

Ortak Vektör Yaklaşımıyla Ses Verisinde Gürültü Giderme

Mehmet Hakan Durak

**YÜKSEK LİSANS TEZİ**

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Temmuz 2015

Denoising Speech Signal using Common Vector Approach

Mehmet Hakan Durak

**MASTER OF SCIENCE THESIS**

Department of Electrical Electronics Engineering

July 2015

Ortak Vektör Yaklaşımıyla Ses Verisinde Gürültü Giderme

Mehmet Hakan Durak

Eskişehir Osmangazi Üniversitesi  
Fen Bilimleri Enstitüsü  
Lisansüstü Yönetmeliği Uyarınca  
Elektrik-Elektronik Mühendisliği Anabilim Dalı  
Telekomünikasyon ve Sinyal İşleme Bilim Dalında  
YÜKSEK LİSANS TEZİ  
Olarak Hazırlanmıştır

Danışman: Yrd. Doç. Dr. Erol Seke

Bu Tez ESOGU BAP tarafından “2015-758” no’lu proje çerçevesinde desteklenmiştir.

Temmuz 2015

## ONAY

Elektrik-Elektronik Mühendisliđi Anabilim Dalı Yüksek Lisans öđrencisi Mehmet Hakan Durak'ın YÜKSEK LİSANS tezi olarak hazırladıđı “Ortak Vektör Yaklaşımıyla Ses verisinde gürültü giderme” başlıklı bu çalıřma, jürimizce lisansüstü yönetmeliđin ilgili maddeleri uyarınca deđerlendirilerek oybirliđi ile kabul edilmiřtir.

**Danıřman** : Yrd. Doç. Dr. Erol Seke

**İkinci Danıřman** : -

**Yüksek Lisans Tez Savunma Jürisi:**

**Üye** : Yrd. Doç. Dr. Erol Seke

**Üye** : Doç. Dr. Rıfat Edizkan

**Üye** : Yrd. Doç. Dr. Semih Ergin

**Üye** : Yrd. Doç. Dr. Kemal Özkan

**Üye** : Yrd. Doç. Dr. Muammer Akçay

<p>Fen Bilimleri Enstitüsü Yönetim Kurulu'nun ..... tarih ve ..... sayılı kararıyla onaylanmıřtır.</p> <p>Prof. Dr. Hürriyet ERŐAHAN Enstitü Müdürü</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------

## ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Yrd. Doç. Dr. Erol Seke danışmanlığında hazırlamış olduğum “Ortak Vektör Yaklaşımıyla Ses Verisinde Gürültü Giderme” başlıklı YÜKSEK LİSANS tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel etik ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim.  
13/07/2015

Mehmet Hakan Durak

İmza

## ÖZET

Bu çalışmada, tek kanallı sistemler için alt uzay tabanlı ortak vektör yaklaşımı (OVY) kullanarak konuşma verisinde gürültü giderme yöntemi geliştirilmiştir. OVY'ye dayalı bu yöntem gürültülü konuşma verisini pencereleme ayırarak frekans alanına dönüştürdükten sonra faz bilgisini saklı tutarak genlik bilgisinde sınıflar oluşturur. Oluşturulan sınıflarda ortak ve farklılık vektörleri bulunur ve farklılık vektöründe gürültü giderilir. Gürültüsü giderilmiş farklılık vektörü ortak vektörle birleştirilerek saklanan faz bilgisiyle toplanır ve zaman alanına dönüştürülür. Önerilen bu yöntem farklı gürültü çeşitleri ve seviyelerine sahip büyük bir veri tabanında test edilmiştir. Yapılan testler sonucunda yöntemin tüm gürültü çeşitleri ve seviyelerinde kullanılması için bir parametre belirlenmiştir. Test sonuçları literatürde bulunan beş farklı yöntemle sekiz farklı objektif değerlendirme yöntemine göre kıyaslanmıştır. Değerlendirme sonuçlarına göre geliştirilen yöntem yüksek başarı oranına sahiptir. Önerilen yöntem günlük hayattaki sesli iletişim sistemlerinde kullanılmak için ideal ve geliştirmelere açık bir yöntemdir.

**Anahtar Kelimeler:** Gürültü giderme, ses verisi iyileştirme, ortak vektör yaklaşımı, altuzay tabanlı yöntem.

## SUMMARY

In this thesis, a single-channel noise reduction method based on subspace using common vector approach (CVA) is developed for speech data. Based on CVA, classes are constructed from windowed speech samples in magnitude data after noisy speech data is transformed to frequency domain and phase data of this signal is kept. Common vector and difference vector of the classes are created and denoising is performed on difference vectors. Magnitude data is then reconstructed sum of the denoised difference vectors and common vectors. After this process, sum of the magnitude and phase data is transformed to time domain. The proposed method was tested in a large database with different noise types and levels. An optimal parameter of the method for all noise types and levels was determined the results of the tests. The test results of the proposed method and five different enhancement methods in the literature were compared according to eight different objective evaluation criteria. According to the evaluation results, developed method has a high success rate. The proposed method is ideal for every voice communication systems and open to improvement in the future.

**Keywords:** Denoising, speech enhancement, common vector approach, subspace based method

## TEŞEKKÜR

Tez konunun belirlenmesinde, araştırma aşamasında, karşılaştığım sorunlarda yönlendirmesinde ve tezimin tamamlanmasında destek olan tez danışmanım Sayın Yrd. Doç. Dr. Erol SEKE'ye, tezimin her aşamasında yardımlarını esirgemeyen değerli hocam Yrd. Doç. Dr. Kemal ÖZKAN'a, 2015-758 numaralı "Ortak vektör yaklaşımıyla ses işaretinde gürültü giderme" adlı proje kapsamında gerekli malzeme ve cihaz desteği sağlayan Bilimsel Araştırma Projeleri (BAP) kuruluna ve tez süresince maddi destek/burs sağlayan TÜBİTAK 2210-C burs programına teşekkür ederim.

Ayrıca bu süreçte her türlü destekleriyle yanımda olan aileme ve arkadaşlarıma teşekkür ederim.



## İÇİNDEKİLER

### Sayfa

<b>ÖZET .....</b>	<b>vi</b>
<b>SUMMARY .....</b>	<b>vii</b>
<b>TEŞEKKÜR .....</b>	<b>viii</b>
<b>İÇİNDEKİLER .....</b>	<b>ix</b>
<b>ŞEKİLLER DİZİNİ .....</b>	<b>xi</b>
<b>ÇİZELGELER DİZİNİ .....</b>	<b>xii</b>
<b>SİMGELER VE KISALTMALAR DİZİNİ .....</b>	<b>xiii</b>
<b>1. GİRİŞ .....</b>	<b>1</b>
<b>2. LİTERATÜR ARAŞTIRMASI .....</b>	<b>5</b>
<b>2.1. Spektral Çıkarma Yöntemi .....</b>	<b>8</b>
2.1.1. Spektral çıkarma yönteminin dezavantajları .....	11
<u>2.1.1.1. Müzikal Gürültü</u> .....	11
<u>2.1.1.2. Bozulma</u> .....	12
<b>2.2. Gürültü tahmini yöntemleri .....</b>	<b>12</b>
2.2.1. Ses aktivitesi tespiti (VAD) .....	12
2.2.2. Spektral güç yoğunluğu ile gürültü tahmini .....	14
<b>2.3. Wiener Süzgeci ile Gürültü Giderme .....</b>	<b>15</b>
<b>2.4. İstatistiksel Model Tabanlı Yöntemler – Doğrusal Olmayan Tahmin Etme .....</b>	<b>17</b>
2.4.1. Maksimum olabilirlik kestirimi (MLE) .....	17
2.4.2. Bayesian kestirimi .....	17
2.4.3. En küçük karesel ortalama hata (MMSE) kestirimi .....	18
<b>2.5. Altuzay Tabanlı Yöntemler .....</b>	<b>18</b>
2.5.1. SVD tabanlı yöntemler .....	20
2.5.2. KLT tabanlı yöntemler .....	21
<u>2.5.2.1. Zaman alanı sınırlı (TDC) kestirici</u> .....	22
<u>2.5.2.2. Spektral alanı sınırlı (SDC) kestirici</u> .....	23
<b>3. ORTAK VEKTÖR YAKLAŞIMI .....</b>	<b>24</b>
<b>3.1. Yeterli Veri Durumu .....</b>	<b>24</b>
<b>3.2. Yetersiz Veri Durumu .....</b>	<b>25</b>

## İÇİNDEKİLER (devam)

	<u>Sayfa</u>
<b>4. ORTAK VEKTÖR YAKLAŞIMIYLA SES VERİSİNDE GÜRÜLTÜ GİDERME .....</b>	<b>29</b>
<b>4.1. Motivasyon.....</b>	<b>29</b>
<b>4.2. Önerilen Yöntem .....</b>	<b>30</b>
<b>5. GERÇEKLEŞTİRME VE PERFORMANS ÖLÇÜMÜ .....</b>	<b>34</b>
<b>5.1. Veri Tabanı.....</b>	<b>34</b>
<b>5.2. Gerçekleştirme .....</b>	<b>35</b>
5.2.1. Pencere boyutunun gürültü gidermeye etkisi .....	36
5.2.2. Örtüşme oranının gürültü gidermeye etkisi.....	39
5.2.3. Arama yapılan komşuluğun gürültü gidermeye etkisi .....	41
<b>5.3. Performans ölçümü.....</b>	<b>42</b>
<b>5.4. Değerlendirme Kriterleri .....</b>	<b>42</b>
5.4.1. Sinyal / gürültü oranı (SNR) .....	42
5.4.2. Frekans ağırlıklı parçalı SNR (fwSNRseg).....	43
5.4.3. Algısal konuşma kalitesinin değerlendirilmesi (PESQ).....	43
5.4.4. Logaritmik benzerlik oranı (LLR) .....	44
5.4.5. Ağırlıklı spektral eğim mesafesi ölçümü (WSSD).....	44
5.4.6. Itakura-Saito mesafe ölçümü (IS) .....	44
5.4.7. Cepstrum mesafe ölçümü (CEP).....	45
5.4.8. Bileşik ölçümler .....	45
<b>5.5. Elde Edilen Sonuçların Değerlendirilmesi.....</b>	<b>46</b>
<b>6. SONUÇ VE ÖNERİLER.....</b>	<b>64</b>
<b>KAYNAKLAR DİZİNİ .....</b>	<b>66</b>

