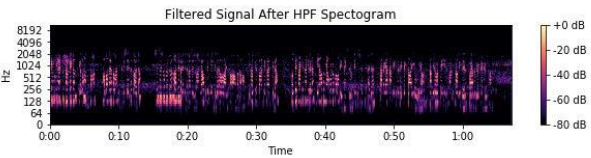
Şekil 22 Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı



Şekil 23 Zaman Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret

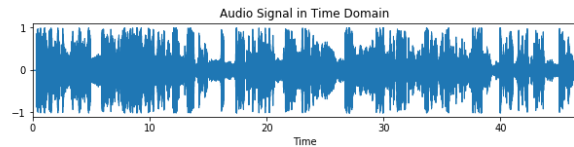


Şekil 24 Frekans Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret

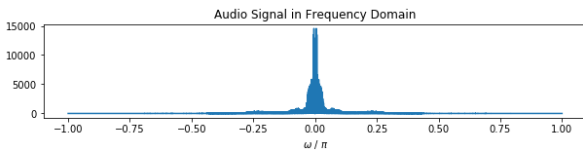


Şekil 25 Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı

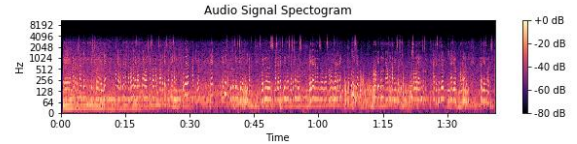
Örnek Ses 3'ün Grafikleri:



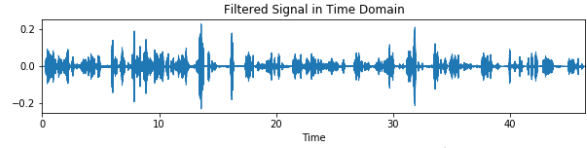
Şekil 26 Zaman Uzayında Ses İşareti



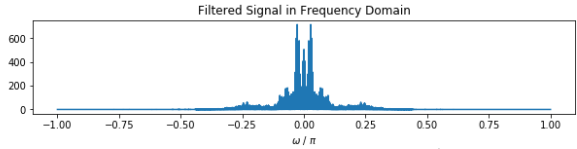
Şekil 27 Frekans Uzayında Ses İşareti



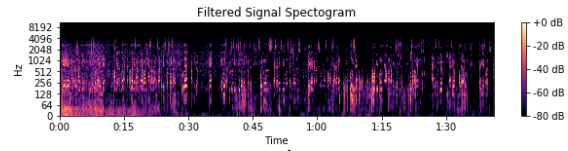
Şekil 28 Ses İşaretinin Spektrogramı



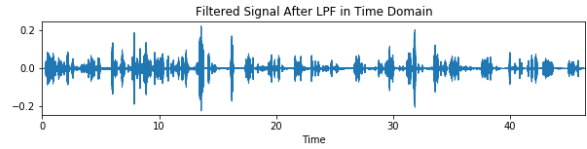
Şekil 29 Zaman Uzayında Filtrelenmiş İşaret



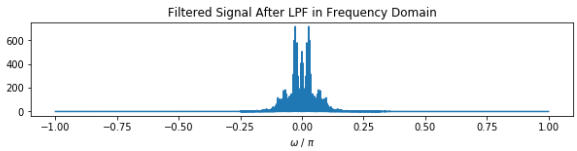
Şekil 30 Frekans Uzayında Filtrelenmiş İşaret



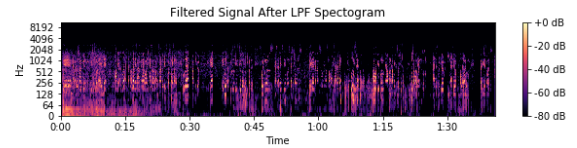
Şekil 31 Filtrelenmiş İşaretin Spektrogramı



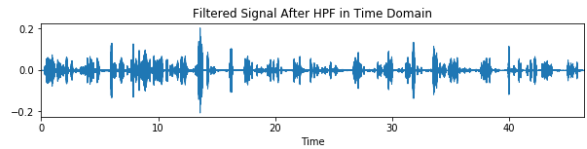
Şekil 32 Zaman Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



