



**ÇOK KANALLI ORTAMLARDA  
GÜRÜLTÜ AZALTMA**

**Ramazan ÇOLAK**

**Yüksek Lisans Tezi  
Elektronik ve Haberleşme Mühendisliği  
Anabilim Dalı  
Danışman: Doç. Dr. Rafet AKDENİZ  
Mayıs 2019**

**T.C.**  
**TEKİRDAĞ NAMIK KEMAL ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS TEZİ**

**ÇOK KANALLI ORTAMLARDA GÜRÜLTÜ AZALTMA**

**RAMAZAN ÇOLAK**

**ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ ANABİLİM DALI**

**DANIŞMAN: Doç. Dr. RAFET AKDENİZ**

**TEKİRDAĞ – 2019**

**©Her Hakkı Saklıdır**



Bu tez NKÜBAP tarafından NKÜBAP.06.YL.18.156 numaralı proje ile desteklenmiştir.

Doç. Dr. Rafet AKDENİZ danışmanlığında, Ramazan ÇOLAK tarafından hazırlanan “Çok Kanallı Ortamlarda Gürültü Azaltma” isimli bu çalışma aşağıdaki jüri tarafından Elektronik ve Haberleşme Mühendisliği Anabilim Dalı’nda Yüksek Lisans Tezi olarak oy birliği ile kabul edilmiştir.

Juri Başkanı: Prof. Dr. Hafız ALİSOY

*İmza:*

Üye: Prof. Dr. Erdem UÇAR

*İmza:*

Üye: Doç. Dr. Rafet AKDENİZ (Danışman)

*İmza:*

Fen Bilimleri Enstitüsü Yönetim Kurulu adına

Doç. Dr. Bahar UYMAZ  
Enstitü Müdürü

## ÖZET

Yüksek Lisans Tezi

### ÇOK KANALLI ORTAMLARDA GÜRÜLTÜ AZALTMA

**Ramazan ÇOLAK**

Tekirdağ Namık Kemal Üniversitesi  
Fen Bilimleri Enstitüsü  
Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Rafet AKDENİZ

Bu tezde, geliştirilen bir VAD algoritması kullanılarak; Wiener filtre, konuşma bozukluğu ağırlıklı Wiener filtre, uzamsal tahmin filtresi, minimum varyans gürültüsüz yanıt filtresi gibi çok kanallı gürültü azaltma algoritmalarının verimliliği karşılaştırılmıştır. Algoritmalarda dört farklı ses sinyali kullanılmakla birlikte, bu ses sinyallerine farklı açılarda eklenen üç farklı gürültü tipi kullanılmıştır. Filtre katsayılarını hesaplamadan önce gürültülü konuşma ve yalnızca gürültülü bölgelerin tespiti bir ses etkinliği algılama algoritması kullanılarak yapılmıştır. Bu algoritmada gürültülü konuşma ve yalnızca gürültülü bölgeleri ayırt edebilmek için; kısa süreli enerji, periyodiklik ve spektral düzlük gibi özellikler kullanılmıştır. Gürültülü konuşma ve gürültülü bölgelerin tespiti yapıldıktan sonra yukarıda bahsedilen algoritmalarla filtre katsayıları hesaplanmıştır. Son olarak da hesaplanan filtre katsayıları ile girişteki referans mikrofonunun frekans bileşenleri çarpılarak her algoritma için iyileştirilmiş sinyaller elde edilmiştir. Algoritmaların performansını hesaplamak için objektif bir metot olan bölümsel sinyal gürültü oranı ve subjektif bir metot olarak da ortalama yargı değeri kullanılmıştır. Gürültü azaltma açısından konuşma bozukluğu ağırlıklı Wiener filtrenin performansı daha iyi olsa da konuşmanın anlaşılabilirliği ve algısal kalite açısından SP ve MVDR filtreleri ile daha iyi sonuçlar alınmıştır.

**Anahtar Kelimeler:** ses etkinliği algılama, adaptif filtre.

2019, 93 sayfa

## **ABSTRACT**

MSc. Thesis

NOISE REDUCTION IN MULTI CHANNEL MEDIUM

**Ramazan ÇOLAK**

Tekirdağ Namık Kemal University  
Graduate School of Natural and Applied Sciences  
Department of Electronics and Communications Engineering

Supervisor: Assoc. Prof. Dr. Rafet AKDENİZ

In this thesis, using a developed VAD algorithm; the efficiency of multichannel noise reduction algorithms such as Wiener filter, Speech Distortion Weighted Wiener Filter, Spatial Prediction Filter, and Minimum Variance Distortionless Response (MVDR) was compared. While four different audio signals are used in the algorithms, three different types of noise are added to these audio signals at different angles. Before calculating the filter coefficients, noisy speech and only the detection of noisy areas were performed using a voice activity detection algorithm. To be able to distinguish between noisy speech and only noisy areas in this algorithm; Short Time Energy, periodicity and spectral flatness are used. After noisy speech and noisy areas were determined, the coefficients of the filter were calculated with the algorithms mentioned above. Finally, the calculated filter coefficients and the frequency components of the reference microphone at the input are multiplied to obtain improved signals for each algorithm. Segmental SNR is an objective method for calculating the performance of algorithms and MOS is used as a subjective method. Although the performance of the SDW wiener filter is better in terms of noise reduction, better results are obtained with SP and MVDR filters in terms of speech intelligibility and perceptual quality.

**Keywords :** voice activity detection, adaptive filter.

**2019, 93 pages**

# İÇİNDEKİLER

## Sayfa

<b>ÖZET</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>İÇİNDEKİLER</b> .....	iii
<b>ÇİZELGE DİZİNİ</b> .....	vi
<b>ŞEKİL DİZİNİ</b> .....	vii
<b>SİMGELER VE KISALTMALAR DİZİNİ</b> .....	x
<b>TEŞEKKÜR</b> .....	xii
<b>1. GİRİŞ</b> .....	1
<b>2. KURAMSAL TEMELLER</b> .....	7
2.1 Konuşma Sinyallerinde Temel Kavramlar .....	7
2.1.1 Ses.....	7
2.1.2 Gürültü.....	7
2.1.3 Konuşmanın oluşumu .....	8
2.1.3.1 Akciğer ve nefes borusu .....	8
2.1.3.2 Gırtlak.....	8
2.1.3.3 Ses sistemi .....	9
2.1.4 Konuşmanın akustik analizi .....	9
2.1.5 İşitme sisteminin özellikleri.....	10
2.2 Konuşma Analiz Yöntemleri .....	10
2.2.1 Dönüşüm (Transform) temelli metotlar.....	11
2.2.1.1 Laplace dönüşümü.....	11
2.2.1.2 Fourier dönüşümü.....	12
2.2.1.3 Z dönüşümü .....	13
2.2.1.4 Dalgacık (Wavelet) dönüşümü .....	14
2.2.2 Model temelli metotlar .....	14
2.2.2.1 Doğrusal tahmin (Linear prediction) .....	14
2.2.2.2 Kalman filtresi .....	15
2.2.2.3 Adaptif filtre .....	15
2.2.3 Bayesian tahmini temelli metotlar.....	15
2.2.4 Nöron elementlerinin katmanlı ağları temelli metotlar .....	16
2.3 Ses Etkinliği Algılama (VAD, Voice Activity Detection) .....	16
2.3.1 Kısa süreli enerji (STE, Short time energy) .....	17
2.3.2 Periyodiklik (Periodicity) .....	18
2.3.3 Spektral düzlük (SF, Spectral flatness).....	20

2.3.4 Sıfır geiř oranı (ZCR, Zero crossing rate) .....	21
2.3.5 En baskın frekans bileřeni (Most dominant frequency component) .....	21
2.3.6 Yksek- alak frekans oranı (High-low frequency rate) .....	22
2.4 Filtreler .....	22
2.4.1 Sonlu drt yanıtlı (FIR, Finite Impulse Response) filtreler.....	23
2.4.1.1 Dođrudan biim yapısı (Direct form) .....	23
2.4.1.2 Art arda biim yapısı (Cascade form) .....	23
2.4.1.3 Dođrusal faz biim yapısı (Linear phase form).....	24
2.4.1.4 Frekans rnekleme biim yapısı (Frequency sampling form).....	24
2.4.2 Sonsuz drt yanıtlı (IIR, Infinite Impulse Response) filtreler .....	25
2.4.2.1 Dođrudan biim yapısı (Direct form) .....	25
2.4.2.2 Art arda biim yapısı (Cascade form) .....	26
2.4.2.3 Paralel biim yapısı .....	26
2.4.3 Adaptif Filtreler .....	27
2.4.3.1 ok kanallı wiener filtre (MWF, Multichannel wiener filter).....	28
2.4.3.2 Konuřma bozukluđu ađırlıklı (SDW, Speech distortion weighted) wiener filtre .....	29
2.4.3.4 Uzamsal tahmin (SP, Spatial prediction) filtresi .....	30
2.4.3.5 Minimum varyans grltsz yanıt (MVDR, Minimum variance distortionless response) filtresi .....	31
2.5 Performans lm Parametreleri .....	32
2.5.1 Ortalama yargı deđeri (MOS, Most opinion score).....	32
2.5.2 P.800.....	33
2.5.3 P.805.....	33
2.5.4 P.830.....	33
2.5.5 Blmlenmiř sinyal grlt oranı (Segmental SNR).....	33
2.5.6 Konuřma kalitesinin algısal deđerlendirmesi (PESQ, Perceptual evaluation speech quality).....	34
2.5.7 Algısal analiz lm sistemleri (PAMS, Perceptual analysis measurement system).....	35
2.5.8 Algısal konuřma kalite lm (PSQM, Perceptual speech quality measurement).....	35
<b>3. MATERYAL ve YNTEM</b> .....	<b>36</b>
3.1 Materyal.....	36
3.2 Temel Kavramlar .....	38
3.2.1 rtřme (Overlapping).....	38
3.2.2 Pencereleme tekniđi.....	39
3.2.3 Sesli erevelerin eklenmesi (Overlap-Add) Metodu .....	40
3.3 Yntem .....	41
3.3.1 VAD algoritması .....	41
3.3.2 Adaptif filtre tasarımı .....	42



3.3.2.1 Kısa süreli fourier dönüşümü (STFT, Analiz).....	43
3.3.2.2 Özilinti matrislerinin güncellenmesi .....	44
3.3.2.3 Filtre katsayılarının hesaplanması .....	45
3.3.2.4 Filtreleme.....	46
3.3.2.5 Ters Fourier dönüşümü (Sentez) .....	46
<b>4. ARAŞTIRMA BULGULARI .....</b>	<b>47</b>
4.1 VAD Sonuçları .....	47
4.2 SNR Sonuçları .....	53
4.3 MOS Sonuçları .....	78
<b>5. TARTIŞMA ve SONUÇ .....</b>	<b>79</b>
<b>KAYNAKLAR.....</b>	<b>81</b>
EK 1: VAD algoritmasının MATLAB programı aracılığıyla yazılan kaynak kodu .....	84
EK 2: Filtre katsayılarının hesaplanması için gerekli kaynak kod .....	87
EK 3: İstenilen SNR değeri için katsayı hesaplanmasına ilişkin kaynak kod .....	91
<b>ÖZGEÇMİŞ .....</b>	<b>93</b>

## ÇİZELGE DİZİNİ

	<b><u>Sayfa</u></b>
Çizelge 2.1 : Ses kalitesi ile MOS arasındaki ilişki.....	33
Çizelge 3.1 : Filtre Katsayıları formül tablosu .....	46
Çizelge 4.1 : sp09.wav ve sp11.wav ses sinyallerinin VAD sonuçları .....	49
Çizelge 4.2 : spkadin.wav ve sperkek.wav ses sinyallerinin VAD sonuçları.....	50
Çizelge 4.3 : sp09.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları .....	53
Çizelge 4.4 : sp11.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları .....	54
Çizelge 4.5 : sp09.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları .....	57
Çizelge 4.6 : sp11.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları .....	58
Çizelge 4.7 : spkadin.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları .....	61
Çizelge 4.8 : sperkek.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları .....	62
Çizelge 4.9 : spkadin.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları ...	65
Çizelge 4.10 : sperkek.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları....	66
Çizelge 4.11 : sperkek.wav ve spkadin.wav ses sinyallerinin MOS sonuçları.....	78

## ŞEKİL DİZİNİ

### Sayfa

Şekil 1.1 : Adaptif filtre blok diyagramı .....	1
Şekil 1.2 : Tek kanallı gürültü azaltma blok şeması .....	2
Şekil 1.3 : Çok kanallı gürültü azaltma blok şeması .....	2
Şekil 2.1 : Merhaba kelimesinin zaman düzlemi gösterimi .....	10
Şekil 2.2 : Sinyal işleme metotları .....	11
Şekil 2.3 : Üç tane sinyalin zaman ve frekans düzlemindeki gösterimi.....	13
Şekil 2.4 : Kalman filtre yapısı .....	15
Şekil 2.5 : VAD blok şeması.....	17
Şekil 2.6 : Temiz bir konuşma sinyali ve onun STE'si.....	18
Şekil 2.7 : Gürültülü bir konuşma sinyali ve onun STE'si .....	18
Şekil 2.8 : Temiz bir konuşma sinyali ve onun periyodiklik grafiği.....	19
Şekil 2.9 : Gürültülü bir konuşma sinyali ve onun periyodiklik grafiği .....	19
Şekil 2.10: Temiz bir konuşma sinyali ve onun SF grafiği.....	20
Şekil 2.11: Gürültülü bir konuşma sinyali ve onun SF grafiği .....	20
Şekil 2.12: Sıfır geçiş noktası gösterimi .....	21
Şekil 2.13: Doğrudan biçim yapısı.....	23
Şekil 2.14: Art arda biçim yapısı.....	24
Şekil 2.15: Doğrusal faz biçim yapısı .....	24
Şekil 2.16: Frekans örnekleme biçim yapısı.....	24
Şekil 2.17: Doğrudan biçim I yapısı.....	25
Şekil 2.18: Doğrudan biçim II yapısı.....	26
Şekil 2.19: Art arda biçim yapısı .....	26
Şekil 2.20: Paralel biçim yapısı .....	27
Şekil 2.21: Adaptif filtre blok şeması.....	27
Şekil 2.22: Çok kanallı gürültü azaltma yapısının blok diyagramı .....	28
Şekil 2.23: Performans ölçüm yöntemlerinin gösterimi.....	32
Şekil 3.1 : Algoritma testinde kullanılan temiz konuşma sinyalleri .....	36
Şekil 3.2 : Algoritma testinde kullanılan gürültü sinyalleri .....	37
Şekil 3.3 : Ses kayıtlarının yapıldığı stüdyonun tasarımı.....	38
Şekil 3.4 : Çerçeveler arasındaki örtüşmenin gösterimi.....	39
Şekil 3.5 : Hanning pencere dürtü tepkisi gösterimi .....	40
Şekil 3.6 : Ses çerçevelerinin eklenmesinin gösterimi.....	40
Şekil 3.7 : Bütün algoritmanın blok diyagramı .....	43
Şekil 4.1 : sp09.wav sinyaline ait mükemmel VAD sonucunun gösterimi.....	47
Şekil 4.2 : sp11.wav sinyaline ait mükemmel VAD sonucunun gösterimi.....	47
Şekil 4.3 : sperkek.wav sinyaline ait mükemmel VAD sonucunun gösterimi.....	48
Şekil 4.4 : spkadin.wav sinyaline ait mükemmel VAD sonucunun gösterimi.....	48

Şekil 4.5 : Tüm konuşmaların açılara göre ortalama VAD sonuçları .....	51
Şekil 4.6 : Tüm konuşmaların gürültü tipine göre ortalama VAD sonuçları .....	51
Şekil 4.7 : sp09.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları .....	55
Şekil 4.8 : sp11.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları .....	55
Şekil 4.9 : sp09.wav ve sp11.wav ses sinyallerinin önerilen VAD kullanılarak algoritmalara göre elde edilen SNR sonuçları .....	56
Şekil 4.10 : sp09.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD ile elde edilen SNR sonuçları .....	59
Şekil 4.11 : sp11.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD ile elde edilen SNR sonuçları .....	59
Şekil 4.12 : sp09.wav ve sp11.wav ses sinyallerinin mükemmel VAD kullanılarak algoritmalara göre elde edilen SNR sonuçları .....	60
Şekil 4.13 : spkadin.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları .....	63
Şekil 4.14 : sperkek.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları .....	63
Şekil 4.15 : spkadin.wav ve sperkek.wav ses sinyallerinin önerilen VAD kullanılarak algoritmalara göre elde edilen SNR sonuçları .....	64
Şekil 4.16 : spkadin.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD kullanılarak elde edilen SNR sonuçları .....	67
Şekil 4.17 : sperkek.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları .....	67
Şekil 4.18 : spkadin.wav ve sperkek.wav ses sinyallerinin önerilen VAD kullanılarak algoritmalara göre elde edilen SNR sonuçları .....	68
Şekil 4.19 : 0 dB Havaalanı gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	69
Şekil 4.20 : 5 dB Havaalanı gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	69
Şekil 4.21 : 10 dB Havaalanı gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	70
Şekil 4.22 : 0 dB Araba gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	70
Şekil 4.23 : 5 dB Araba gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	71
Şekil 4.24 : 10 dB Araba gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	71
Şekil 4.25 : 0 dB Beyaz gürültü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	72

Şekil 4.26 : 5 dB Beyaz gürültü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	72
Şekil 4.27 : 10 dB Beyaz gürültü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	73
Şekil 4.28 : 0 dB Havaalanı gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	73
Şekil 4.29 : 5 dB Havaalanı gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	74
Şekil 4.30 : 10 dB Havaalanı gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	74
Şekil 4.31 : 0 dB Araba gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	75
Şekil 4.32 : 5 dB Araba gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	75
Şekil 4.33 : 10 dB Araba gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	76
Şekil 4.34 : 0 dB Beyaz gürültü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	76
Şekil 4.35 : 5 dB Beyaz gürültü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	77
Şekil 4.36 : 10 dB Beyaz gürültü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi .....	77

## SİMGELER VE KISALTMALAR DİZİNİ

### Simgeler

$A_m$	: Aritmetik Ortalama
$\text{Arg min}$	: Bir fonksiyonu minimize edecek argumanlar
$F\{.\}$	: Fourier operatörü
$\Psi$	: Dalgacık Fonksiyonu
$G_m$	: Geometrik Ortalama
$K$	: Ölçekleme Faktörü
$\text{Log}$	: Logaritma
$P$	: Gürültü Enerjisinin Sabit Çarpanı
$\text{SF}$	: Spektral Düzlük (Spectral Flatness)
$C$	: Periyodiklik
$R(\tau)$	: Özilinti fonksiyonu
$R_x$	: Konuşma Özilinti Matrisi
$R_y$	: Gürültülü Konuşma Özilinti Matrisi
$R_v$	: Gürültü Özilinti Matrisi
$\varepsilon$	: Beklenti operatörü
$\alpha_n$	: Gürültü Özilinti Matrisi İçin Unutma Faktörü
$\alpha_y$	: Gürültülü Konuşma Özilinti Matrisi İçin Unutma Faktörü
$\mu$	: Gürültü Ekleme Katsayısı
$\tau$	: Gecikme Değeri

## **Kısaltmalar**

A	: Average
dB	: Decibel
FFT	: Fast Fourier Transform
HR0	: Hit Rate Silence
HR1	: Hit Rate Speech
Hz	: Hertz
ITU	: International Telecommunication Union
kHz	: Kilohertz
MOS	: Most Opinion Score
MMSE	: Minimum Mean Squared Error
MSE	: Mean Squared Error
MVDR	: Minimum Variance Distortionless Response
PAMS	: Perceptual Analysis Measurement System
PESQ	: Perceptual Evaluation Speech Quality
PSQM	: Perceptual Speech Quality Measurement
SD	: Speech Distortion
SDW	: Speech Distortion Weighted
SF	: Spectral Flatness
SNR	: Signal to Noise Ratio
SP	: Spatial Prediction
STE	: Short Time Energy
STFT	: Short Time Fourier Transform
VAD	: Voice Activity Detection
ZCR	: Zero Crossing Rate

## TEŞEKKÜR

Yüksek lisans tezi çalışmalarım sırasında kıymetli bilgi, birikim ve tecrübeleri ile bana yol gösterici ve destek olan değerli danışman hocam Doç. Dr. Sayın Rafet AKDENİZ'e, sonsuz teşekkür ve saygılarımı sunarım.

Lisans ve yüksek lisans eğitimim boyunca yardım, bilgi ve tecrübeleri ile bana sürekli destek olan bölümdeki tüm hocalarıma teşekkürü bir borç bilirim.

Simülasyon çalışmalarım esnasında ses kayıtlarını elde edebilmem için stüdyosunu kullandığım ve de çalıştığım kurum olan TRT'ye teşekkür ederim.

Çalışmalarım boyunca manevi destekleriyle beni hiçbir zaman yalnız bırakmayan her konuda arkamda olduğunu bildiğim eşime de çok teşekkür ederim.

Mayıs, 2019

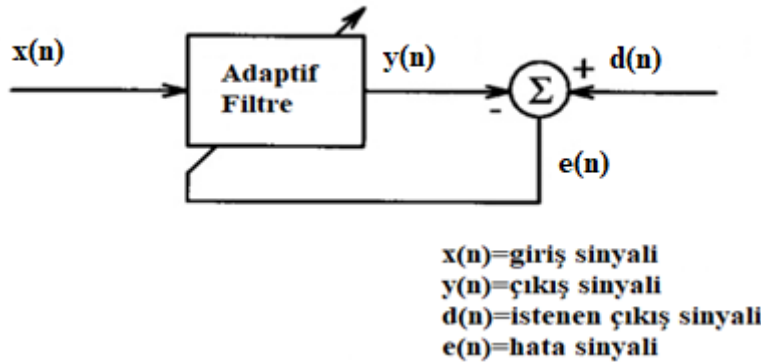
Ramazan ÇOLAK  
Elektronik ve Haberleşme Mühendisi



## 1. GİRİŞ

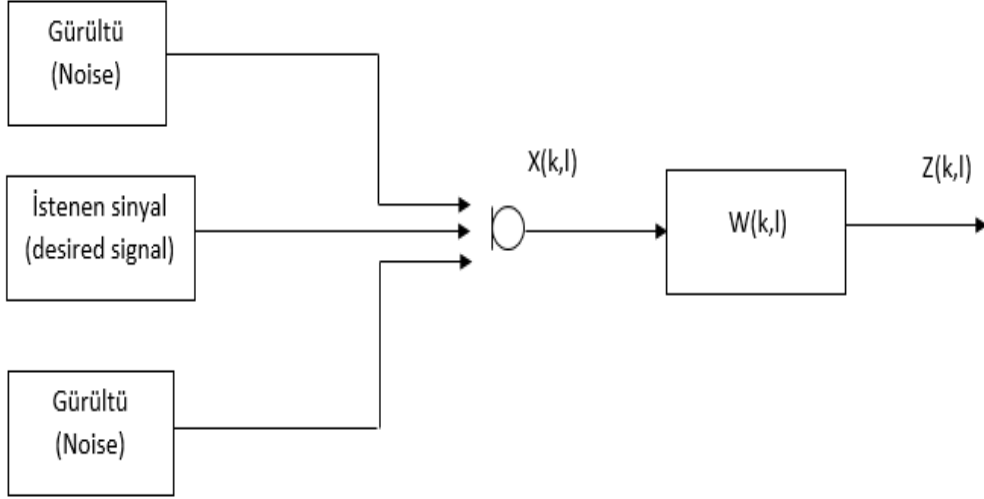
İletişim günlük hayatta insanlar için çok önemli bir yer tutmaktadır. Sağlıklı bir iletişimin yapılmasında en önemli engellerden biri gürültüdür. Çevremizdeki cihazların, insanların, doğal seslerin (su, rüzgâr) varlığıyla birlikte gürültü seviyesi insan sağlığını, iletişimin niteliğini olumsuz anlamda ciddi olarak etkileyecek boyuta gelmiştir. Bu durum iletişim sektörünün büyük bir problemi olmakla birlikte işitme sorunu yaşayan insanların kullandığı cihazlar nedeniyle sağlık sektörünü de etkilemektedir.

Gürültü genel anlamda insanı rahatsız eden her türlü ses olarak tanımlanabilir. Bazen frekans spektrumunda belirli bir frekans bölgesinde bulunurken bazen frekans spektrumundaki yeri zamanla değişebilir. Frekans bölgesinin bir bölgesindeki sabit frekanstaki gürültüyü azaltabilmek için sabit bir frekansı zayıflatabilen filtre (bant durduran filtre) kullanılabilir ancak gürültü frekansı zamanla değişiyorsa bunun için adaptif bir filtre kullanılmalıdır. Adaptif filtre, gürültünün karakteristiğine göre filtre katsayılarını sürekli güncelleyen bir filtredir.

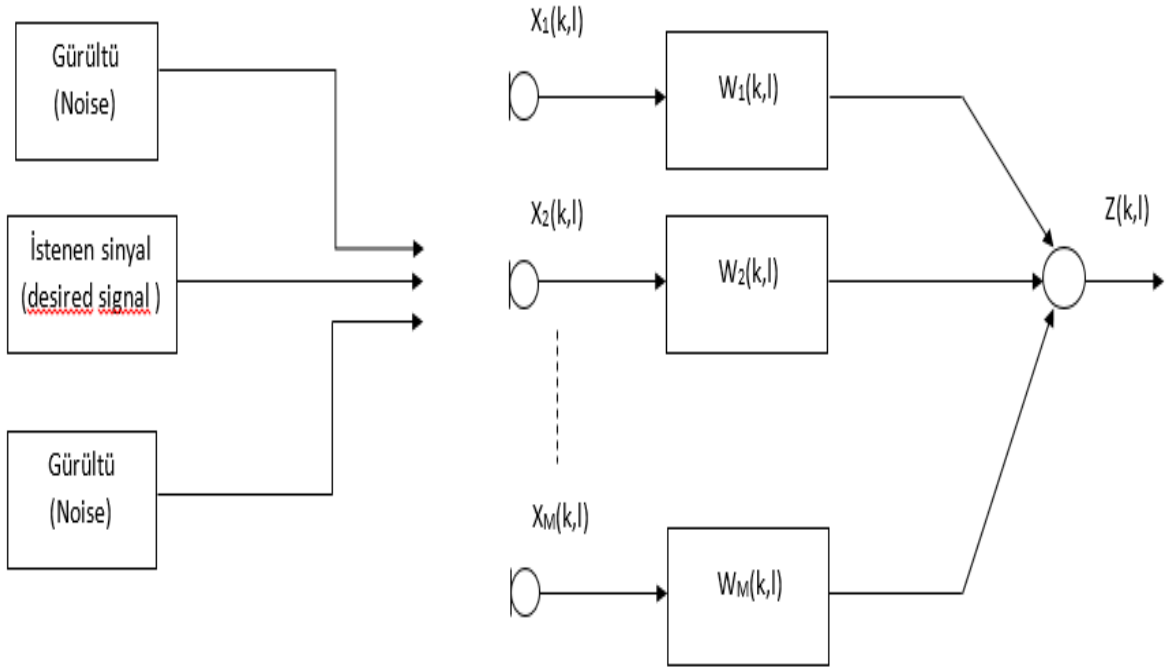


Şekil 1.1. Adaptif filtre blok diyagramı

Şekil 1.1'deki blok şemadan da anlaşıldığı üzere adaptif filtre; gürültülü giriş sinyali ile olması istenen temiz ses sinyalini referans alarak, bu iki sinyal arasındaki hatayı en aza indirmeye çalışan yinelemeli bir filtre yapısıdır. İlk zamanlarda gürültü azaltma metodlarında tek mikrofonlu çalışmalar yapılmıştır (Benesty ve Huang 2011, Huang ve Benesty 2012, Yousheng ve Jianwen 2014). İlerleyen zamanlarda alınan sinyaller arasındaki ilintiden faydalanılarak birden fazla mikrofon ile yapılan gürültü azaltma metodlarından daha iyi sonuçlar alınmıştır (Benesty ve ark. 2011, Modhave ve ark. 2016, Modhave ve ark. 2017).



**Şekil 1.2.** Tek kanallı gürültü azaltma blok şeması



**Şekil 1.3.** Çok kanallı gürültü azaltma blok şeması

Şekil 1.2 ve Şekil 1.3'te tek ve çok kanallı gürültü azaltma algoritmalarının blok yapısı gösterilmiştir. Tek kanallı gürültü azaltma algoritmalarında tüm ses bir mikrofon aracılığıyla alınır ve o mikrofon sinyali üzerinden gürültünün analizi yapılır ve filtre katsayıları hesaplanır. Çok kanallı algoritmalarda ise ortamdaki ses birden fazla mikrofon ile alınır. Bu mikrofonlardan birisi (genelde birinci mikrofon) referans mikrofonu seçilir. Tüm

mikrofonlardan alınan sinyallerin her zaman aralığı ve bu zaman aralıklarının her frekans bileşeni için özilinti matrisleri oluşturulur. Böylelikle filtre katsayı matrisleri oluşturulurken tüm mikrofonlardan alınan sinyaller hesaba katılır ayrıca mikrofon sayısı ile algoritmanın performansı doğru orantılıdır. Tüm mikrofonlar için ayrı ayrı filtre katsayıları hesaplandıktan sonra belirlenen referans mikrofondan alınan sinyali yine o mikrofonun filtre katsayıları ile çarpılıp gürültüsü azaltılmış sinyal elde edilir.

Ses etkinliği algılama (Voice Activity Detection, VAD) uygulaması kullanılarak, tasarlanan adaptif filtrelerin karşılaştırılması amacıyla yapılan bu çalışma altı bölümden oluşmaktadır. Birinci bölümde; konuyla ilgili genel bilgi verilmiştir. İkinci bölümde ise konuyla ilgili bu zamana kadar yapılan adaptif filtre çalışmalarına dair bir literatür özeti ile kuramsal bilgilere yer verilmiştir. Üçüncü bölümde; tez çalışması esnasından kullanılan materyallere ve nasıl bir yöntem kullanıldığına dair bilgi verilmiştir. Dördüncü bölümde; bu çalışmada elde edilen araştırma bulguları detaylı bir şekilde açıklanmıştır. Beşinci bölümde; elde edilen sonuçlara ve önerilere yer verilmiştir. Altıncı bölümde ise literatür çalışması esnasında kullanılan kaynaklara yer verilmiştir.

Uzun yıllardır iletişim sektörünün gelişimiyle paralel ilerleyen bu alanda birçok çalışma yapılmıştır. Bu tez çalışması iki temel bölüme ayrılıp çalışılmıştır:

- İkili VAD
- Adaptif filtre tasarımı

Adaptif filtre algoritmalarında filtre katsayılarını belirlemek için ses sinyalinin yalnızca gürültülü bölgeleri ve gürültülü konuşma bölgelerinin tespiti çok önemlidir. Bu bölgelerin tespitini yapmak için VAD kullanılır. VAD uygulamalarında kullanılan belli başlı özellikler vardır. Bu özelliklerden bir ya da birkaç tanesi kullanılarak verimli bir VAD algoritması oluşturulabilir.

Kısa süreli enerji (Short Time Energy, STE) özelliği yalnız başına özellikle düşük sinyal gürültü oranı (Signal to Noise Ratio, SNR ) değerlerinde verimli bir VAD için istenilen sonuçları vermemektedir. Bu nedenle STE'nin dışında frekans düzleminde spektral düzlük (Spectral Flatness, SF) ve en baskın frekans bileşeni (Most Dominant Frequency Component) kullanılarak düşük SNR değerlerinde daha iyi sonuçlar veren bir VAD algoritması geliştirilmiştir (Moattar ve Homayounpour 2009).

Sakhnov ve ark. (2009) tarafından geliştirilen farklı bir algorithmada ise, sesin periyodik olma özelliği ve yüksek-düşük frekans oranı kullanılarak verimli bir VAD algoritması geliştirilmesi amaçlanmıştır.

Sakhnov ve Verteletskaya (2010) tarafından geliştirilen başka bir algoritmada ise periyodik olma özelliği, yüksek düşük frekans oranının yanı sıra STE özelliği de kullanılarak yeni uygulanması kolay bir algoritma oluşturulmuştur.

Peak Valley Difference (maximum minimum farkı) olarak isimlendirilen farklı bir özellik ile birlikte STE, maksimum frekans bileşeni ve SF özellikleri kullanılarak Moattar ve ark. (2010) tarafından yeni bir VAD algoritması sunulmuştur. Bir başka çalışmada zaman düzleminde ve frekans düzleminde VAD kavramları kullanılarak lineer enerji temelli detektör, bulanık mantık (Fuzzy Logic) ve yapay sinir ağları temelli VAD performansları karşılaştırılmıştır (Bharath ve ark. 2016).

Zaw ve War (2017) tarafından yapılmış olan bir çalışmada; STE, sıfır geçiş oranı (Zero Crossing Rate, ZCR) spektral dağıntı (Spectral Entropy) ve doğrusal tahmini hata (Linear Prediction Error, LPE) gibi özellikler kullanılarak ses sinyalinin her çerçevesi için bir  $D$  bileşik parametresi hesaplanarak ve her ses çerçevesinin  $D/D_{max}$  olarak belirlenen eşik seviyesinin üstünde olup olmamasına göre konuşmanın varlığı tespit edilmiştir.

Yukarıdaki çalışmalardan faydalanılarak STE, SF ve periyodiklik özellikleri kullanılarak bir VAD algoritması tasarlanmıştır. Sonraki aşamada adaptif filtre tasarımı için filtre katsayılarının belirlenmesi gerekir. Literatürde bu konuda yapılmış birçok çalışma vardır. Bu tez çalışmasında Wiener filtre temelli filtre yapıları üzerinde çalışılmıştır. Bu tez çalışmasında faydalanılan bazı gürültü azaltma algoritmaları aşağıda verilmiştir:

Meyer ve Simmer (1997) tarafından araç içindeki gürültüyü azaltmaya yönelik çalışma yapılmıştır. Öncelikle zaman düzlemindeki sinyal frekans düzlemine çevrilip, alçak geçiren filtre ile spektral çıkarma (Spectral Subtraction) ve yüksek geçiren filtre ile Wiener filtre katsayılarını hesaplanıp birleştirilmesiyle yeni algoritma üretilmiştir. Bu yeni metot ile geleneksel spektral çıkarma ve Wiener algoritmalarından daha iyi sonuçlar elde edilmiştir.

Rama Rao ve ark. (2008) tarafından önerilen bu çalışmada iki aşamalı bir hibrit sistem geliştirilmiştir. Spektral kazanç hesaplamalarında, hesaplama kolaylığı için frekans indeksleri homojen olmayan bir biçimde dağıtılmıştır. Bu sistemin ilk aşamasında Ephraim-Malah kazanç fonksiyonuna oluşturulan yumuşatılmış karar kazanç mekanizması uygulanmıştır. İkinci aşamada ise gürültü azaltma için psikoakustik maskeleye eşiği kullanılmıştır. Önerilen bu yöntem değiştirilmiş spektral çıkarma (Spectral Subtraction) ve spektral ağırlıklandırma (Spectral Weighting) algoritmalarıyla karşılaştırılmış ve sabit olmayan gürültülü ortamlarda kullanılmaya uygun olduğu değerlendirilmiştir.

Cornelis ve ark. (2009) tarafından yazılan bu makalede frekans düzleminde Wiener filtre temelli iki algoritmanın karşılaştırılması yapılmıştır. Spatial Prediction (uzamsal tahmin)

ve Speech Distortion Weighted Multichannel Wiener Filter (konuşma bozulması ağırlıklı çok kanallı wiener filtre) olarak belirtilen bu iki filtre tipinin SNR ve konuşma bozukluğu (Speech Distortion, SD) performansları değerlendirilmiştir. Her iki performans ölçütüne göre de uzamsal tahmin filtresinin daha iyi sonuçlar verdiği tespit edilmiştir.

Benesty ve ark. (2011) tarafından önerilen zaman düzleminde çok kanallı gürültü azaltma algoritmasında gürültüsüz mikrofon dizilerindeki sinyali iki bileşene ayrılmıştır. İlk bileşen hedef konuşma sinyali ile tamamen tutarlı, ikinci bileşen ise girişim kalıntılarıdır. Bu ayrıştırma tekniği bilenen zaman düzlemi filtreleri minimum varyans gürültüsüz yanıt (Minimum Variance Distortionless Response, MVDR) Wiener, Çok kanallı bağımsız parametre (Trade off) ve doğrusal olarak kısıtlanmış minimum varyans (Linearly constrained minimum variance, LCMV) filtrelerine uygulanmıştır. Klasik filtrelere göre değiştirilmiş bu filtrelerle daha iyi sonuçlar elde edilmiştir.

Benesty ve Huang (2011) tarafından var olan tek kanal gürültü azaltma algoritmalarının aksine ses çerçeveleri arasındaki ilişki dikkate alınmıştır. Oluşturulan bu yeni model frekans düzleminde MVDR filtre için uygulanmış ve oldukça iyi sonuçlar elde edilmiştir.

Chen ve ark. (2011) tarafından bu çalışmada zaman düzleminde filtrelenmiş ses sinyalini ayrıştırmak için bir öneri sunulmuştur. Filtrelenmiş ses sinyali iki ilişkisiz bileşene ayrılmıştır. Bu yeni ayrıştırma metodu Maksimum SNR, Wiener, MVDR ve Trade off filtreleri ile denenmiştir. Deneysel sonuçlar ve teorik analizler Maksimum SNR, Wiener ve Trade off filtrenin bir ölçekleme çarpanı kullanılarak MVDR filtreyle özdeşleştirildiği saptanmıştır ancak bu ölçekleme parametresi konuşma sinyalinde bozukluğa neden olacağından, konuşma iyileştirme uygulamalarında MVDR filtre tavsiye edilmiştir.

Tek kanallı gürültü azaltma uygulamalarında yıllardır birçok çalışma yapılmış ve çok iyi sonuçlar elde edilmiştir. Ancak bu çalışmalarda ses sinyali üzerinde işlemler yapılırken sadece o an ki ses çerçevesi (frame) üzerinde işlemler yapılmış ve ses çerçeveleri arasındaki ilişki görmezden gelinmiştir. Huang ve Benesty (2012) tarafından yapılan zaman düzleminde tek kanallı gürültü azaltma çalışmasında ses çerçeveleri arasındaki ilişki dikkate alınarak yeni bir doğrusal model geliştirilmiştir. Geliştirilen bu model Wiener, MVDR filtrelerine uyarlanmıştır. Simülasyon sonuçları değerlendirildiğinde umut verici sonuçlar elde edilmiştir.