## ŞEKİLLER DİZİNİ

<u>Sekil</u>	<u>Sayfa</u>
1.1. Konuşma verisinde genel olarak gürültü giderme ve uygulama alanları şeması .....	3
2.1. Gürültü giderme yöntemleri genel şeması (Garg ve Gupta, 2013) .....	5
2.2. Spektral Çıkarma yöntemi genel şeması (Loizou, 2007) .....	10
2.3. Temiz (a), gürültülü (b), gürültüsü giderilmiş (c) konuşma verisinin frekans gösterimi (Commins, 2005) .....	11
2.4. Konuşma verisinde konuşma olmayan anların belirlenmesi (Commins, 2005).....	13
4.1. Önerilen yöntemin akış şeması .....	31
4.2. Belirli bir komşulukta sınıf oluşturma işlemi.....	32
4.3. Hamming penceresi ağırlıkları .....	32
5.1. Pencere boyutu 40 örnek (4 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları.....	36
5.2. Pencere boyutu 8 örnek (0.8 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları.....	37
5.3. Pencere boyutu 80 örnek (8 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları.....	38
5.4. Farklı gürültü seviyelerinde pencere boyutu - sinyal / gürültü oranı (SNR) grafiği .....	39
5.5. Örtüşme oranı %70 olarak seçildiğinde gürültü giderme işlemi sonuçları .....	40
5.6. Komşuluk 20 örnek (0.2 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları.....	41
5.7. Tüm gürültü çeşitlerinde ortalama SNR grafiği .....	48
5.8. Tüm gürültü çeşitlerinde ortalama fwSNRseg grafiği .....	50
5.9. Tüm gürültü çeşitlerinde ortalama PESQ grafiği .....	52
5.10. Tüm gürültü çeşitlerinde ortalama LLR grafiği .....	54
5.11. Tüm gürültü çeşitlerinde ortalama WSSD grafiği.....	56
5.12. Tüm gürültü çeşitlerinde ortalama IS grafiği .....	58
5.13. Tüm gürültü çeşitlerinde ortalama CEP grafiği .....	60
5.14. Tüm gürültü çeşitlerinde ortalama sinyal bozulması (SIG) grafiği.....	62
5.15. Tüm gürültü çeşitlerinde ortalama gürültü bozulması (BAK) grafiği.....	62
5.16. Tüm gürültü çeşitlerinde ortalama toplam kalite (OVL) grafiği .....	63

## ÇİZELGELER DİZİNİ

### Cizelge

### Sayfa

5.1. NOIZEUS veri tabanı içerisindeki cümle metinleri ve özellikleri .....	35
5.2. SNR kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları .....	47
5.3. fwSNRseg kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları .....	48
5.4. PESQ kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları .....	51
5.5. LLR kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları.....	53
5.6. WSSD kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları .....	55
5.7. IS kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları.....	57
5.8. CEP kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları.....	59
5.9. Bileşik ölçümlere göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları .....	61

## SİMGELER VE KISALTMALAR DİZİNİ

<u>Simgeler</u>	<u>Açıklama</u>
$s(n)$	Temiz konuşma sinyali
$d(n)$	Bağımsız gürültü işareti
$y(n)$	Gürültülü konuşma verisi
$w(n)$	Pencere fonksiyonu
$ Y(k) ^a$	Genlik spektrumu
$\varphi(k)$	Faz değeri
$\hat{s}(n)$	Temizlenmiş sinyal
$Y_k(m)$	Gürültülü sinyalin ortalama gücü
$R$	Otokorelasyon matrisi
$w$	Wiener filtresi
$r$	Özilişki vektörü
$\theta$	Temiz sinyaline genlik spektrumu
$\hat{\theta}_{MLE}$	$\theta$ 'nin maksimum olabilirlik kestirimi
$\hat{X}_p$	Toeplitz veri matrisi
$r$	Artık sinyal
$r_x$	Sinyaldeki bozulma
$r_w$	Artık gürültü
$\bar{\varepsilon}_x^2$	$r_x$ vektörünün enerjisi
$\bar{\varepsilon}_w^2$	$r_w$ vektörünün enerjisi
$H_{TDC}$	TDC filtresi
$m$	Öznelik vektör sayısı
$n$	Öznelik vektörlerin boyutu
$\Phi$	Kovaryans matrisi
$\lambda$	Kovaryans matrisinin özdeğerleri
$\mathbf{a}_i$	Öznelik vektörü
$\mathbf{a}_j$	Bir örüntü sınıfına ait fark vektörleri
$P^\perp$	İzdüşüm matrisi

## SİMGELER VE KISALTMALAR DİZİNİ (devam)

<b><u>Simgeler</u></b>	<b><u>Açıklama</u></b>
$\mathbf{a}_c$	Temiz sinyal için doğrusal öngörü vektörü
$\mathbf{a}_p$	İyileştirilmiş sinyal için doğrusal öngörü vektörü
$R_c$	Temiz sinyal için otokorelasyon matrisi
$\sigma_c^2$	Temiz sinyalin LPC kazancı
$\sigma_p^2$	İyileştirilmiş sinyalin LPC kazancı
$d_{cep}$	Cepstrum mesafesi
$c_x(k)$	Temiz sinyalin cepstrum katsayıları
$c_{\hat{x}}(k)$	İyileştirilmiş sinyalin cepstrum katsayıları
$C_{sig}$	Sinyal bozulması
$C_{bak}$	Gürültü bozulması
$C_{ovl}$	Toplam kalite ölçümü

<b><u>Kısaltmalar</u></b>	<b><u>Açıklama</u></b>
DFT	Discrete Fourier Transform
EVD	Eigen Value Decomposition
fwSNRseg	Frequency weighted segmental SNR
IDFT	Inverse Discrete Fourier Transform
IS	Itakura-Saito
KLT	Karhunen-Loève dönüşümü
LLR	Logarithmic Likelihood Ratio
MLE	Maximum Likelihood Estimation
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
OVY	Ortak Vektör Yaklaşımı
PDF	Probability Density Function
PESQ	Perceptually Evaluation Speech Quality
QSVD	Quotient Singular Value Decomposition
SDC	Spectral Domain Constraint

**SİMGELER VE KISALTMALAR DİZİNİ (devam)**

<b><u>Simgeler</u></b>	<b><u>Açıklama</u></b>
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
TBA	Temel Bileşen Analizi
TDC	Time Domain Constraint
TDK	Türk Dil Kurumu
VAD	Voice Activity Detection
WSSD	Weighted Spectral Slope Distance
ZCR	Zero Crossing Rate

## 1. GİRİŞ

İnsan, tarihinden bu zamana kadar çevresiyle hep iletişim kurmuştur. İletişim duygu, düşünce veya bilgilerin akla gelebilecek her türlü yolla başkalarına aktarılmasıdır. Geçmişte insanoğlunun hayatını sürdürebilmesi, şimdi ise ek olarak hayat standartlarını yükseltmesi için iletişim insan hayatında çok önemli bir yere sahiptir. Bireyler birbirleriyle yazılı veya sözlü olarak iletişim kurarlar. Konuşma ve konuşmanın içindeki anlamlı sözcükler karşıdaki kişinin bizi anlaması için referanstır. Mevlana “Ne kadar bilersen bil, anlattıkların karşıdakinin anlayabildiği kadardır.” sözüyle iletişimde iki tarafın önemini vurgulamıştır. İletişimin sağlıklı kurulabilmesi için konuşma bilgisinin karşı tarafa tamamen ulaşması ve bu bilgiyi karşı tarafın tamamen anlayıp algılaması gerekmektedir.

Bilim ve teknolojinin hızını artırarak gelişmesi şüphesiz insanoğlunun hayatını önemli ölçüde değiştirmiştir. Gelişen teknolojiler her açıdan çeşitliliği artırmış ve bu çeşitliliği bize sunmuştur. Bilgi iletişim teknolojileri hayatımıza birçok kolaylık getirmiştir. Akıllı telefonlar, ses tanıma sistemleri, işitme cihazları, yüksek kaliteli müzik sistemleri, sesli bankacılık, sesli imza, otomatik yanıt, akıllı evler ve sayısız birçok yenilik çevremizde kurduğumuz iletişimi en üst düzeyde kurmamızı sağlayan araçlar teknolojinin gelişmesiyle hayatımızda var olmuştur. Var olan teknoloji gün geçtikçe gelişmekte ve yeni araçlar, sistemler vb. eklenerek bilgiye erişme ve paylaşma gün geçtikçe kolaylaşmaktadır. Asırlar öncesine gitmeden 10-15 sene öncesini dahi düşünürsek sadece mesaj yollama ve sesli arama özelliğine sahip olan telefonlarımız şuan görüntülü arama özelliğine sahip olan, konuşmamızı yazılı mesaja çeviren, sesli komut alan hatta sorduğumuz sorulara cevap verebilen akıllı telefonlara dönüşmüştür. Bundan 15 sene sonrası için bile herhangi bir öngörü imkânsız iken, 1-2 sene sonrasını öngörmek çok zor bir durumdur. Teknoloji bu kadar hızlı bir şekilde gelişmekteyken yapılan konuşmanın karşı taraf veya cihaz tarafından zamanında ve doğru anlaşılmasının önemi buna bağlı olarak artmaktadır. Böylelikle insan sesinin ve konuşmanın iletişimin temelini oluşturduğunu söyleyebiliriz.



Bahsettiğimiz tüm bu sesli iletişim sistemlerinin olmazsa olmazı ve aynı zamanda iletişimin temeli olan konuşma verisi çevreden ve diğer sistemlerden gürültüye maruz kalmaktadır. Bilgi iletişim sistemlerinde gürültü, dış kaynaklardan ve/veya ortamdan kaynaklanan, bilgi taşıyan veriye olumsuz etki edip anlaşılabilirliğini bozan ve kalitesini düşüren istenmeyen sinyallerdir. Ses sinyali üzerinde işlem yapan veya bu veriyi giriş olarak alan tüm sistemlerde gürültünün varlığı sistemin tam olarak çalışmasını engeller veya işlevini yerine getirmesini engeller. Bu sebeple konuşma verisinde herhangi bir sebeple var olan gürültüyü gidermek büyük önem taşımaktadır.

Son yıllarda artan teknolojik gelişmelerle birlikte konuşma verisinde gürültü giderme işleminin önemi de artmaktadır. Söz konusu sistemlerin tam olarak çalışması ve işlevini yerine getirmesi için konuşma verisinde gürültü giderme işlemi elzem olduğundan bununla ilgili yapılan çalışmalar durmaksızın devam etmektedir.