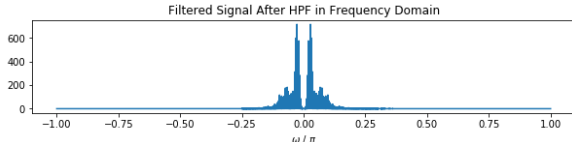
Şekil 33 Frekans Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



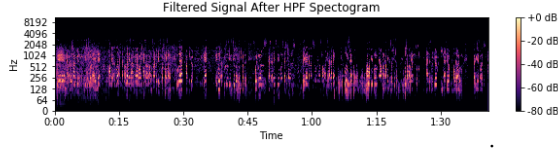
Şekil 34 Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı



Şekil 35 Zaman Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret

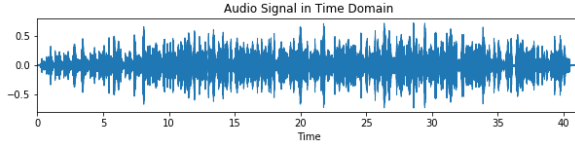


Şekil 36 Frekans Uzakında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret

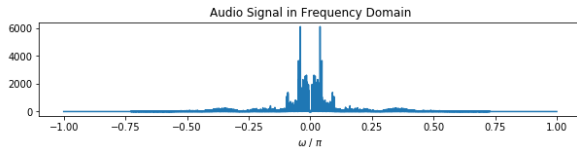


Şekil 37 Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı

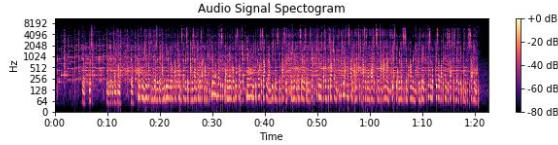
Örnek Ses 4'ün Grafikleri:



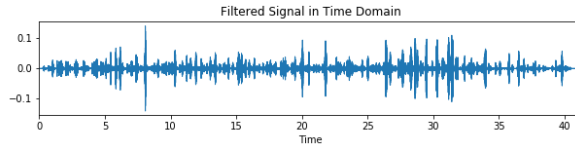
Şekil 38 Zaman Uzakında Ses İşareti



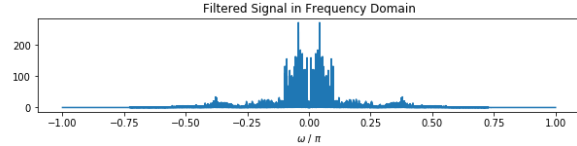
Şekil 39 Frekans Uzakında Ses İşareti



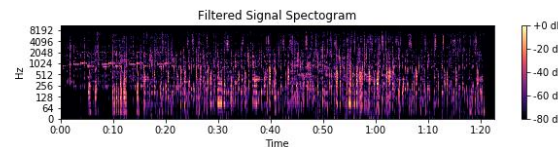
Şekil 40 Ses İşaretinin Spektrogramı



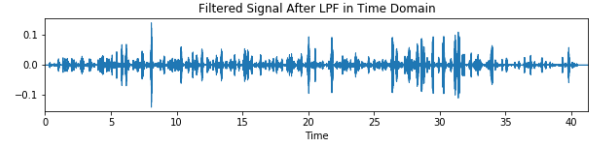
Şekil 41 Zaman Uzakında Filtrelenmiş İşaret



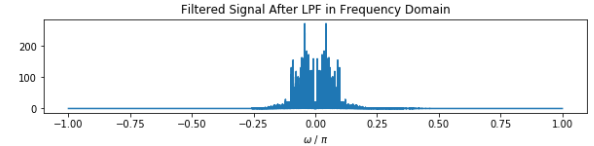
Şekil 42 Frekans Uzakında Filtrelenmiş İşaret