Yousheng ve Jianwen (2014) tarafından geliştirilen bu algoritmada Subspace filtrenin gürültü azaltmadaki avantajından ve Wiener filtrenin kararlı karakteristiğinden dolayı bu iki filtrenin birleşimi ile yeni bir filtre türetilmiştir. SNR ve SD ile performans değerlendirilirken, her iki performans değerlendirmesine göre de yeni algoritma klasik Subspace ve Wiener filtreye göre daha iyi sonuçlar verdiği görülmüştür.

Modhave ve ark. (2016) tarafından yapılan bu çalışmada işitme cihazlarında kullanılmak üzere, çok kanallı Wiener filtre uygulaması geliştirilmiştir. Çoklu giriş çoklu çıkış (multi-input multi-output) adı verilen bu sistem ile tek kanallı Wiener filtre uygulamalarıyla karşılaştırıldığında daha iyi sonuçlar elde edilmiştir.

Genelde konuşma iyileştirme algoritmalarında girişteki konuşma sinyalinin seviyesi ile filtrelenmiş konuşma sinyalinin seviyesi eşit değildir. Long ve ark. (2016) tarafından bu probleme yönelik hem tek kanallı konuşma iyileştirme algoritmaları için hem de çok kanallı konuşma iyileştirme algoritmaları için ikişer tane kısıtlanmış Wiener filtre tasarlanmıştır. Bu kısıtlanmış Wiener kazançları ile girişteki konuşma sinyali ile tahmin edilen konuşma sinyali arasındaki karesel hataların ortalaması (Minimum Squared Error, MSE) minimum yapılmaya çalışılmıştır. Deneysel sonuçlara bakıldığında gürültü sinyalinin genliği ile iyileştirilmiş sinyalin genliği benzer sonuçlar vermiştir.

Jin ve ark. (2017) tarafından yapılan bu çalışmada araç kullanma esnasında gelen çevre gürültülerini azaltmak için çalışılmıştır. İlk önce belirli bir frekansın altını zayıflatmak için gürültülü sinyale yüksek geçiren filtre uygulanmıştır. Daha sonra adaptif filtre katsayılarının daha doğru hesaplanabilmesi için bir engelleyici matris kullanılmıştır. Bu sayede filtre katsayıları hesaplanırken yalnızca gürültü bileşenleri dikkate alınmıştır. Son olarak da önerilen geliştirilmiş yan lop iptal edici (Generalized Sidelobe Canceller, GSC) ile çok bantlı spektral çıkarma (Multiband Spectral Subtraction) özellikleri birleştirilmiştir. Elde edilen sonuçlarla algoritmanın hem GSC hem de geliştirilmiş GSC'den daha iyi sonuçlar verdiği gösterilmiştir.

Modhave ve ark. (2017) tarafından yapılan bu çalışmada işitme cihazlarında kullanılması için yeni birçok kanallı konuşma iyileştirme algoritması önerilmiştir. Bu algorithmada tüm gürültü bileşenleri bir matris yapısı olarak düşünülmüştür. Çoklu konuşma sinyallerinde her konuşma sinyali bir gürültü sinyali ile bozulmaya uğramıştır. Wiener matris yapısı ile matrise girilen gürültü bileşenleri değerlendirilerek her konuşma sinyali için bir filtre katsayısı türetilmiştir. Bu sistem ile hem tek kanallı hem de çok kanallı Wiener filtre uygulamalarından daha iyi verim alınmıştır.

## 2. KURAMSAL TEMELLER

### 2.1 Konuşma Sinyallerinde Temel Kavramlar

#### 2.1.1 Ses

Genel olarak ses; akciğerlerden gelen havanın vücudun ses ile ilgili organlarının biçimlendirmesiyle oluşan kulak ile işitilebilen titreşimlerdir. Akciğerlerdeki havanın nefes borusundan dışarı çıkarken ses tellerini titreştirmesiyle bir başkasının işitebileceği anlamlı sesler oluşur. Oluşan bu sesler yine kulağımıza gelip kulak zarımızı titreştirir ve titreşen kulak zarı beyine gönderilmek üzere bir mesaj oluşturur ve ses karşıdaki kişi tarafından da algılanır.

Belirli seslerin bir araya gelmesiyle belirli bir dile mensup insanların birbirini anlayabileceği konuşma sinyalleri oluşur. Bu konuşma sinyalleri belirli bir düzende bir araya gelerek konuşmacının düşüncesini temsil eder. Her dilin konuşma sinyallerinin bir araya gelme düzeni ve kuralları farklıdır ve bu dil biliminin (linguistics) uğraş alanıdır.

#### 2.1.2 Gürültü

Gürültü en genel tanımıyla iletişimde istenilen ses sinyalin anlaşılmasını olumsuz olarak etkileyen istenmeyen sinyallerdir. Gürültünün termal, elektromanyetik akustik, elektrostatik gibi bazı çeşitleri olmakla birlikte bu tez çalışmasında akustik gürültü azaltılmasıyla ilgili çalışmalar yapılmıştır. Akustik gürültü hareketli titreşimli kaynaklardan yayılır. Arka planda konuşan insan sesi, trafik gürültüsü, bilgisayar fanı, rüzgâr, yağmur gibi sesler akustik gürültü kapsamına girer. Frekans spektrumundaki karakteristiğine göre de gürültü bazı alt kategorilere ayrılır (Vageshi 2006).

- **Beyaz gürültü (White noise):** Frekans spektrumunun tüm frekanslarını eşit yoğunlukta içeren gürültü tipidir.
- **Sınırlı bant beyaz gürültü (Band limited white noise):** Frekans spektrumunun belirli bir aralığında bulunan eşit yoğunluklu gürültüdür.
- **Renkli gürültü (Coloured noise):** Beyaz olmayan gürültü tipidir. Yani frekans spektrumunda frekansların yoğunlukları değişiklik gösterebilir.
- **Darbant gürültü (Narrowband noise):** Frekans spektrumunun az bir bölümünde var olan gürültüdür. Örnek olarak 50 Hz şebeke geriliminin frekansı bu tip bir gürültüdür.

### 2.1.3 Konuşmanın oluşumu

Konuşma nörolojik sinyalle başlar. Aktarılmak istenen düşünce nörolojik sinyallerin kontrolündeki ses üretim sistemi tarafından kodlanan akustik ses basınç dalgası şeklinde dinleyicinin işitme sistemi vasıtasıyla işitilir. Tekrar nörolojik sinyallere çevrilerek beyinin işitme bölgesi tarafından algılanır. Ses üretimi esnasında konuşmacının sesi aynı zamanda konuşmacı tarafından da işitilir ve bu bilgi ses üretiminde geri besleme olarak kullanılır. Konuşma üretimindeki temel bileşenler;

- Akciğer
- Nefes borusu
- Gırtlak
- Boğaz
- Ağız oyuğu
- Burun oyuğu

olarak sınıflandırılabilir.

Genellikle boğaz ve ağız birlikte değerlendirilir ve ses sistemi (Vocal tract) adını alır. Akciğer ve nefes borusu ses üretim sisteminin besleme kaynağı veya güç merkezi olarak değerlendirilebilir. Gırtlak, temel ses üretim mekanizmasını ihtiva eder (Özkaya 1994).

#### 2.1.3.1 Akciğer ve nefes borusu

Konuşma oluşması için gereken havayı akciğerler üretir ve nefes borusu aracılığıyla diğer organlara iletilir. Üretilen havayı tutma, boşaltma gibi işlemler diyafram aracılığıyla kontrol edilir.

#### 2.1.3.2 Gırtlak

Gırtlak, ses telleri, ses tellerini kontrol eden kaslar ve kıkırdak yapılardan oluşan karmaşık bir sistemdir. Ses telleri gırtlığın önü ve arkası arasında gerilmiş durumda bulunan et katmanlardır. Ses tellerinin gerginliği gırtlığın kıkırdak yapısının etrafındaki kaslarla kontrol edilir (Özkaya 1994).

Ses tellerinin akustik ve biyolojik olmak üzere iki görevi vardır. Biyolojik fonksiyonu nefes borusunu kapatarak solunum sistemini tıkanmalara karşı korumak, göğüs ve karında hava basıncının oluşmasını sağlamaktır. Ses telleri bu fonksiyonları ile ıkınma, öksürme veya



hapşırma sırasında gerekli hava basıncının oluşumunu sağlar. Ses tellerinin akustik fonksiyonu ise konuşma için temel uyarı kaynağı olmasıdır (Özkaya 1994).

### **2.1.3.3 Ses sistemi**

Bu terim genellikle ses tellerinden sonra konuşmanın oluşumuna katkı sağlayan, boğaz ve ağız oyuğunda yer alan her şeyi nitelemek için kullanılır. Ses sistemi aşağıdaki bölümlere ayrılır (Özkaya 1994).

1. Gırtlak kapağı ile ses telleri arasındaki bölüm (Gırtlak yutağı)
2. Dilin arkasında gırtlak kapağı ile yumuşak damak arasındaki bölüm (Ağız yutağı)
3. Yumuşak damak ile burun oyuğu arasındaki bölüm (Burun yutağı)
4. Sert damak, dil ve dudaklarla sınırlanan yumuşak damağın ilerisindeki bölüm (Ağız oyuğu)

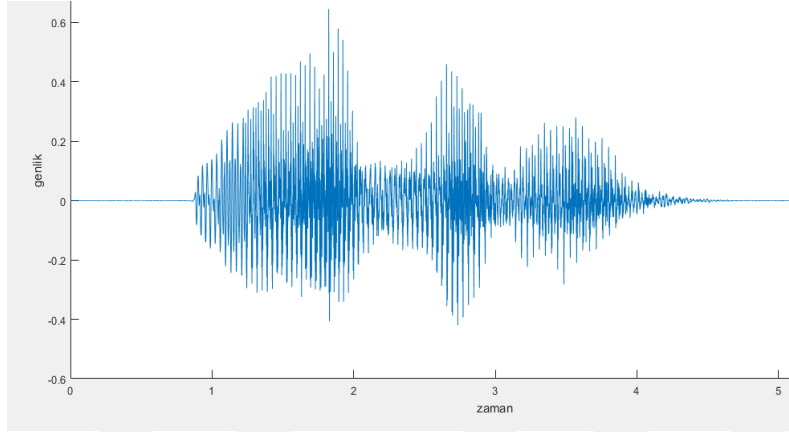
Ses sisteminin ses ile ilgili fonksiyonu sesin renklendirilmesi ve telaffuzun sağlanmasıdır. Ses sistemi aynı zamanda konuşma seslerinin yayınlandığı temel noktaları da ihtiva eder. Gırtlak kapağı hariç renklendirme ve telaffuz işine katılan temel bölümler dil, dudak ve alt çenedir.

Akustik dalga ses sisteminden geçerken, frekans içeriği sistemdeki oyuklarda oluşan titreşimle değiştirilir. Bu titreşimler ses sistemindeki çeşitli bölgelerin şekillerine bağlıdır. Dil hareket ettirilerek ağız oyuğundaki ve ağız yutağındaki boşlukların şekilleri değiştirilir (Özkaya 1994).

### **2.1.4 Konuşmanın akustik analizi**

Konuşma sinyalleri zamana bağlı olarak değişen sinyallerdir. Bu sebeple konuşma sinyalleri üzerinde işlem yapılırken çok küçük zaman dilimlerine ayrılıp, bu küçük ses çerçeveleri üzerinde işlem yapılır. Bu kısa süreli ses çerçevelerinde konuşma sinyalleri benzer özellik gösterirler. Konuşma sinyallerinde hem sesli hem de sessiz harfler olduğu için, sinyal analizinde farklılık gösterirler. Sesli harfler akciğerlerden çıkan havanın herhangi bir engelle karşılaşmaması nedeniyle genliği daha yüksek, sessiz harfler ise engellerle karşılaştığı için genliği daha düşüktür. Şekil 2.1’de örnek olarak “merhaba” kelimesinin zaman düzleminde gösterimi verilmiştir. Şekilde görüldüğü gibi üç ayrı yüksek genlikli bölüm, bu kelimenin üç tane sesli harfine karşılık gelir. Her bir yüksek genlikli bölümün başlangıcında bulunan düşük

genlikli kısımlar ise sesli harfler öncesindeki sessiz harflere karşılık gelir. Konuşma sinyallerinin bant genişliği yaklaşık olarak 5 KHz civarındadır ve frekans yoğunluğu da genel olarak 1 KHz altındadır.



**Şekil 2.1.** Merhaba kelimesinin zaman düzlemi gösterimi

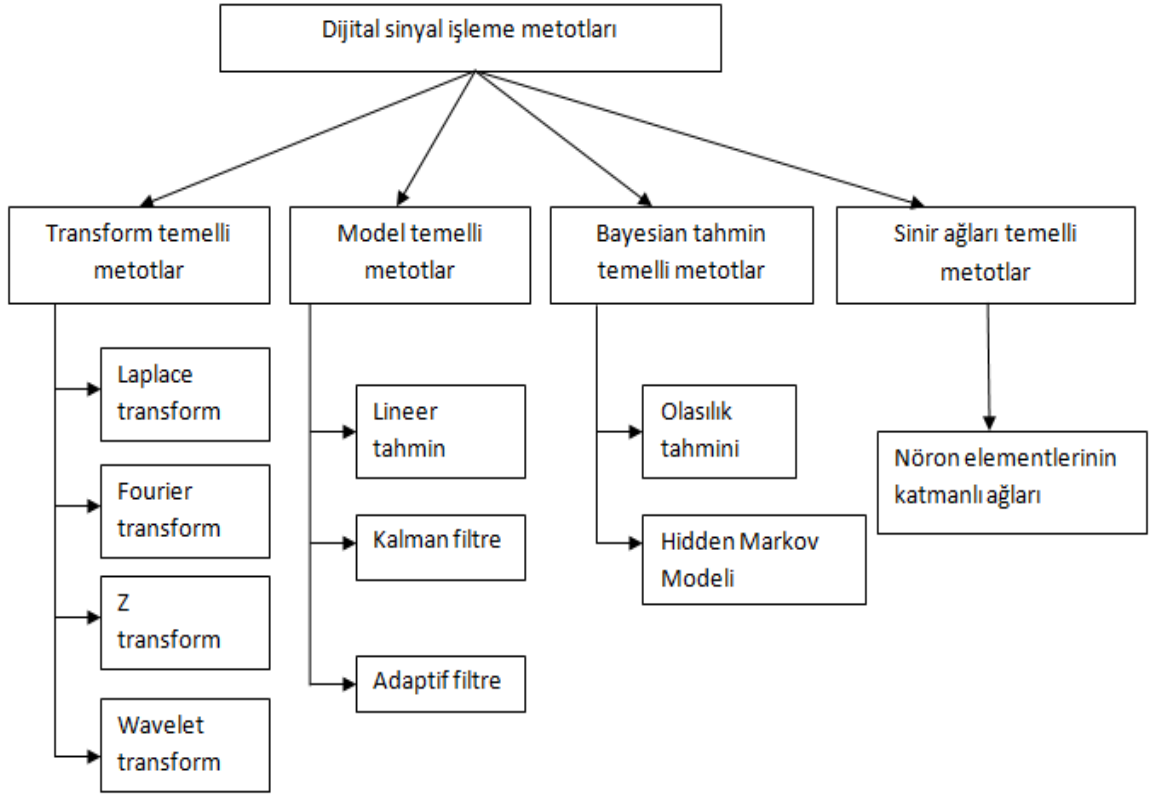
### **2.1.5 İşitme sisteminin özellikleri**

İşitme sistemi dış kulak, orta kulak ve iç kulak olmak üzere üç bölümü vardır. Dış kulağın görevi; ortamdaki ses dalgalarını toplayıp orta kulağa iletmektir. Orta kulak ise aldığı bu ses dalgalarının enerjisini değiştirerek iç kulağa iletir. İç kulak ise gelen ses dalgalarını beyinin algılayacağı forma dönüştürüp beyine gönderir.

İnsan kulağı 20 Hz-20 KHz arasındaki ses sinyallerini duyabilir ancak bu aralık yaşın ilerlemesiyle daha da daralır. Konuşma sinyalleri de yaklaşık olarak 150 Hz-2 KHz frekans aralığında olmakla birlikte 2 KHz'in üstünde de her insana özgü konuşma tonunu belirleyen konuşma sinyalinin harmonik bileşenleri vardır. Algılanan sesin şiddeti ise 0-140 dB arasındadır. Uzun süreli yüksek ses işitme organlarına kalıcı hasarlar verebilir.

### **2.2 Konuşma Analiz Yöntemleri**

Sinyal işleme süreçleri kullanılan metotlara bağlı olarak dört kategoriye ayrılır. Bunlar dönüşüm temelli metotlar, model temelli metotlar, Bayesian tahmini metotlar ve sinir ağları metotlarıdır. Şekil 2.2'de bu metotlar gösterilmiştir (Vageshi 2006).



Şekil 2.2. Sinyal işleme metotları (Vageshi 2006).

## 2.2.1 Dönüşüm (Transform) temelli metotlar

Bu dönüşümün amacı; girişinde belirlenen bir sinyal dizisini, bir sinyal veya bir sistem olarak tanımlamaktır. Dönüşüm temelli metotlar; Laplace dönüşümü, Fourier dönüşümü, Z dönüşümü ve Wavelet dönüşümüdür (Vageshi 2006).

### 2.2.1.1 Laplace dönüşümü

Laplace dönüşümü ile diferansiyel denklemler içeren sistemler daha basit cebirsel denklemlere dönüştürülür ve hesaplama kolaylığı sağlanır.  $x(t)$  gibi zamana bağlı sürekli bir sinyalin Laplace dönüşümü Eşitlik (2.1)'de gösterilmiştir.

$$X(s) = L\{x(t)\} = \int_{t=-\infty}^{\infty} x(t)e^{-st} dt \quad (2.1)$$

Eşitlik (2.1)'de  $s=j\omega$  için integral ifadesi Fourier dönüşümünü verir.  $s=\sigma+j\omega$  olması durumunda ise bu ifadeye Laplace dönüşümü denir.

$$\begin{aligned}
X(s) &= \int_{t=-\infty}^{\infty} x(t)e^{-st} dt \xrightarrow{s=\sigma+jw} X(\sigma+jw) = \int_{t=-\infty}^{\infty} x(t)e^{-(\sigma+jw)t} dt \\
X(\sigma+jw) &= \int_{-\infty}^{\infty} x(t)e^{-\sigma t} e^{-jw t} dt = \int_{-\infty}^{\infty} [x(t)e^{-\sigma t}] e^{-jw t} dt
\end{aligned}
\tag{2.2}$$

Bu durumda eşitliğin sağ tarafının  $x(t)e^{-\sigma t}$  ifadesinin Fourier dönüşümüne eşit olduğu görülür. Sonuç olarak bakıldığında Laplace dönüşümü karmaşık  $s$  düzleminde  $s=jw$  ekseninde hesaplandığında Fourier dönüşümü ile aynı sonucu verir.

### 2.2.1.2 Fourier dönüşümü

Fourier dönüşümü ile herhangi bir dalga formu sinüs ve kosinüs fonksiyonlarının bir bileşeni olarak ifade edilir. Fourier dönüşümü ile zaman düzleminde frekans düzlemine geçilip bu formda işlemler ve değerlendirmeler yapılabilir. Örneğin zaman düzleminde konuşma sinyalinde sadece zamana göre genliğinin değişimi görülebilir ancak frekans düzleminde işlem yapıldığında o sinyalin frekans spektrumu incelenebilir ve buna göre sinyal ile ilgili işlemler yapılabilir. Bu çalışmada sinyal işleme süreçlerinde ayrık zamanlı işaretler kullandığı için Fourier dönüşümü de ayrık zamanlı olarak aşağıdaki gibi verilebilir:

$$X(e^{jw}) = F\{x(n)\} = \sum_{n=-\infty}^{\infty} x(n)e^{jwn}
\tag{2.3}$$

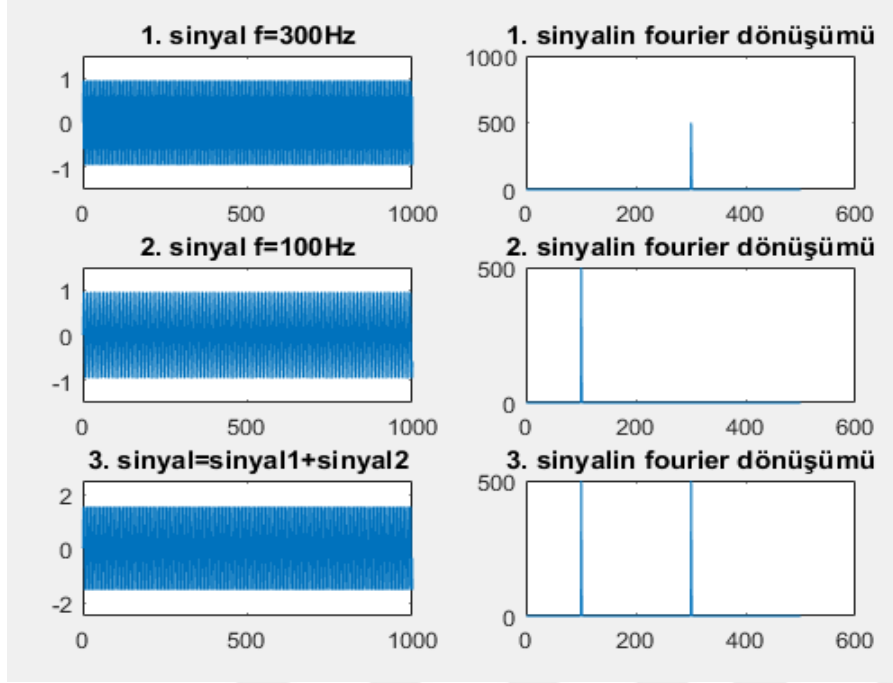
$F\{.\}$  operatörü ile  $x(n)$  ayrık zamanlı işareti,  $w$ 'nin karmaşık değerli ve sürekli bir fonksiyonu olan  $X(e^{jw})$ 'ye dönüşür.  $X(e^{jw})$ ,  $2\pi$  ile periyodik olduğundan temel frekans aralığında kalan örnekler yeterli olacaktır.  $0 < w < 2\pi$  aralığında  $N$  tane eşit aralıklı örnek alındığında ( $\Delta w = 2\pi/N$ ), yukarıdaki denklemi tekrardan değerlendirecek olursak;

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{j2\pi kn/N}, \quad n = 0, 1, 2, \dots, N-1
\tag{2.4}$$

Frekans düzleminde zaman düzlemine de dönüşüm ters Fourier dönüşümü ile mümkündür. Eşitlik (2.5)'te ters Fourier dönüşüm formülü verilmiştir.

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)e^{-j2\pi kn/N}, \quad n = 0, 1, 2, \dots, N-1
\tag{2.5}$$

Bu çalışmada Fourier dönüşüm işlemleri Matlab programında FFT komutu ile gerçekleştirildi. Aşağıdaki örnek olarak ses sinyalinden bir kesit ve onun Fourier dönüşümü verilmiştir.



Şekil 2.3. Üç tane sinyalin zaman ve frekans düzlemindeki gösterimi

Şekil 2.3'te sol tarafta zaman düzleminde işaretler varken sağ tarafta bu işaretlerin frekans düzleminde gösterimleri verilmiştir. Görüldüğü gibi zaman düzlemindeki işaretlerden bir anlam çıkarmak çok güçken sağ tarafta verilen frekans düzlemi gösteriminde sinyallerin frekansı rahatlıkla görülebilir.

### 2.2.1.3 Z dönüşümü

Z dönüşümü, Laplace dönüşümünün sürekli zaman sinyal analizindeki işlevini ayrık zamanlı sinyal analizinde yerine getirir. Fourier dönüşümünün genelleştirilmiş halidir. Aşağıda zaman düzleminde ayrık bir  $x[n]$  işaretinin z dönüşümü formülü verilmiştir.

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} \quad (2.6)$$

Eşitlik (2.6)'daki ayrık zamanlı Fourier dönüşümü formülünü inceleyecek olursak  $z=e^{j\omega}$  olduğunda Z dönüşümü ile Fourier dönüşümü aynı sonucu verir.

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \quad (2.7)$$

#### 2.2.1.4 Dalgacık (Wavelet) dönüşümü

Dalgacık dönüşümü bir işaretin tüm zaman aralığında bir dalgacık fonksiyonunun ölçeklendirilmesi ve kaydırılmasıyla elde edilen dalgacıkların toplanmasıyla elde edilir. Özellikle sismik verilerin incelenmesinde kullanılır. Dalgacık dönüşümü matematiksel olarak Eşitlik (2.8)'deki gibi ifade edilir:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi \left( \frac{t-b}{a} \right) dt \quad (2.8)$$

Bu ifadede a, ölçekleme parametresini; b, dönüşüm parametresini; x(t), zaman düzlemindeki işareti;  $\Psi$ , dalgacık fonksiyonunu; W(a,b) ise x(t) işaretinin sürekli dalgacık dönüşümünü belirtir.

#### 2.2.2 Model temelli metotlar

Sinyal işlemede model tabanlı metotlar parametrik model kullanır. Parametrik model, sinyalin geçmiş zamanki yapısından gelecek zamanı öngörmesidir. Bu metotlardan bazıları aşağıda açıklanmıştır (Vageshi 2006).

##### 2.2.2.1 Doğrusal tahmin (Linear prediction)

Bu yöntemde lineer bir sistemin çıkışını belirlemek için, sistemin giriş bilgilerini kullanarak tahmini katsayılar bulunur.

$$y[n] = \sum_{k=0}^p a_k x[n-k] - \sum_{k=1}^q b_k y[n-k] \quad (2.9)$$

Burada a ve b tahmin katsayılarını, x[n] girişteki ayrık zamanlı sinyali, y[n] ise çıkıştaki ayrık zamanlı sinyali belirtir. Yukarıdaki eşitlik tipik bir IIR filtrenin formülüdür. Lineer tahmin metodunda en yaygın olarak kullanılan metot tüm kutup olarak isimlendirilen a katsayılarının sıfır olduğu yani çıkışın önceki değerlerinden, çıkışın o anki değerinin bulunduğu yöntemdir.

$$y[n] = - \sum_{k=1}^q b_k y[n-k] \quad (2.10)$$

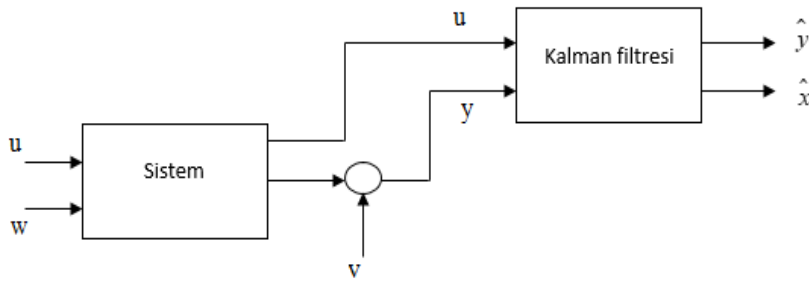
Tahmin edilen çıkış ile gerçek çıkış değerinin arasındaki fark ile hata sinyali  $e[n]$  bulunur.

$$e[n] = y[n] - \hat{y}[n] \quad (2.11)$$

Hata sinyalinin enerjisinin minimize edilmesiyle ise en uygun tahmin katsayıları  $b_k$  bulunur.

### 2.2.2.2 Kalman filtresi

Bu filtre yönteminde de temel esas en uygun filtre katsayısını tahmin edebilmektir. Kalman filtresinde sistemin girişine ve ölçülen gürültülü çıkışa bakarak sistemin gerçek çıkışını bulmaya çalışır. Bu yapının blok şeması Şekil 2.4'teki gibi gösterilir.



Şekil 2.4. Kalman filtre yapısı

Burada  $w$  işlem gürültüsü,  $v$  ölçüm gürültüsü,  $u$  sistemin gerçek girişi, ölçülen gürültülü çıkış  $y$ , tahmin edilen çıkış  $\hat{y}$ , sistemin durumu için yapılan tahmin  $\hat{x}$ 'dir (Kasnakoğlu 2014).

### 2.2.2.3 Adaptif filtre

Adaptif filtreler zamanla değişen gürültülü ortamlarda çıkış sinyalini kullanarak girişteki temiz konuşma sinyali elde edilmeye çalışılır. Adaptif filtre konusu filtreler başlığı altında detaylı olarak ele alınmıştır.

### 2.2.3 Bayesian tahmini temelli metotlar

Tahmin teorisinde gözlemlenen bozulmuş bir sinyalden, bilinmeyen parametreler elde edilmeye çalışılır. Tahminlerin doğruluğu kullanılan verilerin doğruluğuna bağlıdır. Bayesian temelli tahmin edicilere örnek olarak; maximum a posteriori (MAP), maximum likelihood

(ML), minimum mean square error (MMSE) verilebilir. Hidden Markov modeli de Bayesian temelli bir metottur (Vageshi 2006).

#### **2.2.4 Nöron elementlerinin katmanlı ağları temelli metotlar**

İnsan beyninin çalışma yapısını taklit eden sinir ağları metotları sinyal işleme süreçlerinde Bayesian tahmin metotlarını tamamlamak için kullanılır. Girişte birçok veri ile eğitimi yapılan bu sistemde, Bayesian temelli bir tahmin edici kullanılıp, sisteme girilen bir verinin eğitildiği verilere benzetilerek tespit edilmesi sağlanır (Vageshi 2006).

#### **2.3 Ses Etkinliği Algılama (VAD, Voice Activity Detection)**

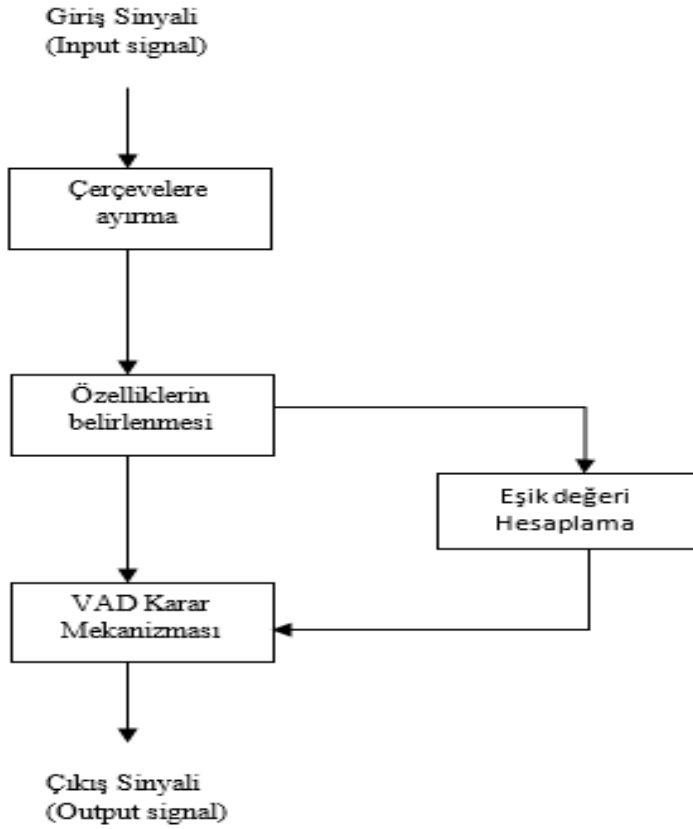
Temel prensip olarak VAD, ses sinyalinin özelliklerinin çıkarılması ve bu özelliklerin ses sinyalinin yalnızca gürültülü bölümlerinden belirlenen eşik değeri ile karşılaştırılmasıyla sesli bölümü ve gürültülü bölümü tespit etmeye yarayan bir uygulamadır. Bu uygulama; konuşma iyileştirme (speech enhancement), konuşma tanıma (speech recognition) ve konuşma kodlama (speech coding) gibi sinyal işleme süreçlerinin ilk aşamasında kullanılır ve sürecin performansını doğrudan etkiler.

VAD algoritmalarında algoritmanın verimliliği, sağlamlığı, basitliği gibi kavramlar algoritmanın kullanılabilirliği ile doğrudan bağlantılıdır. Sesin varlığını ya da yokluğunu tespit etmek için kullanılan özelliklerden bazıları aşağıda sıralanmıştır:

- STE (Short Time Energy - kısa süreli enerji),
- Periodicity (Periyodiklik),
- SF (Spectral Flatness - frekans spektrumundaki düzlük),
- ZCR (Zero Crossing Rate - sıfır geçiş oranı),
- Most dominant frequency component (en baskın frekans bileşeni),
- High-low frequency rate (yüksek- alçak frekans oranı).

Bu özelliklerin dışında da kullanılan birçok özellik olmakla birlikte yukarıdaki özellikler sıklıkla karşılaşılan özelliklerdir. Şekil 2.5'te temel bir VAD'ın blok şeması gösterilmiştir.





Şekil 2.5. VAD blok şeması

Konuşma sinyallerinde yaptığımız işlemleri, seslerin birbirlerine benzerlik göstermesinden ötürü çok küçük zaman aralıklarında yaparız. VAD uygulamalarında kullanılan özelliklerden bazıları aşağıda açıklanmıştır:

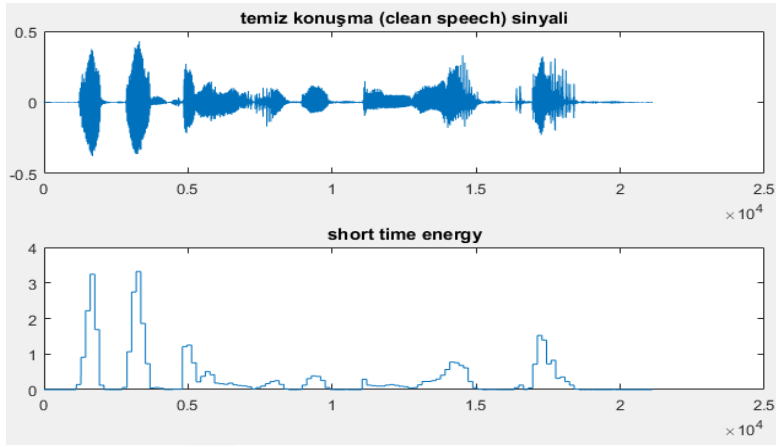
### 2.3.1 Kısa süreli enerji (STE, Short time energy)

Konuşma sinyallerinde zamana bağlı olarak konuşmanın genliği değişiklik göstermektedir. Konuşmanın olduğu bölgelerde STE, konuşmanın olmadığı bölgelere göre daha yüksektir. Bu da bize konuşma bölgeleri ile ilgili önemli ipuçları verir. STE'nin formülü aşağıdaki gibidir:

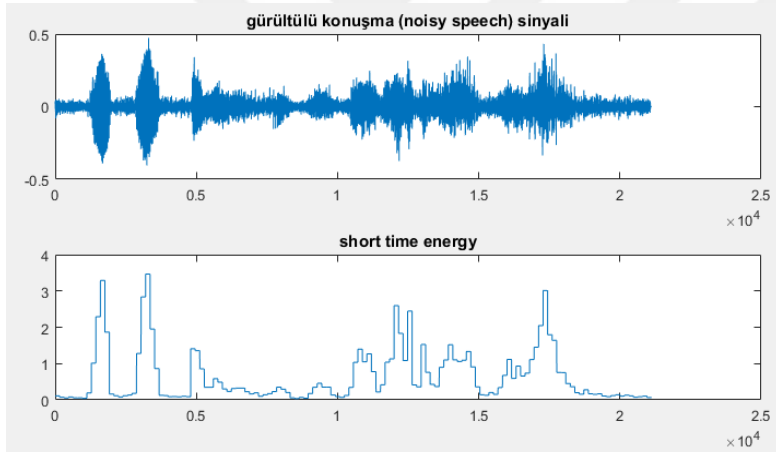
$$E(n) = \sum_{i=1}^N X_n^2(i) \quad (2.12)$$

Burada N, Ses çerçevesinin uzunluğunu, X(i); Orijinal konuşma sinyalini ve E(n) ise ses çerçevesinin enerjisini göstermektedir. Görüldüğü üzere her ses örneğinin karelerinin toplamı alınarak, bir sesli çerçevenin enerjisi hesaplanabilir. Şekil 2.6 ve Şekil 2.7'de sırasıyla

temiz bir konuşma sinyali ve bu konuşma sinyalinin STE'si ve gürültülü bir konuşma sinyali ile onun STE'si gösterilmiştir.



Şekil 2.6. Temiz bir konuşma sinyali ve onun STE'si



Şekil 2.7. Gürültülü bir konuşma sinyali ve onun STE'si

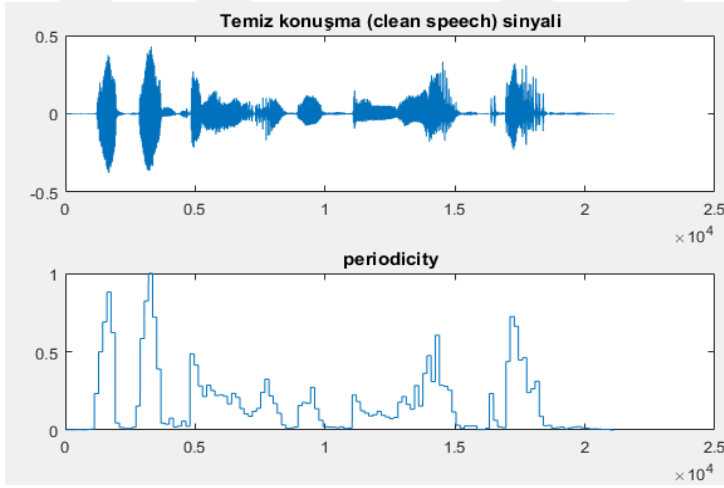
### 2.3.2 Periyodiklik (Periodicity)

Periyodiklik bizim için konuşma, müzik gibi ses işaretlerinde sesli bölgeyi tespit etmede önemli bir özelliktir. Sinyalin periyodik olma özelliği, pitch estimation (en kısa tekrarlanabilir aralık) ile belirlenebilir. Pitch Estimation değerinin hesaplanabilmesi için normalize edilmiş özilinti fonksiyonu  $R(\tau)$  aşağıdaki formülle hesaplanabilir.