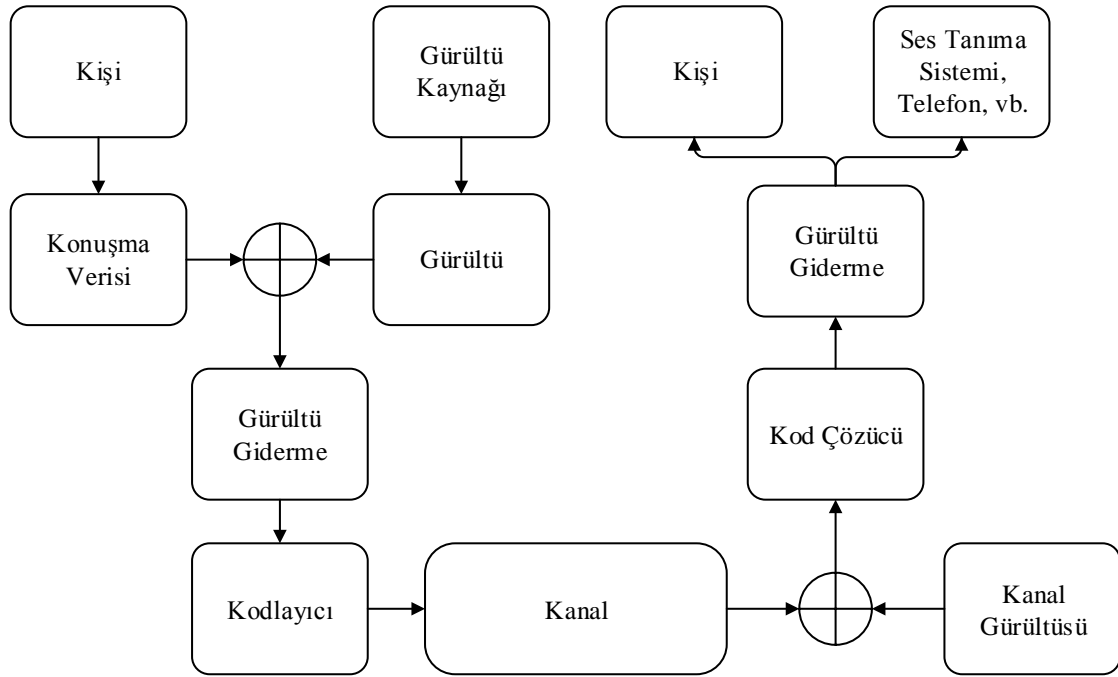
Sesli iletişim sistemlerinde ve teknolojiye sesin olduğu her yerde, ses anlaşılabilir olsa da gürültüsüz olması istenir. İster kodlanmış isterse kodlanmamış olsun konuşma verisinin mümkün olduğunca gürültüden ayıklanmış olması dinleme kalitesi ve tanımlama açısından önemlidir. Bu sebeple tasarlanan tüm sistemlerde sesin iyileştirilmesine ihtiyaç vardır. Bu ihtiyaçtan yola çıkarak bu tezde yapılan çalışmanın amacı;

- Gürültülü bir ortamda yapılan konuşmanın gürültüsünü yeni bir yöntemle gidermek ve anlaşılabilirliğini yükseltmek
- Yöntemin gerçekleşmesinden sonra çeşitli nesnel ölçütlere göre kıyaslamak

olarak belirlenmiştir. Bunlara ek olarak çok kanallı sistemlerde uygulananlardan daha karmaşık olan tek mikrofondan alınan konuşma verisinde gürültü giderilmesi amaçlanmıştır. Böylelikle başka hiçbir bilgiye gerek olmadan tek bir kanaldan alınan konuşma verisi üzerinde belirlenen yöntemle gürültü giderme işlemi yapılmıştır.

Sesle ilgili olan çoğu sistemde, alınan veri doğrudan kullanıma uygun değildir. Birçok bozucunun ve gürültünün etki ettiği ses verisi sistemlerin doğru ve tam çalışmasına, amacını yerine getirmesine engeldir. Bu sebeple konuşma verisi hangi alanda kullanılırsa

kullanılsın gürültü giderme işlemi Şekil 1.1.' de gösterildiği gibi belli bir sürece veya sisteme girmeden önce ve/veya sonra bir ön işlem gerekmektedir. Bu ön işlemten sonra kullanılma alanına göre belirli işlevi yerine getirmek için kullanılır. Tez kapsamında yapılan çalışmanın kaynağında bu motivasyon vardır. Hangi sebeple olursa olsun kullanılacak konuşma verisinin gürültüsü giderilmiş, ses kalitesi ve anlaşılabilirliği yüksek olması istendiği için literatürde denenmemiş yeni bir yaklaşımla bir yöntem geliştirilmesi planlanmıştır.



Şekil 1.1. Konuşma verisinde genel olarak gürültü giderme ve uygulama alanları şeması

Konuşma verisinde gürültü giderme işlemi uygulama alanı olarak sesli iletişimin ve konuşma verisinin olduğu geniş çaplı bir kapsam alanına sahiptir. Mevcut teknolojide ve bilişim sektöründe farklı alanlarda birçok uygulaması mevcuttur. Bazı örnekler aşağıda verilmiştir.

- Mobil Telefonlar ve Haberleşme: Haberleşmede en büyük problem verideki (iletişim kanalındaki veya ortamdaki) gürültüdür. Konuşma verisindeki bu gürültüyü gidermek için

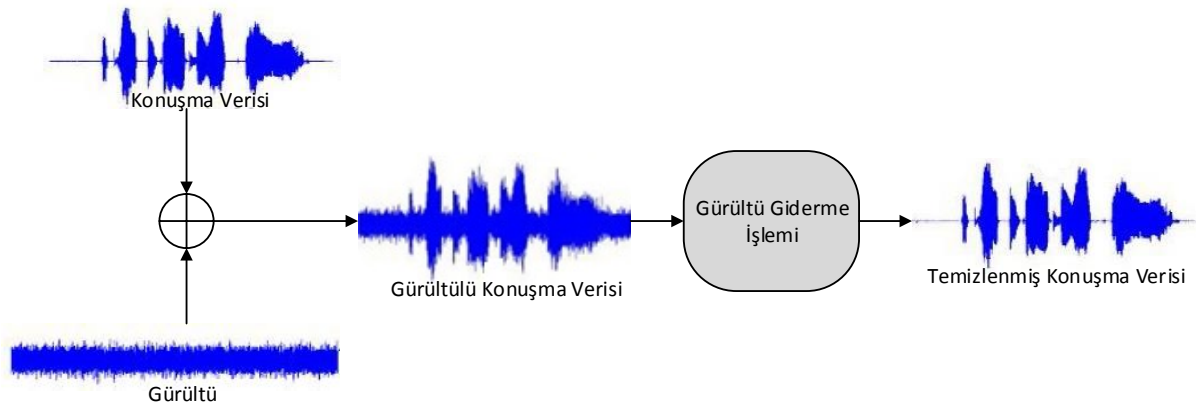
mobil telefon üreten firmaları, GSM firmaları ve arařtırmacılar birçok algoritma geliřtirmiřtir ve çalıřmalar devam etmektedir.

- Ses Tanıma ve sesli yanıt sistemleri: Biliřim sektöründe en çok uygulamanın olduđu alan ses tanıma ve sesli yanıt sistemleridir. Bu sistemlerle herhangi bir insan olmadan otomatik olarak karřıdaki kiřiyle iletiřim kuran insansı robotlar/sistemler tasarlanmaya çalıřılmaktadır. Bunun için sađlıklı iletiřim kurulması gerektiđinden gürültü giderme iřlemine ihtiyaç vardır. Gelecekte yapay zekâyla birlikte öğrenme yetisi olan ve dıř dünyayla rahatça iletiřim kurabilen robotlar hayatımızda olacaktır.
- İřitme Cihazları ve kulaklık, mikrofon sistemleri: İřitme cihazları kullananlar genellikle belli miktarda iřitme kaybı yařayan ve iřitme problemi olan kiřilerdir. Kulaklık, mikrofon sistemleri ise belirli amaçlarla (telefonla konuřma, müzik dinleme vb. ) kullanılan içerisinde gürültülü sesin olabildiđi sistemlerdir. Bu sistemlerde de gürültülü konuřma verisini temizlemek sistemin kalitesi ve kullanılabilirliđi açısından önemlidir.

## 2. LİTERATÜR ARAŞTIRMASI

İnsan sesinin doğru olarak anlaşılabilmesi için gürültüden arındırılması gerekmektedir. Gürültü giderme işlemi ne kadar iyi olursa anlaşılma ve dolayısıyla kullanıldığı alana göre sistem performansı o kadar artacaktır. Bilgi iletişim alanında gürültü temizleme ve gürültünün bastırılması yaklaşık yarım asırlık bir arka plana sahip olup hâlâ güncelliğini korumaktadır (Garg ve Gupta, 2013). Haberleşme sistemlerinden akıllı telefonlara, ses tanıma sistemlerinden işitme cihazlarına kadar geniş bir yelpazede gelişmeler devam etmekte olup bu alandaki teknolojinin yeniliklerine her geçen gün yeni cihazlar, sistemler eklenmektedir. Bu bölümde, konuşma verisinde gürültü gidermeyle ilgili literatüre geçmiş ve bu alanın yapıtaşları olan çalışmalar hakkında kısa bilgiler verilecektir.

Konuşma verisinde gürültü giderme yöntemlerinin ortak özelliği temel olarak konuşmanın kalitesini ve anlaşılabilirliğini artırmak ve bu işlemi en az kayıpla yapmaktır (Garg ve Gupta, 2013). Genel olarak gürültü giderme işlemi Şekil 2.1’de gösterildiği gibi gerçekleştirilmektedir.



Şekil 2.1. Gürültü giderme yöntemleri genel şeması (Garg ve Gupta, 2013)

Şuana kadar yapılmış olan çalışmalar için en iyi performansı gösteren yöntemin varlığından söz edilemez. Çünkü farklı ortamlardaki gürültüler aynı karakteristiği göstermez ve her bireyin konuşma şekli, ses tonu, vurgulamaları farklıdır. Yapısal olarak gürültüler aşağıdaki gibi sınıflandırılabilirler (Boll, 1979).

- Çevresel gürültü: Konuşmadan bağımsız olarak oluşan çevreden kaynaklanan gürültüdür. Beyaz gürültü olarak da adlandırılır. Bulunulan ortama göre değişiklik gösterebilen bu gürültü çeşidine cadde kalabalığının veya fabrikadaki çalışanların oluşturduğu sesler örnek olarak gösterilebilir. Gürültünün genliği (dB) kaynağına göre değişiklik gösterir.
- İstenmeyen konuşmalar: Ortamda bulunan diğer kişilerin oluşturduğu gürültü çeşididir. Bu gürültüler gürültüsü giderilmek istenen sinyale karışabilmektedir. Karakteristik olarak konuşma verisine benzediği için arındırılması zor bir gürültü çeşididir.
- Akustik yankı ve geri beslemesi: Bulunulan ortama göre konuşmaya ek olarak yankı ve bu yankının geri beslemesi söz konusu olabilmektedir. Giderilmesi zor bir gürültü çeşididir.
- Ani ve zamanla değişen gürültü: Aniden oluşup sona eren ve zamana göre sürekli değişen gürültü çeşididir. Yere düşen bir nesne veya kapanan bir kapı ani gürültüye örnek kaynak çeşidi olarak verilebilir.