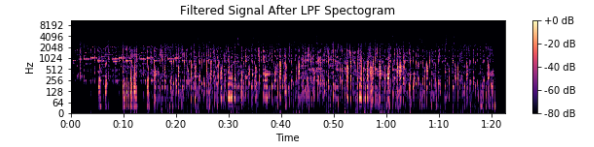
Şekil 43 Filtrelenmiş İşaretin Spektrogramı



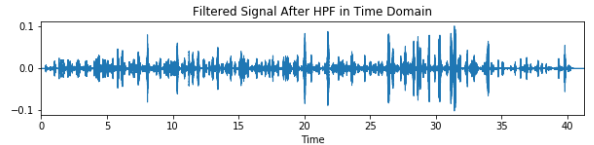
Şekil 44 Zaman Uzakında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



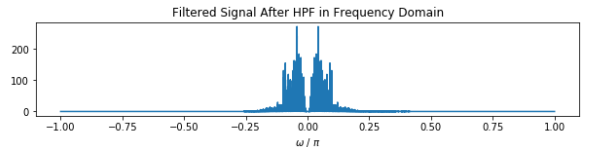
Şekil 45 Frekans Uzakında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



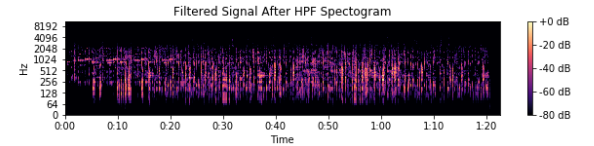
Şekil 46 Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı



Şekil 47 Zaman Uzakında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret

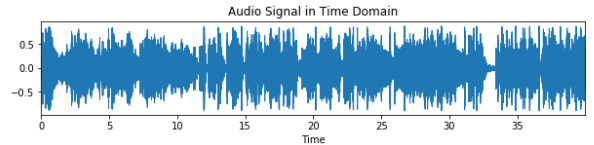


Şekil 48 Frekans Uzakında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret

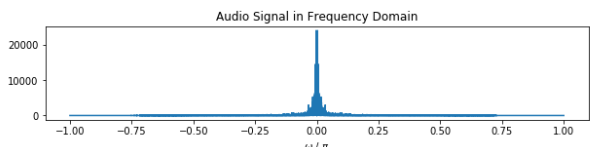


Şekil 49 Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı

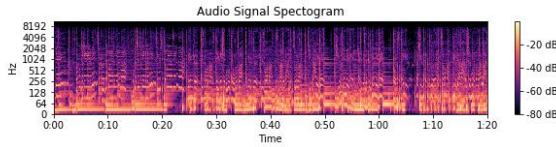
Örnek Ses 5'in Grafikleri:



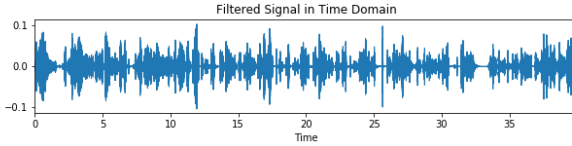
Şekil 50 Zaman Uzakında Ses İşareti



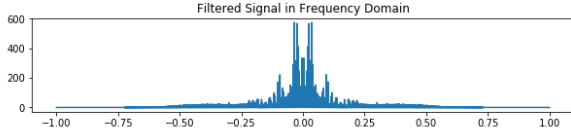
Şekil 51 Frekans Uzakında Ses İşareti



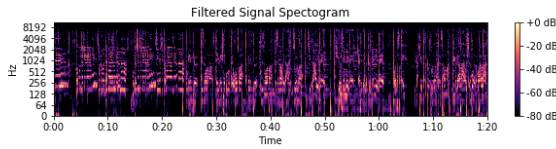
Şekil 52 Ses İşaretinin Spektrogramı



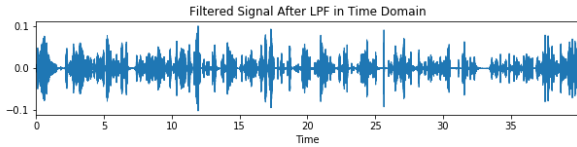
Şekil 53 Zaman Uzayında Filtrelenmiş İşaret