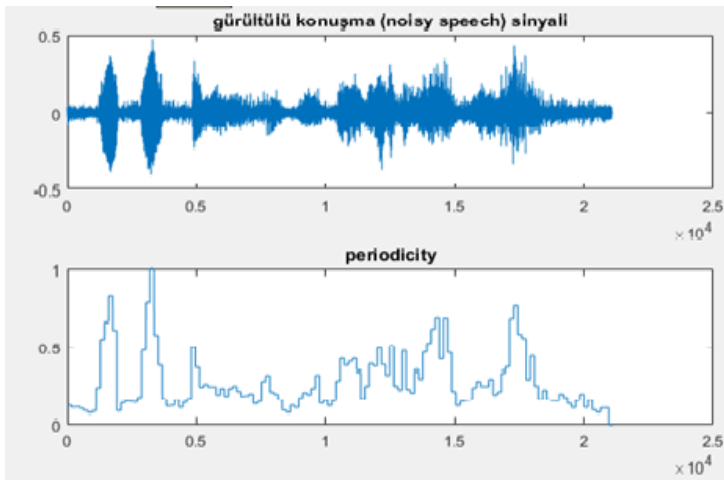
$$R(\tau) = \frac{\sum_{n=0}^{N-m-1} x(n)x(n+\tau)}{\sqrt{\sum_{n=0}^{N-m-1} x^2(n+\tau)}}, \quad T_{\min} < \tau < T_{\max}, \quad C=(R(\tau))_{\text{Max}} \quad (2.13)$$

Burada  $\tau$ , gecikme değerini göstermektedir.

Yukarıdaki formülde  $x(n)$ ;  $n=0,1\dots N$  aralığında giriş sinyal çerçevesinin uzunluğunu temsil etmektedir.  $\tau$  değerinin  $T_{\min}$ 'den  $T_{\max}$ 'a kadar değerlerini kullanarak özilinti fonksiyonu hesaplanır. Şekil 2.8 ve Şekil 2.9'da sırasıyla temiz bir konuşma sinyali ve bu konuşma sinyalinin periyodikliği ve gürültülü bir konuşma sinyali ve onun periyodikliği gösterilmiştir.



Şekil 2.8. Temiz bir konuşma sinyali ve onun periyodiklik grafiği



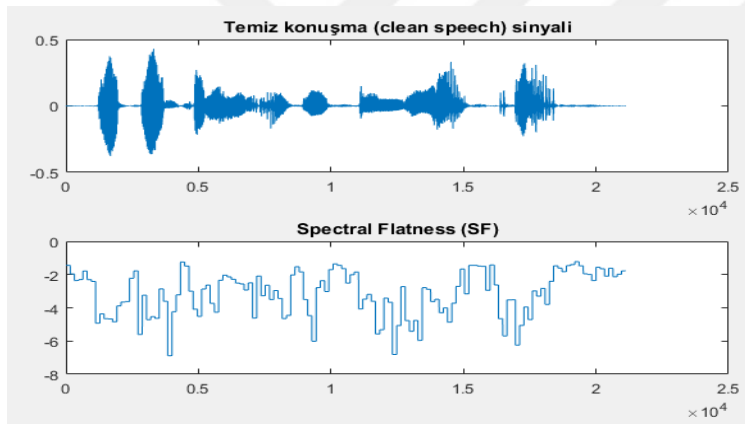
Şekil 2.9. Gürültülü bir konuşma sinyali ve onun periyodiklik grafiği

### 2.3.3 Spektral düzlük (SF, Spectral flatness)

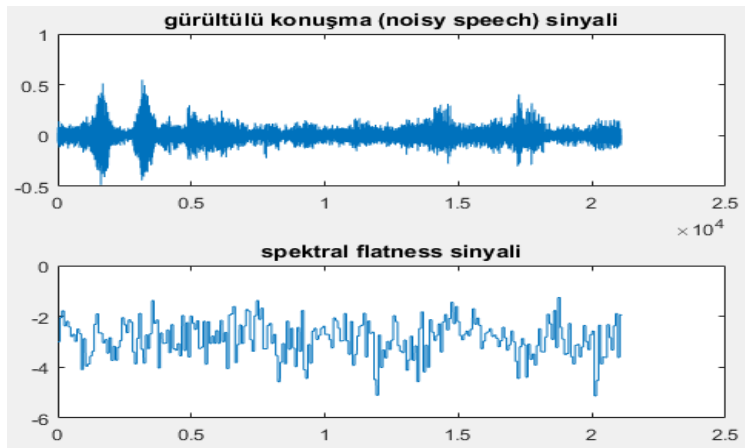
SF frekans düzleminde kullanılan bir özelliktir. Sadece gürültünün olduğu bölgelerde SF değeri sifira yaklaşıma eğilimindedir. Konuşmanın olduğu bölgelerde  $-\infty$  (eksi sonsuz)'a gitme eğilimindedir. Aşağıdaki formülü kullanarak dB cinsinden SF'yi hesaplayabiliriz.

$$SF_{dB}=10*\log_{10}\left(G_m/A_m\right) \quad (2.14)$$

Burada  $G_m$  Geometrik ortalama ve  $A_m$  Aritmetik ortalamayı göstermektedir. Bunun için ses sinyalinin çerçeveslere ayırıp her çerçeve için ayrı ayrı FFT alındıktan sonra her çerçeve için aritmetik ortalama ve geometrik ortalama değerleri hesaplanır. Eşitlik (2.14) ile sesin her çerçevesi için SF hesaplanmış olur. Şekil 2.10 ve Şekil 2.11'de sırasıyla temiz bir konuşma sinyali ve bu konuşma sinyalinin SF grafiği ile gürültülü bir konuşma sinyali ve onun SF grafiği gösterilmiştir.



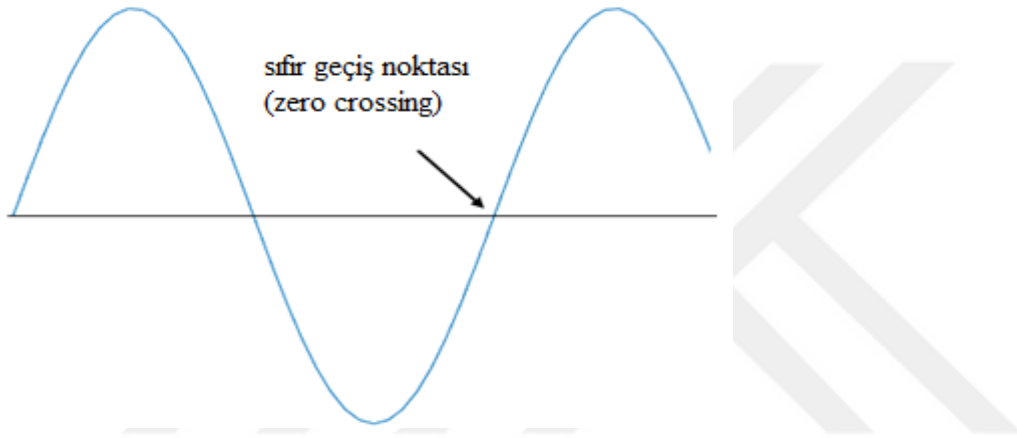
Şekil 2.10. Temiz bir konuşma sinyali ve onun SF grafiği



Şekil 2.11. Gürültülü bir konuşma sinyali ve onun SF grafiği

### 2.3.4 Sıfır geiş oranı (ZCR, Zero crossing rate)

Konuşma sinyallerinde sesli ve sessiz bölümün ayırt edilmesinde sıkça kullanılan bir uygulamadır. Adından da anlaşılacağı üzere zaman düzleminde bir sinyalin genliğinin belirli bir zaman aralığında sıfırdan geiş sayısını hesaplamak için kullanılır. Ses sinyalinin genliği konuşma olduğu zamanlarda sıkça sıfırın altına ve üstüne çıkar. Aynı şekilde konuşmanın olmadığı zaman aralıklarında ise daha az sıfır seviyesinden geer. Bu sebeple konuşma sinyallerinde sesli ve sessiz bölgelerin ayırt edilmesinde ZCR önemli bir araçtır. Şekil 2.12’de sıfır geiş noktası sinyal üzerinde gösterilmiştir.



Şekil 2.12. Sıfır geiş noktası gösterimi

$$Z_n = \sum_{m=-\infty}^{\infty} |\text{sgn}[x(m)] - \text{sgn}[x(m-1)]| w(n-m)$$
$$\text{sgn}[x(n)] = \begin{cases} 1, & x(n) \geq 0 \\ -1, & x(n) < 0 \end{cases} \quad (2.15)$$
$$w(n) = \begin{cases} \frac{1}{2N}, & 0 \leq n \leq N-1 \\ 0, & \text{diğer} \end{cases}$$

### 2.3.5 En baskın frekans bileşeni (Most dominant frequency component)

En baskın frekans bileşeni, herhangi bir sinyalin frekans düzleminde en yüksek genliğe sahip bileşeninin olduğu frekans değeridir. En baskın frekans bileşenini bulmak için birçok yöntem olmasına karşın, FFT bu uygulama için kullanılan en iyi yöntemdir. En baskın frekans bileşenini bulmak için;

1. Ses sinyali küçük zaman aralıklarına (ses çerçeveleri) ayrılır.

2. Her zaman aralığında N nokta FFT uygulanıp sinyalin spektral güç yoğunluğu bulunur.
3. Her zaman aralığı için gücün maksimum olduğu değer tespit edilir ve bu değere karşılık gelen frekans bileşeni, en baskın frekans bileşenidir.

Gürültü sinyalinde her ses çerçevesinde en baskın frekans bileşeni birbirine yakın sonuçlar verir. Bu durum göz önünde bulundurularak, her sesli çerçevede baskın olan frekans tespit edildikten sonra, ses çerçevesinin gürültü ya da gürültünün de bulunduğu konuşma sinyali olup olmadığı tespit edilmeye çalışılır.

### **2.3.6 Yüksek- alçak frekans oranı (High-low frequency rate)**

Konuşma sinyalinin enerjisinin en fazla olduğu bölge frekans spektrumunun yaklaşık olarak 100 Hz ile 2 KHz arasındaki bölgesidir. Gürültü ise bu anlamda daha farklı olabilir. Örneğin bazen sadece frekans spektrumunun alt frekanslarında yer alabilir bazen de frekans spektrumunun tamamına yayılabilir. Gürültünün özellikleri sabit olarak varsayıldığında frekans karakteristiği zaman ile çok fazla değişmeyecektir ve bu sinyal işleme için önemli bir ipucudur. Bu durum dikkate alınarak frekans spektrumunda belirli bir noktanın referans alınıp, bu frekans değerinin üst ve altındaki frekansların enerjilerinin oranının hesaplanmasıyla sinyalin gürültü ya da gürültülü konuşma olduğu konusunda fikir edinilebilir. Çünkü yukarıda da bahsedildiği gibi konuşma sinyalinin alt frekanslarda olduğu düşünülürse; konuşmanın olduğu bölgelerde alt frekansların enerjisi sadece gürültülü bölgelere daha fazladır. Bundan dolayı yüksek-alçak frekans oranı VAD uygulamalarında kullanılır.

## **2.4 Filtreler**

Sinyal filtreleme ile içinde hem istenen hem de istenmeyen bileşenleri bulunan bir sinyalden istenen sinyalin çıkarılması (sinyal ayırma) ve bozulmuş bir sinyali iyileştirme (sinyal restorasyon) gibi işlemler yapılır. Birçok filtre çeşidi olmakla birlikte bu filtreleri sınıflandırmanın da birçok yolu vardır. Darbe cevabı türüne göre filtre yapıları IIR (Infinite Impulse Response-sonsuz dürtü yanıtı) ve FIR (Finite Impulse Response- sonlu dürtü yanıtı) olarak ikiye ayrılır. Bu filtrelerde filtre çıkışı, giriş sinyaline uygulanan lineer bir fonksiyon olarak tanımlanabilir. Adaptif filtre ise, sinyal işlemede en iyi sinyali elde edebilmek için sinyal zaman içinde değişse bile filtre katsayılarını sinyalin istenmeyen bileşenlerine göre değiştirebilen filtredir. Aşağıdaki bu filtreler ile ilgili bilgilere yer verilmiştir.

### 2.4.1 Sonlu dürtü yanıtı (FIR, Finite Impulse Response) filtreler

FIR filtrelerde çıkış; filtre girişinin o an ki ve önceki değerlerinin bileşkesi olarak hesaplanır. FIR filtre için transfer fonksiyonu ve fark denklemi aşağıdaki gibidir:

$$H(z) = \sum_{n=0}^{M-1} b_n z^{-n} \quad (2.16)$$

$$y[n] = \sum_{k=-M_1}^{M_2} b_k x[n-k] \quad (2.17)$$

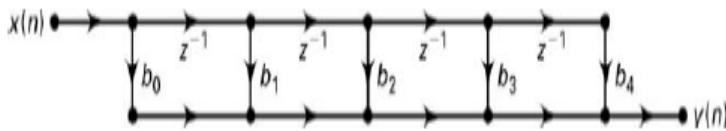
Örnek olarak üç terimli ileri beslemeli bir filtrenin fark denklemi aşağıdaki gibidir:

$$y[n] = \frac{1}{2} (x[n+1] + x[n] + x[n-1]) \quad (2.18)$$

İleri beslemeli olması filtre çıkışının o an ki ya da geçmiş değerlerinden birisinin girişe eklenmemesidir yani filtre çıkışı sadece filtre giriş sinyali bileşenlerinden oluşmaktadır. FIR filtre tasarımı için aşağıdaki yapılar kullanılır (Ingle ve Proakis 2011).

#### 2.4.1.1 Doğrudan biçim yapısı (Direct form)

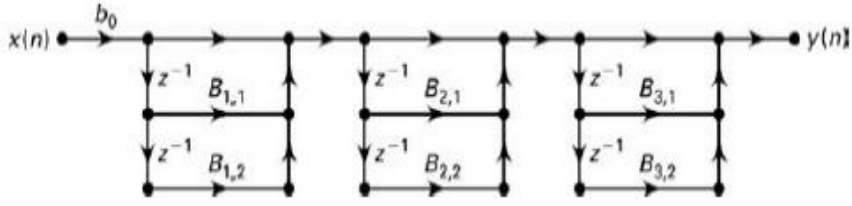
Bu yapıda FIR filtre için verilen fark denklemi olduğu gibi uygulanır. Akış diyagramı Şekil 2.13'teki gibidir.



Şekil 2.13. Doğrudan biçim yapısı (Ingle ve Proakis 2011).

#### 2.4.1.2 Art arda biçim yapısı (Cascade form)

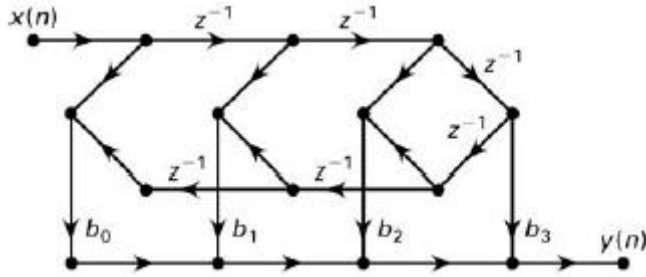
Bu yapıda  $H(z)$  sistem fonksiyonu ikinci derece faktörlere faktörlendirilmiştir. Daha sonra da art arda bağlanarak uygulanmıştır (Ingle ve Proakis 2011). Bu yapının akış diyagramı Şekil 2.14'te gösterilmiştir.



Şekil 2.14. Art arda biçim yapısı (Ingle ve Proakis 2011).

### 2.4.1.3 Doğrusal faz biçim yapısı (Linear phase form)

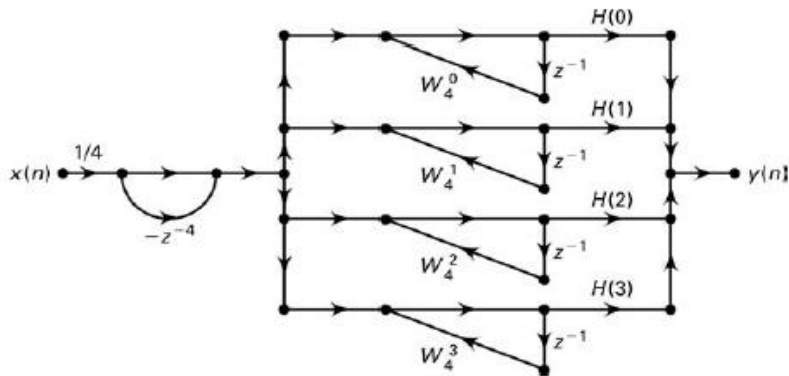
Bir FIR filtresinin doğrusal faz tepkisi olduğunda, dürtü yanıtı belli simetri koşulları gösterir. Bu yapıda çarpımları yaklaşık yarısı kadar azaltmak için bu simetri ilişkiler kullanılır (Ingle ve Proakis 2011). Doğrusal faz biçim yapısının akış diyagramı Şekil 2.15'te verilmiştir.



Şekil 2.15. Doğrusal faz biçim yapısı (Ingle ve Proakis 2011).

### 2.4.1.4 Frekans örnekleme biçim yapısı (Frequency sampling form)

Bu yapı  $h(n)$  dürtü yanıtının ayrık zamanlı Fourier dönüşümünü temel olarak alır ve bu da paralel yapıya yol açar (Ingle ve Proakis 2011). Akış diyagramı Şekil 2.16'daki gibidir.



Şekil 2.16. Frekans örnekleme biçim yapısı (Ingle ve Proakis 2011).



## 2.4.2 Sonsuz dürtü yanıtı (IIR, Infinite Impulse Response) filtreler

IIR filtrelerde FIR filtreden farklı olarak filtre çıkışı, sadece filtre giriş sinyaline değil, filtre çıkışının da önceki değerlerine bağlıdır. Bu filtre yapısına geri beslemeli filtre denir. IIR filtrenin transfer fonksiyonu ve fark denklemi Eşitlik (2.19) ve Eşitlik (2.20)'de verilmiştir:

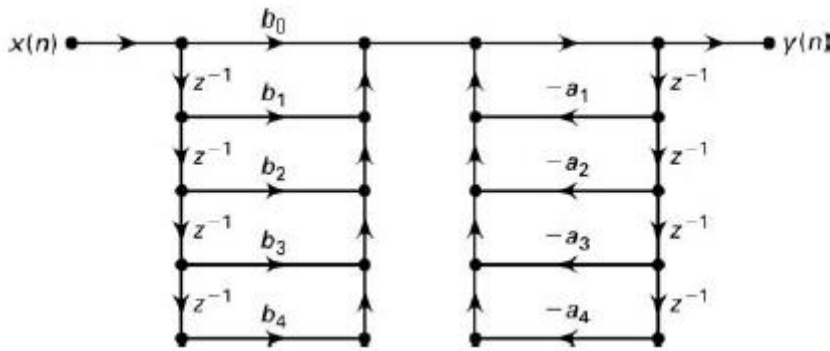
$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{n=0}^M b_n z^{-n}}{\sum_{n=0}^N a_n z^{-n}} \quad (2.19)$$

$$y[n] = \sum_{m=0}^M b_m x[n-m] - \sum_{m=1}^N a_m y[n-m] \quad (2.20)$$

Burada  $a_0$  katsayısı 1 olarak varsayılır. IIR filtre tasarımında farklı yapılar kullanılabilir:

### 2.4.2.1 Doğrudan biçim yapısı (Direct form)

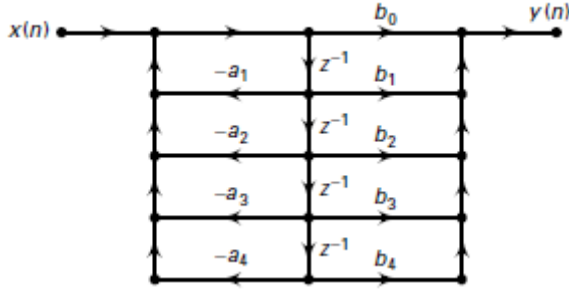
Bu biçimde yukarıdaki denklem verildiği gibi uygulanır. Bu filtrenin pay ve payda olmak üzere iki bölümü vardır ve bu uygulamanın doğrudan biçim I ve doğrudan biçim II olarak iki tipi vardır.  $M=N=4$  için doğrudan biçim I yapısının akış diyagramı Şekil 2.17'deki gibidir.



Şekil 2.17. Doğrudan biçim I yapısı (Ingle ve Proakis 2011).

Şekil 2.17'deki blok şemada görüldüğü gibi filtre ileri besleme ve geri besleme olarak iki bölümden oluşmaktadır. Çıkış hem kendisinin hem de girişin o anki ve önceki değerlerinden etkilenmiştir.

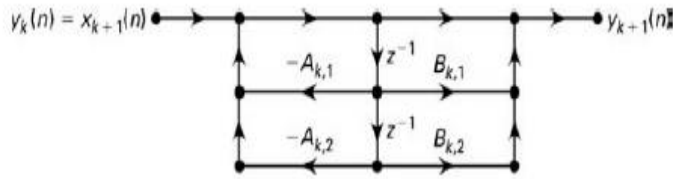
Doğrudan biçim I yapısında kaskat bağlı her iki kol için dört tane olmak üzere sekiz tane gecikme noktası vardır. Doğrudan biçim II yapısı ile gecikme noktası sayısı dörde düşürülür (Ingle ve Proakis 2011).



Şekil 2.18. Doğrudan biçim II yapısı (Ingle ve Proakis 2011).

#### 2.4.2.2 Art arda biçim yapısı (Cascade form)

Bu yapıda  $H(z)$  transfer fonksiyonu biquad isimli ikinci dereceden daha küçük bölümlere ayrılır. Daha sonra sistem fonksiyonu bu biquadların bir temsili olarak gösterilir. Her biquad doğrudan bir biçimde uygulanır. Art arda biçim yapısı aşağıda gösterilmiştir (Ingle ve Proakis 2011).



Şekil 2.19. Art arda biçim yapısı (Ingle ve Proakis 2011).

#### 2.4.2.3 Paralel biçim yapısı

Bu yapı da art arda biçim yapısına benzer ancak çarpanlara ayırma işleminden sonra transfer fonksiyonu  $H(z)$ 'yi daha küçük ikinci derece kısımların toplamı olarak temsil eden kesirli ifadelerin genişletilmiş hali kullanılır. Paralel biçim yapısının gösterimi Şekil 2.20'de verilmiştir (Ingle ve Proakis 2011).



Burada  $w_m$  filtre katsayısını belirtir. Görüldüğü üzere yukarıdaki formül M kademeli bir FIR filtrenin aynısıdır. Burada FIR filtreden farklı olarak katsayılar zamanla güncellenir.  $y[n]$  tahmin edilen sinyalin kalitesi hata sinyalinden ( $e[n]=x^s[n]-y[n]$ ) belirlenir. Filtre katsayıları Minimum Mean-Squared Error (MMSE) metoduyla minimize edilmeye çalışılır.

$$E\{e^2[n]\} = E\{(x^s[n]-y[n])^2\} \quad (2.22)$$

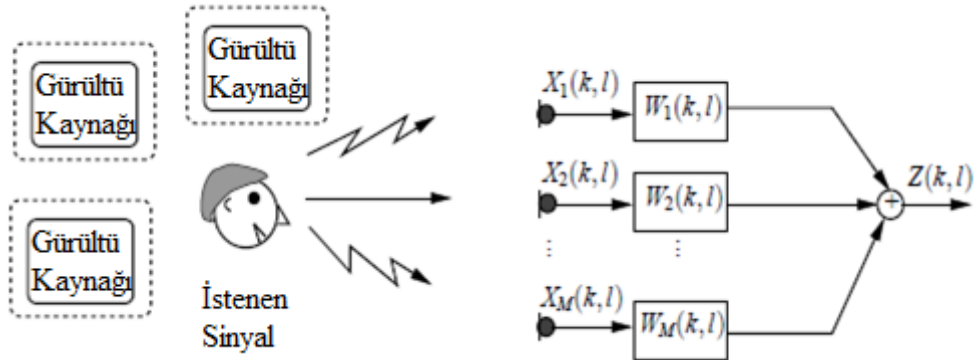
Burada en uygun filtre katsayıları, karesel hatanın filtre katsayılarına göre türevi alınarak hesaplanır.

$$\frac{\partial}{\partial w_m} E\{e^2[n]\} = 0, \quad m = 0, 1, 2, \dots, M-1 \quad (2.23)$$

Bu tez çalışmasında uygulanan çok kanallı adaptif filtre uygulamalarından bazıları aşağıdaki gibidir (Ngo 2011).

#### 2.4.3.1 Çok kanallı wiener filtre (MWF, Multichannel wiener filter)

Daha önce bahsedildiği gibi çok kanallı Wiener filtre uygulamalarında hataların karesinin ortalamasının minimize (MMSE) edilmesi ile filtre çıkışındaki hata minimum seviyede tutulmaya çalışılır. Şekil 2.22'de çok kanallı Wiener filtre blok şeması gösterilmiştir.



Şekil 2.22. Çok kanallı gürültü azaltma yapısının blok diyagramı (Ngo 2011).

$$\mathbf{W}_{\text{MMSE}}(k,l) = \arg \min_{\mathbf{W}(k,l)} \mathcal{E}\{|X_1(k,l) - \mathbf{W}^H(k,l)\mathbf{Y}(k,l)|^2\} \quad (2.24)$$

Eşitlik (2.24)'te Wiener filtre katsayılarının hesaplanması için temel formül verilmiştir.  $X_1(k,l)$  istenen temiz sinyali temsil eder ve mikrofonlardan biri (genelde birinci mikrofon) referans kabul edilerek o mikrofondan alınır.  $X_1(k,l)$  konuşma sinyalindeki gecikme sıfır olarak alındı. Eşitlik (2.24) aşağıdaki gibi yeniden yazılabilir.

$$\begin{aligned} J_{\text{MMSE}}(k,l) &= \varepsilon\{|X_1(k,l) - \mathbf{W}^H(k,l)\mathbf{Y}(k,l)|^2\} \\ &= \varepsilon\{X_1(k,l)X_1^H(k,l)\} + \varepsilon\{\mathbf{W}^H(k,l)\mathbf{Y}(k,l)\mathbf{Y}^H(k,l)\mathbf{W}(k,l)\} \\ &\quad - \varepsilon\{X_1(k,l)\mathbf{Y}^H(k,l)\mathbf{W}(k,l)\} - \varepsilon\{\mathbf{W}^H(k,l)\mathbf{Y}(k,l)X_1^H(k,l)\} \end{aligned} \quad (2.25)$$

Eşitlik (2.25)'teki ifadenin  $\mathbf{W}(k,l)$ 'ye göre türevi alınıp, sıfıra eşitlenerek filtre katsayıları minimize edilir.

$$\frac{\partial J_{\text{MMSE}}(\mathbf{W}(k,l))}{\partial(\mathbf{W}(k,l))} = -2\varepsilon\{\mathbf{Y}(k,l)X_1^H(k,l)\} + 2\varepsilon\{\mathbf{Y}(k,l)\mathbf{Y}^H(k,l)\mathbf{W}(k,l)\} \quad (2.26)$$

Konuşma ve gürültü sinyallerin ilişkisiz olduğu varsayılırsa;

$$\varepsilon\{\mathbf{V}(k,l)X_1(k,l)\} = 0 \quad (2.27)$$

Tahmin edilen konuşma özilinti vektörü şöyle yazılabilir.

$$\varepsilon\{X(k,l)X_1^H(k,l)\} = \varepsilon\{\mathbf{Y}(k,l)X_1^H(k,l)\} - \varepsilon\{\mathbf{V}(k,l)\mathbf{V}_1^H(k,l)\} \quad (2.28)$$

Son olarak Eşitlik (2.26)'nın çözülmesiyle MWF formülü aşağıdaki gibi verilir:

$$\mathbf{W}_{\text{MMSE}}(k,l) = [\mathbf{R}_x(k,l) + \mathbf{R}_v(k,l)]^{-1} \mathbf{R}_x(k,l) \mathbf{e}_1 \quad (2.29)$$

Burada  $\mathbf{e}_1$ ,  $M \times 1$  mikrofon sayısı kadar uzunluktaki birim matrisin ilk sütunudur,  $\mathbf{e}_1 = [1 \ 0 \ \dots \ 0]^T$ . Tahmin edilen konuşma sinyali ise referans olarak belirlenen mikrofon sinyalinin elde edilen Wiener katsayıları ile çarpılmasıyla aşağıdaki gibi elde edilir (Ngo 2011).

$$\mathbf{Z}^s(k,l) = \mathbf{W}_{\text{MMSE}}^H(k,l)\mathbf{Y}(k,l) \quad (2.30)$$

#### 2.4.3.2 Konuşma bozukluğu ağırlıklı (SDW, Speech distortion weighted) wiener filtre

Klasik çok kanallı Wiener filtreden türetilen bu filtre yapısında bir  $\mu$  parametresiyle konuşma bozukluğu ile gürültü azaltma arasında bir ilişki sağlanır.  $\mu$  değeri arttığında filtre

çıkışı daha gürültüsüz olur ancak konuşma bozukluğu daha fazla olur. Eğer konuşma ve gürültü sinyallerinin ilişkisiz olduğu varsayılırsa MMSE kriterine göre aşağıdaki gibi yazılır:

$$\varepsilon\{|X_1(k,l) - \mathbf{W}^H(k,l)\mathbf{Y}(k,l)|^2\} \quad (2.31)$$

Bu denklemi aşağıdaki gibi ayrıştırabiliriz:

$$\varepsilon\{|X_1(k,l) - \mathbf{W}^H(k,l)\mathbf{X}(k,l)|^2\} + \varepsilon\{|\mathbf{W}^H(k,l)\mathbf{V}(k,l)|^2\} \quad (2.32)$$

İlk terim konuşma bozukluğunun enerjisini, ikinci terim ise mikrofondan alınan gürültünün enerjisini temsil eder. Yukarıdaki denkleme bir  $\mu$  parametresi eklenerek konuşma bozukluğu ile gürültü azaltılması arasında bir ilişki sağlanır.

$$\mathbf{W}_{MWF\mu}(k,l) = \arg \min_{\mathbf{W}(k,l)} \varepsilon\{|X_1(k,l) - \mathbf{W}^H(k,l)\mathbf{X}(k,l)|^2\} + \mu\varepsilon\{|\mathbf{W}^H(k,l)\mathbf{V}(k,l)|^2\} \quad (2.33)$$

Bu denklemin çözülmesi ile SDW çok kanallı Wiener filtre katsayıları aşağıdaki hesaplanır:

$$\mathbf{W}_{MWF\mu}(k,l) = [\mathbf{R}_x(k,l) + \mu\mathbf{R}_v(k,l)]^{-1}\mathbf{R}_x(k,l)e_1 \quad (2.34)$$

Yukarıdaki formülden de anlaşıldığı gibi  $\mu=1$  olduğunda klasik Wiener filtresinin formülü elde edilir.  $\mu>1$  olduğunda filtre çıkışındaki gürültü daha da azalır ancak aynı zamanda daha fazla konuşma bozulmasına neden olur (Spriet ve ark 2005).

#### 2.4.3.4 Uzamsal tahmin (SP, Spatial prediction) filtresi

Cornelis ve ark. (2009) tarafından zaman düzleminde uzamsal tahmin temelli bu metod frekans düzlemine aktarılarak çalışılmıştır. Bu yaklaşımda referans mikrofonun frekans bileşeni referans kabul edilip diğer mikrofonlarla ilişkisini gösteren bir vektör oluşturulur.

$$\mathbf{X} = \begin{bmatrix} \mathbf{H}_{1,ref} \\ \mathbf{H}_{2,ref} \\ \cdot \\ \cdot \\ \mathbf{H}_{N,ref} \end{bmatrix} X_{ref} = \mathbf{H} X_{ref} \quad (2.35)$$

Bu şekilde konuşma bileşenlerinin uzamsal bir tahmin vektörü oluşturularak SP filtre tasarlanır. H vektörünü Wiener filtredeki minimize etme problemi aşağıdaki gibi yazılabilir:

$$\min_{\mathbf{H}} \varepsilon\{(\mathbf{X} - \mathbf{H} X_{ref})^H (\mathbf{X} - \mathbf{H} X_{ref})\} \quad (2.36)$$

Buradan da H vektörü aşağıdaki gibi bulunur:

$$\mathbf{H}(k,l) = \frac{1}{\mathbf{u}^H \mathbf{R}_x(k,l) \mathbf{u}} \mathbf{R}_x(k,l) \mathbf{u} \quad (2.37)$$

Bu formüle göre H;  $\mathbf{R}_x$  özilinti matrisinin ilk sütununun,  $\mathbf{R}_x$  özilinti matrisinin ilk elemanına bölünmesiyle hesap edilir. Burada u, bir girişi 1 diğer girişleri 0 olan uzunluğu ise mikrofon sayısı kadar olan bir vektördür.  $\mathbf{R}_x$  ise konuşma özilinti matrisidir. Eşitlik (2.35) kullanılarak konuşma bozukluğu hatası  $E_x$  aşağıdaki gibi yazılabilir:

$$E_x = (\mathbf{W} - \mathbf{u})^H \mathbf{X} = (\mathbf{W}^H \mathbf{H} - 1) \mathbf{X}_{\text{ref}} \quad (2.38)$$

Burada konuşma bozukluğu  $E_x$ 'in 0 olabilmesi için  $\mathbf{W}^H \mathbf{H} = 1$  olmalıdır. Bu optimizasyon problemine göre Wiener katsayıları şu şekilde hesap edilir:

$$\mathbf{W}_{SP}(k,l) = \frac{1}{\mathbf{H}^H(k,l) \mathbf{R}_v^{-1}(k,l) \mathbf{H}(k,l)} \mathbf{R}_v^{-1}(k,l) \mathbf{H}(k,l) \quad (2.39)$$

#### 2.4.3.5 Minimum varyans gürültüsüz yanıt (MVDR, Minimum variance distortionless response) filtresi

Capon (1969) tarafından önerilen Beamforming (hüzme şekillendirici) tekniğine dayalı bu yaklaşımda filtre katsayılarının hesaplanması için hedef konuşma sinyalinin bir kılavuz vektörü oluşturulur ve her zaman aralığında değişen kılavuz vektörüyle filtre katsayıları hesap edilir. MWF'de ise oluşturulan özilinti matrislerine dayalı bir yaklaşım söz konusudur. MVDR filtrede çıkıştaki hata minimize edilmeye çalışılırken, hedef konuşma sinyalinin de bozulmaması istenir. Bu filtrenin optimizasyon formülü aşağıdaki gibidir:

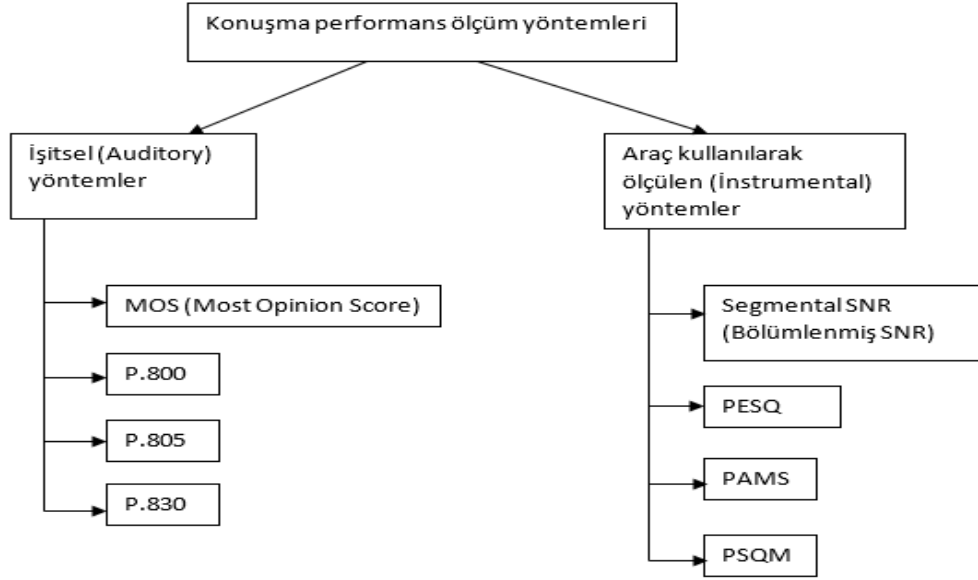
$$\min_{\mathbf{W}(k,l)} \mathbf{W}^H(k,l) \mathbf{R}_x(k,l) \mathbf{W}(k,l), \quad \mathbf{e}^H(k,l) \mathbf{W}(k,l) = 1 \quad (2.40)$$

Burada  $\mathbf{e}^H(k,l)$  kılavuz vektörüdür. Filtre katsayıları vektörü ile kılavuz vektörü elemanlarının çarpımlarının 1 olması prensibinden yola çıkarak bilinen kılavuz vektöründen bilinmeyen filtre katsayıları vektörü hesaplanmaya çalışılır. Eşitlik (2.40)'a en yakın çözüm Capon (1969) tarafından verilmiştir.

$$\mathbf{W}_{MVDR}(k,l) = \frac{1}{\mathbf{e}^H(k,l) \mathbf{R}_x^{-1}(k,l) \mathbf{e}(k,l)} \mathbf{R}_x^{-1}(k,l) \mathbf{e}(k,l) \quad (2.41)$$

## 2.5 Performans Ölçüm Parametreleri

Sinyal işleme alanında performansı değerlendirmek için kullanılan birtakım yöntemler vardır. Bu yöntemlerden bazıları objektif bazıları ise sübjektiftir. Ölçülmek istenen ses sinyali bu ölçüm yöntemleri ile kalite ve anlaşılabilirlik yönlerinden incelenip değerlendirilir.



Şekil 2.23. Performans ölçüm yöntemlerinin gösterimi

Genel olarak işitsel ve araç kullanılarak yapılan bu ölçümlerden bazıları Şekil 2.23'te gösterilmiştir. İşitsel yöntemler gerçek ortamlarda, kişiler tarafından yapılan değerlendirmelerdir. Bu değerlendirmeler aslında en doğru sonucu vermesi ile birlikte bazı zorlukları vardır. Kişileri bir araya getirme, harcanan zaman gibi dezavantajları gidermek için de araç kullanılarak yapılan yöntemler kullanılarak işitsel yöntemlere paralel sonuçlar alınmaya çalışılmıştır.

### 2.5.1 Ortalama yargı değeri (MOS, Most opinion score)

MOS, sinyal işleme uygulamalarında değerlendirme için kullanılan sübjektif bir yöntemdir. Belirli sayıda kişiye orijinal ve iyileştirilmiş seslerin dinletilip, bu kişilerden birden beşe kadar bir not vermesi istenmesine dayalı bir yöntemdir. Bu notların tanımlandığı değerlendirme tablosu Çizelge 2.1'de gösterilmiştir (Sümer 2004).