Bunun için araştırmacılar varsayımlar ve kısıtlar belirlerler. Bu varsayımlar ve kısıtlara göre yöntem geliştirilir. Bazı yöntemler belli gürültü çeşitlerinde iyi performans gösterirken farklı gürültü çeşitlerinde iyi performans göstermeyebilir. Yöntemlerin performansını etkileyen birçok sebep vardır. Bunlar aşağıdaki gibi sınıflandırılabilir (Ephraim, 1992)

- Ses kaynaklarının sayısı
- Gürültülü sinyal hakkındaki bilgi miktarı
- Gürültü karakteristiğinin zamanla değişmesi/değişmemesi
- Gürültü yapısı

Bunlarla birlikte gürültü giderme yöntemleri kanal sayısına (bir/iki/çok kanallı), çalışma alanına (zaman/frekans), algoritma türüne (uyarlamalı/uyarlamasız) göre sınıflandırılırlar (Ephraim, 1992). Kanal sayısına göre bu yöntemler şu şekilde sınıflandırılırlar.

- Çoklu Mikrofon sistemleri: Bu sistemler çoklu sinyal girişinin bulunması, uyarlamalı gürültü temizleme cihazlarında referans gürültü kullanılması, istenilmeyen gürültü bileşenlerinin engellenmesi için faz düzenlemesi yapılması, faz düzenlemesi ve gürültü yok etme katlarının birlikte tasarlanabilmesi açısından avantajlıdır (Gonzalez ve Rodriguez, 1996). Bu sistemlerin düzenli olmayan gürültülere karşı tek mikrofonlu sistemlere göre daha iyi performans gösterdiği söylenebilir. Başka bir kanalın bulunmasından dolayı uyarlamalı gürültü yok etme yöntemi konuşma verisinde gürültü giderme yöntemleri arasında performans olarak iyi konumdadır. Bu yöntemde referans yolu olan başka bir kanal yardımıyla gürültünün bir örneği sisteme dâhil edilir. Daha sonra uyarlamalı bir filtreden geçirilen bu referans gürültü sinyali ana sinyalden çıkartılarak temiz sinyal edilir.

Çoklu mikrofon sistemleri tek kanallı sistemlere göre uygulaması daha kolay ve başarılı sonuçlar veriyor olsa da, günlük hayatta uygulama alanı olarak sık karşılaşılan durum tek mikrofonlu sistemlerdir. Bu tez kapsamında yapılacak çalışmalar da tek mikrofonlu sistemlere dayalı olduğundan dolayı bu sistemlerdeki yapılmış çalışmalar anlatılacaktır.

- Tek mikrofonlu sistemler: İşitme cihazları, mobil iletişim sistemleri gibi günlük hayatta en çok karşılaşılan sistemlerdir. Bu sistemlerde başka bir sinyal giriş kanalı yoktur. Çoklu mikrofon sistemlerine kıyasla ucuz ve basit sistemlerdir. Fakat bu sistemler gürültüyü temizleme açısından bilgi miktarı az olduğu için ve referans sinyali olmadığından uygulamaları en zor olan sistemlerdir. Genellikle bu sistemlerde istenmeyen gürültünün ve konuşma verisinin farklı istatistikleri kullanılır. Yöntemlerin çoğu gürültünün konuşma süresince sabit olduğunu varsaydığından bu sistemlerin performansı gerçek hayattaki sabit olmayan gürültünün varlığından dolayı sınırlıdır. Düşük sinyal/gürültü oranlarında bu sistemlerin performansı hızlı bir şekilde düşer.

Performansın kalitesi açısından önemli iki parametre ses kalitesi ve anlaşılabilirliktir. Bu iki parametre arasında dengenin sağlanması gerekmektedir. Son yıllarda birçok çalışma yapılan ve çalışmaların devam ettiği bir konudur. Bu konuda yapılan çalışmaların birkaçı bu bölümde anlatılacaktır.

Tek kanallı sistemler için literatürde bulunan konuşma verisinde gürültü giderme yöntemlerini temel olarak aşağıdaki gibi sınıflandırabiliriz (Liao ve Gregory, 1999).

- Spektral çıkarma yöntemleri
- Gürültü tahmini yöntemleri
- Wiener filtresi yöntemi
- İstatistiksel model tabanlı yöntemler
- Altuzay tabanlı yöntemler

## 2.1. Spektral Çıkarma Yöntemi

Literatüre konuşma verisinde gürültü giderme ile ilgili geçen yöntem olan spektral çıkarma, gürültülü sinyal spektrumundan tahmin edilen gürültü spektrumunun çıkarılması olarak tanımlanan bir yöntemdir (Berouti vd., 1979) Gürültü spektrumu genel olarak konuşma sinyalinin olmadığı zaman tahmin edilmektedir. Ayrıca gürültünün sabit veya zamanla değişimin az olduğu kabul edilmektedir (Loizou, 2007).

Temiz konuşma sinyali  $s(n)$  ve sesten bağımsız gürültü işareti  $d(n)$  olarak gösterilirse, gürültülü konuşma verisi

$$y(n) = s(n) + d(n) \quad (2.1)$$

şeklinde ifade edilebilir. Normalde konuşma verisinin genliği sabit olmamasına rağmen burada sabit olarak kabul edilmiştir. Konuşma verisi kısa zaman aralıklarına bölündüğünden gürültü ve konuşmanın da olduğu tüm sinyaller pencerelerle çoklandırılır. Hamming veya Hanning ağırlıklandırma kullanılarak oluşturulan pencerelenmiş sinyal aşağıdaki eşitlikteki gibi gösterilebilir.

$$\begin{aligned}
y_w(n) &= w(n)y(n) \\
&= w(n)[s(n) + d(n)] \\
&= s_w(n) + d_w(n)
\end{aligned} \tag{2.2}$$

Pencereleme işlemi sonunda  $y_w(n)$  işaretini frekans alanına dönüştürürsek genlik veya güç spektrumunu

$$|Y(k)|^a \approx |S(k)|^a + |D(k)|^a \tag{2.3}$$

olarak gösterebiliriz. Eşitlik 2.3 genel bir denklemdir. Bu denklemde a yerine 1 veya 2 yazarak genlik veya güç spektrumu elde edilir. Burada  $Y(k)$  işaretinin Ayırık Frekans Dönüşümü (DFT) alındığında gürültülü konuşma sinyali

$$\begin{aligned}
Y(k) &= \sum_{n=0}^{N-1} y(n)e^{-j\frac{2\pi kn}{N}} = |Y(k)|e^{-j\varphi(k)} \\
\varphi(k) &: Y(k)'nin faz değeri
\end{aligned} \tag{2.4}$$

şeklinde ifade edilir. Gürültü spektrumunu doğrudan bulmak mümkün olmadığından, konuşmanın olmadığı bir andaki güç spektrumu  $\widehat{D}(k)$  hesaplanır ve spektrumu kontrol etmek için  $E$  parametresi kullanılır. Tam çıkarma için bu parametre 1 seçilir, aşırı çıkarmalar için 1'den büyük değerler seçilir. Gürültünün konuşmadan bağımsız olduğu kabul edilerek gürültülü konuşma verisinden Eşitlik 2.5'teki gibi çıkarılır.

$$|\hat{S}(k)|^a = |Y(k)|^a - E|\widehat{D}(k)|^a \tag{2.5}$$

Gürültü spektrumu hesaplamalarındaki hatalardan dolayı hesaplamalarda bazen negatif değerler görülebilir. Bu değerler yarım dalga doğrultma yapılarak sıfıra çekilir. Bu işlemle beraber düzenlenen spektrum

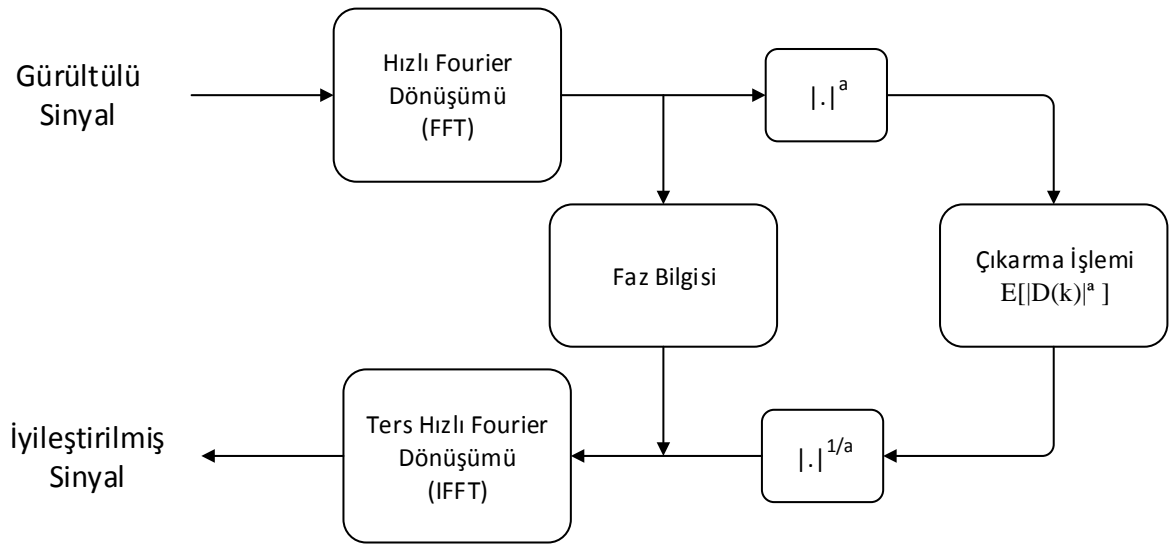
$$|\hat{S}(k)|^a = \begin{cases} |\hat{S}(k)|^a, & |\hat{S}(k)|^a > 0 \\ 0, & |\hat{S}(k)|^a \leq 0 \end{cases} \tag{2.6}$$



şeklinde gösterilebilir. Faz bilgisiyle tekrardan birleştirilerek ve bunun Ters Ayrık Fourier Dönüşümü (IDFT) alınarak gürültüsü temizlenmiş sinyali elde edebiliriz.

$$\hat{s}(n) = IDFT(|\hat{S}(k)|e^{j\varphi(k)}) \quad (2.7)$$

Spektral çıkarma yöntemini ana hatlarıyla şöyle özetleyebiliriz. Gürültüsü giderilecek olan konuşma sinyali Fourier dönüşümüyle frekans alanına aktarılır. Faz bilgisi ayrıldıktan sonra genlik veya güç spektrumu elde edilir ve sonra Eşitlik 2.5'teki çıkarma işlemi uygulanır. Yapılan işlemlerin tersi yapılarak tekrardan zaman alanındaki gürültüsü giderilmiş konuşma verisi elde edilir. Bu yöntemin şeması aşağıda verilmiştir.



Şekil 2.2. Spektral Çıkarma yöntemi genel şeması (Loizou, 2007)

Spektral güç çıkarma denklemi, Eşitlik (2.5)'de  $a=2$  ve  $E=1$  alınarak elde edilir.

$$|\hat{S}(k)|^2 = |Y(k)|^2 - |\hat{D}(k)|^2 \quad (2.8)$$

Spektral genlik çıkarma denklemi, Eşitlik (2.5)'de  $a=1$  ve  $E=1$  alınarak elde edilir.