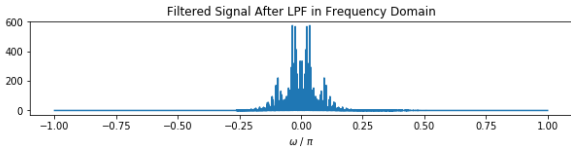
Şekil 54 Frekans Uzayında Filtrelenmiş İşaret



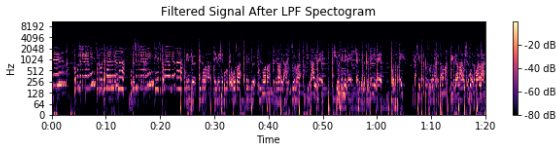
Şekil 55 Filtrelenmiş İşaretin Spektrogramı



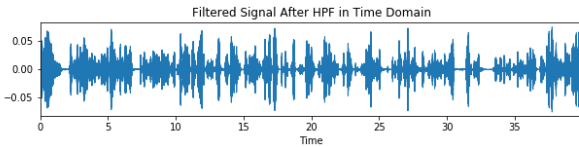
Şekil 56 Zaman Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



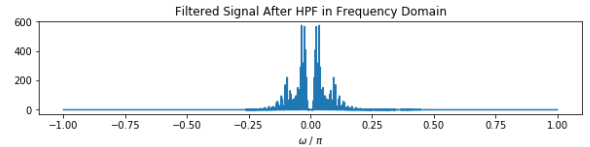
Şekil 57 Frekans Uzayında Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



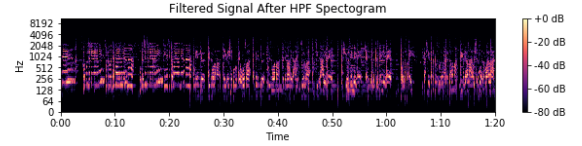
Şekil 58 Alçak Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı



Şekil 59 Zaman Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



Şekil 60 Frekans Uzayında Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaret



Şekil 61 Yüksek Geçiren Filtre Çıkışındaki Filtrelenmiş İşaretin Spektrogramı

3. Sonuç ve Yorum

Kullandığımız gürültüyü giderme yöntemi ile gürültü başarılı bir şekilde giderilmiştir bunun yanında insan sesinde bozulma gözlemlenmiştir. Bu bozulma kişinin konuşmasını anlaşılmaz yapmamaktadır. Fakat kişinin sesi robotik bir ses gibi duyulmaktadır.

Gürültü saç kurutma makinesi sesi gibi sabit bir gürültü ise gürültü bastırma işleminden sonra filtre çıkışındaki ses daha az bozulmaya uğrayarak orijinal insan sesine daha yakın bir sonuç vermiştir.

Projeye hazırlık aşamasında birçok kaynak taraması yaparak yeni kavramlar öğrendik. STFT, cosine distance, Nearest Neighbour interpolasyon, hard mask, soft mask vb. matematiksel işlemleri öğrenerek uygulandık. Spektrogram ile analiz yapmayı öğrendik. Librosa[8] kütüphanesi ile ses analizi yapmayı öğrendik.

Kaynaklar

1. https://ie.utcluj.ro/files/acta/2004/Number%206/Paper05_Draghiciu.pdf
2. <https://core.ac.uk/download/pdf/42553879.pdf>
3. Emad M. Grais, Gerard Roma, Andrew J.R. Simpson, Mark D. Plumbley (May 26 2016). "Single Channel Audio Source Separation using Deep Neural Network Ensembles" Centre for Vision, Speech and Signal Processing, University of Surrey Guildford, UK
4. Buades, A., Coll, B., & Morel, J. M. (2005, June). A non-local algorithm for image denoising. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 2, pp. 60-65). IEEE.
5. Rafii, Z., & Pardo, B. (2012, October). "Music/Voice Separation Using the Similarity Matrix." International Society for Music Information Retrieval Conference, 2012.
6. <https://towardsdatascience.com/importance-of-distance-metrics-in-machine-learning-modelling-e51395ffe60d>
7. <https://pnsn.org/spectrograms/what-is-a-spectrogram>
8. <https://librosa.org/librosa/>