**Çizelge 2.1.** Ses kalitesi ile MOS arasındaki ilişki (Sümer 2004).

MOS	Ses kalitesi	Değerlendirme
5	Mükemmel	Mükemmel bir ses kaydı
4	İyi	Anlaşılır, doğal, telefon kaydı kalitesinde bir ses kaydı
3	Orta	Haberleşme kalitesinde ancak duymak için çaba gerekiyor
2	Zayıf	Düşük kalitede, konuşmanın anlaşılması zorlaşıyor
1	Kötü	Konuşma anlaşılmıyor

### 2.5.2 P.800

Uluslararası iletişim birliğinde (ITU, International Telecommunication Union) iletişim sektörünün standartlaştırılması konusunda çalışmalar yapılır. Konuşma kalitesinin ölçümü için yapılan bu test 1996 yılında oluşturulmuştur. Gerçek kişilerin beş dakikada civarında yaptığı telefon görüşmelerinde dinleme ve konuşma kalitesi açısından değerlendirme yapılır. Değerlendirme birden beşe kadar not verilmek suretiyle yapılır (ITU-T Rec. P.800).

### 2.5.3 P.805

Bu yöntemde P.800 yöntemindeki gibi konuşma kalitesinin ölçümü yedi farklı ölçekleme parametresi kullanılarak yapılır. Bunun dışında deneklerden ses sinyalindeki bozulmanın ne olduğunu açıklaması istenir.

### 2.5.4 P.830

ITU tarafından yayınlanan bu yöntem genel olarak P.800 ile aynı olmakla birlikte, P.800 yöntemi darbant (Narrowband) iletim sistemlerinde kullanılırken P.830 yöntemi geniş bant (Wideband) iletim sistemlerinde kullanılmaktadır. Burada darbant iletim sistemiyle kastedilen örnekleme frekansı 8 Khz olan sistemlerdir. Geniş bant sistemlerde ise örnekleme frekansı 16 Khz olarak kabul edilir (ITU-T Rec. P.830).

### 2.5.5 Bölümlenmiş sinyal gürültü oranı (Segmental SNR)

Elektronikte sıkça kullanılan bir kavram olan SNR, temel olarak işaretin gürültüye oranı demektir. Konuşma sinyallerinde elde edilen verimi ölçmek için kullanılan bu terim temiz konuşma sinyalinin, sentezlendiği gürültü sinyaline oranı olarak hesaplanır. Ancak temiz

konuşma sinyalinde, konuşmanın olmadığı bölümler temiz sinyalinin gücünün ortalamasını düşüreceğinden yanlış değerlendirme yapılmasına sebep olabilir. Bu sebeple konuşma sinyalinde sadece konuşmanın olduğu bölümler değerlendirmeye alınır (Zelinski ve Noll 1977). Bölümlenmiş SNR aşağıdaki formülle hesaplanır.

$$SNR_{seg} = \frac{1}{L} \sum_{l=0}^{L-1} 10 \log_{10} \left( \frac{\sum_{k=1}^M x(lM+k)^2}{\sum_{k=1}^M [y(lM+k) - x(lM+k)]^2} \right) \quad (2.42)$$

Burada M: bir sesli çerçevedeki örnek sayısını, L: sesli çerçeve sayısını, y: gürültülü konuşma sinyalini, x:temiz konuşma sinyalini ifade eder.

Hesaplanan SNR değerinin birimi yoktur yani sadece temiz konuşma sinyalinin gürültü sinyaline oranıdır. Ses sinyallerinde desibel (dB) kavramı kullanılır Bu oran desibel olarak aşağıdaki formülle verilebilir.

$$SNR_{dB} = 10 * \log_{10} (SNR) \quad (2.43)$$

### 2.5.6 Konuşma kalitesinin algısal değerlendirmesi (PESQ, Perceptual evaluation speech quality)

ITU tarafından 2001 yılında yayınlanan P.862 numaralı bu yöntem konuşmanın kalitesini değerlendirmede en sık kullanılan yöntemlerden biridir. -0.5-4.5 arasında verilen PESQ değerleri 1-4.5 arasında ölçeklenir. Bu ölçekleme aşağıdaki formülle yapılmıştır.

$$y = 0,999 + \frac{4,999 - 0,999}{1 + e^{-1,4945x + 4,6607}} \quad (2.44)$$

Eşitlik (2.44)'te x, (-0,5)-(4,5) arasında bulunan saf PESQ değeridir. PESQ konuşma iletim sistemine girilen temiz konuşma sinyalinin alıcı kısmında ne kadar bozulmaya uğradığını tespit etmek için kullanılır. Bu açıdan bakıldığında aslında kullanıldığı alan konuşma iyileştirme algoritmaları değildir. Konuşma iyileştirme algoritmalarında PESQ ile değerlendirme yapılırken sisteme temiz konuşma sinyali referans olarak girilir. Bozulmuş sinyal olarak ise konuşma iyileştirme algoritmasının girişine uygulanan ses sinyali uygulanır. Elde edilen PESQ değeri giriş PESQ değeri olarak isimlendirilir. Aynı şekilde temiz konuşma sinyali referans sinyal olarak girilip, konuşma iyileştirme algoritmasının çıkışındaki sinyal ise bozulmuş sinyal olarak uygulanır ve böylece çıkış PESQ değeri olarak isimlendirilir. Giriş ve

çıkış PESQ değerlerinin olumlu ya da olumsuz olarak değişimi ile ses sinyalinin algısal olarak değerlendirilmesi hakkında fikir edinilir (ITU-T Rec. P.862).

### **2.5.7 Algısal analiz ölçüm sistemleri (PAMS, Perceptual analysis measurement system)**

Hollier ve ark. (1994) tarafından çalışılan sisteme giren işitilebilir hataların bir tarifi yapıldı. Orijinal ve bozulmuş sinyalin arasındaki farklara hata yüzeyi ismi verildi. Kısaca hata entropisi, birinin birkaç hata tanımlayıcısını çıkarmasını ve hataların toplam miktarının ölçülmesini sağlar. Bu tanım ile konuşma iyileştirme algoritmalarının kalite tahminleri için kullanılan ilk modele öncülük edilmiştir.

Bu versiyonun bütün iletim hattının kalite tahminlerine genişletilmesi için elektro akustik ara yüzeyleri içeren farklı gelişmeler yapıldı (Rix ve Hollier 2000). Bu kalite modeline PAMS adı verildi.

### **2.5.8 Algısal konuşma kalite ölçümü (PSQM, Perceptual speech quality measurement)**

Konuşma kalite ölçümü için geliştirilen bu yöntem PSQM adı verilmiştir (Beerends ve Stemerdink 1994). Bu yöntemde psikoakustik model kullanılmıştır. Ayrıca bu modelde algısal dönüşüm konuşma sinyali için optimize edilmiştir. Algısal modelin konuşma kalitesini modelleyemediği ve bu sebeple bilişsel modele ihtiyaç duyduğu varsayılmıştır.

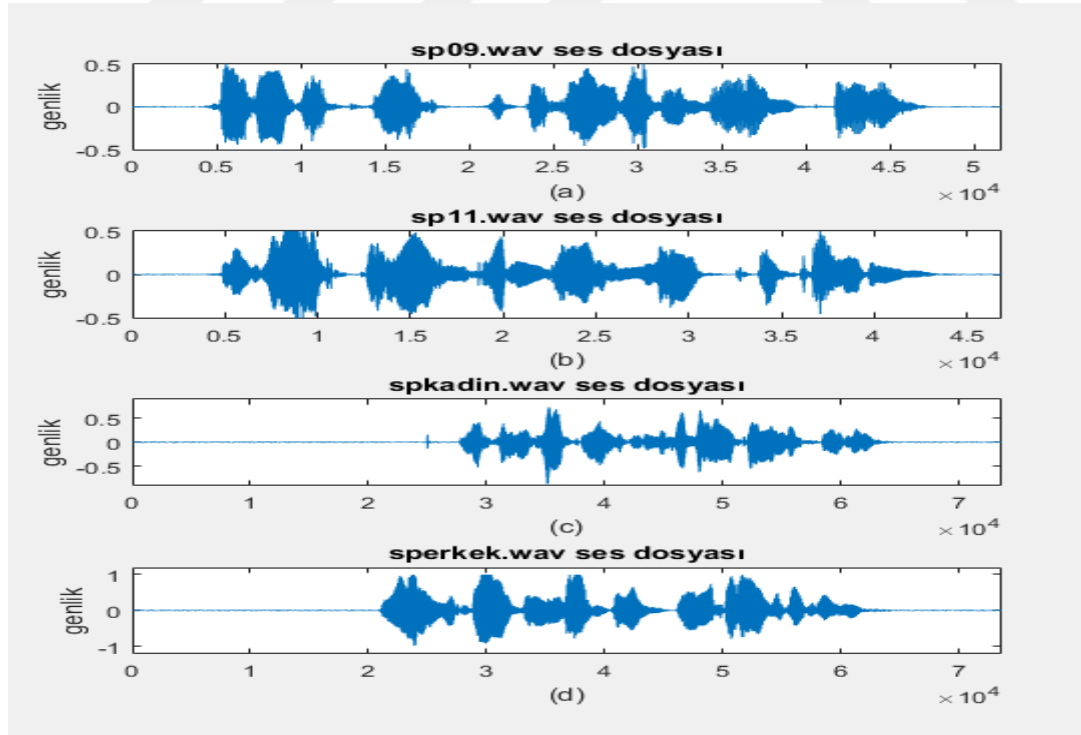
Bu modelin ITU tarafından değerlendirilmesinden sonra 1996 yılında standart bir konuşma kalite ölçüm aracı olarak ITU-T Rec. P.861 adıyla yayınlanmıştır.

### 3. MATERYAL ve YÖNTEM

#### 3.1 Materyal

Bu çalışmada; Intel Core i5-5200U CPU 2.2 GHz işlemcili, 8 GB RAM bellek ve 64 bit Windows 10 işletim sistemine sahip bir bilgisayar, MATLAB R2016a programı, ses kablosu, hoparlör kablosu, üç adet mikrofon, üç adet mikrofon sehpası, hoparlör sehpası, hoparlör, ses kartı ve Adobe Audition ses düzenleme programı kullanılmıştır. Çalışmada kullanılan ses dosyaları yabancı dilde konuşan bir kadın ve bir erkek konuşmacı sesleridir (Anonim 1969). Ayrıca stüdyo ortamında Türkçe konuşan bir erkek ve bir kadın sesleri kaydedilmiştir. Kullanılan gürültü ses dosyaları ise AURORA veri tabanından alınmıştır (Anonim 2018).

Şekil 3.1'deki sp09.wav ve sp11.wav ses dosyaları İngilizce konuşan bir kadın ve bir erkeğe ait olmakla birlikte spkadin.wav ve sperkek.wav ses dosyaları ise Türkçe konuşan bir kadın ve bir erkeğe aittir. Bu tez çalışmasında uygulanan çok kanallı gürültü azaltma algoritmasında üç mikrofon kullanıldığı için Şekil 3.1'deki dört adet ses sinyali üç mikrofonla bir metre uzaklığa yerleştirilen bir hoparlör aracılığıyla okutulup, her ses sinyalinin üç mikrofonla kaydedilen ses dosyaları ses kartı vasıtasıyla bilgisayara kaydedilmiştir.



Şekil 3.1. Algoritma testinde kullanılan temiz konuşma sinyalleri

Kullanılan cümlelerin içeriği aşağıdaki gibidir:

sp09.wav: “Hurdle the pit with the aid of a long pole.”

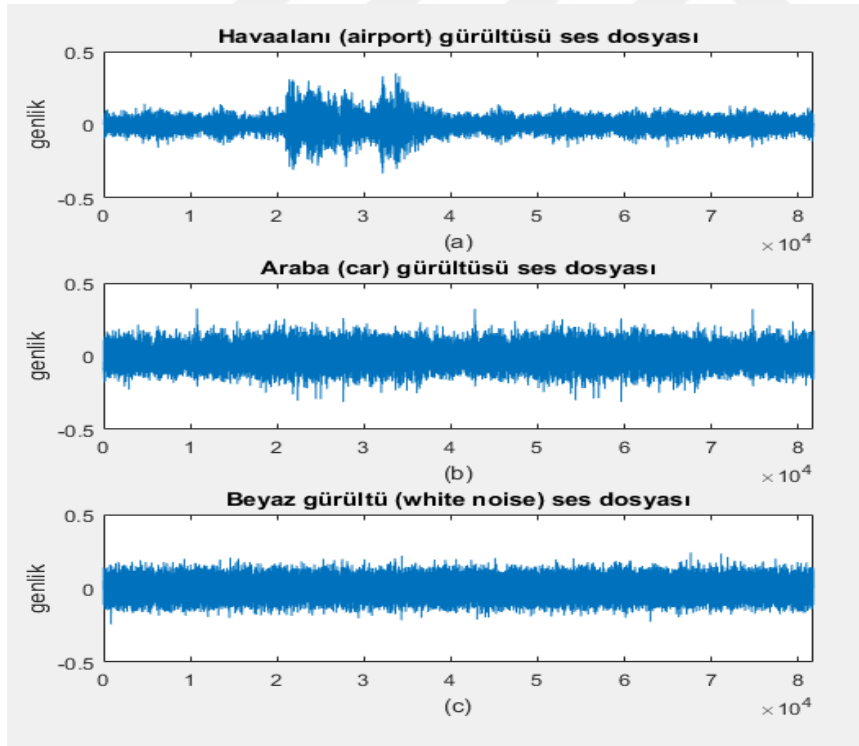
sp11.wav: “He wrote down a long list of items.”

spkadin.wav: “Yapılacaklar için, eve erken gitmeliyim.”

sperkek.wav: “Yarın saat 14.00’te Ankara’ya gideceğim.”

Kullanılan gürültü sinyalleri de üç mikrofonla 45, 90, 135 ve 180 derece olmak üzere dört farklı açıda kaydedilmiştir. 90 ve 180 derecelik açılarda mikrofonlar ve gürültünün okunduğu hoparlör arası mesafe bir metre iken 45 ve 180 derecelik açılarda 1,41 metredir.

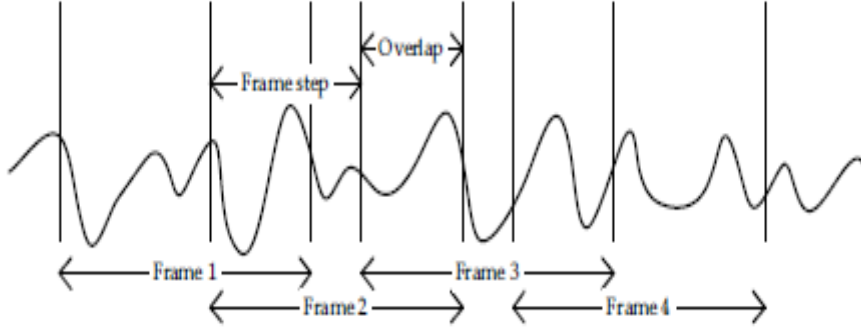
Burada kullanılan gürültülerde ilki olan havaalanı gürültüsü, bir havaalanının bekleme salonundaki ortam gürültüsüdür. Kayıt esnasında uçak bekleyen yolcular için yapılan anons da kaydedildiğinden kaydın içinde genlik olarak kaydın geneline göre yüksek olan bir bölüm göze çarpmaktadır. Yine araba gürültüsü de yolculuk esnasında bir aracın içinden alınmış gürültüdür. Bakıldığı zaman gürültünün genliği zamanla çok fazla değişmemiştir. Son gürültü ise Matlab programı vasıtasıyla yapay olarak oluşturulmuş beyaz gürültüdür.



Şekil 3.2. Algoritma testinde kullanılan gürültü sinyalleri



durum bir örnekle şöyle açıklanabilir. Her sesli çerçeve (frame) için 100 örnek olduğu ve %50 örtüşme (overlap) olduğu varsayılırsa; birinci sesli çerçeve 0-100 arasındaki örnekler, ikinci sesli çerçeve 50-150 arasındaki örnekler, üçüncü sesli çerçeve 100-200 arasındaki örnekler şeklinde ses sinyalinin sonuna kadar örtüşme uygulanır. Şekil 3.4’de örtüşme ile ilgili görsel yer verilmiştir.



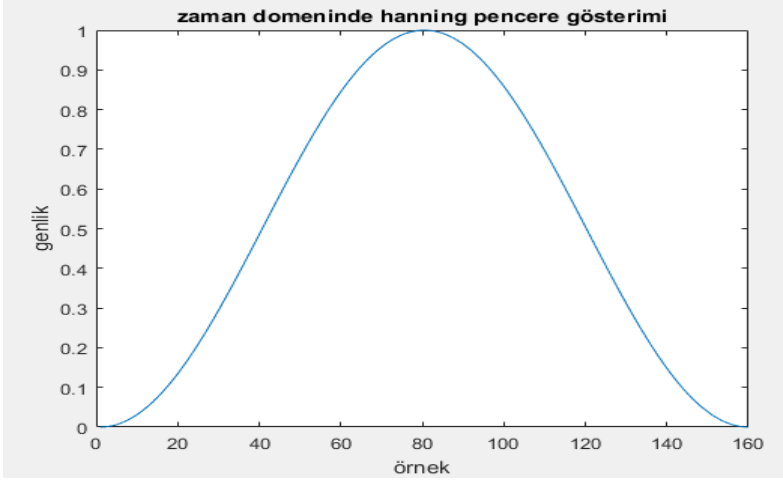
**Şekil 3.4.** Çerçeveler arasındaki örtüşmenin gösterimi

### 3.2.2 Pencereleme tekniği

Pencereleme yöntemi sinyal işleme süreçlerinde zaman düzleminde sinyalin belirli bir bölümünü işlemek için kullanılır. Sinyal işlemede kullanılan pencere fonksiyonlarından bazıları Kaiser, Hamming, Hanning, Blackman ve Dikdörtgen (Rectangular) pencere fonksiyonlarıdır. Bu tez çalışmasında Hanning pencere fonksiyonu kullanılmıştır. Hanning pencere fonksiyonunun katsayılarının hesaplanması için kullanılan formül Eşitlik (3.1)’de verilmiştir. Ayrıca Hanning penrece fonksiyonuna ait görsel Şekil 3.5’te verilmiştir.

$$w(n) = 0.5(1 - \cos(2\pi \frac{n}{N})), 0 \leq n \leq N \quad (3.1)$$

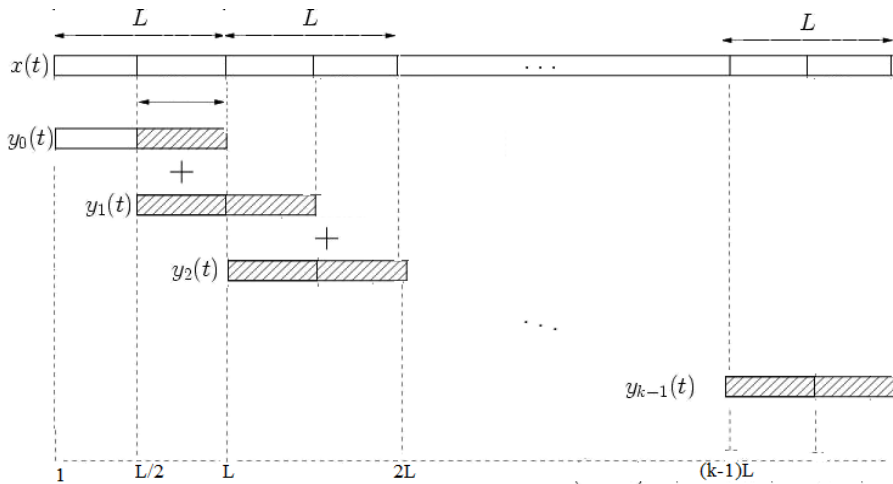
$w(n)$ : pencere katsayıları, pencere uzunluğu  $L$ :  $N+1$



Şekil 3.5. Hanning pencere dürtü tepkisi gösterimi

### 3.2.3 Sesli çerçevelerin eklenmesi (Overlap-Add) Metodu

Bu teknik, örtüşme (overlapping) uygulanan ses sinyalinin, frekans düzleminde işlemleri tamamlandıktan sonra zaman düzleminde tekrar elde edilmesi için kullanılır. Bitişik sesli çerçevelerde örtüşme oranına göre önceki sesli çerçevenin sonundaki örnekler ile sonraki çerçevenin başındaki örneklerin toplanmasıyla sinyalin son hali elde edilir. %50 örtüşme oranına sahip bir sinyale Overlapp-add metodunun uygulandığı görsel Şekil 3.6’te gösterilebilir.



Şekil 3.6. Ses çerçevelerinin eklenmesinin gösterimi

Burada  $L$  sesli çerçeve uzunluğudur. Şekil 3.6’da görüldüğü gibi  $y_0(t)$ ’den başlayıp  $y_{k-1}(t)$ ’ye giden ses çerçevelerinin örnek sayılarının %50’si olacak şekilde önceki sesli çerçevenin ikinci



yarısı ile sonraki çerçevenin ilk yarısı toplanarak  $x(t)$  sinyali yeniden elde edilmiştir. Bu tez çalışmasında örtüşme oranı %50 olarak alınmıştır fakat bu oran daha farklı değerlerde de olabilir.

### 3.3 Yöntem

Bu bölümde VAD algoritması ve filtre tasarımından bahsedilmiştir.

#### 3.3.1 VAD algoritması

VAD tasarımında kısa süreli enerji, periyodiklik ve spektral düzlük olmak üzere üç özellik kullanılmıştır. Öncelikle VAD algoritmasının girişindeki ses sinyali örtüşme işlemi uygulanır. Kısa süreli enerjinin eşik seviyesini hesaplamak için ses sinyalinin ilk 20 çerçevesi kullanılır. Her ses çerçevesi 128 ses örneğinden oluşur. Periyodiklik hesabı için sesli çerçevelerin 16. örneğinden, 64. örneğine kadar olan bölümü kullanılmıştır. Her ses çerçevesi için maksimum  $R(\tau)$  değeri bize o ses çerçevesinin periyodikliğini (C) verir.  $C=1$  ise tamamen periyodik,  $C=0$  ise sinyal rastgele bir sinyaldir diyebiliriz. Ses sinyalinin frekans düzlemine aktarılmasıyla spektral düzlük hesaplanır sonra yine kısa süreli enerji hesaplamada olduğu gibi ilk 20 sesli çerçeve içinde maksimum değer bulunarak spektral düzlük eşik seviyesi hesaplanır. Tüm bu özelliklerin hesaplanmasından sonra aşağıdaki algoritma kullanılarak ses sinyalinin her çerçevesi için 1 ya da 0 olmak üzere konuşma olup olmadığı tespit edilmiştir. VAD vektörü aşağıdaki algoritma ile tespit edilmiştir:

- 
1. Ses çerçevesinin boyutu 128 örnek olarak alınmıştır.
  2. Kullanılan özellikler için eşik değerleri harici olarak belirlenmiştir.
    - STE için K sabiti ( $K>1$ )
    - Periyodiklik eşik değeri
    - SF\_eşik değeri
  3. for  $i=1$  den çerçeve sayısına kadar
    - 3.1. STE(i)'yi hesapla
    - 3.2. Periyodikliği(i)'yi hesapla
    - 3.3 Her çerçeve için FFT uygula.
      - 3.3.1. Her çerçeve için SF değerini hesapla.
    - 3.4. STE, Periyodiklik ve SF eşik değeri belirlendikten sonra, ilk 20 çerçevenin enerjisinin ortalaması gürültü eşik değeri olarak, yine ilk 20 çerçevenin maksimum SF değeri ise gürültünün SF'si olarak değerlendirildi ve  $STE_{noise}$  ile  $SF_{noise}$  değeri hesaplanır.
      - 3.4.1.  $SF_{noise}=\max(SF(1,1:20))$ ; olarak girildi

3.4.2.  $STE_{noise} = STE(1,1:20)/20$ ;

3.5.  $Sayaç=0$

- Eğer  $STE(i) \geq K \cdot STE_{noise}$  değeri,  $sayaç=sayaç+1$ ;
- Eğer  $SF(i) - SF_{noise} \leq SF_{esik}$  ise  $sayaç=sayaç+1$ ;
- Eğer  $Periodicity(i) \geq 0.35$  ise  $sayaç=sayaç+1$ ;

3.6 Eğer  $sayaç \geq 1$  ise; Ses çerçevesi gürültülü konuşmadır;

Eğer  $sayaç \geq 1$  değilse; Ses çerçevesi gürültüdür;

3.7. Eğer ses çerçevesi gürültü ise aşağıdaki formülle STE'nin eşik değeri güncellendi.  
 $P=0.15$  kabul edildi.

$$STE_{esik_{yeni}} = (1-p) \cdot STE_{esik_{eski}} + p \cdot STE_{noise}$$

4. Ardışık 10 çerçeveden daha az sessiz çerçeve varsa bir olarak değiştir.

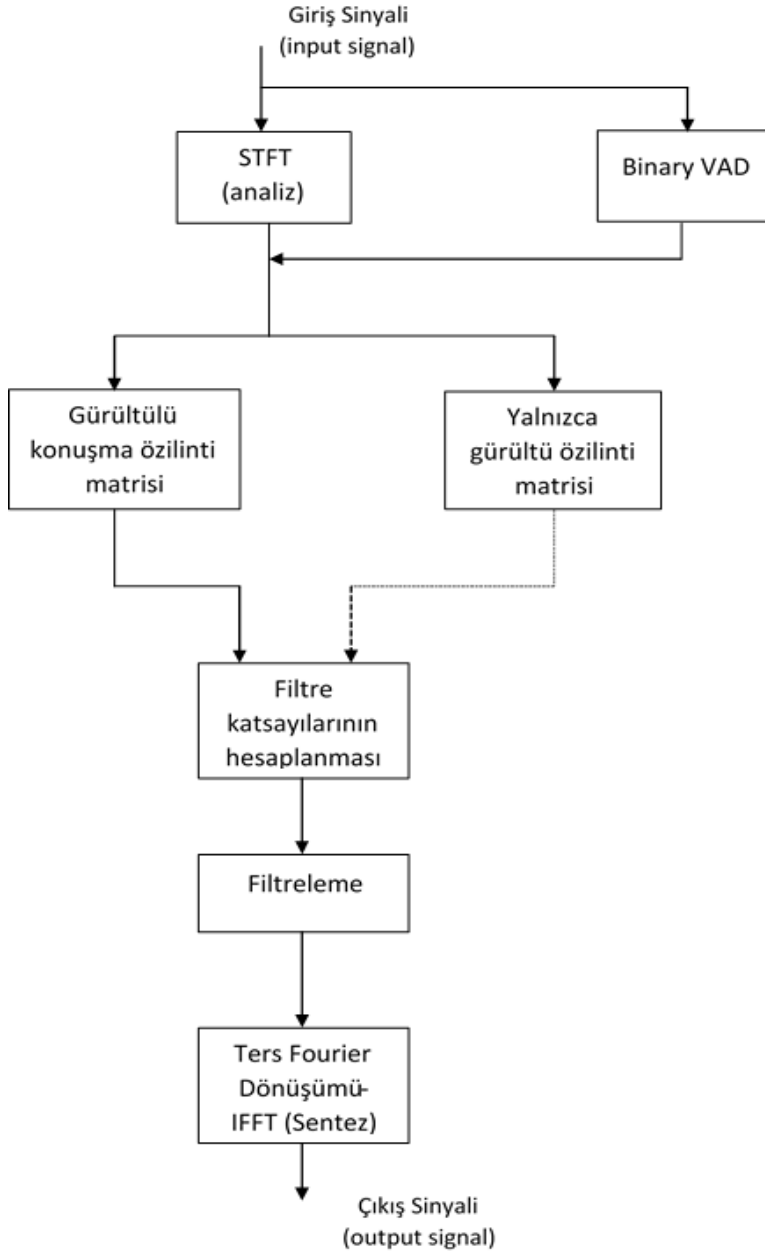
5. Ardışık beş çerçeveden daha az konuşma çerçeve varsa sıfır olarak değiştir.

---

Ses sinyaline yukarıdaki algoritmanın uygulanmasıyla 0 ve 1 değerlerinin bulunduğu sesli çerçeve sayısı kadar uzunlukta bir vektör elde edilir. VAD'ın sesli ve gürültülü bölgeleri doğru tespiti algoritmanın devamında filtre katsayılarının doğru hesaplanabilmesi için önemlidir (Çolak ve Akdeniz 2018).

### 3.3.2 Adaptif filtre tasarımı

Ses sinyalinde işlemler çok küçük zaman aralıklarında yapılır (örneğin 10 msn gibi). Bu çalışmadaki gürültü azaltma algoritmasının akış diyagramına göre yapılan ilk işlem; zaman düzleminde konuşma sinyalinin tamamını küçük çerçevelere ayırıp, örtüşme işlemi uygulanarak FFT ile zaman düzleminden frekans düzlemine aktarmaktır. Tüm ses çerçeveleri frekans düzlemine aktarıldıktan sonra her ses çerçevesi için VAD algoritmasından 0 (konuşma yok) ya da 1 (konuşma var) şeklinde çıkan değerlere göre konuşma ve gürültü özilinti matrisleri belirlenmiştir. Gürültülü konuşma ve gürültü özilinti matrisleri VAD değerlerine göre belirli bir formül ile güncellenmiştir. Bu matrislerin aldığı değerlere göre her ses çerçevesinin filtre katsayıları hesaplanmıştır. Bu işlemin ardından her çerçeveye filtreleme işlemi yapıldıktan sonra ters Fourier dönüşümü uygulanarak ses sinyali tekrar zaman düzlemine aktarılmıştır. Son olarak da Overlapp-add metodu uygulanarak zaman düzleminde iyileştirilmiş sinyal elde edilmiştir. Akış diyagramı Şekil 3.7'de verilmiştir:



Şekil 3.7. Bütün algoritmanın blok diyagramı

### 3.3.2.1 Kısa süreli fourier dönüşümü (STFT, Analiz)

Analiz aşamasında zaman düzleminde ses sinyaline %50 örtüşme yöntemi uygulanarak küçük çerçevelere ayrıldıktan sonra bu sesli çerçeveler Hanning pencere fonksiyonu ile çarpılmıştır. Çerçeve boyutu 128 örnek olarak alınmıştır. Daha sonra bu sesli çerçevelere FFT uygulanarak, zaman düzlemindeki sesli çerçeveler frekans düzlemine aktarılmıştır. Bu işlem için Matlab programında FFT fonksiyonu kullanılmıştır.

### 3.3.2.2 Özilinti matrislerinin güncellenmesi

Öncelikle frekans düzlemine aktarılan ses sinyalleri aşağıdaki gibi ifade edilir:

$$Y_i(k,l) = X_i(k,l) + V_i(k,l) \quad (3.2)$$

$Y_i(k,l)$ : mikrofondan alınan sinyal,  $X_i(k,l)$ : sinyalin konuşma bileşeni,  $V_i(k,l)$ : sinyalin gürültü bileşeni,  $k$ : frekans indeksi,  $l$ : sesli çerçevenin indeksi,  $i$ : mikrofonsayı.

Kullanılan mikrofonsayı kadar ses sinyali vektörleri aynı matris içine toplanmıştır.

$$\mathbf{Y}(k,l) = [Y_1(k,l) \ Y_2(k,l) \ \dots \ Y_i(k,l)]^T \quad (3.3)$$

Burada üst indis olan  $T$ , matrisin transpozisini belirtir.  $i$  ise kullanılan mikrofonsayıdır. Özilinti matrisleri aşağıdaki gibidir:

$$\mathbf{R}_y(k,l) = \varepsilon\{\mathbf{Y}(k,l)\mathbf{Y}^H(k,l)\} \quad (3.4)$$

$$\mathbf{R}_x(k,l) = \varepsilon\{\mathbf{X}(k,l)\mathbf{X}^H(k,l)\} \quad (3.5)$$

$$\mathbf{R}_v(k,l) = \varepsilon\{\mathbf{V}(k,l)\mathbf{V}^H(k,l)\} \quad (3.6)$$

Burada  $\mathbf{R}_y(k,l)$ : gürültülü konuşmanın özilinti matrisi,  $\mathbf{R}_x(k,l)$ : gürültüsüz konuşma sinyalinin özilinti matrisi,  $\mathbf{R}_v(k,l)$ : gürültü özilinti matrisi, üst indis olan  $H$ , matrisin Hermitian transpozisidir. Hermitian transpoze işleminde karmaşık sayıların eşleniği alınır.  $\varepsilon\{\cdot\}$  ifadesi beklenti operatörünü temsil eder.

Sonraki aşamada gürültülü konuşma ve gürültü ilinti matrisleri belirlenme işlemine geçilmiştir. Gürültülü konuşma ilinti matrisinin hesaplanması için  $VAD=1$  olduğu sesli çerçevelerin belirli bir sayıda çerçeve alınarak ortama gürültülü konuşma özilinti matrisi hesaplanmıştır. Gürültü özilinti matrisinin belirlenmesi için ise sinyalin yalnızca gürültü bölgelerinde yani  $VAD=0$  olan çerçevelerden belirli bir sayıda kullanılarak ortalama bir gürültü özilinti matrisi elde edilmiştir. Gürültülü konuşma ve gürültü özilinti matrislerinin ortalaması elde edildikten sonra iteratif bir formül ile ses sinyalinin her çerçevesi için  $VAD$  değerleri dikkate alınarak ilinti matrisleri güncellenmiştir.

$$\mathbf{H}_0(k,l): \mathbf{Y}(k,l) = \mathbf{V}(k,l) \quad (3.7)$$

$$\mathbf{H}_1(k,l): \mathbf{Y}(k,l) = \mathbf{X}(k,l) + \mathbf{V}(k,l) \quad (3.8)$$

Burada  $\mathbf{H}_0(k,l)$  ve  $\mathbf{H}_1(k,l)$  konuşma sinyalinin olmadığı ve olduğu durumları temsil eder. Kullanılan VAD algoritmasıyla  $\mathbf{H}_0(k,l)$  ve  $\mathbf{H}_1(k,l)$  tespiti yapıldıktan sonra aşağıdaki iteratif formüllerle ses sinyalinin her frekans indeksi için  $\mathbf{R}_y(k,l)$  ve  $\mathbf{R}_v(k,l)$  özilinti matrisleri hesaplanmıştır.

$$\mathbf{H}_0(k,l) : \begin{cases} \mathbf{R}_v(k,l+1) = \alpha_n \mathbf{R}_v(k,l) + (1-\alpha_n) \mathbf{Y}(k,l) \mathbf{Y}^H(k,l) \\ \mathbf{R}_y(k,l+1) = \mathbf{R}_y(k,l) \end{cases} \quad (3.9)$$

$$\mathbf{H}_1(k,l) : \begin{cases} \mathbf{R}_y(k,l+1) = \alpha_y \mathbf{R}_y(k,l) + (1-\alpha_y) \mathbf{Y}(k,l) \mathbf{Y}^H(k,l) \\ \mathbf{R}_v(k,l+1) = \mathbf{R}_v(k,l) \end{cases} \quad (3.10)$$

Burada  $\alpha_n$  ve  $\alpha_y$  gürültü ve gürültülü konuşma matrislerinin unutma faktörleridir. Yukarıdaki formüller ile girilen ses sinyalinin her örneği için ilinti matrisleri hesaplandıktan sonra temiz konuşma sinyalinin tahmini ilinti matrisi aşağıdaki gibi hesaplanmıştır:

$$\mathbf{R}_x(k,l) = \mathbf{R}_y(k,l) - \mathbf{R}_v(k,l) \quad (3.11)$$

Yukarıda da görüldüğü üzere özilinti matrislerinin güncellenmesi VAD vektörünün değerlerine göre olduğu için, elde edilen sonuçların doğruluğu için VAD algoritmasının verimliliği çok önemlidir (Ngo 2011).

### 3.3.2.3 Filtre katsayılarının hesaplanması

Özilinti matrisleri belirlendikten sonra filtre katsayılarını hesaplamak için aşağıdaki filtre katsayıları hesaplama yöntemleri ile filtre katsayıları hesaplanmıştır. Girişteki mikrofonların her biri için ayrı ayrı filtre katsayıları hesaplanmıştır.

**Çizelge 3.1.** Filtre Katsayıları formül tablosu

Klasik Wiener Filtre	$\mathbf{W}_{\text{MMSE}}(k,l) = [\mathbf{R}_x(k,l) + \mathbf{R}_v(k,l)]^{-1} \mathbf{R}_x(k,l) e_1$
SDW Wiener Filtre	$\mathbf{W}_{\text{MWF}\mu}(k,l) = [\mathbf{R}_x(k,l) + \mu \mathbf{R}_v(k,l)]^{-1} \mathbf{R}_x(k,l) e_1$
SP Filtre	$\mathbf{W}_{\text{SP}}(k,l) = \frac{1}{\mathbf{H}^H(k,l) \mathbf{R}_v^{-1}(k,l) \mathbf{H}(k,l)} \mathbf{R}_v^{-1}(k,l) \mathbf{H}(k,l)$
MVDR Filtre	$\mathbf{W}_{\text{MVDR}}(k,l) = \frac{1}{e^H(k,l) \mathbf{R}_x^{-1}(k,l) e(k,l)} \mathbf{R}_x^{-1}(k,l) e(k,l)$

### 3.3.2.4 Filtreleme

Filtreleme aşamasında filtreleme çıkışındaki sinyali elde etmek için filtre katsayıları ile girişteki referans olarak seçilen mikrofonun katsayıları çarpılmıştır.

$$Z(k,l)=\mathbf{W}^H(k,l)\mathbf{Y}(k,l) \quad (3.12)$$

Burada  $Z(k,l)$  filtrelenmiş sinyaldir.