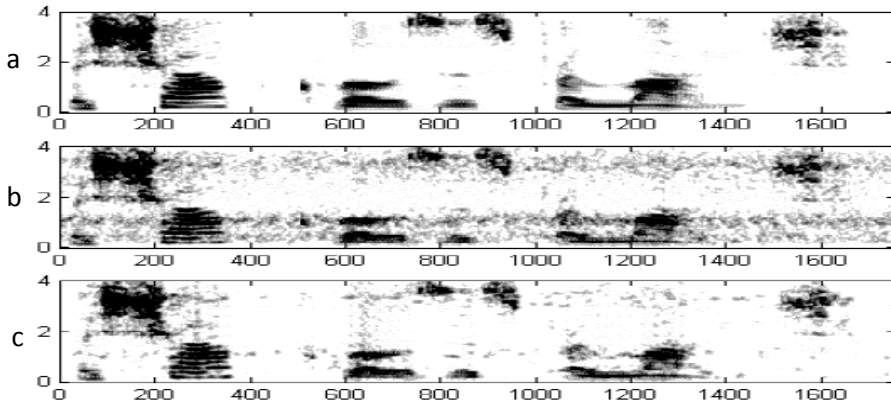
$$|\hat{S}(k)| = |Y(k)| - |\hat{D}(k)| \quad (2.9)$$

### 2.1.1. Spektral çıkarma yönteminin dezavantajları

Spektral çıkarma yöntemi uygulama olarak kolay ve verimli olmasına rağmen bazı dezavantajları bulunmaktadır. Yöntemin performansı tahmin edilen gürültünün spektrumuna bağlı olduğundan değişen çevre şartlarında gürültüyü tam olarak tahmin edebilmek bu yöntemin en zor kısmıdır. Gürültü tam olarak tahmin edilmediğinde iki önemli sorun oluşmaktadır. Bunlar müzikal gürültü ve bozulmadır.

#### 2.1.1.1. Müzikal Gürültü

Gürültü tahminindeki hatalardan dolayı kısa süreli sinüs dalga şekilleri oluşur ve ton şeklinde müzikal bir gürültü oluşturur. Rahatsız edici bir ses olduğu için bununla ilgili çok sayıda çalışma yapılmıştır ve çalışmalar devam etmektedir (Loizou, 2007).



Şekil 2.3. Temiz (a), gürültülü (b), gürültüsü giderilmiş (c) konuşma verisinin frekans gösterimi (Commins, 2005)

Örnek olarak bir konuşma verisinin ve bunun spektral çıkarma işlemi sonucunda gürültüsü giderilmiş konuşma verisinin frekans alanındaki gösterimi Şekil 2.3.'te verilmiştir. Burada temiz veride olmayan frekans bileşenlerinin olduğu görülmektedir. Yapılan çalışmalara rağmen tek mikrofona gürültü giderme yöntemlerinde müzikal gürültü tamamen giderilememektedir. Fakat sesteki bozulma ve gürültü giderme miktarı arasında uygun bir denge sağlanarak problemin etkisi azaltılmaya çalışılmaktadır (Kamath, 2001).

### **2.1.1.2. Bozulma**

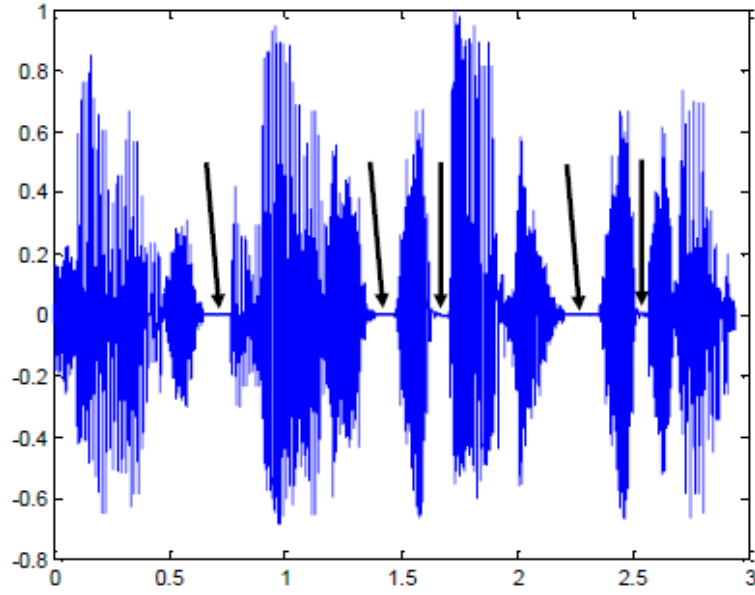
Spektral çıkarmanın diğerk bir dezavantajı bozulmadır. Müzikal gürültünün oluşmasında etken olan gürültü tahmini bozulmanın da etkenidir. Gürültü temizlenirken konuşma verisine ait olan frekans bileşenlerinin de temizlenmesiyle oluşan durumdur. Müzikal gürültü ve bozulmayı aynı anda gidermek mümkün değildir. Literatürde müzikal gürültüyü belli bir seviyede tutup bozulmayı gidermeye yönelik çalışmalar mevcuttur. (Commins, 2005)

## **2.2. Gürültü tahmini yöntemleri**

Herhangi bir gürültü giderme işleminde gürültünün karakteristiğinin belirlenmesi çok önemlidir. Gürültü tahmin etme yöntemleri literatürde fazlasıyla mevcuttur. Temel olarak konuşmanın olmadığı zamanlarda bulunan sesler gürültü olarak kabul edilmektedir. Fakat gürültünün değişken genliklerde olduğu durumlarda bu yaklaşım iyi sonuçlar vermemektedir. Gürültü tahmini yöntemlerinde ağırlıklı olarak Ses Aktivitesi Tespiti (Voice Activity Detection) algoritmaları kullanılmaktadır. (Commins, 2005)

### **2.2.1. Ses aktivitesi tespiti (VAD)**

Ses aktivitesi tespiti (VAD) algoritmalarında ortak özellik konuşmanın olmadığı anları kestirmektir. Bunun için çeşitli yöntemler geliştirilmiştir. Zero Crossing Rate (ZCR) olarak adlandırılan bir VAD yönteminde, parçalara ayrılan konuşma verisinde ses sinyalinin genliğinin kaç defa x-ksenini kestiği hesaplanır. (Commins, 2005)



Şekil 2.4. Konuşma verisinde konuşma olmayan anların belirlenmesi (Commins, 2005)

Şekil 2.4.'te gösterildiği gibi konuşmanın olmadığı anlarda x-ksenini kesme sayısının az olduğu varsayılmıştır. Dolayısıyla ayrılan parçalarda hesaplanan ZCR değeri belli bir eşik değerinin altındaysa o kısımlar gürültü olarak kabul edilmiştir (Commins, 2005).

Bir başka algoritmada ise Rabiner (1975) gürültülü sinyalin kısa süreli enerjisini kullanarak konuşmanın parçalanmış kısımlarında konuşmanın var olup olmadığını kestirmektedir. Yani ZCR algoritmasına benzer bir yaklaşımla konuşma süresince enerji hesaplanmakta ve hesaplanan bu enerji belli bir eşik değerinin altındaysa, gürültü olarak kabul edilmektedir. Eğer eşik değerinin üzerindeyse gürültülü konuşma verisi olarak düşünülmektedir. Belli parçalara ayrılan sinyalin  $m$  kısımlı olmak üzere,  $m$ . kısmın hesaplanan enerjisi  $Y(m)$ , eşik değeri ZCR için  $\alpha$ , Rabiner için  $\delta$  olmak üzere VAD genel olarak Eşitlik 2.10'daki gibi tanımlanabilir. Burada gürültülü konuşma verisini '1', gürültüyü '0' olarak temsil edilmektedir

$$VAD = \begin{cases} 1 & ZCR(m) > \alpha \text{ ve } Y(m) < \delta \\ 0 & \text{diğer} \end{cases} \quad (2.10)$$

Literatürde VAD ile ilgili karşılaşılan birçok zorluk bulunmaktadır (Martin, 2001). Konuşma verisinde konuşmanın olmadığı zamanların çok az olduğu veya konuşma duraksamalarının çok kısa olduğu durumlarda algoritmalar gürültüyü tahmin edemeyebilirler. Bu durumlarda müzikal gürültü ve bozulma gibi sorunlar ortaya çıkmaktadır. Sinyal/gürültü oranının düşük olduğu konuşma verilerinde de algoritma gürültüyü tahmin edemeyebilir. Bundan dolayı iyileştirme sırasında gürültüsü giderilmiş konuşma verisinde bazı bileşenlerde bastırılma sorunu ortaya çıkmaktadır.

### 2.2.2. Spektral güç yoğunluğu ile gürültü tahmini

VAD kullanmadan gürültü tahmini yapan uygulamalar da mevcuttur. Bunlardan biri Doblinger (1995) tarafından yapılan uygulamadır. Doblinger, konuşmanın ayrılan her bir kısmı için gürültü gücünün minimum olduğu anları takip eder (Commins, 2005). Doblinger'in yaptığı bu uygulamayı genel olarak özetlersek,

- Gürültülü konuşma verisi alt bantlara ayrılır ve bu alt bantlar parçalara bölünür. Burada,  $k$ . alt bandın  $m$ . parçasında gürültülü sinyalin gücü  $Y_k(m)$  olarak tanımlanır.
- Gürültülü sinyal gücünün tahmini,

$$Y_k(m) = \alpha Y_k(m-1) + (1 - \alpha) Y_k(m) \quad (2.11)$$

olarak hesaplanır. Burada,  $Y_k(m)$  gürültülü sinyalin ortalama gücü,  $\alpha$  düzeltme katsayısı olarak tanımlanır.

- Tahmin edilen  $m$ . kısımdaki gürültü gücü,

$$W_k(m) = \begin{cases} \gamma W_k(m-1) + \frac{1-\gamma}{1-\beta} (Y_k(m) - \beta Y_k(m-1)), & W_k(m-1) < Y_k(m) \\ W_k(m-1) & , \text{ diğ}er \end{cases} \quad (2.12)$$

olarak hesaplanır. Parametreler deneysel uygulamalar sonucunda  $0.7 \leq \alpha \leq 0.9$ ,  $\beta = 0.96$ ,  $\gamma = 0.998$  değerlerinde seçilmiştir.

Martin (2001), Doblinger'in yapmış olduğu çalışmaya ek olarak sabit olan  $\alpha$  değerinin zamanla değişmesini sağlamıştır. Eşitlik 2.13'te bu değişim gösterilmiştir.

$$\alpha_k(m) = \frac{|Y_k(m-1)|^2}{|W_k(m)|^2} \quad (2.13)$$

Martin'in geliştirdiği uygulamada  $\alpha_k$  devamlı güncellenmekte ve gürültünün tahmin edilen gücü sürekli değişmektedir. Bu uygulamayla 30 kısma ayrılan gürültülü sinyal için gürültü tahmini iyi sonuçlar vermektedir (Rangachari vd., 2004).