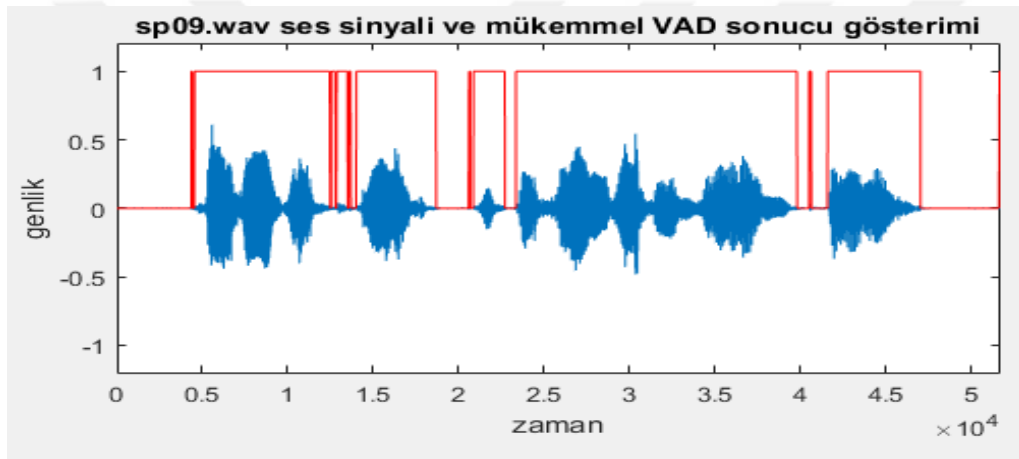
### 3.3.2.5 Ters Fourier dönüşümü (Sentez)

Elde edilen filtre çıkış sinyali frekans düzleminde ve örtüşme yöntemi uygulanmış sinyaldir. Öncelikle bu sinyalin her ses çerçevesi ters Fourier dönüşümü ile zaman düzlemine aktarılır. Bu işlem Matlab'ta IFFT fonksiyonu ile yapılabilir. Zaman düzlemine aktarılan sinyalin ses çerçevelerine analiz bölümünde de bahsedilen Hanning pencere fonksiyonu uygulanmıştır. Son olarak da örtüşme metodu uygulanmış bu ses çerçevelerine Overlapp-Add metodu uygulanır. Elde edilen sinyalin uzunluğu tekrar giriş sinyalinin boyutu ile eşit hale getirilmiştir.

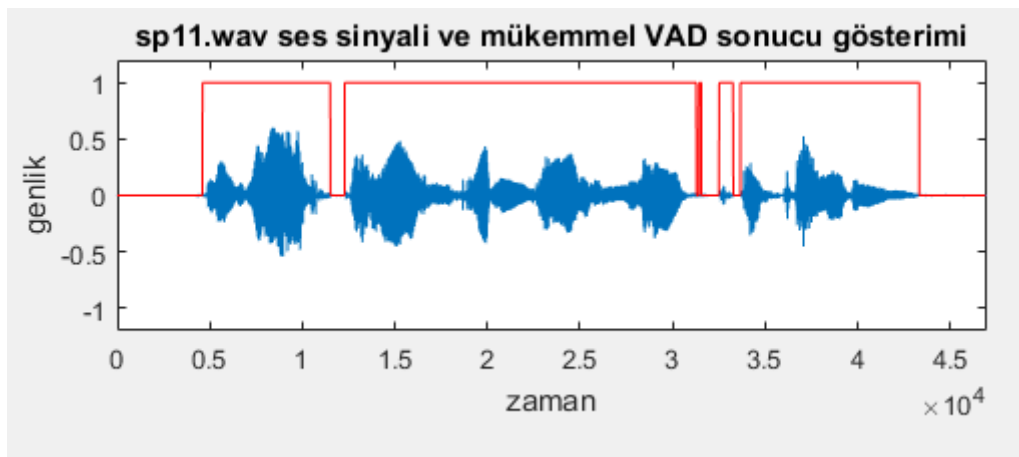
## 4. ARAŞTIRMA BULGULARI

### 4.1 VAD Sonuçları

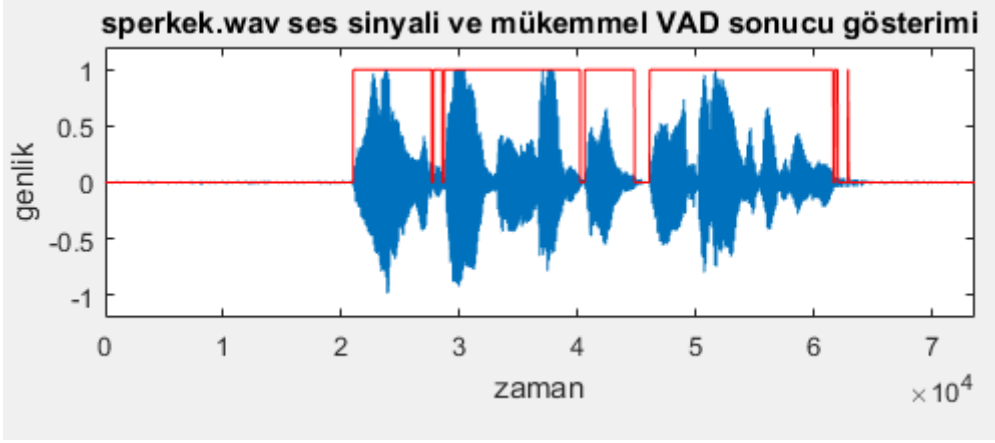
Konuşma seslerinin VAD sonuçlarını bulabilmek için öncelikle bu konuşma sinyallerinde, konuşma olan bölgelerin tam olarak tespiti yapılmıştır. Aslında bu VAD algoritmasının hedeflediği mükemmel sonuçtur. Öncelikle mükemmel VAD sonuç vektörleri bulunup, daha sonra gürültülü konuşma sinyalleri için bulunan VAD vektörleriyle karşılaştırılarak yüzde olarak doğruluğu tespit edilmiştir. Dört temiz ses sinyaline ilişkin mükemmel VAD tespitlerinin bulunduğu görseller Şekil 4.1, Şekil 4.2, Şekil 4.3 ve Şekil 4.4'te gösterilmiştir.



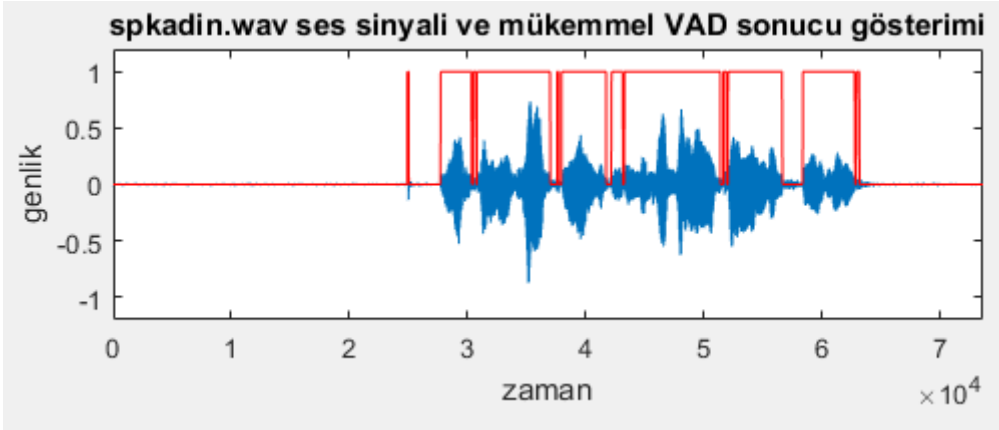
Şekil 4.1. sp09.wav sinyaline ait mükemmel VAD sonucunun gösterimi



Şekil 4.2. sp11.wav sinyaline ait mükemmel VAD sonucunun gösterimi



Şekil 4.3. sperkek.wav sinyaline ait mükemmel VAD sonucunun gösterimi



Şekil 4.4. spkadin.wav sinyaline ait mükemmel VAD sonucunun gösterimi

Dört konuşmacı ve bu konuşmacılara 0 dB, 5 dB ve 10 dB olmak üzere üç farklı SNR değerlerinde eklenen havaalanı (airport), araba (car), beyaz gürültü (White noise) kullanılarak dört farklı açıda oluşturulan ses dosyaları kullanılarak elde edilen VAD sonuçları Çizelge 4.1 ve Çizelge 4.2’de verilmiştir.

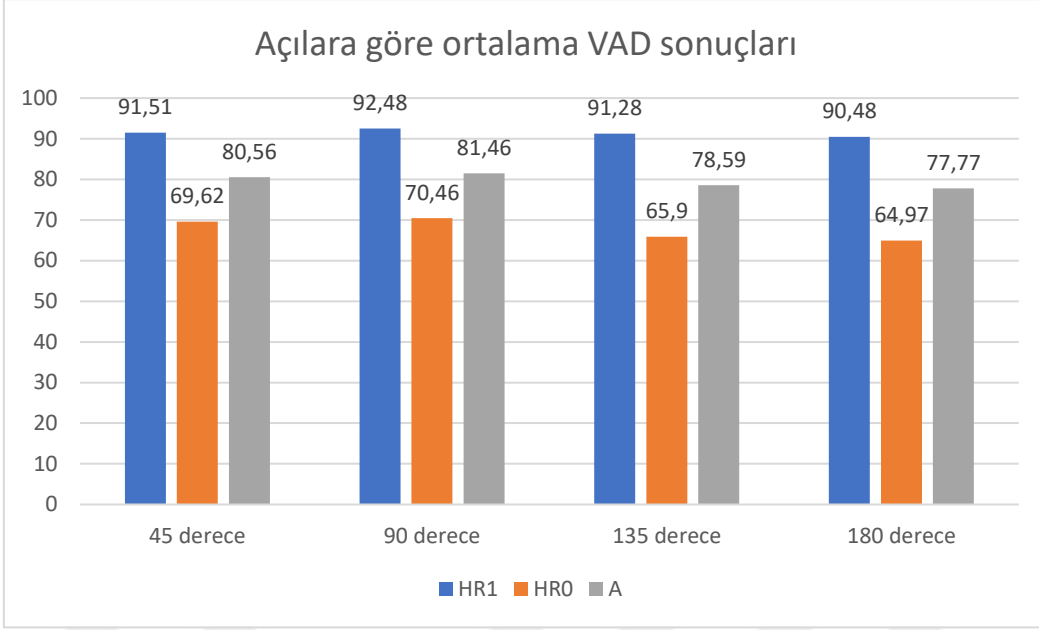


**Çizelge 4.1.** sp09.wav ve sp11.wav ses sinyallerinin VAD sonuçları

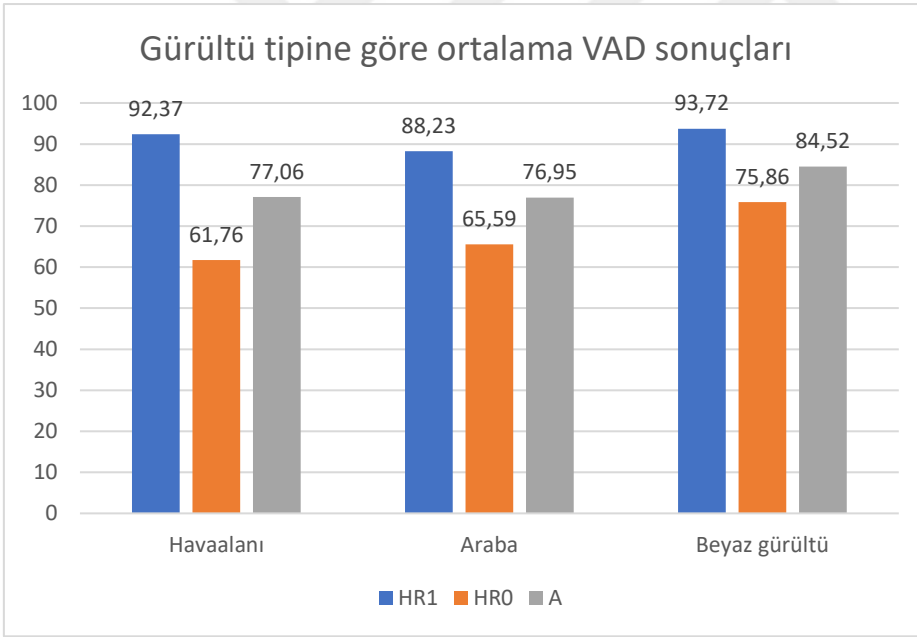
Ses	Gürültü Tipi	SNR	45 Derece			90 Derece			135 Derece			180 Derece		
			HR1 (%)	HR0 (%)	A (%)	HR1 (%)	HR0 (%)	A (%)	HR1 (%)	HR0 (%)	A (%)	HR1 (%)	HR0 (%)	A (%)
Sp09.wav	Havaalanı (Airport)	0 dB	93,04	65,17	79,1	94,58	62,76	78,67	87,62	73,69	80,7	94,2	53,79	73,99
		5 dB	92,65	80,34	86,49	95,74	65,17	80,45	92,84	67,59	80,2	91,65	65,52	78,5
		10 dB	96,52	80	88,26	98,26	70	84,12	96,32	66,9	81,6	93,42	73,45	83,43
	Araba (Car)	0 dB	76,21	74,83	75,51	91,49	54,14	72,81	86,07	51,03	68,6	85,88	55,52	70,69
		5 dB	90,33	70,69	80,5	92,46	70,34	81,4	88,97	64,14	76,6	88,59	69,66	79,12
		10 dB	81,04	83,79	82,41	91,49	78,62	85,05	91,3	63,79	77,5	93,62	58,97	78,29
	Beyaz Gürültü (White noise)	0 dB	90,14	66,55	78,34	85,69	77,24	81,46	93,23	53,45	73,3	86,07	73,1	79,58
		5 dB	95,55	73,45	84,5	88,97	93,79	91,38	92,26	80,69	86,5	89,94	76,9	83,41
		10 dB	95,55	94,14	94,84	96,91	91,72	94,31	96,32	93,1	94,7	95,55	85,17	90,36
Sp11.wav	Havaalanı (Airport)	0 dB	87,28	59,81	73,54	87,86	50,93	69,39	88,44	40,19	64,3	87,48	32,24	59,85
		5 dB	89,02	67,76	78,38	89,4	57,01	73,2	90,37	43,46	66,9	94,22	38,32	66,26
		10 dB	98,46	78,5	88,48	92,68	71,5	82,08	97,69	54,21	75,9	97,5	49,07	73,28
	Araba (Car)	0 dB	83,24	65,42	74,32	79	61,21	70,1	80,35	61,21	70,8	82,47	62,62	72,54
		5 dB	86,32	67,29	76,8	90,94	66,36	78,64	83,04	65,42	74,2	85,93	76,64	81,28
		10 dB	91,91	82,24	87,07	91,91	82,24	87,07	90,17	65,89	78	92,29	83,64	87,96
	Beyaz Gürültü (White noise)	0 dB	91,33	50,47	70,89	95,95	67,76	81,85	94,03	56,54	75,3	64,35	94,39	79,37
		5 dB	99,23	64,49	81,85	99,81	80,84	90,32	95,38	75,23	85,3	96,53	72,43	84,48
		10 dB	98,27	84,11	91,18	99,04	92,52	95,78	96,72	79,44	88,1	96,34	87,38	91,86

**Çizelge 4.2.** spkadin.wav ve sperkek.wav ses sinyallerinin VAD sonuçları

VAD algoritmasının sonuçları														
			45 Derece			90 Derece			135 Derece			180 Derece		
Ses	Gürültü tipi	SNR	HR1 (%)	HR0 (%)	A (%)	HR1 (%)	HR0 (%)	A (%)	HR1 (%)	HR0 (%)	A (%)	HR1 (%)	HR0 (%)	A (%)
Spkadin.wav	Havaalanı (Airport)	0 dB	89,72	42,73	66,22	87,3	43,19	65,24	90,73	42,27	66,5	85,48	42,88	64,18
		5 dB	89,11	62,02	75,56	87,3	71,67	79,48	93,35	49,77	71,6	89,72	51,15	70,43
		10 dB	96,77	75,34	86,05	96,37	69,98	83,17	94,76	79,17	87	93,15	62,48	77,81
	Araba (Car)	0 dB	83,87	52,68	68,27	88,1	75,8	81,95	73,79	70,75	72,3	84,68	48,09	66,38
		5 dB	90,12	66	78,06	91,94	66,16	79,04	86,9	66,92	76,9	88,91	52,68	70,79
		10 dB	95,16	86,98	91,07	91,94	84,38	88,15	95,97	73,35	84,7	94,96	80,25	87,6
	Beyaz Gürültü (White noise)	0 dB	92,14	69,37	80,75	91,33	57,89	74,6	94,76	53,75	74,3	91,53	58,96	75,24
		5 dB	96,17	72,74	84,45	95,77	83,61	89,69	94,96	74,43	84,7	94,96	78,25	86,6
		10 dB	97,98	86,52	92,25	98,99	92,34	95,66	97,78	91,27	94,5	98,99	91,12	95,05
Sperkek.wav	Havaalanı (Airport)	0 dB	90,98	54,73	72,85	90,16	46,57	68,36	86,89	55,84	71,4	85,41	47,87	66,63
		5 dB	92,3	63,08	77,68	95,41	72,91	84,16	92,95	62,15	77,6	92,3	62,3	77,3
		10 dB	94,92	81,45	88,18	96,89	86,27	91,57	97,05	86,83	91,9	97,54	84,6	91,07
	Araba (Car)	0 dB	83,61	46,75	65,17	88,36	52,32	70,33	82,3	64,19	73,2	91,31	54,73	73,02
		5 dB	91,48	47,12	69,29	89,51	53,25	71,37	84,26	69,57	76,9	91,15	51,39	71,26
		10 dB	93,44	56,96	75,2	94,43	63,45	78,93	90,98	70,32	80,6	92,95	59	75,97
	Beyaz Gürültü (White noise)	0 dB	88,85	64,94	76,89	90	52,5	71,25	91,48	60,3	75,9	83,61	43,6	63,6
		5 dB	95,9	77,18	86,54	94,43	84,42	89,42	97,05	66,23	81,6	87,58	81,45	84,41
		10 dB	96,07	90,72	93,39	99,02	85,9	92,345	99,18	79,41	89,3	97,05	79,59	88,32



**Şekil 4.5.** Tüm konuşmaların açılara göre ortalama VAD sonuçları



**Şekil 4.6.** Tüm konuşmaların gürültü tipine göre ortalama VAD sonuçları

Burada HR1 doğru gürültülü konuşma tespiti (Hit rate 1), HRO doğru gürültü tespiti (Hit rate 0), A ise ortalama (Average)'dır.

VAD algoritmasının verimliliğini değerlendirmek için kullanılan parametreler aşağıda açıklanmıştır:

1. Doğru gürültülü konuşma tespiti (HR1): Bu parametre ses sinyalinde doğru tespit edilen gürültülü konuşma içeren ses çerçevelerini belirtir.
2. Doğru gürültü tespiti (HR0): Bu parametre ses sinyalinde doğru tespit edilen gürültülü ses çerçevelerini belirtir.
3. Hatalı gürültülü konuşma tespiti (HR0\_hata): Bu parametre ses sinyalinde yalnızca gürültü olmasına rağmen VAD karar mekanizmasının yanlış tespiti ile gürültülü konuşma tespit edilmesidir.
4. Hatalı gürültü tespiti (HR1\_hata): Bu parametre ses sinyalinin gürültülü konuşma olmasına rağmen VAD karar mekanizmasının yanlış tespiti ile ses çerçevesinin yalnızca gürültü olarak tespit edilmesidir.

$$\%HR1 = \frac{HR1}{HR1 + HR1\_hata} \cdot 100 \quad (4.1)$$

$$\%HR0 = \frac{HR0}{HR0 + HR0\_hata} \cdot 100 \quad (4.2)$$

Eşitlik (4.1) ve Eşitlik (4.2)'deki formüller kullanılarak VAD sonuçları tespit edilmiştir. Şekil 4.5'teki grafiğe göre gürültünün 45 derece ve 90 derece olduğu pozisyonlarda elde edilen sonuçlar benzerlik göstermektedir. 135 ve 180 derece açıyla eklenen gürültülerin varlığında ise HR1 biraz daha düşük olmakla birlikte HR0 sonucunda belirgin bir düşme vardır.

Şekil 4.6'da ise gürültünün tipine göre tüm konuşmaların ortalama doğruluk tespit değerleri belirtilmiştir. Burada kullanılan gürültülerden en iyi tespit oranı, zamanla çok fazla değişmeyen genliğinden ötürü beklenildiği gibi beyaz gürültüde olmuştur. Havaalanı ve araba gürültülerinde ortalama tespit sonuçları hemen hemen aynıken, havaalanı gürültüsünde HR1 tespit oranının araba gürültüsüne göre fazla olduğu tespit edilmiştir. Buna karşılık araba gürültüsünde ise HR0 tespit oranı havaalanı gürültüsüne göre daha fazladır.

Kullanılan VAD algoritmasıyla birlikte dört algoritma için SNR sonuçları ise Çizelge 4.3 ve Çizelge 4.4'te verilmiştir:

## 4.2 SNR Sonuçları

**Çizelge 4.3.** sp09.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları

SNR Sonuçları								
Ses	Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)				
	Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR	SP
Sp09.wav	Havaalanı (Airport)	45	0 dB	6,61	7,19	7,73	5,68	7,44
		45	5 dB	10,58	10,89	11,09	10,11	11,13
		45	10 dB	14,40	14,82	15,03	14,14	14,91
		90	0 dB	6,08	7,35	8,24	5,17	8,01
		90	5 dB	9,43	9,88	10,25	8,91	9,98
		90	10 dB	13,55	13,80	13,93	13,21	13,14
		135	0 dB	5,52	6,18	6,64	5,07	7,03
		135	5 dB	9,29	9,88	10,23	8,79	10,63
		135	10 dB	13,70	14,43	14,75	13,45	14,84
		180	0 dB	5,83	7,64	8,46	4,54	6,84
		180	5 dB	9,36	10,19	10,90	8,40	10,82
		180	10 dB	13,03	13,47	13,87	11,89	14,88
	Araba (Car)	45	0 dB	1,91	3,36	3,86	0,95	0,38
		45	5 dB	6,47	7,78	8,28	6,03	7,05
		45	10 dB	10,83	11,80	12,20	10,86	11,56
		90	0 dB	2,64	4,12	4,58	1,73	2,26
		90	5 dB	7,54	8,46	8,76	7,44	7,30
		90	10 dB	11,73	12,68	12,96	11,93	11,70
		135	0 dB	4,08	5,22	5,70	3,17	4,12
		135	5 dB	7,83	9,20	9,76	6,50	8,03
		135	10 dB	11,46	13,04	13,72	10,47	12,92
		180	0 dB	2,14	5,65	6,81	2,02	4,13
		180	5 dB	5,90	8,90	9,95	5,88	9,00
		180	10 dB	10,10	12,03	12,91	*	12,85
	Beyaz gürültü (White Noise)	45	0 dB	5,67	8,56	9,17	4,04	4,01
		45	5 dB	10,53	12,41	12,93	9,32	10,34
		45	10 dB	13,82	15,36	15,90	12,37	14,68
		90	0 dB	6,93	8,66	9,20	4,94	4,65
		90	5 dB	10,70	12,27	12,83	9,18	9,28
		90	10 dB	14,30	15,69	16,20	13,29	13,83
		135	0 dB	7,51	9,64	10,31	5,52	7,98
		135	5 dB	10,32	12,11	12,87	8,68	11,44
		135	10 dB	13,57	15,07	15,82	12,33	16,05
		180	0 dB	6,56	11,38	12,24	2,84	6,90
		180	5 dB	9,33	13,64	14,95	5,82	10,60
		180	10 dB	10,36	14,93	16,54	*	13,88

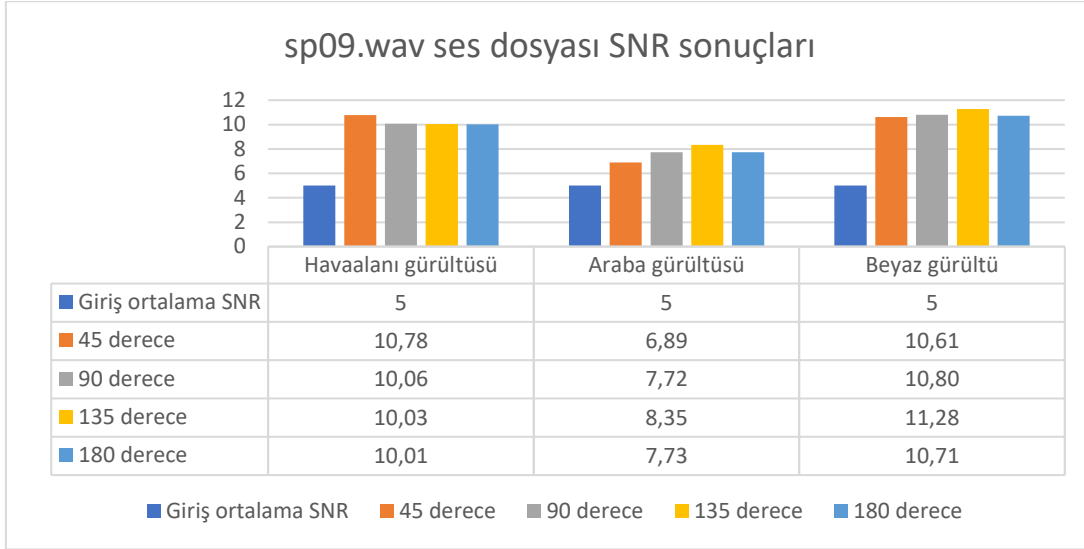
“\*”: iyileştirme yapılmayan sonuçları belirtir.

**Çizelge 4.4.** sp11.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları

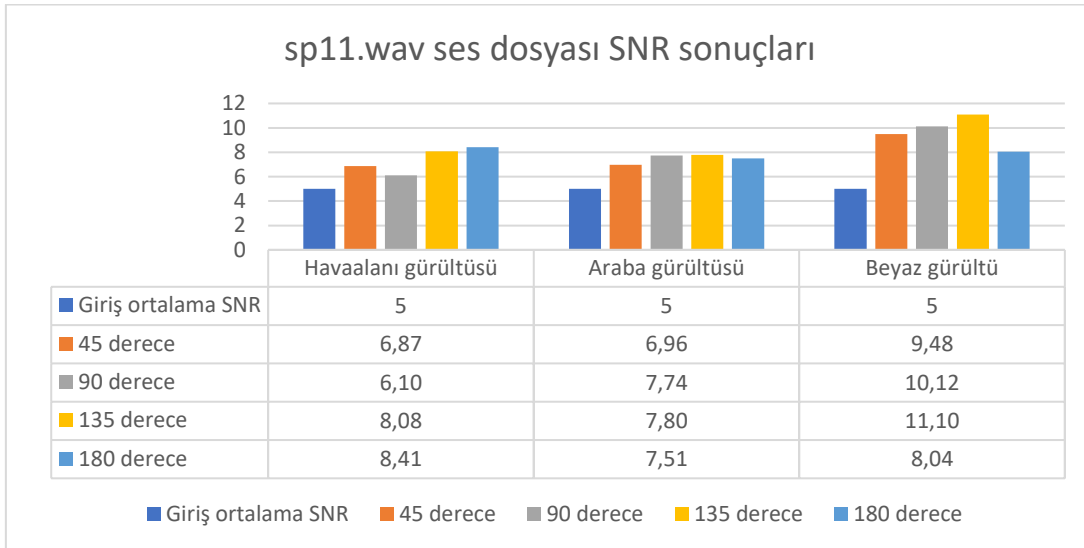
SNR Sonuçları								
Ses	Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)				
	Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR	SP
Sp11.wav	Havaalanı (Airport)	45	0 dB	2,44	3,00	3,20	1,77	*
		45	5 dB	6,90	6,98	7,03	6,20	*
		45	10 dB	11,39	11,28	11,28	11,00	*
		90	0 dB	4,26	4,07	3,82	3,64	1,35
		90	5 dB	6,87	5,81	5,21	6,19	*
		90	10 dB	11,15	10,31	*	10,52	*
		135	0 dB	4,71	4,74	3,83	5,01	1,29
		135	5 dB	8,48	10,34	11,15	7,31	8,92
		135	10 dB	11,88	11,95	11,98	11,47	*
		180	0 dB	*	7,18	8,36	1,84	5,40
		180	5 dB	5,39	9,80	10,52	6,25	8,93
		180	10 dB	*	12,87	13,31	*	11,11
	Araba (Car)	45	0 dB	3,19	5,79	6,71	1,36	1,61
		45	5 dB	5,31	8,27	9,00	*	5,23
		45	10 dB	10,17	11,49	12,03	*	10,37
		90	0 dB	0,27	2,39	3,05	*	*
		90	5 dB	6,07	8,56	9,04	6,50	6,02
		90	10 dB	10,89	12,62	13,06	10,89	11,29
		135	0 dB	2,48	5,60	6,41	3,35	3,85
		135	5 dB	6,61	8,54	9,41	5,44	6,53
		135	10 dB	10,81	12,38	13,17	10,06	12,29
		180	0 dB	0,45	6,64	8,29	1,29	3,79
		180	5 dB	5,36	8,71	9,86	*	7,22
		180	10 dB	*	13,20	13,99	*	11,29
	Beyaz gürültü (White Noise)	45	0 dB	3,24	7,94	9,04	1,64	1,43
		45	5 dB	8,69	12,01	12,77	7,14	7,80
		45	10 dB	13,48	14,91	15,43	12,46	14,18
		90	0 dB	5,76	8,21	8,85	4,46	3,84
		90	5 dB	9,90	11,72	12,19	8,73	8,03
		90	10 dB	13,57	15,30	15,84	12,56	12,87
		135	0 dB	7,17	10,28	11,15	5,77	7,53
		135	5 dB	9,41	11,85	12,79	7,99	10,53
		135	10 dB	13,51	14,94	15,61	12,42	15,60
180		0 dB	0,15	6,22	7,57	0,06	2,87	
180		5 dB	5,20	10,48	12,14	*	8,76	
180		10 dB	10,37	13,40	14,56	*	12,73	

‘\*’: iyileştirme yapılmayan sonuçları belirtir.

Sp09.wav ve sp11.wav ses dosyalarının başındaki sessiz bölge az olduğunda  $R_y$  gürültülü konuşma ve  $R_v$  gürültü özilinti matrislerinin başlangıç değerleri hesaplanırken;  $R_y$  için 50 sesli çerçeve (0,4 sn),  $R_v$  için 40 sesli çerçeve (0.32 sn) kullanılmıştır. Çizelge 4.3 ve Çizelge 4.4 ile ilgili grafikler Şekil 4.7 ve Şekil 4.8 de verilmiştir.



**Şekil 4.7.** sp09.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları

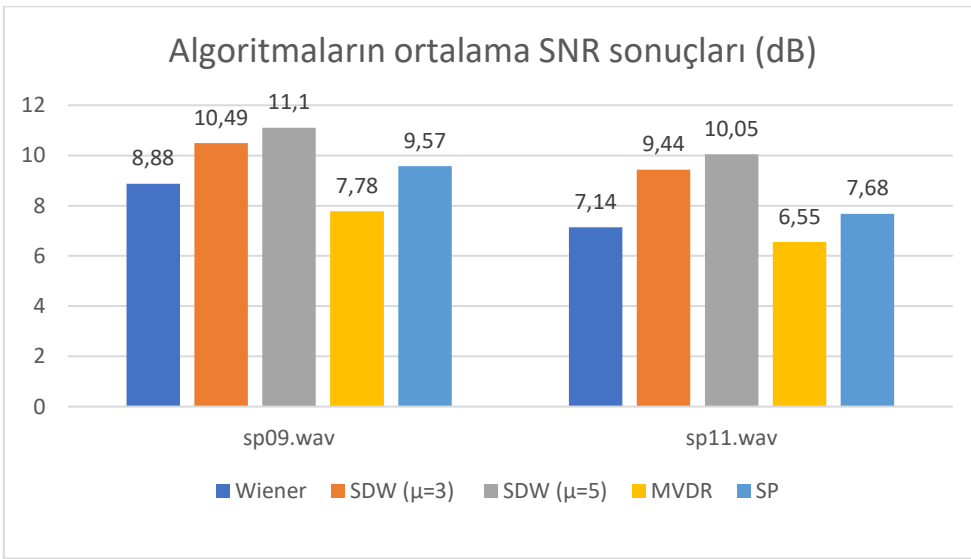


**Şekil 4.8.** sp11.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları

Değerlendirme aşamasında iyileştirme yapılamayan algoritma değerleri dikkate alınmamıştır. Giriş SNR değerleri olan 0, 5 ve 10 dB değerlerinin ortalaması giriş SNR

değeri olarak kabul edilmiştir. Ayrıca çıkış SNR değerlerinin de ortalama değeri çıkış SNR değeri olarak alınmıştır.

Sp09.wav ve sp11.wav ses dosyalarının SNR sonuçlarının açığa ve gürültü tipine göre değişimi incelendiğinde sp09.wav ses dosyasının özellikle havaalanı gürültüsünde sp11.wav ses dosyasından daha iyi sonuçlar verdiği tespit edilmiştir. Gürültü tipine göre bakıldığında ise en iyi SNR sonuçları beyaz gürültü tipinde tespit edilmiştir. SNR açısından değerlendirilirse; her iki konuşma sinyali için de SDW ( $\mu=5$ ) algoritması tarafından en iyi sonuçlar elde edilmiştir. Algoritmanın tipine göre her iki konuşma için elde edilen sonuçlar ise Şekil 4.9’da verilmiştir.



**Şekil 4.9.** sp09.wav ve sp11.wav ses sinyallerinin önerilen VAD kullanılarak algoritmalarla elde edilen SNR sonuçları

Çalışılan ses dosyalarının tamamı önerilen VAD algoritmasıyla tespit edildiğinden, önerilen VAD algoritmasının verimliliğini tespit etmek için aynı algoritma ve ses dosyalarının mükemmel VAD sonuçları ile değerlendirme gerekliliği doğmuştur. Bu sebeple sp09.wav ve sp11.wav dosyalarının mükemmel VAD sonuçları ile tüm algoritmalar tarafından elde edilen sonuçları Çizelge 4.5 ve Çizelge 4.6’da verilmiştir.



**Çizelge 4.5.** sp09.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları

SNR Sonuçları								
Ses	Gürültünün		Giriş SNR değerleri	ÇIKIŞ SNR Değerleri (dB)				
	Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR	SP
Sp09.wav	Havaalanı (Airport)	45	0 dB	6,02	7,34	8,01	5,82	8,52
		45	5 dB	10,21	10,95	11,27	9,93	11,52
		45	10 dB	14,33	14,93	15,18	14,14	15,32
		90	0 dB	1,67	6,75	6,55	8,3	8,19
		90	5 dB	5,39	9,54	9,71	9,17	9,8
		90	10 dB	10,12	13,2	13,55	11,72	12,1
		135	0 dB	2,16	5,47	5,91	4,45	7,12
		135	5 dB	7,22	9,52	9,93	7,82	10,46
		135	10 dB	12,62	13,89	14,27	12,61	14,75
		180	0 dB	2,12	7,44	8,28	3,54	8,46
		180	5 dB	6,28	10,67	11,46	7,84	11,3
		180	10 dB	10,94	14,08	14,65	12,82	15,36
	Araba (Car)	45	0 dB	1,14	3,17	3,90	2,57	2,52
		45	5 dB	6,23	7,51	8,02	6,72	7,07
		45	10 dB	10,67	11,76	12,16	10,97	11,73
		90	0 dB	3,21	4,01	4,31	2,77	2,07
		90	5 dB	7,63	8,50	8,79	7,64	7,47
		90	10 dB	11,87	12,83	13,14	12,05	12,41
		135	0 dB	3,13	4,82	5,36	2,54	3,49
		135	5 dB	7,56	9,03	9,53	7,47	8,69
		135	10 dB	10,94	12,74	13,34	11,98	13,83
		180	0 dB	2,89	5,70	6,35	2,72	4,31
		180	5 dB	6,37	9,09	9,92	6,61	9,01
		180	10 dB	10,14	12,04	12,91	*	13,23
	Beyaz gürültü (White Noise)	45	0 dB	7,17	8,43	8,86	5,99	6,56
		45	5 dB	10,83	12,20	12,65	9,62	11,14
		45	10 dB	13,98	15,49	16,02	12,65	15,26
		90	0 dB	7,12	8,67	9,17	5,51	5,16
		90	5 dB	10,93	12,42	12,95	9,64	9,60
		90	10 dB	14,23	15,67	16,22	13,29	14,16
		135	0 dB	6,85	8,85	9,52	4,93	7,27
		135	5 dB	10,13	12,11	12,91	8,46	11,8
		135	10 dB	13,58	15,13	15,91	12,26	16,42
		180	0 dB	4,85	9,03	10,17	2,37	6,8
		180	5 dB	7,38	10,56	12,19	*	10,48
		180	10 dB	11,01	12,64	13,68	*	14,15

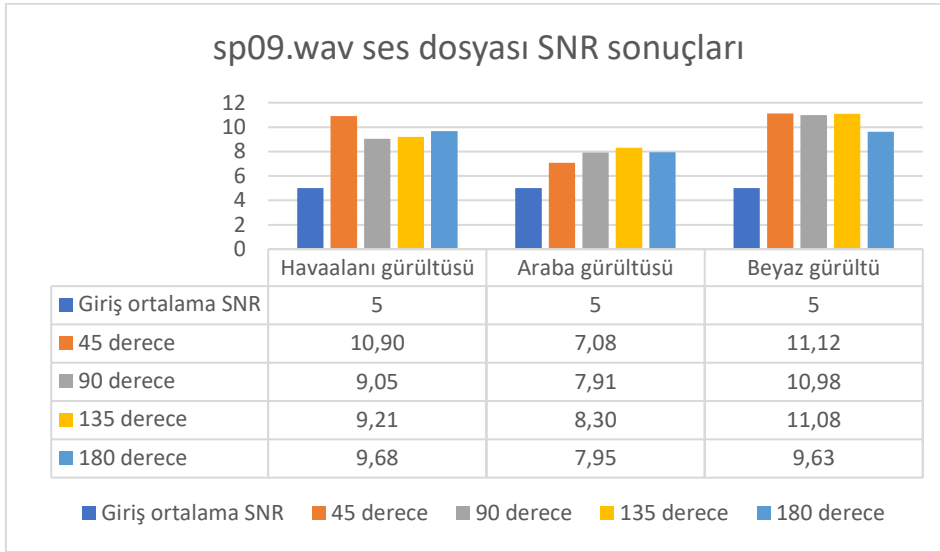
“\*”: iyileştirme yapılmayan sonuçları belirtir.

**Çizelge 4.6.** sp11.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları

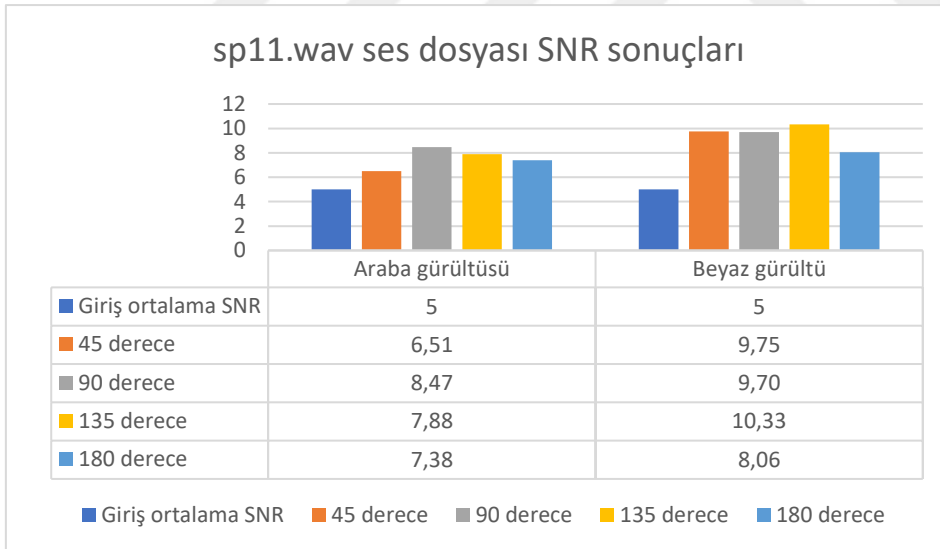
Ses		SNR Sonuçları						
		Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)			
		Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR
Sp11.wav	Havaalanı (Airport)	45	0 dB	0,24	0,90	1,23	*	*
		45	5 dB	*	5,29	5,53	*	*
		45	10 dB	*	*	*	*	*
		90	0 dB	*	*	*	*	*
		90	5 dB	*	*	*	*	*
		90	10 dB	*	*	*	*	*
		135	0 dB	2,50	3,64	4,05	2,63	*
		135	5 dB	5,99	7,13	7,73	5,53	*
		135	10 dB	10,59	11,07	11,58	*	*
		180	0 dB	3,63	4,40	4,80	1,98	2,56
		180	5 dB	8,23	8,26	8,37	6,71	6,30
		180	10 dB	13,19	13,25	13,18	12,33	10,70
	Araba (Car)	45	0 dB	1,14	4,20	4,95	1,18	2,06
		45	5 dB	5,72	7,72	8,38	5,25	6,08
		45	10 dB	10,09	11,53	12,05	*	10,81
		90	0 dB	*	4,87	5,42	4,50	4,46
		90	5 dB	6,17	8,60	9,03	7,59	7,69
		90	10 dB	10,93	12,63	13,05	11,58	12,09
		135	0 dB	1,66	5,14	5,88	4,15	4,90
		135	5 dB	5,98	8,61	9,24	7,84	8,57
		135	10 dB	*	12,27	12,93	10,77	12,35
		180	0 dB	1,39	5,42	6,16	1,51	2,62
		180	5 dB	6,68	9,22	9,74	5,39	6,50
		180	10 dB	10,55	13,19	13,72	*	11,21
	Beyaz gürültü (White Noise)	45	0 dB	5,46	7,28	7,72	4,41	4,24
		45	5 dB	9,25	10,93	11,44	7,88	8,83
		45	10 dB	13,20	14,59	15,09	12,13	13,74
		90	0 dB	5,28	7,66	8,29	3,38	3,07
		90	5 dB	9,47	11,36	11,90	8,12	7,84
		90	10 dB	13,42	15,15	15,69	12,27	12,63
		135	0 dB	5,50	7,81	8,44	4,28	5,92
		135	5 dB	9,19	11,38	12,08	8,10	10,73
		135	10 dB	13,27	14,95	15,65	12,16	15,51
		180	0 dB	1,64	6,52	7,65	1,37	4,38
		180	5 dB	6,43	10,05	11,07	5,29	8,31
		180	10 dB	10,78	13,29	14,24	*	11,82

“\*”: iyileştirme yapılamayan sonuçları belirtir.

Her ne kadar mükemmel VAD sonuçları ile değerlendirilse de iyileştirilemeyen konuşma sinyalleri gözlemlenmiştir. İyileştirilemeyen sinyallerin büyük bir bölümü sp11.wav konuşmasında ve havaalanı gürültüsünde gözlemlenmiştir.



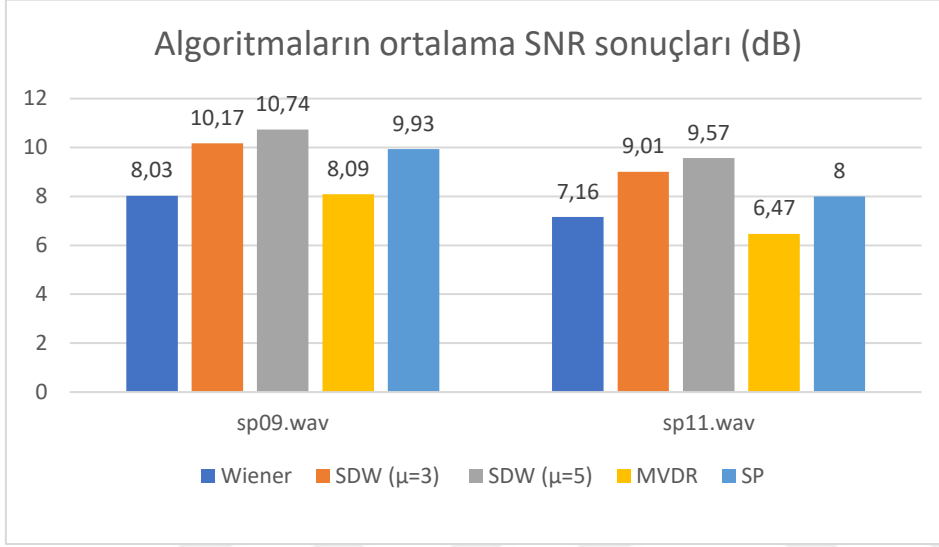
Şekil 4.10. sp09.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD ile elde edilen SNR sonuçları



Şekil 4.11. sp11.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD ile elde edilen SNR sonuçları

Şekil 4.10 ve Şekil 4.11’de mükemmel VAD sonuçları kullanılarak elde edilen SNR sonuçları gösterilmektedir. Sp11.wav ses dosyasında havaalanı gürültüsünün iyileştirilmiş sinyallerinin az olmasından dolayı Şekil 4.11’de gösterilmemiştir. Yine daha önce olduğu gibi hesaplamalarda iyileştirilmemiş sinyaller değerlendirilmeye

alınmamıştır. Mükemmel VAD sonuçları kullanılmasına rağmen özellikle sp11.wav ses dosyasında ve havaalanı gürültüsünde, 90 derecede bütün sinyallerde, 45 ve 135 derece açılarında ise bazı sinyallerde tüm algoritmalarda iyileştirilmiş sinyal alınmamıştır. Bunun dışında sonuçlar açısından değerlendirildiğinde ise önerilen algoritma ile mükemmel VAD sonuçları kullanılarak elde edilen algoritma sonuçları da yakınlık göstermiştir.



**Şekil 4.12.** sp09.wav ve sp11.wav ses sinyallerinin mükemmel VAD kullanılarak algoritmalarla elde edilen SNR sonuçları

$R_v$  ve  $R_y$  özilinti matrislerini hesaplarırken,  $R_v$  başlangıç değeri için 50 sesli çerçeve,  $R_y$  başlangıç değeri için 40 çerçeve kullanılmıştır ve VAD sonuçlarına göre bu değerler güncellenmiştir. Daha fazla ses çerçevesinin ortalaması alınmasının daha iyi sonuçlar getireceği düşünülerek stüdyo ortamında başlangıç boşluğu bir saniyeden fazla olan spkadin.wav ve sperkek.wav ses dosyaları kaydedilmiştir. Bu ses dosyalarının özilinti matrisleri hesaplanırken  $R_v$  başlangıç değeri 100 sesli çerçevenin (0,8 sn) ortalaması alınarak,  $R_y$  başlangıç değeri ise 70 sesli çerçevenin (0,56 sn) ortalaması alınarak hesaplanmıştır. Önerilen VAD ile elde edilen SNR sonuçları Çizelge 4.7 ve Çizelge 4.8’de verilmiştir.

**Çizelge 4.7.** spkadin.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları

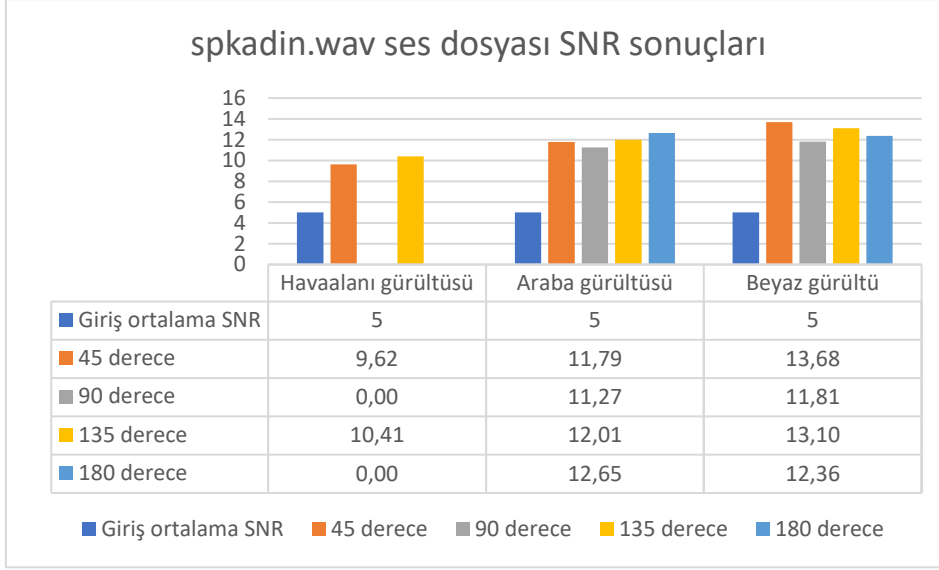
SNR Sonuçları								
Ses	Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)				
	Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR	SP
Spkadin.wav	Havaalanı (Airport)	45	0 dB	5,56	6,86	7,77	3,05	6,88
		45	5 dB	8,91	10,71	11,50	6,50	10,39
		45	10 dB	12,25	13,81	14,68	11,49	13,94
		90	0 dB	*	*	*	*	3,94
		90	5 dB	*	*	*	*	6,09
		90	10 dB	10,08	10,26	10,42	*	*
		135	0 dB	4,71	7,02	8,45	2,21	8,54
		135	5 dB	9,67	11,77	12,93	7,54	12,31
		135	10 dB	13,18	14,87	15,87	12,12	14,93
		180	0 dB	*	0,16	1,07	*	0,86
		180	5 dB	*	7,09	8,16	*	6,64
		180	10 dB	10,56	14,60	15,48	*	12,96
	Araba (Car)	45	0 dB	7,20	10,08	10,70	2,95	3,45
		45	5 dB	11,89	13,59	13,99	9,81	9,13
		45	10 dB	17,43	17,95	18,16	15,51	15,06
		90	0 dB	7,62	10,39	11,25	2,89	3,18
		90	5 dB	12,49	13,51	13,60	7,89	5,93
		90	10 dB	15,21	18,39	18,60	14,47	13,67
		135	0 dB	4,68	10,46	10,91	5,45	5,95
		135	5 dB	*	15,44	15,94	10,19	10,24
		135	10 dB	*	18,93	19,23	14,48	14,28
		180	0 dB	9,36	11,13	11,40	6,27	6,84
		180	5 dB	11,84	15,59	16,06	9,94	11,01
		180	10 dB	13,36	19,28	19,69	13,05	14,95
	Beyaz gürülyü (White Noise)	45	0 dB	10,42	13,13	13,90	7,70	8,42
		45	5 dB	14,50	16,80	17,37	12,68	12,90
		45	10 dB	15,03	16,08	16,46	14,35	15,53
		90	0 dB	8,48	11,45	12,03	5,87	7,03
		90	5 dB	10,87	12,09	12,48	9,64	10,47
		90	10 dB	15,06	15,96	16,22	14,37	15,09
		135	0 dB	10,03	11,94	12,50	5,93	6,30
		135	5 dB	15,04	16,26	16,56	10,63	11,09
		135	10 dB	16,14	16,61	16,71	15,20	15,49
		180	0 dB	8,34	11,97	12,55	5,67	7,69
		180	5 dB	13,03	15,89	16,43	9,37	11,90
		180	10 dB	14,24	15,74	16,04	12,18	14,38

“\*”: iyileştirme yapılamayan sonuçları belirtir.

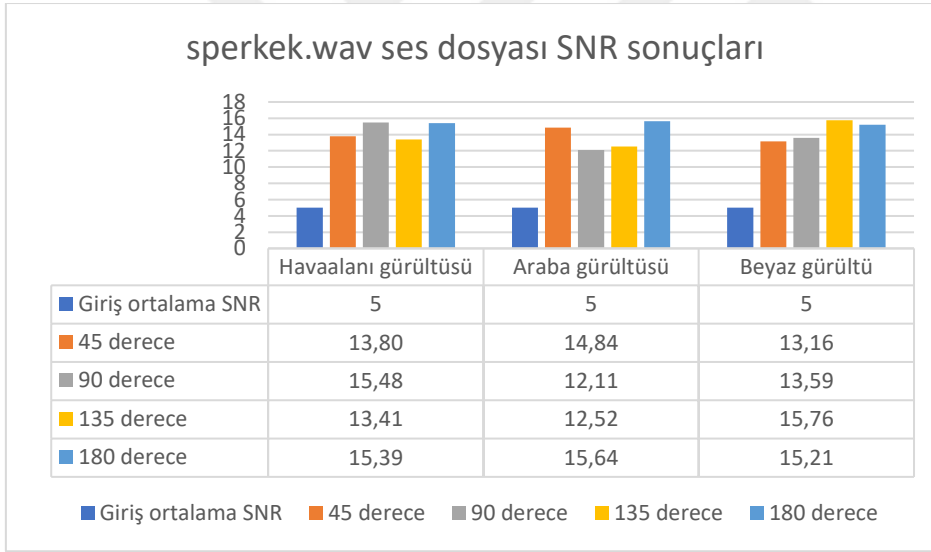
**Çizelge 4.8.** sperkek.wav ses sinyalinin önerilen VAD ile elde edilen SNR sonuçları

Ses		SNR Sonuçları						
		Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)			
		Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR
Sperkek.wav	Havaalanı (Airport)	45	0 dB	9,38	12,87	14,12	5,12	10,04
		45	5 dB	13,00	15,73	16,59	10,71	13,02
		45	10 dB	16,55	18,34	19,09	15,22	17,25
		90	0 dB	11,29	15,23	17,11	7,69	13,74
		90	5 dB	14,32	17,59	18,89	10,63	15,41
		90	10 dB	16,31	18,69	19,98	15,15	20,10
		135	0 dB	11,10	12,71	13,70	7,30	11,93
		135	5 dB	13,58	14,74	15,45	10,40	14,48
		135	10 dB	15,19	15,41	15,78	14,12	15,31
		180	0 dB	12,57	17,92	20,18	9,54	14,53
		180	5 dB	12,79	18,56	20,30	11,57	17,12
		180	10 dB	12,22	15,65	16,98	11,92	18,98
	Araba (Car)	45	0 dB	10,63	15,90	17,78	5,51	7,60
		45	5 dB	15,58	17,73	18,64	9,75	10,23
		45	10 dB	19,47	21,20	21,90	14,93	15,81
		90	0 dB	7,92	11,68	13,25	1,75	2,00
		90	5 dB	13,20	15,48	16,15	8,80	8,86
		90	10 dB	17,52	19,74	20,37	12,67	12,21
		135	0 dB	9,14	11,54	12,49	5,87	6,32
		135	5 dB	12,35	14,11	14,79	8,78	9,55
		135	10 dB	17,31	18,81	19,45	13,33	13,95
		180	0 dB	11,64	16,54	18,63	5,95	7,11
		180	5 dB	15,83	18,78	20,00	12,13	13,48
		180	10 dB	18,50	22,03	22,78	14,58	16,59
	Beyaz gürültü (White Noise)	45	0 dB	10,66	12,83	13,47	8,43	8,21
		45	5 dB	12,20	13,89	14,45	11,19	11,25
		45	10 dB	15,96	16,63	16,77	15,43	16,09
		90	0 dB	9,69	13,83	15,22	6,41	7,50
		90	5 dB	12,78	13,43	13,57	11,73	12,00
		90	10 dB	17,57	18,25	18,36	16,65	16,86
		135	0 dB	17,53	18,25	18,36	16,65	16,86
		135	5 dB	14,13	15,74	16,19	9,79	10,63
		135	10 dB	16,05	17,57	18,13	14,82	15,65
		180	0 dB	10,87	14,09	15,04	6,84	8,40
		180	5 dB	14,91	17,53	18,11	10,98	12,73
		180	10 dB	19,96	22,52	22,83	15,64	17,66

“\*”: iyileştirme yapılamayan sonuçları belirtir.



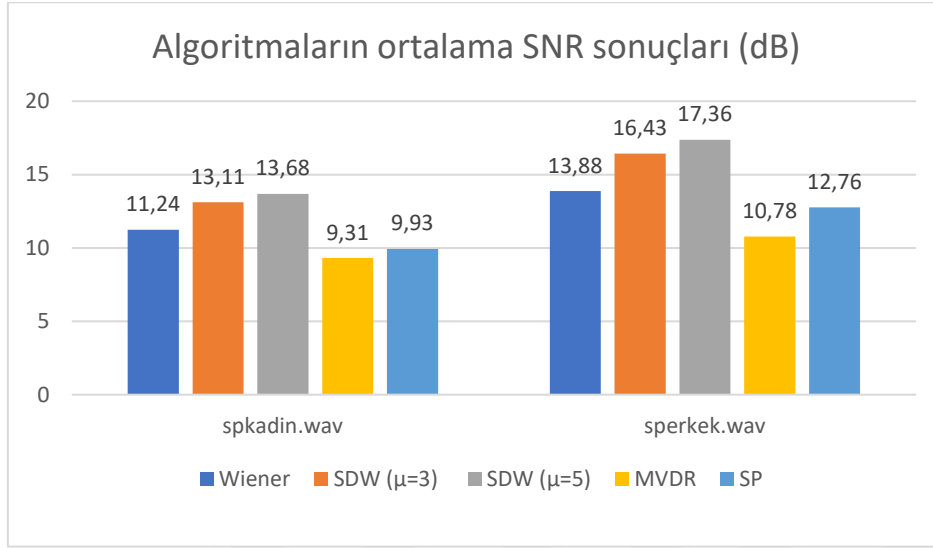
**Şekil 4.13.** spkadin.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları



**Şekil 4.14.** sperkek.wav ses sinyaline ait derece ve gürültüye göre önerilen VAD ile elde edilen SNR sonuçları

Bu ses dosyalarının değerlendirme aşamasında da iyileştirme yapılamayan algoritma değerleri dikkate alınmamıştır. Şekil 4.13 ve Şekil 4.14'te Çizelge 4.7 ve Çizelge 4.8 ile ilgili özetleyici veriler gösterilmiştir. spkadin.wav ve sperkek.wav ses dosyalarının SNR sonuçlarının açığa ve gürültü tipine göre değişimi incelendiğinde spkadin.wav ses dosyasının özellikle havaalanı gürültüsünün 90 ve 180 derecelik açılarında iyileştirme yapılamamıştır. Bunun dışında spkadin.wav ve sperkek.wav ses

dosyalarının tamamında iyileştirme olup, iyileştirilen sinyallerin SNR değeri sp09.wav ve sp11.wav ses dosyalarından daha iyi sonuçlar verdiği tespit edilmiştir. SNR açısından değerlendirilirse; her iki konuşma sinyali için de SDW ( $\mu=5$ ) algoritması tarafından en iyi sonuçlar elde edilmiştir. Algoritmanın tipine göre her iki konuşma için elde edilen sonuçlar ise Şekil 4.15'te verilmiştir.



**Şekil 4.15.** spkadin.wav ve sperkek.wav ses sinyallerinin önerilen VAD kullanılarak algoritmalara göre elde edilen SNR sonuçları

Özilinti matrislerinin başlangıç değerlerinin daha doğru hesaplanabilmesi için fazla ses çerçevesinin ortalaması alınması işlemi ile belirgin derecede olumlu etki gözlemlenmiştir. Daha önce hesaplandığı gibi önerilen VAD algoritmasının spkadin.wav ve sperkek.wav ses dosyalarındaki etkisinin incelenmesi için, bu iki ses dosyasının mükemmel VAD sonuçları ile değerlendirmeleri de yapılmıştır. spkadin.wav ve sperkek.wav ses dosyalarının mükemmel VAD vektörü ile elde edilen SNR sonuçları Çizelge 4.9 ve Çizelge 4.10'da gösterilmiştir.



**Çizelge 4.9.** spkadin.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları

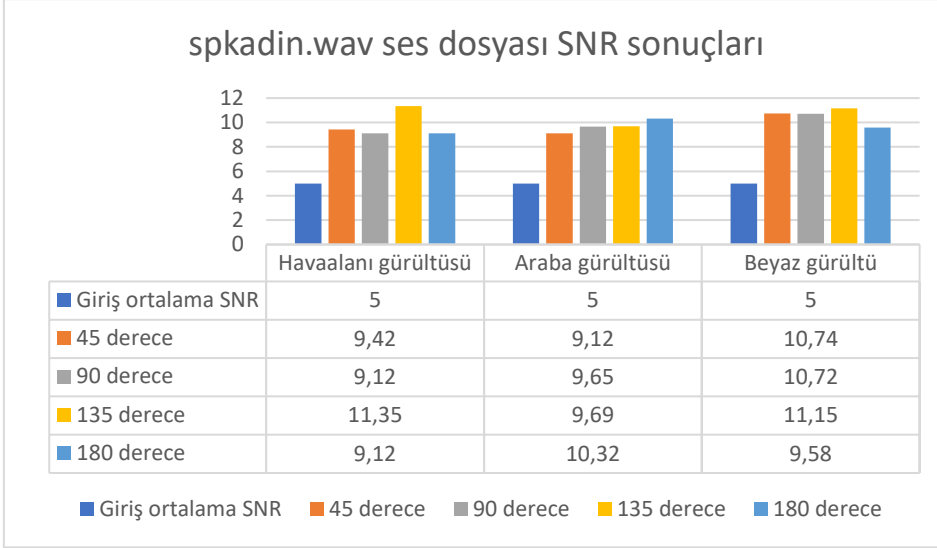
SNR Sonuçları								
Ses	Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)				
	Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR	SP
Spkadin.wav	Havaalanı (Airport)	45	0 dB	3,78	1,03	*	*	*
		45	5 dB	10,05	8,25	7,60	8,90	5,76
		45	10 dB	14,79	13,39	12,93	15,03	11,47
		90	0 dB	3,80	*	*	*	*
		90	5 dB	7,38	6,75	5,77	*	*
		90	10 dB	14,12	11,91	11,45	11,75	*
		135	0 dB	10,67	7,35	6,38	7,43	3,27
		135	5 dB	14,71	10,98	10,04	13,26	7,49
		135	10 dB	18,37	15,35	14,59	18,03	12,35
		180	0 dB	7,07	3,55	1,22	2,01	*
		180	5 dB	12,01	9,63	8,45	7,52	*
		180	10 dB	16,24	13,63	12,58	15,58	*
	Araba (Car)	45	0 dB	5,50	5,15	4,95	3,81	2,57
		45	5 dB	9,75	9,67	9,73	8,98	8,18
		45	10 dB	13,97	13,94	14,02	13,50	13,13
		90	0 dB	6,04	6,24	6,21	4,21	3,60
		90	5 dB	9,87	10,20	10,38	8,88	8,76
		90	10 dB	14,25	14,24	14,40	13,70	13,71
		135	0 dB	5,63	5,47	5,39	4,52	3,75
		135	5 dB	10,32	10,25	10,25	9,52	8,92
		135	10 dB	14,65	14,73	14,76	13,82	13,44
		180	0 dB	6,25	5,86	5,62	5,34	4,32
		180	5 dB	11,37	11,08	10,88	10,06	8,97
		180	10 dB	15,70	15,92	15,78	14,11	13,55
	Beyaz gürültü (White Noise)	45	0 dB	5,60	7,09	7,52	4,81	5,79
		45	5 dB	10,40	11,49	11,83	9,68	10,43
		45	10 dB	14,94	15,83	16,12	14,40	15,17
		90	0 dB	5,69	6,76	6,95	5,01	5,42
		90	5 dB	10,67	11,47	11,65	9,94	10,19
		90	10 dB	15,29	15,96	16,13	14,77	14,94
		135	0 dB	6,58	6,69	6,67	5,66	5,40
		135	5 dB	11,60	11,66	11,69	10,89	10,54
		135	10 dB	16,13	16,32	16,39	15,55	15,41
		180	0 dB	3,60	5,66	5,94	4,04	5,40
		180	5 dB	8,75	10,71	11,05	7,98	9,79
		180	10 dB	13,27	15,12	15,59	12,30	14,50

“\*”: iyileştirme yapılamayan sonuçları belirtir.

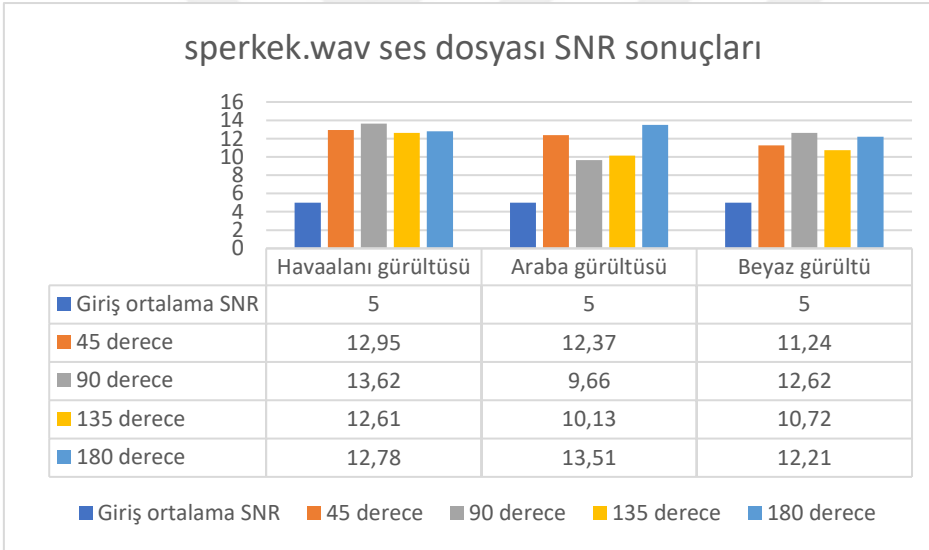
**Çizelge 4.10.** sperkek.wav ses sinyalinin mükemmel VAD ile elde edilen SNR sonuçları

SNR Sonuçları								
Ses	Gürültünün		Giriş SNR Değerleri	ÇIKIŞ SNR Değerleri (dB)				
	Tipi	Açısı		Wiener	SDW ( $\mu=3$ )	SDW ( $\mu=5$ )	MVDR	SP
Sperkek.wav	Havaalanı (Airport)	45	0 dB	8,31	9,77	10,53	7,59	9,77
		45	5 dB	12,09	13,38	14,07	11,51	12,90
		45	10 dB	16,19	17,31	17,87	15,83	17,10
		90	0 dB	9,02	10,71	11,18	8,68	10,82
		90	5 dB	12,42	13,92	14,42	11,92	14,04
		90	10 dB	16,44	17,85	18,39	15,88	18,62
		135	0 dB	8,12	8,96	9,53	7,04	9,50
		135	5 dB	12,13	12,81	13,26	11,30	13,25
		135	10 dB	16,49	17,00	17,29	15,87	16,56
		180	0 dB	7,14	9,17	10,37	5,33	10,09
		180	5 dB	11,53	13,24	14,31	10,35	14,16
		180	10 dB	16,30	17,64	18,47	15,47	18,12
	Araba (Car)	45	0 dB	9,11	8,86	8,94	7,07	5,37
		45	5 dB	13,32	13,09	13,16	12,08	10,81
		45	10 dB	17,21	17,20	17,28	16,41	15,70
		90	0 dB	5,88	5,81	5,98	3,71	2,41
		90	5 dB	10,53	10,34	10,47	9,01	7,97
		90	10 dB	15,20	15,03	15,09	14,17	13,24
		135	0 dB	8,12	5,62	5,28	6,33	2,24
		135	5 dB	12,28	10,30	10,05	10,96	7,77
		135	10 dB	16,15	14,96	14,86	14,84	12,25
		180	0 dB	9,83	9,25	9,35	8,30	6,63
		180	5 dB	14,63	13,86	13,72	13,11	14,33
		180	10 dB	19,16	18,86	18,60	17,05	15,91
	Beyaz gürültü (White Noise)	45	0 dB	5,75	6,88	7,16	5,10	5,97
		45	5 dB	11,01	11,92	12,14	10,51	11,33
		45	10 dB	15,69	16,68	16,88	15,30	16,31
		90	0 dB	7,89	8,55	8,65	7,02	6,87
		90	5 dB	12,73	13,32	13,39	12,00	11,98
		90	10 dB	17,31	17,95	18,06	16,71	16,82
		135	0 dB	5,55	6,11	6,24	4,75	5,18
		135	5 dB	10,83	11,28	11,38	10,10	10,45
		135	10 dB	15,75	16,27	16,35	15,09	15,47
		180	0 dB	7,35	7,97	8,01	6,48	6,40
		180	5 dB	12,43	13,03	13,07	11,37	11,32
		180	10 dB	17,35	18,02	18,08	16,14	16,20

“\*”: iyileştirme yapılamayan sonuçları belirtir.



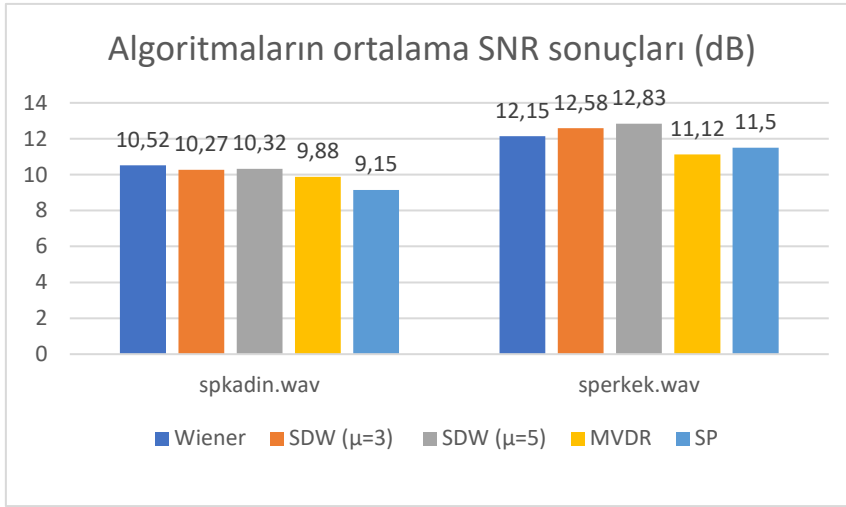
Şekil 4.16. spkadin.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD kullanılarak elde edilen SNR sonuçları



Şekil 4.17. sperkek.wav ses sinyaline ait derece ve gürültüye göre mükemmel VAD kullanılarak elde edilen SNR sonuçları

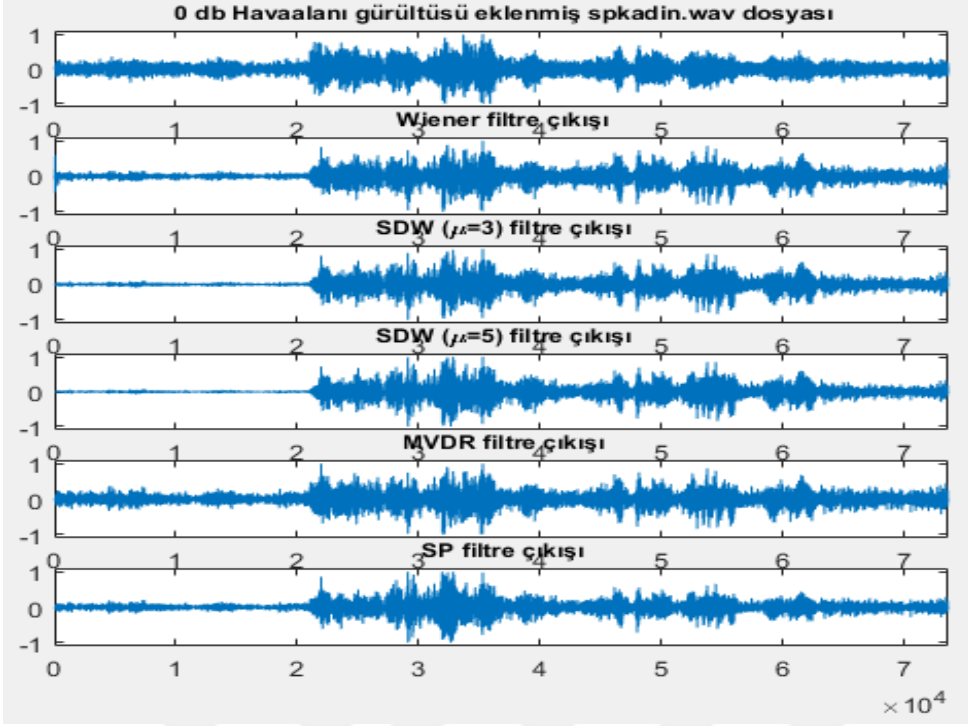
Şekil 4.16 ve Şekil 4.17’de Çizelge 4.9 ve Çizelge 4.10’a ait veriler sunulmuştur. Bu veriler dikkate alındığında spkadin.wav konuşmasında özellikle SP algoritmasıyla havaalanı gürültüsünde iyileştirilemeyen sinyaller olmuştur ve onun dışında az miktarda diğer algoritmalarda da iyileştirilemeyen sinyal olmuştur. Araba gürültüsü ve beyaz gürültüde ise bütün sinyallerde iyileştirme sağlanmıştır. sperkek.wav ses dosyasında ise bütün ses dosyalarında iyileştirme sağlanmıştır. Açılara göre değerlendirildiğinde ise

önemli bir fark gözlemlenmemiştir. Mükemmel VAD sonuçlarıyla elde edilmiş sp09.wav ve sp11.wav ses dosyalarının SNR sonuçlarıyla karşılaştırıldığında da SNR sonuçlarının daha iyi olduğu görülmüştür. SNR açısından değerlendirilirse; spkadin.wav ses dosyasında elde edilen sonuçların ortalaması alındığında, Wiener algoritmasıyla az bir farkla daha iyi sonuçlar elde edilmiştir. sperkek.wav konuşma sinyali için de SDW ( $\mu=5$ ) algoritmasıyla yine az bir farkla daha iyi sonuçlar elde edilmiştir. Algoritmanın tipine göre her iki konuşma için elde edilen sonuçlar ise Şekil 4.18’de verilmiştir.

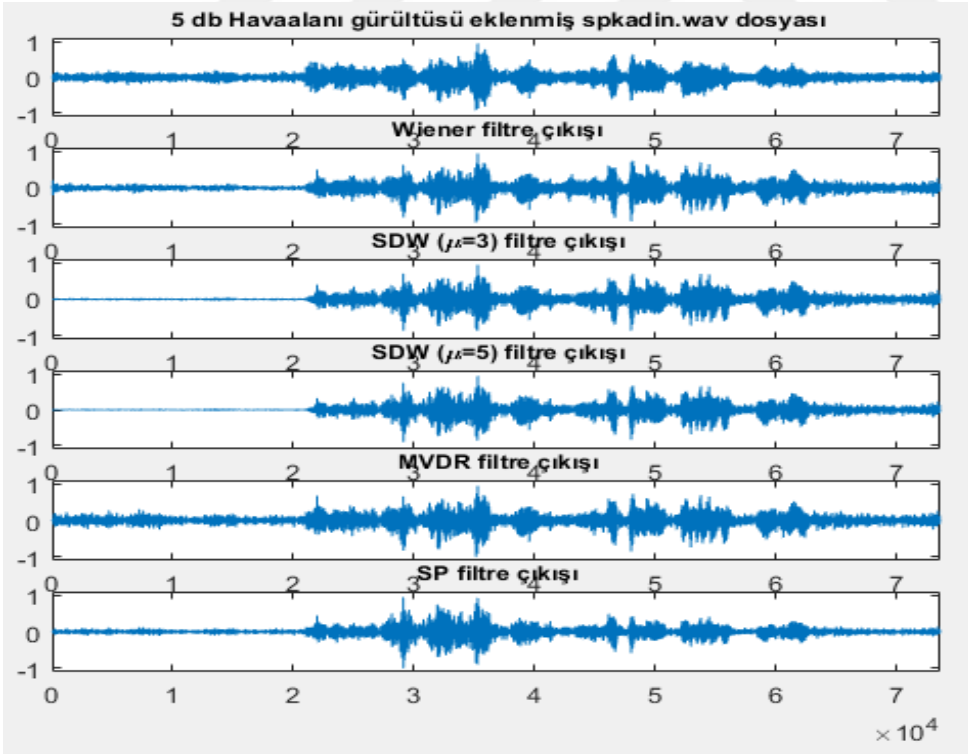


**Şekil 4.18.** spkadin.wav ve sperkek.wav ses sinyallerinin önerilen VAD kullanılarak algoritmalara göre elde edilen SNR sonuçları

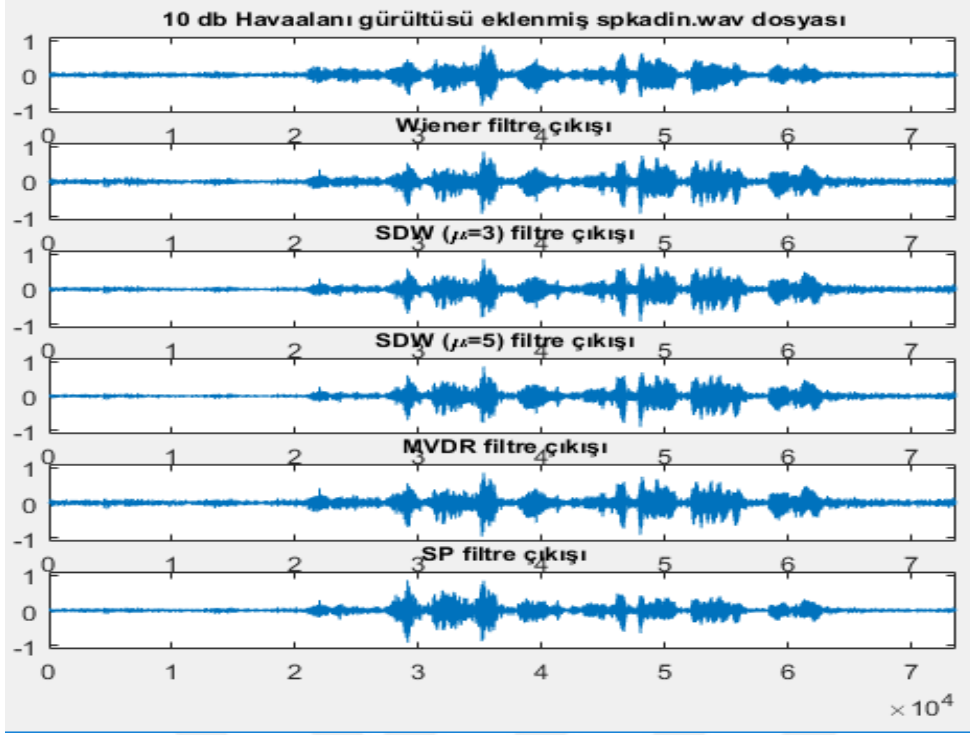
spkadin.wav ve sperkek.wav ses dosyalarının algoritmalara göre iyileştirilme görselleri ile örnekler Şekil 4.19’dan şekil 4.36’ya kadar gösterilmiştir. Her görselde altı tane sinyal bulunmaktadır. Bu sinyallerden; birinci sıradaki giriş sinyalini, ikinci sıradaki Wiener filtre uygulanmış çıkış sinyalini, üçüncü sıradaki SDW ( $\mu=3$ ) filtre uygulanmış çıkış sinyalini, dördüncü sıradaki SDW ( $\mu=5$ ) filtre uygulanmış çıkış sinyalini, beşinci sıradaki MVDR filtre uygulanmış çıkış sinyalini ve altıncı sıradaki sinyal ise SP filtre ile elde edilmiş çıkış sinyalini ifade eder.