### 2.3. Wiener Süzgeci ile Gürültü Giderme

Spektral çıkarma işlemi daha verimli hale getirmek için Lim ve Oppenheim (1979) Wiener filtresi ile gürültü gidermeyi önermişlerdir. Prensip olarak tahmini sinyal ile gerçek sinyal arasındaki ortalama karesel hatayı en küçükmeye dayanan bir yöntemdir. Spektral olarak doğrudan çıkartma yerine bir Wiener kazanç fonksiyonu hesaplanıp gürültülü spektrum ile çarpılarak gürültüye ait frekans bileşenlerinin zayıflatılması ile gürültü giderimi amaçlanmıştır. Burada konuşma sinyalinin ve gürültünün birbirleriyle ilişkisiz olduğu kabul edilir ve gürültülü sinyalin otokorelasyon matrisinin temiz ve gürültü sinyalinin otokorelasyon matrislerinin toplamına eşit olduğu varsayılır. Konuşma verisi  $s(n)$  ve gürültü  $d(n)$  olarak ifade etmek üzere gürültülü konuşma verisi  $y(n)$

$$y(n) = s(n) + d(n) \quad (2.14)$$

olarak ifade edilir. Konuşma verisi ve gürültü ilintisiz olduğundan gürültülü sinyalin otokorelasyon matrisi

$$R_{yy} = R_{ss} + R_{dd} \quad (2.15)$$

şeklinde hesaplanır. Wiener filtresi bu eşitliklere göre

$$w = (R_{yy} + R_{dd})^{-1}r_{ss} \quad (2.16)$$

olarak bulunur. Bu eşitlikte  $r_{ss}$  gürültüsüz sinyalin özilişki vektörüdür. Bu filtre gürültü azaltma için en uygun doğrusal filtredir. Konuşma verisinin güç spektrumu  $P_{SS}(f)$ , gürültünün güç spektrumu  $P_{DD}(f)$  olmak üzere Wiener filtresi frekans alanında

$$W(f) = \frac{P_{SS}(f)}{P_{SS}(f) + P_{DD}(f)} \quad (2.17)$$

şeklinde ifade edilir. Konuşma verisinde gürültü giderilmesi için, önce gürültülü konuşma verisi Hamming veya Hanning fonksiyonlarından biri ile pencereleme yapılır. Pencerelerin uzunluğu genellikle 10-30 ms arasında seçilirken, pencereler arası örtüşme oranı %50-%70 aralığında seçilir. Wiener filtresi

$$S_w(f) = W(f)Y_w(f) \quad (2.18)$$

şeklinde uygulanır. Burada gürültülü sinyal  $Y_w(f)$ 'e filtre uygulanarak gürültüsü giderilmiş  $S_w(f)$  sinyali elde edilir. Filtre çıkışında ters fourier dönüşümü uygulanarak ve pencerelenen konuşma verisi tekrardan birleştirilerek gürültüsü giderilmiş konuşma verisi elde edilir.

Eşitlik 2.17'de verilen Wiener filtresini uygulamak için konuşma verisinin ve gürültünün spektral güç yoğunluğunun tahmin edilmesi gerekmektedir. Konuşma verisinin spektral güç yoğunluğu yerine  $|S_w(f)|^2$  ifadesini, gürültünün spektral güç yoğunluğu yerine  $|D(f)|^2$  olarak tanımlanan zaman-ortalamalı spektrum kullanılabilir. Bu ifadeler Eşitlik 2.17'de yerine yazılırsa

$$W(f) = \frac{|S_w(f)|^2}{|S_w(f)|^2 + |D(f)|^2} \quad (2.19)$$

şeklinde Eşitlik 2.17'deki süzgecin ifadesi yeniden düzenlenmiş olur. Eşitlik 2.18 ve Eşitlik 2.19 birleştirilirse aşağıdaki eşitlik elde edilir.

$$S_w(f) = \frac{|S_w(f)|^2}{|S_w(f)|^2 + |D(f)|^2} X_w(f) \quad (2.20)$$

Eşitlik 2.20'nin birinci çözümü sıfır olurken, ikinci çözümü aşağıda verilmiştir.

$$S_w(f) = \sqrt{|X_w(f)|^2 - |D(f)|^2} \quad (2.21)$$

## 2.4. İstatistiksel Model Tabanlı Yöntemler – Doğrusal Olmayan Tahmin Etme

İstatistiksel model tabanlı yöntemlerde, doğrusal olmayan tahmin etme uygulamaları kullanılır. Bu uygulamalar gürültünün olasılık yoğunluk fonksiyonunu (PDF) ve konuşmanın ayrık fourier dönüşümünü (DFT) kullanır. Literatürde birçok uygulama bulunmaktadır (Loizou, 2007).

### 2.4.1. Maksimum olabilirlik kestirimi (MLE)

Maksimum olabilirlik kestirimi istatistiksel yöntemler arasında en çok kullanılan yöntemdir. Genellikle karmaşık tahmin etme problemlerinde kullanılırlar. Konuşma verisinde gürültülü konuşma genlik spektrumu  $y$ , temiz sinyalin genlik spektrumu  $\theta$  olmak üzere  $p(y; \theta)$   $y$ 'nin olasılık yoğunluk fonksiyonunu ifade etmektedir. Matematiksel olarak  $p(y; \theta)$ 'yi maksimum yapan  $\theta$  değerlerini

$$\hat{\theta}_{MLE} = \operatorname{argmax} p(y; \theta) \quad (2.22)$$

formülünden bulabiliriz. Bu kestirime  $\theta$ 'nin maksimum olabilirlik kestirimi denir.  $\hat{\theta}_{MLE}$ ,  $p(y; \theta)$ 'nin  $\theta$ 'ye göre türevinin sıfıra eşitlenmesi ile hesaplanır. Ayrıca  $p(y; \theta)$ 'nin logaritması alınarakta hesaplamalar yapılır. Bu kestirime ise maksimum logaritmik-olabilirlik kestirimi (log-likelihood) denir. (McAulay ve Malpass, 1980)

### 2.4.2. Bayesian kestirimi

Maksimum olabilirlik kestiriminde  $\theta$ 'nin deterministik fakat bilinmeyen olduğu kabul edilir. Bayesian'da ise  $\theta$  rassal değişken olarak kabul edilir. Bayes teoreminden adını alan bu yöntem  $\theta$  hakkında ön bilgi sahibi olduğundan maksimum olabilirlik kestirimine göre daha iyi performans göstermektedir (Loizou, 2007).



### 2.4.3. En küçük karesel ortalama hata (MMSE) kestirimi

Optimal çıkarımlarda tahmini sinyal ile gerçek sinyal arasındaki ortalama karesel hata minimize edilmiştir (Loizou, 2007).

$$e = E \left\{ (\hat{X}_k - x_k)^2 \right\} \quad (2.23)$$

Eşitlik 2.23'te  $\hat{X}_k$ , tahmini spektral genlik,  $x_k$  ise temiz sinyalin gerçek genlik değerlerini göstermektedir.

$$Bayesian(\hat{X}_k) = \iint (\hat{X}_k - x_k)^2 p(Y, X_k) dY dX_k \quad (2.24)$$

Eşitlik 2.24'te gösterdiği gibi Bayesian yaklaşımında, kestirim birleşik olasılık yoğunluk fonksiyonu  $p(Y, X_k)$ 'ya göre yapılır. Bayesian yaklaşımını  $\hat{X}_k$ 'ya göre minimize edersek Eşitlik 2.25'teki MMSE kestirimini bulmuş oluruz.

$$\hat{X}_k = \int X_k p(X_k|Y) dX_k \quad (2.25)$$

## 2.5. Altuzay Tabanlı Yöntemler

Altuzay tabanlı yöntemler literatürde frekans kestirimi, sistem belirleme, varış istikameti tahmini gibi farklı alanlarda kullanılmışlardır. Konuşma verisinde gürültü giderme geçmişi çok fazla olmayan bir alandır. Altuzay tabanlı yöntemlerde temel olarak gürültülü sinyalin iki farklı altuzaya ayrılması fikrine dayanır. Gürültü-konuşma sinyali toplamı ve konuşma sinyali veya gürültü sinyali iki altuzayı oluşturmaktadır. Gürültü altuzayında istenmeyen gürültü sinyal bileşenlerinin olduğu kabul edilmektedir. Temizlenmiş konuşma sinyalini elde etmek için bu bileşenleri kaldırmak gerekmektedir. Altuzaya ayırma işlemi Tekil Değer Ayrıştırma (SVD) (Dendrinos vd., 1991) veya Öz Değer Ayrıştırma (EVD) ile yapılmaktadır (Ephraim ve Van Trees, 1995).

Dendrinos vd. (1991)'nin önerdiği SVD tabanlı yöntem genel hatlarıyla şöyle anlatılabilir. En küçük tekil değerleri yakınsayan öz vektörler gürültü bilgisini temsil ederken en büyük tekil değerleri yakınsayan öz vektörler konuşma sinyalini temsil eder. Dolayısıyla gürültüsü giderilmiş konuşma sinyali en büyük tekil değerlerden oluşturulur. SVD tabanlı bu yöntem beyaz gürültü varlığında iyi sonuçlar vermektedir. Jensen vd. , Dendrinos vd.'nin yöntemini genişleterek renkli gürültüler için performansını artırmışlardır. Bu yöntem, bölümlü tekil değer ayrıştırma (QSVD) olarak bilinir. Bu yöntemde işlem olarak çok ağır kalmaktadır ve artık gürültüleri kontrol etmede kısıtları bulunmaktadır.

Ephraim ve Van Trees (EV) (1995) tarafından altuzay tabanlı gürültü giderme yöntemlerine farklı bir formül geliştirilerek artık gürültünün sebep olduğu ses sinyalindeki bozulmayı minimize edecek en uygun bir kestirici bulmayı hedeflemişlerdir. Karhunen-Loève dönüşümüyle (KLT) gürültülü sinyali gürültü ve ses sinyali altuzaylarına ayırıp vektör uzayında ayrıştırma işlemi yapmıştır. Gürültü altuzayındaki bileşenler sıfırlanırken konuşmanın olduğu altuzaydaki KLT bileşenlerinin oluşturduğu sinyal kazanç fonksiyonu kullanılarak düzenlenir ve bu sinyale ters KLT uygulanarak gürültüsü giderilmiş konuşma sinyali elde edilir. Objektif olmayan testlere göre altuzay tabanlı bu yaklaşım spektral çıkarma yöntemlerine göre daha iyi sonuç vermektedir. EV'nin oluşturduğu bu formülasyon gürültünün beyaz gürültü olduğuna dayanmaktadır.

Mittal ve Phamdo (2000) ve Rezayee ve Gazor (2001) isimli araştırmacılar EV'nin çalışmasını her bir frekans spektrumunda istenmeyen bileşenlerin bulunduğu renkli gürültü için genişletmişlerdir. Mittal ve Phamdo bunu yapmak için ilk olarak gürültülü konuşma verisini konuşma ağırlıklı bölüm ve gürültü ağırlıklı bölüm olarak iki sınıfa ayırmışlardır ve bu bölümlerde farklı KLT matrisleri kullanarak kestirim yapmaya çalışmışlardır. Objektif sonuçlara göre onların bu yaklaşımı EV'nin yöntemine göre daha iyi sonuçlar vermektedir. Daha sonra Rezayee ve Gazor ise EV'nin yaklaşımını gürültü vektörlerinin KLT kovaryans matrisini köşegen matrisine yakınsayarak renkli gürültüler için de genişletmişlerdir. Bu yaklaşım optimal altı bir kestirim sağlamaktadır fakat sübjektif sonuçlara göre dinleyicilerin büyük çoğunluğu ses kalitesine göre EV'nin kestiricisinden daha iyi olduğunu belirtmişlerdir.

### 2.5.1. SVD tabanlı yöntemler

Tekil Değer Ayırıştırma (SVD) tabanlı konuşma verisinde gürültü giderme yöntemi Dendrinos vd. (1991) tarafından önerilmiştir. Temel olarak veri matrisinin tekil vektörlerinin ve bu tekil vektörlere ait tekil değerlerin bazıları konuşma sinyaline ait iken diğerleri gürültüye ait olduğu kabul edilmektedir. Bu varsayımlarda yüksek tekil değerlere göre ve onların tekil vektörlerine göre sinyal tekrardan oluşturulurken düşük tekil değerlere sahip olan gürültü bilgisi yok sayılır. Yapılan bu işlem gürültünün beyaz gürültü olduğunu kabul etmektedir.

Gürültülü ses sinyali  $Y$ , temiz ses sinyali  $X$  ve gürültü sinyali  $D$  olmak üzere

$$Y = X + D \quad (2.26)$$

olarak tanımlanır. Çerçeve büyüklüğü  $N$ ,  $k = 0, \dots, N - 1$  olmak üzere  $y(k)$  gürültülü ses sinyali ve  $L$  deneysel olarak belirlenmek üzere gürültülü ses sinyali

$$Y = \begin{bmatrix} y(L-1) & y(L-2) & \dots & y(0) \\ y(L) & y(L-1) & \dots & y(1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \dots & y(N-L) \end{bmatrix} \quad (2.27)$$

olarak gösterilir.  $Y$ 'nin SVD'si alınırsa

$$Y = \sum_{i=1}^L \sigma_i u_i v_i^T \quad (2.28)$$

olarak elde edilir. Eşitlik 2.28'de  $i = 1, \dots, L$  'ye kadar küçükten büyüğe sıralanmış tekil değerler  $\sigma_i$ , sol ve sağ tekil vektörler  $u_i$  ve  $v_i$  olarak tanımlanmıştır. Beyaz gürültü için  $X$ 'in SVD'si

$$X = \sum_{i=1}^p \sigma'_i u_i v_i^T \quad (2.29)$$

olarak elde edilir. Burada tekil değerler  $\sigma_i'$  olarak tanımlanır ve gürültülü ses sinyali için tekil değer sayısı  $p \leq L$  olarak varsayılır. Gürültü beyaz olarak kabul edildiğinden temiz ve gürültülü ses sinyalinin kovaryans matrisi aynıdır. Teorem 2.1'e göre minimum norm tahmini  $\hat{Y}_p = \sum_{i=1}^p \sigma_i u_i v_i^T$  elde edilerek hesaplanır.

### Teorem 2.1

Eğer  $Y$   $M \times L$  boyutlarında  $L$  rankına sahip bir matris ise ve  $S_p$ ,  $M \times L$  boyutunda  $Y$  matrisine en yakın seti göstermek üzere  $Y$ 'ye en yakın matris olan  $\hat{Y}_p = \sum_{i=1}^p \sigma_i u_i v_i^T$  şeklinde gösterilir.

Yukarıdaki teorem kullanarak  $p$  rankına sahip matris elde edilir. Fakat ne yazık ki bu matris veri matrisi olarak doğrudan kullanılamamaktadır. Bu sebeple Toeplitz yaplanması gerekir. Köşegen elemanları ortalama değerleriyle değiştirilirse bu problem aşılmaktadır. Böylelikle Toeplitz veri matrisi

$$\hat{X}_p = \begin{bmatrix} \hat{x}_p(L-1) & \hat{x}_p(L-2) & \dots & \hat{x}_p(0) \\ \hat{x}_p(L) & \hat{x}_p(L-1) & \dots & \hat{x}_p(1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}_p(N-1) & \hat{x}_p(N-2) & \dots & \hat{x}_p(N-L) \end{bmatrix} \quad (2.30)$$

olarak gürültüsü giderilmiş  $\hat{x}(k)$ 'ya yakınsayan matris elde edilir.

Burada  $p$ 'nin tahmini çok önemli bir işlemdir. Dendrinos  $p$  rankını basit sıfır çaprazlama yöntemi ile bulmuştur. Bu yöntem temiz sinyal ve onun tekrardan düzenlenen hali arasındaki enerji farkını sıfırlamaya dayanır.

### 2.5.2. KLT tabanlı yöntemler

Ephraim ve Van Trees (EV) tarafından önerilen KLT tabanlı yöntem şüphesiz konuşma verisinde gürültü giderme algoritmaları arasında en önemli yere sahip çalışmalardan bir tanesidir. Gürültünün beyaz gürültü olduğunu varsayarak, EV gürültü

gidermeyi sınırlamalarına göre optimizasyon problemi olarak formüle etmiştir. Sınırlamalarına göre 2 kestirici bulunmaktadır. Bunlar zaman alanı sınırlı kestirici ve spektral alanı sınırlı kestiricidir (Commins, 2005).

### **2.5.2.1. Zaman alanı sınırlı (TDC) kestirici**

Temiz sinyal  $x$  olmak üzere doğrusal kestirici  $\hat{x} = Hy$  olarak gösterebiliriz. Burada  $H$  matrisi  $K \times K$  boyutlarında matristir. Artık sinyal

$$r = \hat{x} - x = (H - I)x + Hw = r_x + r_w \quad (2.31)$$

olarak gösterilebilir. Eşitlik 2.31'de sinyaldeki bozulma  $r_x = (H - I)x$  ve artık gürültü  $r_w = Hw$  olarak ifade edilir.  $r_x$  vektörünün enerjisi

$$\bar{\varepsilon}_x^2 = trE\{r_x r_x^\# \} = tr\{(H - I)R_x(H - I)^\#\} \quad (2.32)$$

olarak hesaplanır. Benzer şekilde  $r_w$  vektörünün enerjisi

$$\bar{\varepsilon}_w^2 = trE\{r_w r_w^\# \} = \sigma_w^2 tr\{HH^\#\} \quad (2.32)$$

olarak hesaplanır. TDC Doğrusal kestirici  $\frac{1}{K}\bar{\varepsilon}_x^2 \leq \alpha\sigma_w^2$  ( $0 \leq \alpha \leq 1$ ) şartını sağlaması kaydıyla *minimum*  $\bar{\varepsilon}_w^2$  olarak formüle edilir. Birkaç hesaplamadan sonra EV TDC filtreyi şu şekilde verirler.

$$H_{TDC} = \frac{R_x}{(R_x + \mu\sigma_w^2 I)} \quad (2.33)$$

KLT uygulayarak elde edilen sinyal  $\hat{x}_{TDC} = H_{TDC}y$  kazanç fonksiyonuyla düzenlendikten sonra ters KLT uygulanıp gürültüsü giderilmiş sinyal oluşturulur.

### **2.5.2.2. Spektral alanı sınırlı (SDC) kestirici**

Spektral alanı sınırlı (SDC) kestiricide temel olarak sinyaldeki bozulmayı minimize etmek esastır. Ses sinyaline benzer bir spektrum yapılabileceği için müzikal gürültü oluşmadan artık gürültü ses sinyali tarafından maskelenebilir. Gürültü sinyalini temizlemek için Eşitlik 2.34'teki gibi bir filtre oluşturulmuştur. Bu filtre uygulama olarak TDC kestiricisi ile benzerdir.

$$\text{minimum}_{\bar{\varepsilon}_x} \begin{cases} E\{|u_k \# r_w|^2\} \leq \alpha_k \sigma_w^2, & k = 1, \dots, M \\ E\{|u_k \# r_w|^2\} = 0 & , \quad k = M + 1, \dots, K \end{cases} \quad (2.34)$$

### 3. ORTAK VEKTÖR YAKLAŞIMI

Ortak Vektör Yaklaşımı (OVY), ses ve görüntü tanıma uygulamalarında kullanılan bir altuzay sınıflandırma yöntemidir. Bir sınıfa ait öznelik vektörlerindeki farklılıklar çıkarıldıktan sonra geriye kalan ve o sınıf için değişmeyen özelliklerin oluşturduğu vektöre ortak vektör denir. Bir sınıfın içerisindeki farklılıklara, o sınıfı oluşturan kişilerin seslerinin kaydedilmesi esnasındaki gürültü, kişilerin konuşma farklılıkları, cinsiyetleri hatta yaşları etkili olabilir (Ergin, 2004). OVY, temel olarak farklılıkların bulunup yok edilmesine ve ortak vektörün bulunmasına dayanmaktadır. OVY, bugüne kadar ses tanıma uygulamalarında (Gülmezoğlu vd., 1999, 2001, 2007), konuşmacı tanıma (Gülmezoğlu ve Barkana, 1998), yüz tanıma (Çevikalp vd., 2005), elektrik motorlarında arıza tespiti (Gülmezoğlu and Ergin, 2007), istenmeyen e-posta tanıma (Günel vd, 2006), görüntüde gürültü giderme (Özkan and Seke, 2015) uygulamalarında kullanılmıştır. OVY'deki en önemli ve kritik kısım, farklılık ve farksızlık altuzaylarının oluşturulması ve bu altuzaylara olan izdüşümlerin hesaplanmasıdır. OVY, öznelik vektörlerinin sayısı ve bu öznelik vektörlerinin boyutlarına göre iki farklı durumda incelenir.

#### 3.1. Yeterli Veri Durumu

Yeterli veri durumunda, sınıfların eğitim kümesindeki öznelik vektör sayısı ( $m$ ), öznelik vektörlerin boyutundan ( $n$ ) daha büyüktür ( $m > n$ ). Bu koşullar altında yetersiz veri durumunda tanımlanan farklılık veya farksızlık altuzayları tanımlamaları yapılamamaktadır (Edizkan, 1999). Bu durumda ortak vektör, kovaryans matrisi kullanılarak bulunabilir. OVY'nin, kovaryans matrisi kullanılarak gerçekleştirilir. Bu kovaryans matrisi

$$\Phi = \sum_{i=1}^m [(\mathbf{a}_i - \mathbf{a}_{ort})(\mathbf{a}_i - \mathbf{a}_{ort})^T] \quad (3.1)$$

Burada,  $\mathbf{a}_{ort}$  örüntünün eğitim kümesindeki bir sınıfa ait öznelik vektörlerinin ortalama vektörünü,  $\Phi$  ise o örüntü sınıfına ait kovaryans matrisini göstermektedir.