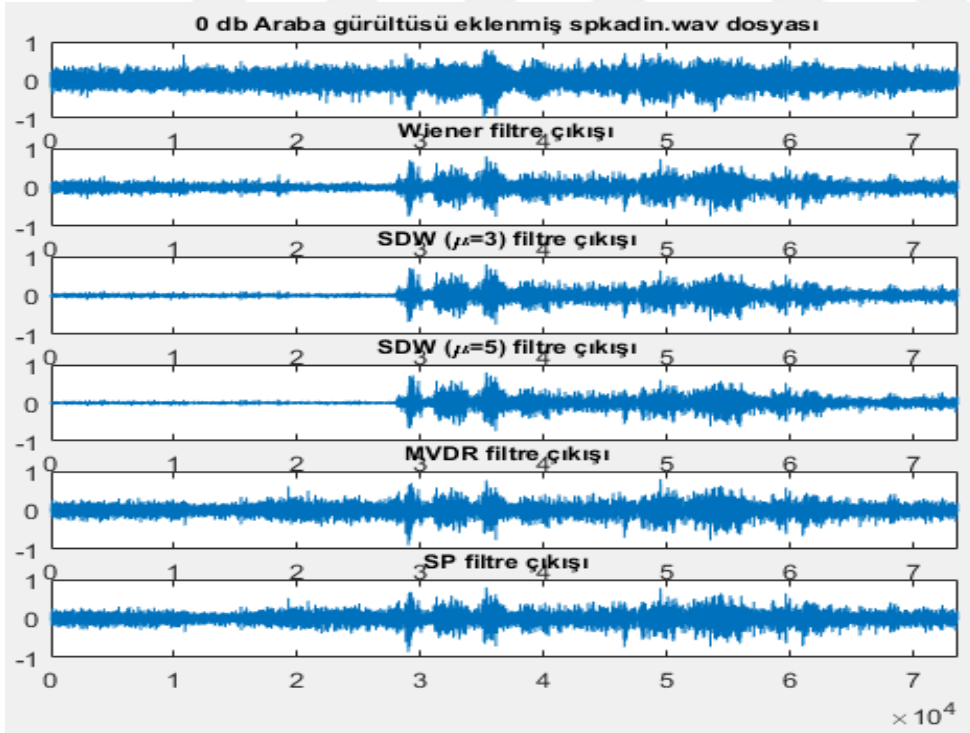
Şekil 4.19. 0 dB Havaalanı gürültüsü eklenmiş spkadin.wav dosyası ve filtrenmiş çıkışlarının gösterimi



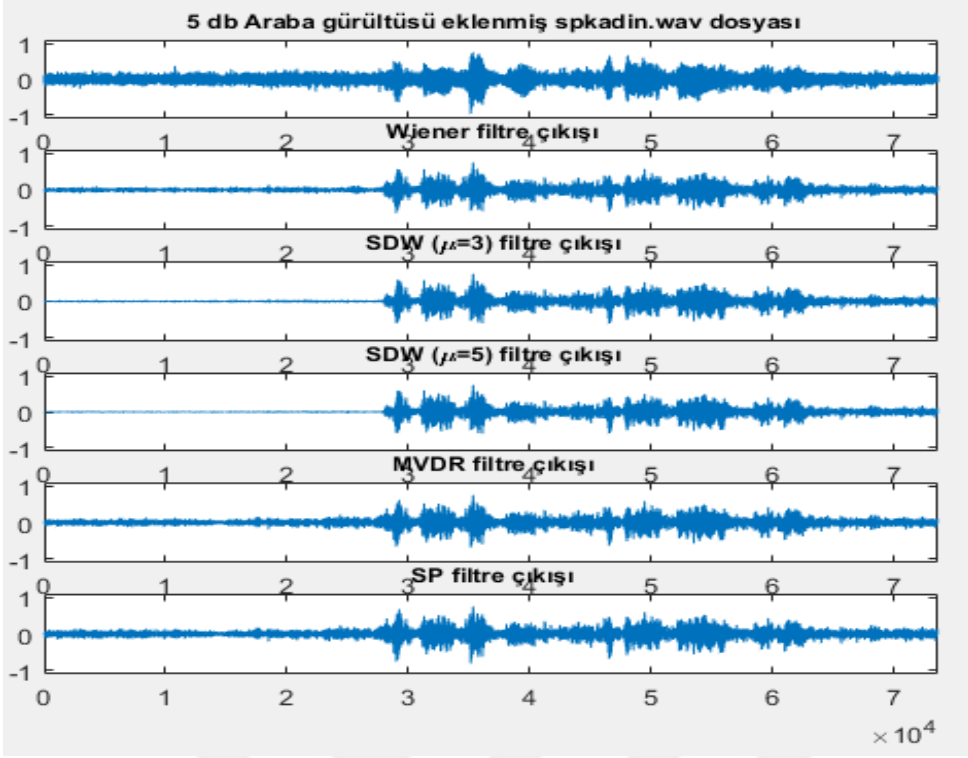
Şekil 4.20. 5 dB Havaalanı gürültüsü eklenmiş spkadin.wav dosyası ve filtrenmiş çıkışlarının gösterimi



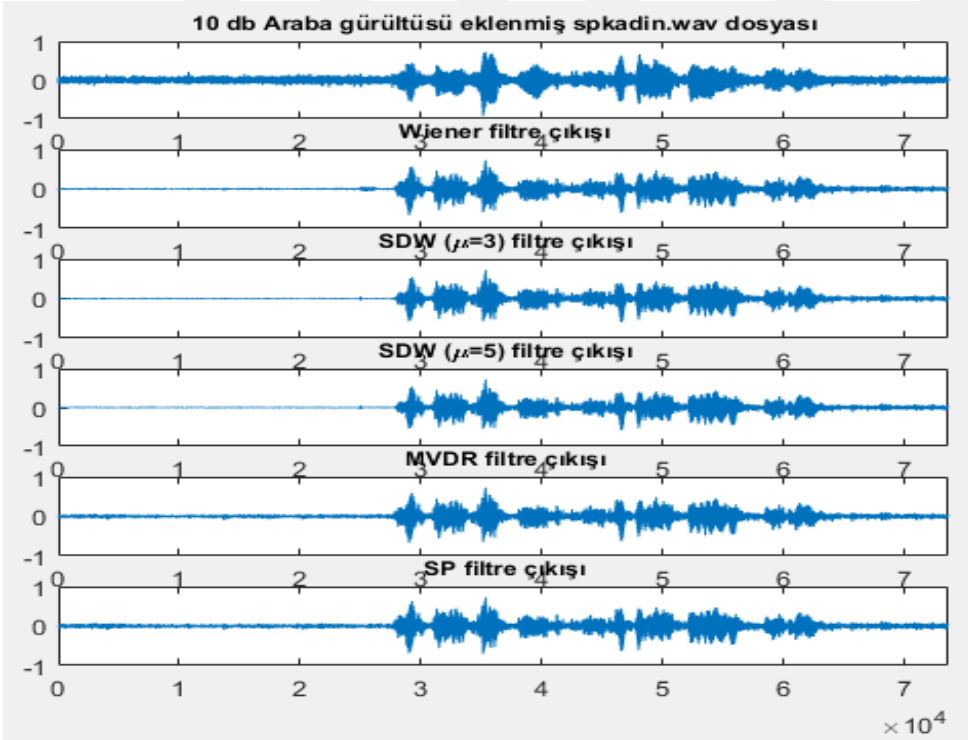
Şekil 4.21. 10 dB Havaalanı gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



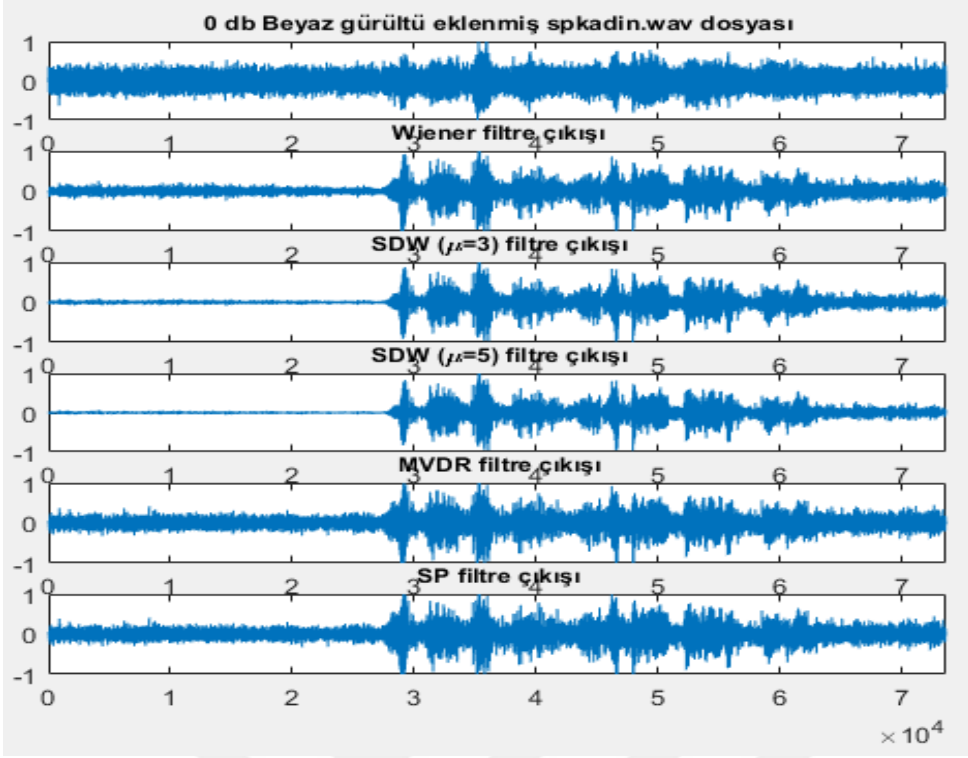
Şekil 4.22. 0 dB Araba gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



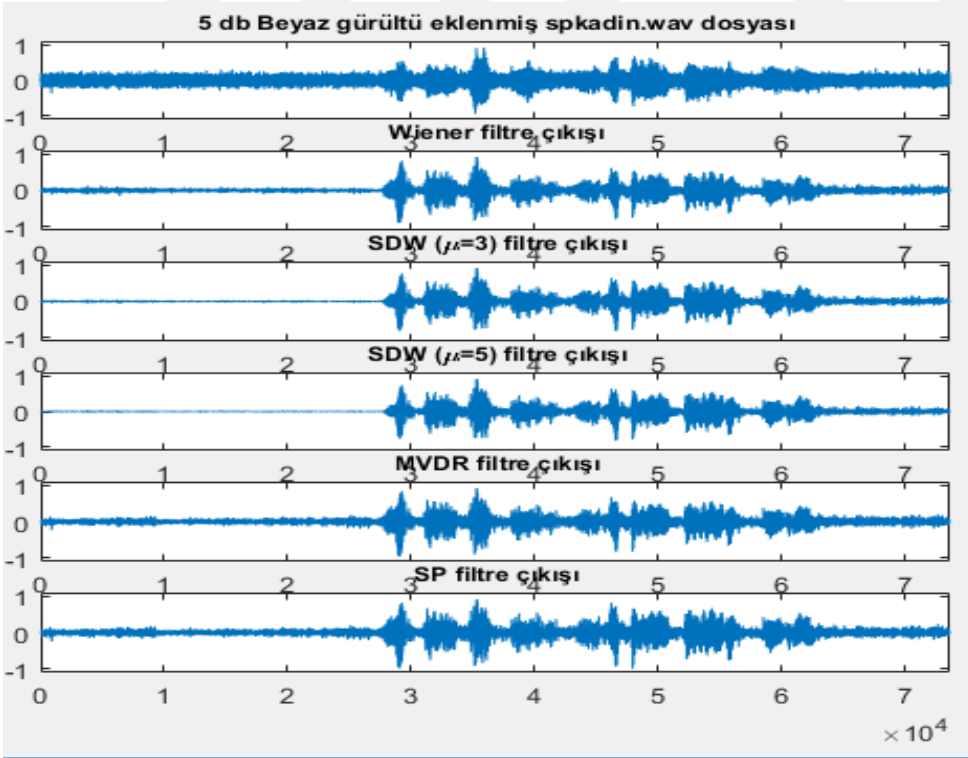
Şekil 4.23. 5 dB Araba gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



Şekil 4.24. 10 dB Araba gürültüsü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi

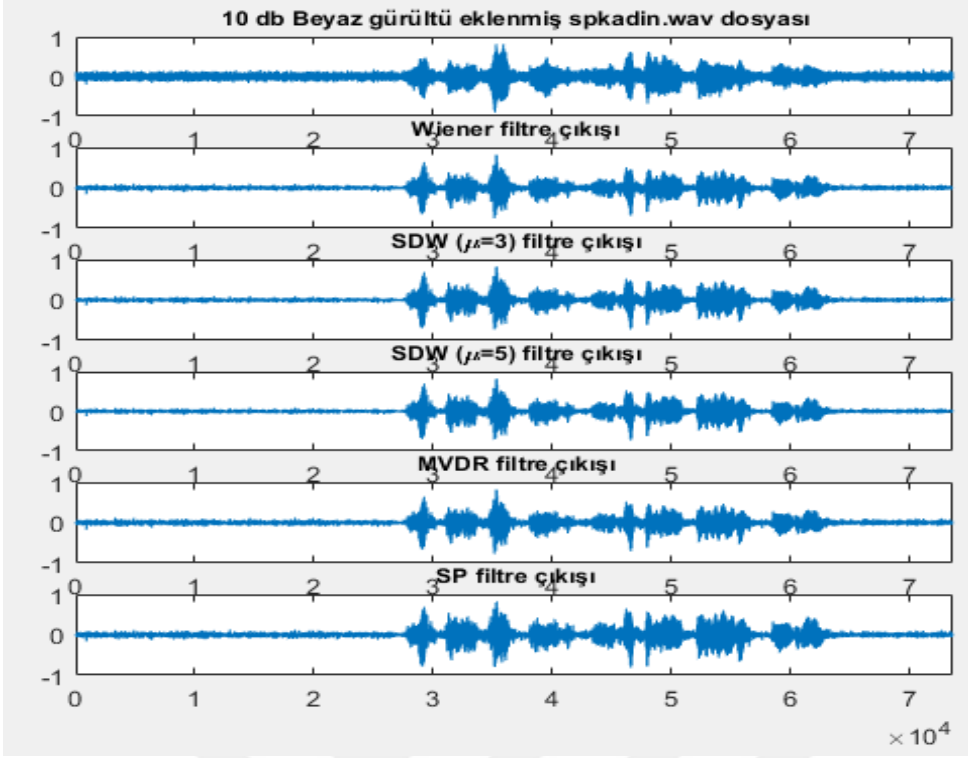


Şekil 4.25. 0 dB Beyaz gürültü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi

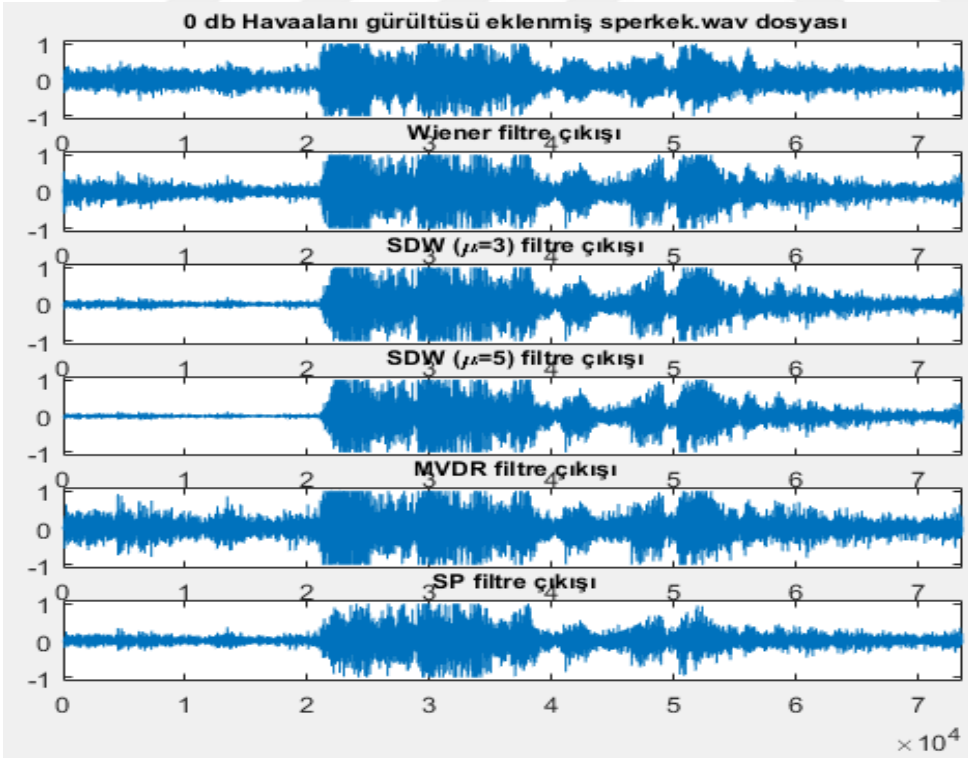


Şekil 4.26. 5 dB Beyaz gürültü eklenmiş spkadin.wav dosyası ve filtrelenmiş çıkışlarının gösterimi

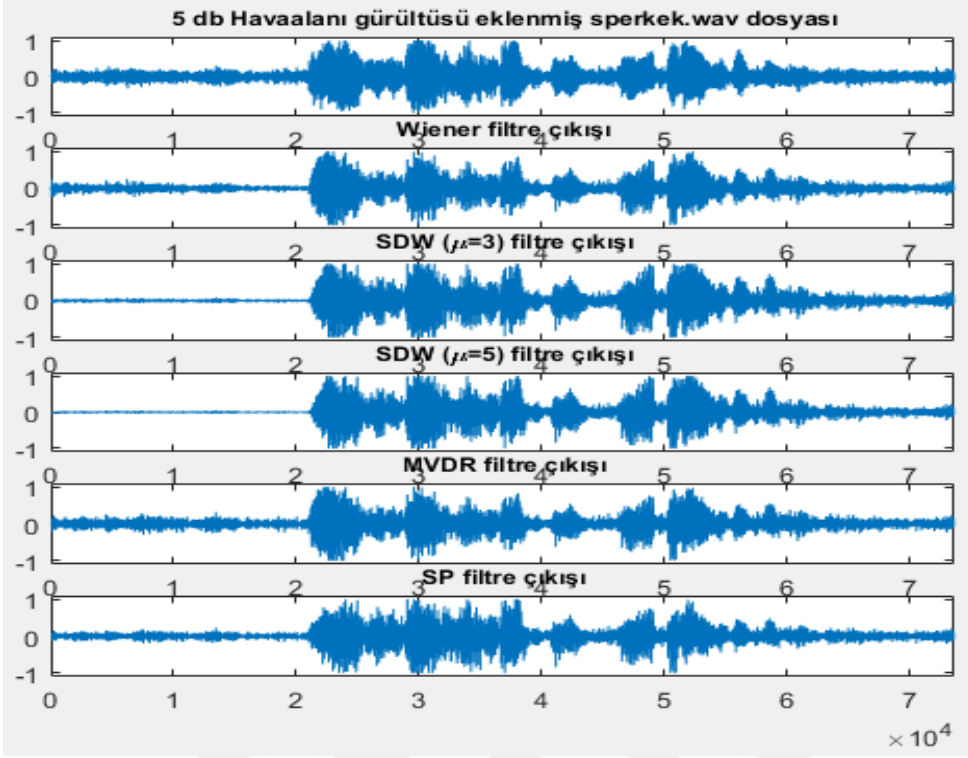




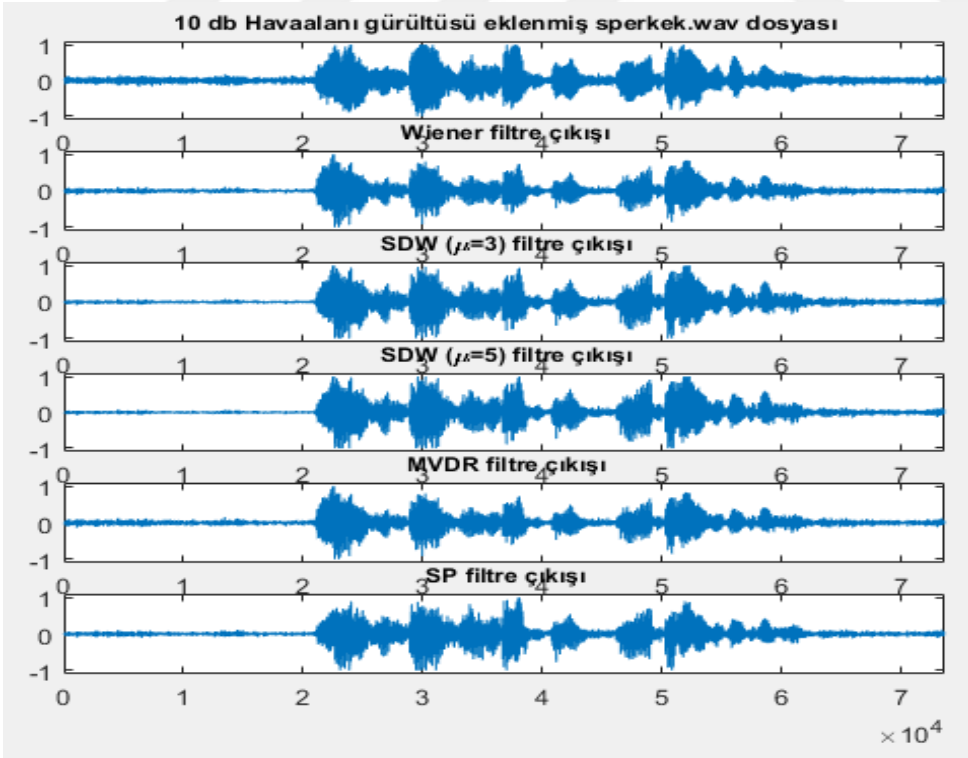
Şekil 4.27. 10 dB Beyaz gürültü eklenmiş spkadin.wav dosyası ve filtrenmiş çıkışlarının gösterimi



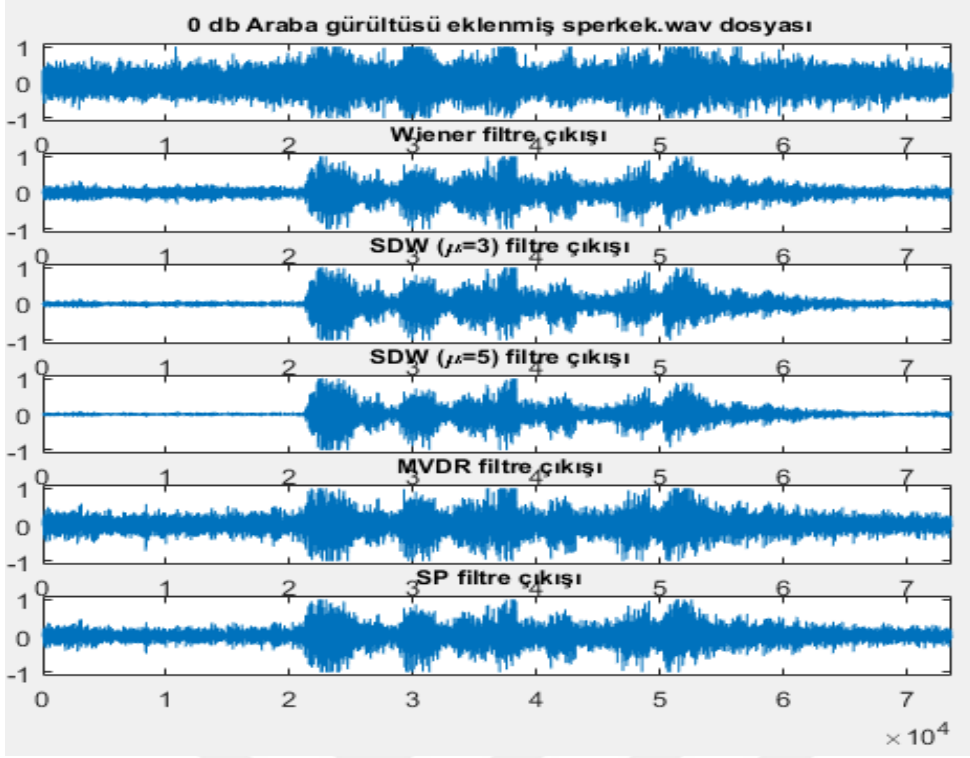
Şekil 4.28. 0 dB Havaalanı gürültüsü eklenmiş sperkek.wav dosyası ve filtrenmiş çıkışlarının gösterimi



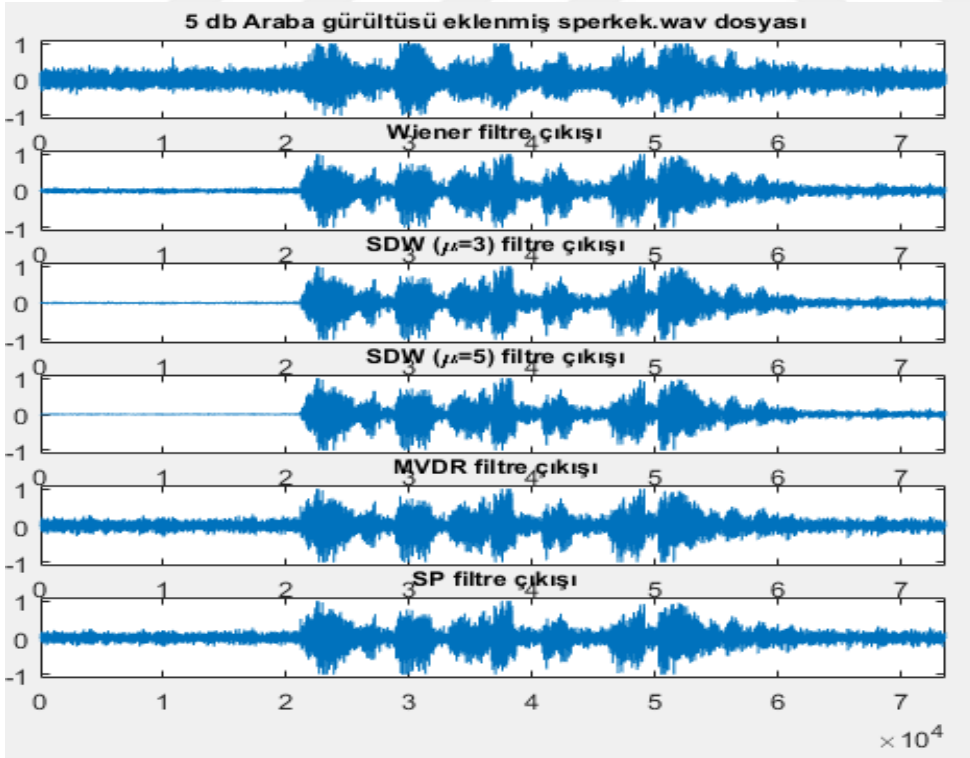
Şekil 4.29. 5 dB Havaalanı gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



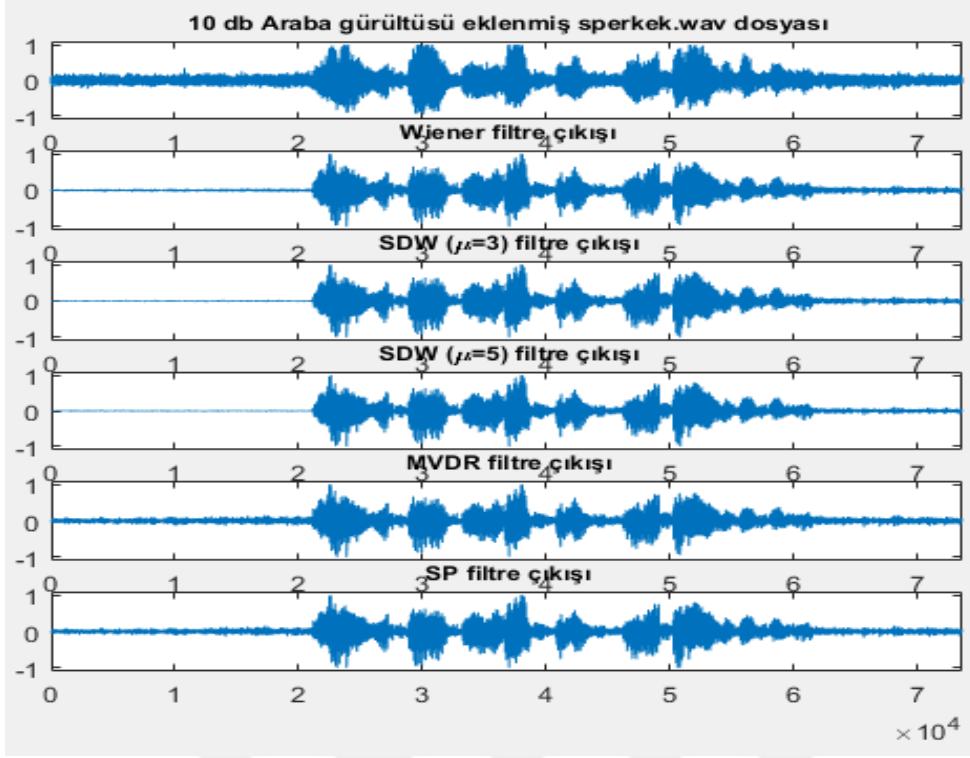
Şekil 4.30. 10 dB Havaalanı gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



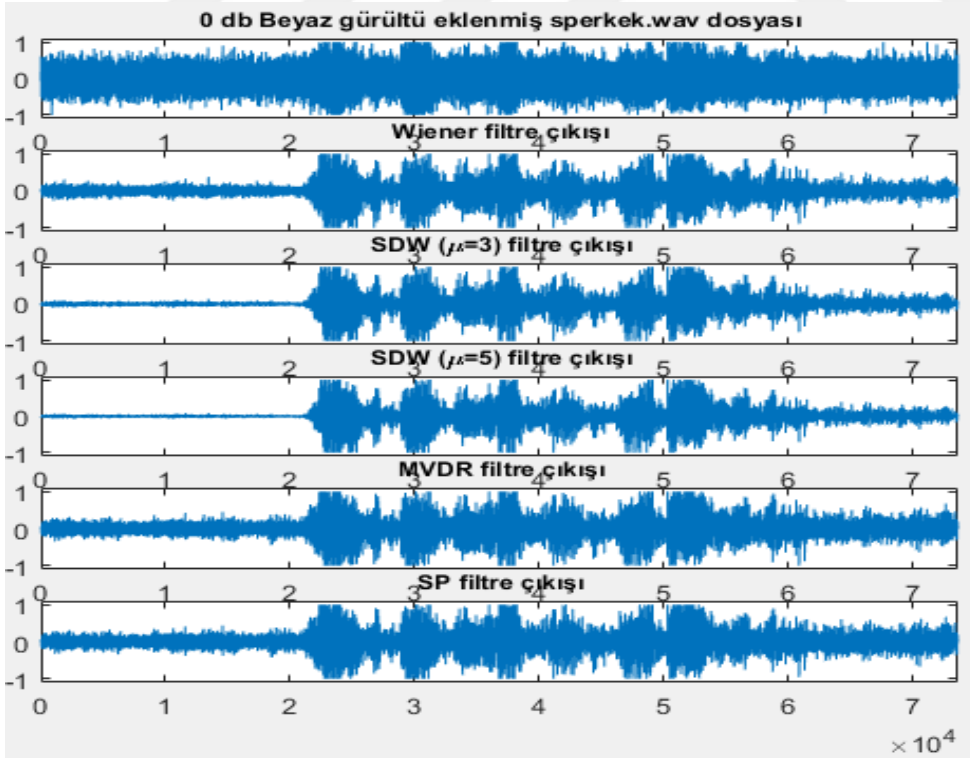
Şekil 4.31. 0 dB Araba gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



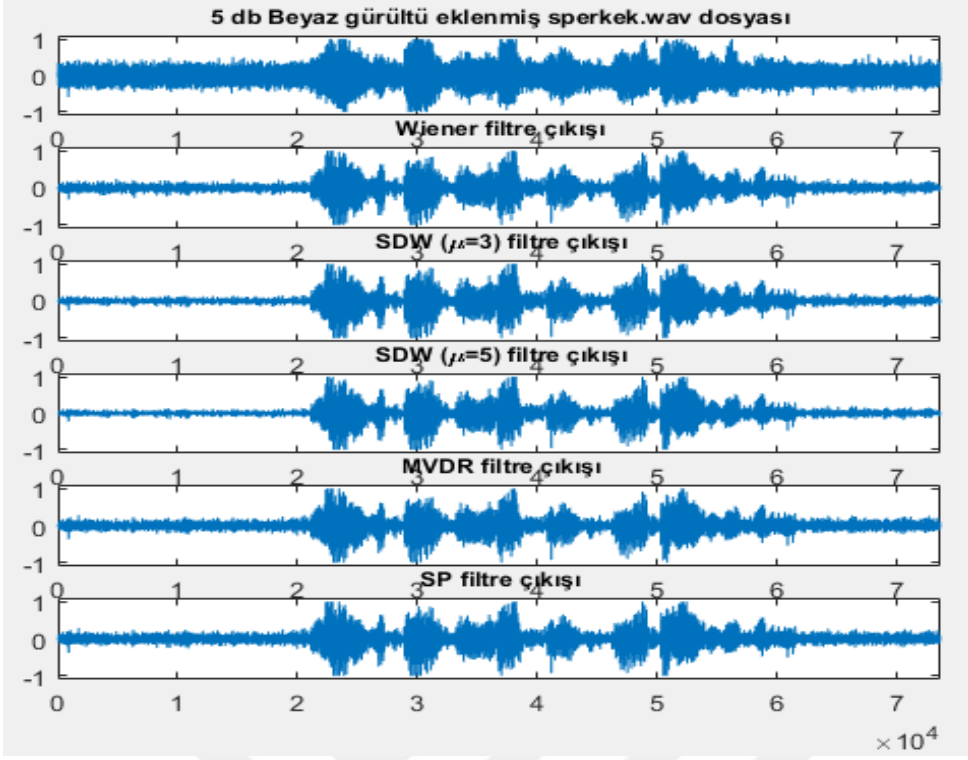
Şekil 4.32. 5 dB Araba gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



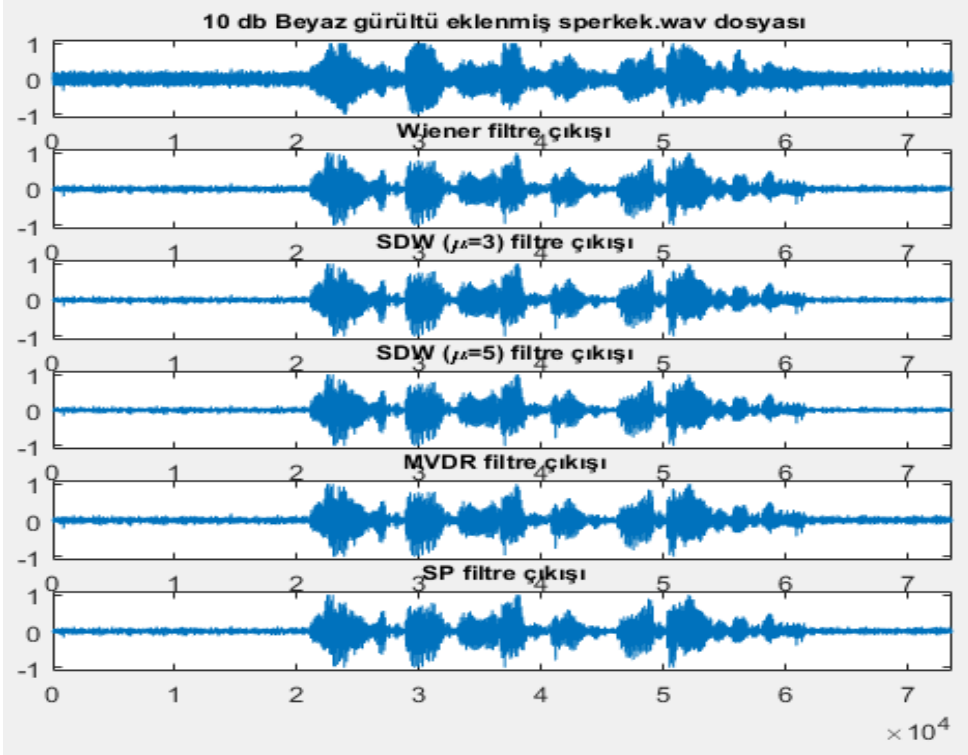
Şekil 4.33. 10 dB Araba gürültüsü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



Şekil 4.34. 0 dB Beyaz gürültü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



Şekil 4.35. 5 dB Beyaz gürültü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi



Şekil 4.36. 10 dB Beyaz gürültü eklenmiş sperkek.wav dosyası ve filtrelenmiş çıkışlarının gösterimi

### 4.3 MOS Sonuçları

spkadin.wav ve sperkek.wav ses dosyalarına; havaalanı gürültüsü, araba gürültüsü ve beyaz gürültünün 45 derecelik açıyla eklenmiş versiyonu TRT’de profesyonel olarak ses alanında çalışan 9 kişiye dinlettirilmiştir. Çizelge 2.1’de verilen MOS değerlendirme tablosu dikkate alınarak 9 kişinin yaptığı değerlendirme ile elde edilen MOS sonuçlarının ortalaması Çizelge 4.11’de gösterilmiştir.

**Çizelge 4.11.** sperkek.wav ve spkadin.wav ses sinyallerinin MOS sonuçları

MOS Değerlendirme Sonuçları					
Ses	Gürültünün tipi	Wiener	SDW ( $\mu=3$ )	MVDR	SP
Sperkek.wav	Havaalanı (Airport)	2,90	3,47	2,97	2,77
	Araba (Car)	2,71	2,88	2,74	2,59
	Beyaz Gürültü (White Noise)	2,81	2,77	2,68	2,70
spkadin.wav	Havaalanı (Airport)	3,25	3,56	3,20	2,78
	Araba (Car)	3,10	3,29	3,10	2,95
	Beyaz Gürültü (White Noise)	2,81	2,79	2,73	2,61

Çizelge 4.11’e göre spkadin.wav ses dosyası ile havaalanı ve araba gürültülerinde sperkek.wav ses dosyasına göre daha iyi sonuçlar elde edilmiştir. Beyaz gürültüde ise her iki ses dosyasında da hemen hemen aynı sonuçlar elde edilmiştir.

## 5. TARTIŞMA ve SONUÇ

İletişim sektöründe konuşma kalitesini etkileyen en temel sorunlardan biri de gürültüdür. Bu gürültü problemi ile başa çıkabilmek için literatürde yapılan birçok çalışma vardır. Bu tez çalışmasında kolay uygulanabilir bir VAD algoritması geliştirilip, literatürde bulunan temel adaptif filtre algoritmaları denenmiştir.

VAD algoritması konuşma iyileştirme algoritmalarında çok önemli bir yer tutmaktadır. Ses sinyallerinde gürültülü konuşma ve gürültünün doğru tespit edilmesi doğru filtre katsayılarının hesaplanabilmesi için çok önemlidir. Önerilen VAD algoritması oluşturulurken STE, SF ve periyodiklik gibi üç temel özellik kullanılmıştır. Geliştirilen algoritma dört farklı açıdan eklenen; havaalanı, araba gürültüsü ve beyaz gürültü ile denenmiştir. Gürültülü konuşmanın tespiti %90'ın üzerinde doğru sonuçlar vermekle birlikte sadece gürültülü bölgelerin tespitinde ortalama %70 civarında doğru tespit yapılmıştır. Bu VAD algoritmasının uygulamadaki verimliliğini tespit etmek için dört farklı konuşma sinyaline hem önerilen VAD ile hem de mükemmel VAD kullanılarak, konuşma iyileştirme algoritmaları uygulanmıştır. Konuşma iyileştirme algoritmalarıyla elde edilen SNR değerleri kıyaslandığında önerilen VAD algoritmasının oldukça iyi sonuçlar verdiği gözlemlenmiştir.

Adaptif filtre olarak Wiener, SDW, MVDR ve SP algoritmaları kullanılmıştır. SDW algoritmasında  $\mu$  parametresinin değişimiyle gürültünün azaltılması arasındaki ilişkinin tespit edilmesi için, her ses sinyalinin iyileştirilmesinde  $\mu=3$  ve  $\mu=5$  olarak alınmıştır. Uygulanan adaptif filtre algoritmalarının sonuçları gözlemlendiğinde, bu sonuçları etkileyen bazı faktörler vardır. Bu faktörler dış faktörler ve iç faktörler olmak üzere ikiye ayrılabilir:

**Dış faktörler:** Konuşma iyileştirme algoritmalarında; ses sinyallerine eklenen gürültünün tipi, ses sinyallerinin ait olduğu kişilerin cinsiyeti sonuçları etkileyen dış faktörlerdir. Gürültünün frekans karakteristiği ile konuşmanın frekans karakteristiği birbirine ne kadar yakın olursa, algoritmanın performansı o kadar düşmektedir. Gürültü ve konuşma sinyallerinin frekans karakteristiği farklılaştığında ise algoritma sonuçları daha iyi sonuçlar vermektedir. Bu sebeple tüm algoritmalarda beyaz gürültü ile en iyi sonuçlar elde edilmiştir. Bunun dışında havaalanı ve kadın seslerinin üst frekansları daha fazla olduğundan tüm algoritmalarda sp11.wav ve spkadin.wav kadın ses sinyallerinin iyileştirmeleri konusunda problemler yaşanmakla birlikte ya çok az iyileştirme olmuş ya

da iyileştirme yapılamamıştır. Oysa aynı gürültü erkek seslerine eklenip konuşma iyileştirme algoritma algoritmalarıyla denendiğinde oldukça iyi sonuçlar elde edilmiştir.

**İç faktörler:** Algoritmaların kendisinden kaynaklanan faktörlerden ilki unutm faktörleridir.  $R_v$  ve  $R_y$  özilinti matrislerinin hesaplanmasında kullanılan bu katsayı değerleri ( $\alpha_n$  ve  $\alpha_y$ ) filtre katsayılarını doğrudan etkiler. Bu sabit değerler en baştan belirlenmiştir. Bu değerlerin bire yakın olması filtre katsayılarının daha yavaş değişmesi, sıfıra yakın olması filtre katsayılarının daha çabuk değişmesi anlamına gelmektedir. Bu tez çalışmasında tüm algoritmalarda unutm faktörleri 0,99 olarak alınmıştır.

Bir diğer önemli faktör de özilinti matrislerinin hesaplanmaya başlanmasından önce  $R_v$  ve  $R_y$  özilinti matrislerinin başlangıç değerlerinin belirlenmesidir. sp09.wav ve sp11.wav ses dosyaları ile uygulamalar yapılırken;  $R_v$  için başlangıç değeri 50 tane gürültülü ses çerçevesinin ortalaması alınarak tespit edilmiştir.  $R_y$  için ise 40 tane gürültülü konuşma çerçevesi kullanılarak başlangıç değeri tespit edilmiştir. Karşılaştırma yapılabilmesi için spkadin.wav ve sperkek.wav ses dosyaları için özilinti matrisleri hesaplanırken bu değerler  $R_v$  için 100,  $R_y$  için 70 olarak alınmıştır. Sonuçlar karşılaştırıldığında spkadin.wav ve sperkek.wav ses dosyaları ile sp09.wav ve sp11.wav ses dosyalarına göre daha iyi sonuçlar elde edilmiştir.

Sonuç olarak bu tez çalışmasında önerilen VAD algoritmasının adaptif filtre uygulamalarında kullanılabilir olduğu gösterilmiştir. Adaptif filtre katsayılarının hesaplanmasında ise özilinti matrislerinin başlangıç değerlerinin mümkün olduğunca fazla ses çerçevesi kullanılarak hesaplanması gerektiği deneysel sonuçlarla ispatlanmıştır.

Algoritmaların sonuçlarına göre Wiener ve SDW algoritmalarında gürültü azaltma açısından SP ve MVDR algoritmalarına göre daha iyi sonuçlar alınmasıyla birlikte; SP ve MVDR algoritmalarında ise konuşma kalitelerinin korunması açısından daha iyi sonuçlar alınmıştır. Bu bağlamda konuşma kalitesinin daha önemli olduğu uygulamalarda SP ve MVDR algoritmaları, gürültü azaltmanın öncelik verildiği uygulamalarda SDW algoritmasının kullanımına öncelik verilebilir.



## KAYNAKLAR

- Anonim (2018). A noisy speech corpus for evaluation of speech enhancement algorithms. <http://ecs.utdallas.edu/loizou/speech/noizeus/> (erişim tarihi, 13.02.2018)
- Anonim (1969). IEEE Recommended Practice for Speech Quality Measurements. IEEE Trans. Audio and Electroacoustics, Vol. 17, 225-246.
- Beerends J, Stemerdink J(1994). A Perceptual Speech-Quality Measure Based on a Psychoacoustic Sound Representation. JAES, Vol. 42,115-123.
- Benesty J, Huang Y (2011). A Single Channel Noise Reduction MVDR Filter. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 273-276.
- Benesty J, Souden M, Chen J (2011). A Study Of Multi Channel Noise Reduction Linear Filters In The Time Domain. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC),1-6.
- Bharath Y, Veena S, Nagalakshmi K, Darshan M, Nagapadma R (2016). Development of Robust VAD Schemes for Voice Operated Switch Application in Aircrafts, 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT),191-195.
- Capon J (1969). High resolution frequency wave number spectrum analysis. Proc. IEEE, Vol. 57,1408-1418.
- Chen J, Huang Y, Benesty J, Gaensler T (2011). On Single Channel Noise Reduction In The Time Domain. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),277-280.
- Cornelis B, Moonen M, Wouters J (2009). Comparasion of Frequency Domain Noise Reduction Strategies Based on Multichannel Wiener Filtering and Spatial prediction. IEEE International Conference on Acoustics, Speech and Signal Processing,129-132.
- Hollier M, Hawksford M, Guard D (1994). Error activity and error entropy as a measure of psychoacoustic significance in the perceptual domain. IEEE Proceedings – Vision, Image and Signal Processing, vol. 141 (3), 203–208.
- Huang Y, Benesty J (2012). A Multi-Frame Approach to the Frequency-Domain Single-Channel Noise reduction Problem. IEEE Transactions On Audio, Speech, And Language Processing, Vol.20, No.4,1256-1269.
- Ingle V ve Proakis J (2011). Digital signal processing using MATLAB. Cengage learning Engineering 3 edition, 672 p, USA.
- ITU-T Rec. P.800 (1996). Methods for Subjective Determination of Transmission Quality. International Telecommunication Union, Switzerland
- ITU-T Rec. P.830 (1996). Subjective performance assessment of telephone-band and wideband digital codecs, International Telecommunication Union, Switzerland.
- ITU-T Rec. P.862 (2001). Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for End-to-end Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs, International Telecommunication Union, Switzerland.

- Jin W, Wei J, Zhong X (2017). Multichannel Speech Enhancement in Driving Environment. 11th International Conference on Signal Processing and Communication Systems (ICSPCS),1-5.
- Kasnakoğlu C (2014). Kalman Filtre Uygulamaları, [https://kasnakoglu.files.wordpress.com/2014/01/ders7\\_ck01.pdf](https://kasnakoglu.files.wordpress.com/2014/01/ders7_ck01.pdf) , (erişim tarihi, 31.12.2018).
- Long T, Benesty J, Chen J (2016). Constrained Wiener Gains and Filters for Single-Channel and Multichannel Noise Reduction. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT),843-847.
- Meyer J, Simmer K (1997). Multi-Channel Speech Enhancement In A Car Environment Using Wiener Filtering And Spectral Subtraction. IEEE International Conference on Acoustics, Speech and Signal Processing Vol. 2,1167-1170.
- Moattar M, Homayounpour M (2009). A Simple But Efficient Real-Time Voice Activity Detection Algorithm. 17th EUSIPCO, 2549-2553.
- Moattar M, Homayounpour M, Kalantari N (2010). A New Approach For Robust Realtime Voice Activity Detection Using Spectral Pattern. International Conference on Acoustic, Speech and Signal Processing IEEE,4478-4481.
- Modhave N, Karuna Y, Tonde S (2016). Design of Multichannel Wiener Filter for Speech Enhancement in Hearing Aids and Noise Reduction Technique. Online International Conference on Green Engineering and Technologies (IC-GET),1-4.
- Modhave N, Karuna Y, Tonde S (2017). Design of Matrix Wiener Filter for Noise Reduction and Speech Enhancement in Hearing Aids. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA),1-5.
- Ngo K (2011). Digital signal processing algorithms for noise reduction dynamic range compression and feedback cancellation in hearing aids. (Doktora tezi), Arenberg Doctoral School of Science Engineering&Technology Faculty of Engineering, Leuven.
- Özkaya A (1994). Gürültülü Konuşma Sinyalinin Gürültüsünün Yok Edilmesi. (Doktora Tezi), A.Ü.F.B.E, Ankara
- Pollak P, Sovka P, Jan Uhlir J (1993). Noise Suppression System For A Car. Third European Conference on Speech Communication and Technology, 3rd European Conference on Speech Communication and Technology- EUROSPEECH,1073-1076.
- Rama Rao V, Rama Murthy M, Sheela K (2008). A new technique for street noise reduction in signal procesing applications, TENCON, IEEE Region 10 Conference,1-5.
- Rix A, Hollier M (2000). "The perceptual analysis measurement system for robust end-to-end speech quality assessment." IEEE International conference on Acoustics speech and signal processing, ICASSP,1-4.
- Sakhnov K, Verteletskaya E, Simak B (2009). Low Complexity Voice Activity Detector Using Periodicity And Energy Ratio. 16th International Conference on Systems, Signals and Image Processing IEEE,1-5.

- Spriet A, Moonen M, Wouters J (2005). Stochastic gradient based implementation of spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction in hearing aids. *IEEE Transactions on Signal Processing*, 53(3), 911–925.
- Sümer N (2004). Yeni Kuşak Ara bağlantı:2G/3G Ara bağlantı. Telekomünikasyon Kurumu Uzmanlık Tezi, Ankara.
- Vageshi S (2006). *Advanced digital signal processing and noise reduction*. Wiley ve Sons, 453 p, İngiltere.
- Verteleetskaya E, Sakhnov K (2010). Voice Activity Detection for Speech Enhancement Applications. *ACTA POLYTECHNICA*, Vol.50, No.4,100-105.
- Yousheng X, Jianwen H (2014). Speech Enhancement Based on Combination of Wiener Filter and Subspace Filter. *International Conference on Audio, Language and Image Processing*,459-463.
- Zaw T, War N (2017). The Combination of Spectral Entropy, Zero Crossing Rate, Short Time Energy, and Linear Prediction Error for Voice Activity Detection. *International Conference of Computer and Information Technology (ICCIT)*,1-5.
- Zelinski R, Noll P (1977). Adaptive transform coding of the speech signals. *IEEE trans on acoustics speech and signal processing ASSP*,299-309.
- Çolak R, Akdeniz R (2018). Ses Sinyalinde Gürültü Saptama İçin Özgün Bir Yaklaşım. *European Journal of Engineering and Applied Sciences*, Vol.1:31-38.

## EK 1: VAD algoritmasının MATLAB programı aracılığıyla yazılan kaynak kodu

```
function [ livad] = binary_vad_2(mic1_ses,fs,k,threshold_SFM );
frame_time=128;%kaç örnek işleneceği belirtildi. 128 in tam katı değilse sinyale sıfırlar eklenir.
kalan = mod(length(mic1_ses),frame_time);
if kalan==0
mic1_ses=mic1_ses;
else
mic1_ses=[mic1_ses;zeros(frame_time-kalan,1)];
end
len = length(mic1_ses);
SP = 0.5; % Shift percentage is 50
wl=frame_time;
%frame lere ayırma işlemi yapılıyor.
overlap = fix((1-SP)*wl);
offset = wl - overlap;
max_m = fix((len-wl)/offset);%overlapping frame sayısı;
frame=zeros(frame_time,max_m);
for m = 0:max_m;
begin = m*offset+1;
iend = m*offset+wl;
speech_mic1 = mic1_ses(begin:iend);
frame(:,m+1)=speech_mic1;
end
%%%%başlangıç short time enerji bulma işlemi
toplam=0;
for i=1:(max_m+1);
ste(i)=sum((frame(:,i).^2));
end
normalize_ste=ste./max(ste);%normalize edilmiş ste katsayıları.
ste_sinyal=[];
for j=1:length(ste);
d=length(ste_sinyal);
ste_sinyal(1+d:d+frame_time,1)=normalize_ste(j);
end
%%%%periyodiklik hesaplama
R_frame=zeros(length(frame(:,1)),length(frame(1,:)));
for x=1:max_m+1;
for lag=16:64;
for n=1:frame_time-lag;
pay_toplam=sum(frame(n,x)*frame(n+lag,x));
payda_toplam=sqrt(sum(frame(n+lag,x)^2));
end
R_frame(lag,x)=pay_toplam/payda_toplam;
end
periodicity(1,x)=max(R_frame(:,x));
end
periodicity=(1/max(periodicity)).*periodicity;
periyodik_sinyal=[];
for j=1:length(periodicity);
d=length(periyodik_sinyal);
periyodik_sinyal(1+d:d+frame_time,1)=periodicity(1,j);
end
%%tüm framelerin fft si alınır.
```

```

for i=1:max_m+1;
    fft_frame(:,i)=abs(fft(frame(:,i)));
end
%%%spectral flatness hesaplama;
for j=1:max_m+1;
    geo_toplam=0;
    arit_toplam=0;
    for i=1:frame_time/2;%sol yanını aldık fftnin.
        geo_toplam=geo_toplam+(log10(fft_frame(i,j)));
        arit_toplam=arit_toplam+fft_frame(i,j);
    end
    geo_toplam=geo_toplam/(frame_time/2);
    geo_orta(1,j)=10^geo_toplam;
    arit_orta(1,j)=arit_toplam/(frame_time/2);
    SFM(1,j)=10*log10(geo_orta(1,j)/arit_orta(1,j));
end
flatness_sinyal=[];
for j=1:length(SFM);
    d=length(flatness_sinyal);
    flatness_sinyal(1+d:d+frame_time,1)=SFM(1,j);
end
min_SFM=max(SFM(1,1:20));
threshold_ste(1,1)=sum(normalize_ste(1,1:20))/20;% ilk 20 frame alındı.
p=0.15;
for i=1:max_m+1;
    sayac=0;
    if normalize_ste(1,i)>=k*threshold_ste(1,i);
        sayac=sayac+1;
    end
    if periodicity(1,i)>=0.35;
        sayac=sayac+1;
    end
    if (SFM(1,i)-min_SFM)<=threshold_SFM;
        sayac=sayac+1;
    end
    if sayac>=1;
        livad(1,i)=1;
        threshold_ste(1,i+1)=threshold_ste(1,i);
    else
        livad(1,i)=0;
        threshold_ste(1,i+1)=(1-p)*threshold_ste(1,i)+p*normalize_ste(1,i);
    end
end
end
%frameleri tamamlama işlemleri
sayac=0;
for i=1:length(livad);
    if livad(1,i)==0;
        sayac=sayac+1;
    else if sayac<10;
        livad(1,i-sayac:i)=1;
        sayac=0;
    else if sayac>=10;
        livad(1,i-sayac:i-1)=0;
        sayac=0;
    end
end

```

```

    end
    end
end
%frameleri tamamlama işlemleri
sayac=0;
for i=1:length(livad);
    if livad(1,i)==1;
        sayac=sayac+1;
    else if sayac<5;
        livad(1,i-sayac:i)=0;
        sayac=0;
    else if sayac>=5;
        livad(1,i-sayac:i-1)=1;
        sayac=0;
    end
end
end
end
end
%%lineer vad çizdirme;
lineervad_sinyal=zeros(length(mic1_ses),1);
oran=floor(length(mic1_ses)/(max_m+1));
for j=1:length(ste);
    d=length(lineervad_sinyal);
    lineervad_sinyal((j-1)*oran+1:(j-1)*oran+oran,1)=livad(1,j);
end
timedomain_enhanced_signal=zeros(length(mic1_ses),1);
for i= 0:max_m-1;
    begin = i*offset+1;
    iend = i*offset+wl;
    timedomain_enhanced_signal(begin:iend,1)=frame(:,i+1);
end
figure(1);
plot(1:length(mic1_ses),mic1_ses,1:length(mic1_ses),lineervad_sinyal,'r');
axis([0 length(mic1_ses) -1.5 1.5]);
xlabel('time');
ylabel('amplitude');
figure(2);
plot(1:length(ste),normalize_ste,1:length(threshold_ste),threshold_ste);
legend('ste sinyal','threshold sinyali');
title('normalize edilmiş short time enerji sinyali');
figure(3);
subplot(211)
plot(1:length(ste),periodicity,1:length(ste),livad,'r');
legend('periyodik sinyal','binary vad sinyali');
axis([0 length(ste) -2 2]);
title('periyodiklik ve binary vad sinyali');
subplot(212)
plot(flatness_sinyal);
title('spektral flatness sinyali');

```

## EK 2: Filtre katsayılarının hesaplanması için gerekli kaynak kod

```
clc; clear all;
%1. mikrofon sinyalleri toplanıyor.
[mic1_speech,fs]=audioread('sp_erkek_mic1.wav');
[mic1_noise,fs]=audioread('whitenoise_45deg_mic1.wav');
mic1_noise=mic1_noise(1:length(mic1_speech));
mic1_noise=SNR*mic1_noise;
mic1_ses=mic1_speech+mic1_noise;
%%2. mikrofon sinyalleri toplanıyor
[mic2_speech,fs]=audioread('sp_erkek_mic2.wav');
[mic2_noise,fs]=audioread('whitenoise_45deg_mic2.wav');
mic2_noise=mic2_noise(1:length(mic2_speech));
mic2_noise=SNR*mic2_noise;
mic2_ses=mic2_speech+mic2_noise;
%%3. mikrofon siyalleri toplanıyor.
[mic3_speech,fs]=audioread('sp_erkek_mic3.wav');
[mic3_noise,fs]=audioread('whitenoise_45deg_mic3.wav');
mic3_noise=mic3_noise(1:length(mic3_speech));
mic3_noise=SNR*mic3_noise;
mic3_ses=mic3_speech+mic3_noise;
%%
frame_len=128;
kalan = mod(length(mic1_ses),frame_len);
if kalan==0
mic1_ses=mic1_ses;
mic2_ses=mic2_ses;
mic3_ses=mic3_ses;
else
mic1_ses=[mic1_ses;zeros(frame_len-kalan,1)];
mic2_ses=[mic2_ses;zeros(frame_len-kalan,1)];
mic3_ses=[mic3_ses;zeros(frame_len-kalan,1)];
end
[livad] = binary_vad_2(mic1_ses,fs,k, threshold_SFM);
%livad frame sayısını overlapp e göre ayarladık.
%% ----- input noisy speech -----
len = length(mic1_ses);
SP = 0.5; % Shift percentage is 50 % Overlap
wl=frame_len;
overlap = fix((1-SP)*wl); % ardışık framalar arası overlapp
offset = wl - overlap;
max_m = fix((len-wl)/offset);%overlapplı frame sayısı;
M=3; %mikrofon sayısı
%-----başlangıçtaki ortalama gürültü hesaplanacak -----
count = 0;
Rn_noise_corr_ort=zeros(M*wl,M);
pencere=hanning(wl);
for i = 0:length(livad)-1;% 100 gurultulu frame hesaplandı.
begin = i*offset+1;
iend = i*offset+wl;
noise_mic1= mic1_ses(begin:iend).*pencere;
fft_noise_mic1=fft(noise_mic1,wl);
noise_mic2= mic2_ses(begin:iend).*pencere;
fft_noise_mic2=fft(noise_mic2,wl);
```

```

noise_mic3= mic3_ses(begin:iend).*pencere;
fft_noise_mic3=fft(noise_mic3,wl);
Rn=[fft(noise_mic1,wl) fft(noise_mic2,wl) fft(noise_mic3,wl)];
sayac=0;
    if count~=100 && livad(1,i+1)==0;
        for j=1:wl;
            Rn_corr(:,j)=Rn(j,:)*Rn(j,:);
            Rn_cor_ort(sayac+1:sayac+M,:)=Rn_corr(:,j);
        sayac=sayac+M;
        end
        Rn_noise_corr_ort=Rn_noise_corr_ort+Rn_cor_ort;
        count = count + 1;
    else if count==100;
        break
    end
end
end
Rn_noise_corr_ort=Rn_noise_corr_ort/count;
%%%%%%baslangicta noisy speech ortalaması hesaplanacak.
Ry_noisyspeech_corr_ort=zeros(M*wl,M);
counter = 0;
for i = 0:length(livad)-1;%70 gurultulu frame hesaplandı.
    begin = i*offset+1;
    iend = i*offset+wl;
    noisy_mic1= mic1_ses(begin:iend).*pencere;
    fft_noisy_mic1=fft(noisy_mic1,wl);
    noisy_mic2= mic2_ses(begin:iend).*pencere;
    fft_noisy_mic2=fft(noisy_mic2,wl);
    noisy_mic3= mic3_ses(begin:iend).*pencere;
    fft_noisy_mic3=fft(noisy_mic3,wl);
    Ry=[fft(noisy_mic1,wl) fft(noisy_mic2,wl) fft(noisy_mic3,wl)];
    sayac=0;
        if counter~=70 && livad(1,i+1)==1 ;
            for j=1:wl;
                Ry_corr(:,j)=Ry(j,:)*Ry(j,:);
                Ry_cor_ort(sayac+1:sayac+M,:)=Ry_corr(:,j);
            sayac=sayac+M;
            end
            Ry_noisyspeech_corr_ort=Ry_noisyspeech_corr_ort+Ry_cor_ort;
            counter = counter + 1;
        else if counter==70;
            break
        end
    end
end
Ry_noisyspeech_corr_ort=Ry_noisyspeech_corr_ort/counter;
for m = 0:max_m;
    begin = m*offset+1;
    iend = m*offset+wl;
    speech_mic1 = mic1_ses(begin:iend);
    speech_mic2 = mic2_ses(begin:iend);
    speech_mic3 = mic3_ses(begin:iend);
    %konuşma segmentlerinin fftlerinin çıkarılması
    recover_speech_1 = pencere.*speech_mic1; % hanning window uygulaması
    fft_speech_mic1(m*wl+1:m*wl+wl,1) = fft(recover_speech_1,wl);

```



```

recover_speech_2 = pencere.*speech_mic2; % hanning window uygulaması
fft_speech_mic2(m*wl+1:m*wl+wl,1) = fft(recover_speech_2,wl);
recover_speech_3 = pencere.*speech_mic3; % hanning window uygulaması
fft_speech_mic3(m*wl+1:m*wl+wl,1) = fft(recover_speech_3,wl);
end
%%%korelasyon matrislerinin belirlenmesi
Ry_cor=zeros(M*length(fft_speech_mic1),M);
Rn_cor=zeros(M*length(fft_speech_mic1),M);
Rx_cor=zeros(M*length(fft_speech_mic1),M);
Ry_cor(1:M*wl,:)=Ry_noisyspeech_corr_ort;
Rn_cor(1:M*wl,:)=Rn_noise_corr_ort;
alfa_y=0.99;
alfa_n=0.99;
for i=1:length(livad)-1;
    if livad(1,i)==1;
        sayac=1;
        for j=1:M:M*wl;
            x=[fft_speech_mic1((i-1)*wl+1:(i-1)*wl+wl,:) fft_speech_mic2((i-1)*wl+1:(i-1)*wl+wl,:)
fft_speech_mic3((i-1)*wl+1:(i-1)*wl+wl,:)];
            Ry_cor(i*(3*wl)+j:i*(3*wl)+j+M-1,:)=alfa_y*Ry_cor((i-1)*(3*wl)+j:(i-1)*(3*wl)+j+M-1,:)+((1-alfa_y)*(x(sayac,:)*x(sayac,:)));
            Rn_cor(i*(3*wl)+j:i*(3*wl)+j+M-1,:)= Rn_cor((i-1)*(3*wl)+j:(i-1)*(3*wl)+j+M-1,:);
            sayac=sayac+1;
        end
    else if livad(1,i)==0;
        sayac=1;
        for j=1:M:M*wl;
            n=[fft_speech_mic1((i-1)*wl+1:(i-1)*wl+wl,:) fft_speech_mic2((i-1)*wl+1:(i-1)*wl+wl,:)
fft_speech_mic3((i-1)*wl+1:(i-1)*wl+wl,:)];
            Rn_cor(i*(3*wl)+j:i*(3*wl)+j+M-1,:)=alfa_n*Rn_cor((i-1)*(3*wl)+j:(i-1)*(3*wl)+j+M-1,:)+((1-alfa_n)*(n(sayac,:)*n(sayac,:)));
            Ry_cor(i*(3*wl)+j:i*(3*wl)+j+M-1,:)= Ry_cor((i-1)*(3*wl)+j:(i-1)*(3*wl)+j+M-1,:);
            sayac=sayac+1;
        end
    end
end
end
Rx_cor=Ry_cor-Rn_cor;
%%%FİLTRE KATSAYILARININ BELİRLENMESİ (Wiener ve SDW)
e1=zeros(M,1);
e1(1,1)=1;
Wiener_cooficient=zeros(length(Ry_cor),1);
enhanced_signal_fft=zeros(length(fft_speech_mic1),1);
mu=1; %Gürültü katsayı parametresi (1 den büyük olursa SDW algoritması olur)
for i=1:M:length(Ry_cor);
    Wiener_cooficient(i:i+(M-1),:)=((Rx_cor(i:i+(M-1),:)+mu*Rn_cor(i:i+(M-1),:))\Rx_cor(i:i+(M-1),:))*e1;
End
%%%FİLTRE KATSAYILARININ BELİRLENMESİ (SPATIAL PREDİCTION)
for i=1:M:length(Ry_cor);
    H(i:i+(M-1),:)=(Rx_cor(i:i+(M-1),1)/Rx_cor(i,1));
    Wiener_cooficient(i:i+(M-1),:)=(Rn_cor(i:i+(M-1),:)\H(i:i+(M-1),:))*1/((H(i:i+(M-1),:)*inv(Rn_cor(i:i+(M-1),:)))*H(i:i+(M-1),:)));
end
end

```

```

%% FİLTRE KATSAYILARININ BELİRLENMESİ (MVDR)
for i=1:M:length(Ry_cor);
    E(i:i+(M-1),:)=(Rx_cor(i:i+(M-1),1)/Rx_cor(i,1));
    Wiener_coeficient(i:i+(M-1),:)=(Ry_cor(i:i+(M-1),:)\E(i:i+(M-1),:))*1/((E(i:i+(M-1),:))'*inv(Ry_cor(i:i+(M-1),:)))*E(i:i+(M-1),:));
end
% YUKARIDAKİ FİLTRE KATSAYI HESAPLAMA ALGORİTMALARINDAN BİRİ SEÇİLİP WIENER_COOFİCİENT VEKTÖRLERİ HESAPLANIP AŞAĞIDAKİ SATIRDAN DEVAM EDİLİR.
for i=0:length(fft_speech_mic1)-1;
    wiener_mic1(i+1,1)=Wiener_coeficient(M*i+1,1);
    wiener_mic2(i+1,1)=Wiener_coeficient(M*i+2,1);
    wiener_mic3(i+1,1)=Wiener_coeficient(M*i+3,1);
end
wiener_coeficient_mic1=wiener_mic1';
wiener_coeficient_mic2=wiener_mic2';
wiener_coeficient_mic3=wiener_mic3';
for i=1:length(fft_speech_mic1);
    Z1(i,1)=fft_speech_mic1(i,1)*wiener_coeficient_mic1(1,i);
    Z2(i,1)=fft_speech_mic2(i,1)*wiener_coeficient_mic2(1,i);
    Z3(i,1)=fft_speech_mic3(i,1)*wiener_coeficient_mic3(1,i);
    enhanced_signal_fft(i,1)=Z1(i,1);
end
    %%% sinyali zaman domenine çevirme işlemi
recover_signal=zeros(length(mic1_ses),1);
for i= 0:max_m;
    begin = i*offset+1;
    iend = i*offset+w1;
    timedomain_enhanced_signal(i*w1+1:i*w1+w1,1)=pencere.*real(ifft(enhanced_signal_fft(i*w1+1:i*w1+w1,1)));
    recover_signal(begin:iend,1)=recover_signal(begin:iend,1)+timedomain_enhanced_signal(i*w1+1:i*w1+w1,1);
end
recover_signal=(max(mic1_ses)/max(recover_signal))*recover_signal;

```

### EK 3: İstenilen SNR değeri için katsayı hesaplanmasına ilişkin kaynak kod

```
SNR=input('SNR değerini giriniz:');
if SNR==1;
wl=128;%
frame_len=128;
kalan = mod(length(mic1_speech),frame_len);
if kalan==0
mic1_noise=mic1_noise;
mic1_speech=mic1_speech;
mic2_noise=mic2_noise;
mic2_speech=mic2_speech;
mic3_noise=mic3_noise;
mic3_speech=mic3_speech;
else
mic1_noise=[mic1_noise;zeros(frame_len-kalan,1)];
mic1_speech=[mic1_speech;zeros(frame_len-kalan,1)];
mic2_noise=[mic2_noise;zeros(frame_len-kalan,1)];
mic2_speech=[mic2_speech;zeros(frame_len-kalan,1)];
mic3_noise=[mic3_noise;zeros(frame_len-kalan,1)];
mic3_speech=[mic3_speech;zeros(frame_len-kalan,1)];
end
isnr_pay=0;
isnr_payda=0;
for i = 1:length(lineer_vad)-1;
if lineer_vad(1,i)==1
isnr_pay=isnr_pay+sum((mic1_speech((i-1)*wl+1:(i-1)*wl+wl,1).^2));
isnr_payda=isnr_payda+sum((mic1_noise((i-1)*wl+1:(i-1)*wl+wl,1).^2));
end
end
SNR=isnr_pay/(isnr_payda*10^(0/10));%ana algoritmada katsayı olarak %kullanılır
SNR=sqrt(katsayi);
mic1_noise=katsayi*mic1_noise;
mic2_noise=katsayi*mic2_noise;
mic3_noise=katsayi*mic3_noise;
mic1_ses=mic1_speech+mic1_noise;
mic2_ses=mic2_speech+mic2_noise;
mic3_ses=mic3_speech+mic3_noise;
elseif SNR==2;
wl=128;%10 msn1 olarak belirlendi.
frame_len=128;
kalan = mod(length(mic1_speech),frame_len);
if kalan==0
mic1_noise=mic1_noise;
mic1_speech=mic1_speech;
mic2_noise=mic2_noise;
mic2_speech=mic2_speech;
mic3_noise=mic3_noise;
mic3_speech=mic3_speech;
else
mic1_noise=[mic1_noise;zeros(frame_len-kalan,1)];
mic1_speech=[mic1_speech;zeros(frame_len-kalan,1)];
mic2_noise=[mic2_noise;zeros(frame_len-kalan,1)];
mic2_speech=[mic2_speech;zeros(frame_len-kalan,1)];
```

```

mic3_noise=[mic3_noise;zeros(frame_len-kalan,1)];
mic3_speech=[mic3_speech;zeros(frame_len-kalan,1)];
end
isnr_pay=0;
isnr_payda=0;
for i = 1:length(lineer_vad)-1;
if lineer_vad(1,i)==1
isnr_pay=isnr_pay+sum((mic1_speech((i-1)*wl+1:(i-1)*wl+wl,1).^2));
isnr_payda=isnr_payda+sum((mic1_noise((i-1)*wl+1:(i-1)*wl+wl,1).^2));
end
end
SNR=isnr_pay/(isnr_payda*10^(5/10));
SNR=sqrt(katsayi);
mic1_noise=katsayi*mic1_noise;
mic2_noise=katsayi*mic2_noise;
mic3_noise=katsayi*mic3_noise;
mic1_ses=mic1_speech+mic1_noise;
mic2_ses=mic2_speech+mic2_noise;
mic3_ses=mic3_speech+mic3_noise;
elseif SNR==3;
wl=128;%10 msn1 olarak belirlendi.
frame_len=128;
kalan = mod(length(mic1_speech),frame_len);
if kalan==0
mic1_noise=mic1_noise;
mic1_speech=mic1_speech;
mic2_noise=mic2_noise;
mic2_speech=mic2_speech;
mic3_noise=mic3_noise;
mic3_speech=mic3_speech;
else
mic1_noise=[mic1_noise;zeros(frame_len-kalan,1)];
mic1_speech=[mic1_speech;zeros(frame_len-kalan,1)];
mic2_noise=[mic2_noise;zeros(frame_len-kalan,1)];
mic2_speech=[mic2_speech;zeros(frame_len-kalan,1)];
mic3_noise=[mic3_noise;zeros(frame_len-kalan,1)];
mic3_speech=[mic3_speech;zeros(frame_len-kalan,1)];
end
isnr_pay=0; isnr_payda=0;
for i = 1:length(lineer_vad)-1;
if lineer_vad(1,i)==1
isnr_pay=isnr_pay+sum((mic1_speech((i-1)*wl+1:(i-1)*wl+wl,1).^2));
isnr_payda=isnr_payda+sum((mic1_noise((i-1)*wl+1:(i-1)*wl+wl,1).^2));
end
end
SNR=isnr_pay/(isnr_payda*10^(10/10));
SNR=sqrt(katsayi);
mic1_noise=SNR*mic1_noise;
mic2_noise=SNR*mic2_noise;
mic3_noise=SNR*mic3_noise;
mic1_ses=mic1_speech+mic1_noise;
mic2_ses=mic2_speech+mic2_noise;
mic3_ses=mic3_speech+mic3_noise;
end

```

## ÖZGEÇMİŞ

Ramazan ÇOLAK, 12.10.1985 tarihinde Çorlu'da doğdu. 1996 yılında Gazi İlkokulundan, 1999 yılında Çorlu Lisesi Ortaokulundan ve 2002 yılında Çorlu Mehmet Rüştü Uzel Endüstri Meslek Lisesinden mezun oldu. 2006 yılında Kocaeli Üniversitesi Elektronik Öğretmenliği lisans programını bitirdi. 2014 yılında Mühendislik Tamamlama Programı kapsamında Namık Kemal Üniversitesi, Elektronik ve Haberleşme Mühendisliği lisans programını bitirdi. Şubat 2017'de başladığı Namık Kemal Üniversitesi Elektronik ve Haberleşme Mühendisliği Yüksek Lisans eğitimine halen devam etmektedir. 2009 Şubat ayından itibaren atandığı Türkiye Radyo Televizyon Kurumunda mühendis olarak görevini sürdürmektedir.