Hesaplanan kovaryans matrisinin özdeğer-özvektör ayrıştırması yapılır ve bu işlemden sonra  $n$  adet sıfırdan büyük özdeğer elde edilir. Bu özdeğerler, büyükten küçüğe doğru sıralanır. Özdeğerler içerisinde en büyük  $k$  tane özdeğer aşağıdaki eşitsizlik yardımı ile seçilir.

$$\frac{(\sum_{i=k+1}^n \lambda_i)}{(\sum_{i=1}^n \lambda_i)} \leq Y \quad (3.2)$$

Seçilen özdeğer sayısı  $k$ , özdeğerler  $\lambda_i$ , özdeğer seçiminde kullanılacak olan sabit yüzde değeri ise  $Y$  ile ifade edilmektedir (Swets ve Weng, 1996). Sabit yüzde değer, ne kadar büyükse seçilen özdeğer sayısı o kadar az, ne kadar küçükse özdeğer sayısı o kadar fazladır.

Kovaryans matrisinin  $k$  adet özdeğerine karşılık gelen  $k$  tane özvektör vardır. Bu özvektörler, örüntü sınıfının farklılık altuzayını gererler. Geriye kalan  $(n - k)$  tane özvektör de o örüntü sınıfının farksızlık altuzayını gerecektir. Bu durumda örüntü sınıfına ait herhangi bir  $\mathbf{a}_i$  öznitelik vektörünün, o örüntü sınıfına ait farklılık altuzayına izdüşümü ( $\mathbf{a}_{i, fark}$ ) şu şekilde hesaplanır. (Gülmezoğlu vd., 2007):

$$\mathbf{a}_{i, fark} = \sum_{i=1}^k [(\mathbf{a}_i^T \mathbf{u}_i) \mathbf{u}_i] \quad (3.3)$$

Burada,  $\mathbf{u}_i$ 'ler örüntü sınıfının farklılık altuzayını geren özvektörlerdir. Örüntü sınıfına ait ortak vektör ise, ortalama vektörü yardımı ile aşağıdaki gibi hesaplanır.

$$\mathbf{a}_{ort} = \sum_{i=k+1}^n [(\mathbf{a}_{ort}^T \mathbf{u}_i) \mathbf{u}_i] \quad (3.4)$$

### 3.2. Yetersiz Veri Durumu

Yetersiz veri durumunda, eğitim kümesindeki her bir sınıf için öznitelik vektörü sayısı ( $m$ ), türetilen öznitelik vektörlerinin boyutundan ( $n$ ) küçük ya da eşittir. ( $m \leq n$ )



Bir örüntü sınıfının eğitim kümesinde bulunan öznitelik vektörleri sırasıyla  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$  olsun. Doğrusal bağımsız olarak kabul edilen bu öznitelik vektörleri şu şekilde yazılabilir.

$$\mathbf{a}_i = \mathbf{a}_{i,fark} + \mathbf{a}_{ortak} + \boldsymbol{\varepsilon}_i \quad i = 1, 2, \dots, m \quad (3.5)$$

Burada,  $\mathbf{a}_{i,fark}$  vektörü, örüntü sınıfındaki herhangi bir öznitelik vektörünün diğer öznitelik vektörleri ile arasındaki farklarının yanı sıra, faz ve geçici farklılıkları gösterir.  $\mathbf{a}_{ortak}$ , bir örüntü sınıfına ait ortak vektörü ve  $\boldsymbol{\varepsilon}_i$  hata vektörünü gösterir. Eşitlik 3.1'de,  $(2m + 1)$  tane bilinmeyen vektörle beraber,  $m$  tane vektör denklemi vardır. Bu sebeple  $\mathbf{a}_{ortak}$ ,  $\mathbf{a}_{i,fark}$  ve  $\boldsymbol{\varepsilon}_i$  için sonsuz sayıda çözüm olduğu söylenebilir.  $\mathbf{a}_{ortak}$ 'ı elde etmek için, öncelikle, bir örüntü sınıfına ait fark vektörlerini ( $\mathbf{b}_j$ ) tanımlayalım.

$$\mathbf{b}_j = \mathbf{a}_{j+1} - \mathbf{a}_1 \quad j = 1, 2, \dots, m - 1 \quad (3.6)$$

Burada,  $\mathbf{a}_1$ , örüntü sınıfında seçilen bir referans vektördür.

Bu tanımlamanın ardından, Gram-Schmidt dikgenleştirme yöntemi kullanılarak fark vektörlerinden birimdik taban vektör kümesi  $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{m-1}\}$  oluşturulur.

$$\begin{aligned} \mathbf{d}_1 &= \mathbf{b}_1 & \mathbf{z}_1 &= \frac{\mathbf{d}_1}{\|\mathbf{d}_1\|} = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|} \\ \mathbf{d}_2 &= \mathbf{b}_2 - \langle \mathbf{b}_2, \mathbf{z}_1 \rangle \mathbf{z}_1 & \mathbf{z}_2 &= \frac{\mathbf{d}_2}{\|\mathbf{d}_2\|} \\ & \vdots & & \vdots \\ & \vdots & & \vdots \\ \mathbf{d}_{m-1} &= \mathbf{b}_{m-1} - \sum_{i=1}^{m-2} \langle \mathbf{b}_{m-1}, \mathbf{z}_i \rangle \mathbf{z}_i & \mathbf{z}_{m-1} &= \frac{\mathbf{d}_{m-1}}{\|\mathbf{d}_{m-1}\|} \end{aligned} \quad (3.7)$$

Bu birimdik taban vektörlerin gerdiği altuzay, o örüntü sınıfı için farklılık altuzayıdır. Örüntü sınıfına ait herhangi bir  $\mathbf{a}_i$  öznitelik vektörünün, o örüntü sınıfına ait farklılık altuzayına izdüşüm hesaplanırsa,  $\mathbf{a}_{i,fark}$  bulunmuş olur.

$$\mathbf{a}_{i,fark} = \langle \mathbf{a}_i \mathbf{z}_1 \rangle \mathbf{z}_1 + \langle \mathbf{a}_i \mathbf{z}_2 \rangle \mathbf{z}_2 + \cdots + \langle \mathbf{a}_i \mathbf{z}_{m-1} \rangle \mathbf{z}_{m-1} \quad \forall i = 1, 2, \dots, m \quad (3.8)$$

Bu hesaplamadan sonra  $\mathbf{a}_{ortak}$ , herhangi bir öznitelik vektörünün taban vektör kümesi üzerindeki izdüşümlerinin kendisinden çıkartılmasıyla elde edilir. (Gülmezoğlu vd, 1999,2001)

$$\mathbf{a}_{ortak} = \mathbf{a}_i - \mathbf{a}_{i,fark} \quad \forall i = 1, 2, \dots, m \quad (3.9)$$

$\mathbf{a}_{ortak}$ , 'i' indisinden ve referans vektöründen bağımsızdır.

Gram-Schmidt dikgenleştirme yönteminden farklı olarak altuzay tekniğiyle de ortak ve farklılık vektörü bulunabilir. Bu teknikte ortak vektör ve özvektörün kovaryans matrisi arasındaki ilişki kullanılarak ortak vektör hesaplanır.

Eğitim setindeki öznitelik vektörleri Eşitlik 3.5.'te verildiği gibi olmak üzere, her bir sınıf için kovaryans matrisi

$$\Phi = \sum_{i=1}^m [(\mathbf{a}_i - \mathbf{a}_{ortak})(\mathbf{a}_i - \mathbf{a}_{ortak})^T] \quad (3.10)$$

olarak hesaplanır. Öznitelik altuzayı farklılık  $B$  ve farksızlık  $B^\perp$  altuzay olmak üzere birbirine dik iki altuzaya ayrılır. Özvektörler büyükten küçüğe doğru ( $\lambda_n \geq \lambda_{n-1} \geq \dots, \lambda_{n-m+2} \geq \lambda_{n-m+1} \geq \dots, \lambda_2 \geq \lambda_1$ ) sıralandığında  $n - m + 1$  tanesinin sıfır olduğu görülür. Farklılık  $B$  ve farksızlık  $B^\perp$  altuzayları sırasıyla sıfır olmayan ve sıfır olan özdeğerleri ilgili özvektörleri kapsar. Bir sınıf için ortak vektör herhangi bir öznitelik vektörünün farksızlık altuzayına izdüşümü ile

$$\mathbf{a}_{ortak} = P^\perp \mathbf{a}_i \quad \forall i = 1, 2, \dots, m \quad (3.11)$$

olarak bulunur. İzdüşüm matrisi

$$P^\perp = \sum_{j=1}^{n-m+1} \mathbf{u}_j \mathbf{u}_j^T \quad (3.12)$$

olarak bulunur. Ortak vektör yaklaşımı dışındaki yöntemler kovaryans matrisinin tersini hesaplamayı gerektirir (Zhang vd., 2010). Fakat yetersiz veri durumu için bütün sınıflarda kovaryans matrisinin tersi alınamaz. Ortak vektör yaklaşımı bu duruma görüldüğü gibi bir çözüm getirmiştir.

## 4. ORTAK VEKTÖR YAKLAŞIMIYLA SES VERİSİNDE GÜRÜLTÜ GİDERME

Üçüncü bölümde ortak vektör yaklaşımının ses ve görüntü tanıma uygulamalarında kullanılan altuzay tabanlı bir yöntem olduğunu ve konuşmacı tanıma, ses tanıma, yüz tanıma gibi uygulamalarının olduğu bahsedilmişti. Bu bölümde ise ortak vektör yaklaşımıyla gürültülü bir konuşma verisinde gürültünün giderilmesi detaylı bir şekilde açıklanacaktır. Yöntemin motivasyonu, matematiksel olarak detaylandırılması ve akış şeması verilecektir.

### 4.1. Motivasyon

Bugüne kadar konuşma verisinde gürültü giderme ile alakalı olarak birçok çalışma bulunmasına rağmen hiçbir çalışma insan kulağı kadar gürültülü ortama adapte olma özelliği göstermez. Gürültü karakteristiğinin düzenli olduğu beyaz gürültü gerçek hayatta karşılaşılma sıklığı az olan bir gürültü çeşididir. Buna karşın karakteristiği düzensiz olan ve literatüre renkli gürültü olarak geçmiş olan gürültüler sıklıkla karşılaşılan gürültü çeşididir. Son zamanlarda yapılan çalışmalar bu gürültü çeşidine yöneliktir.

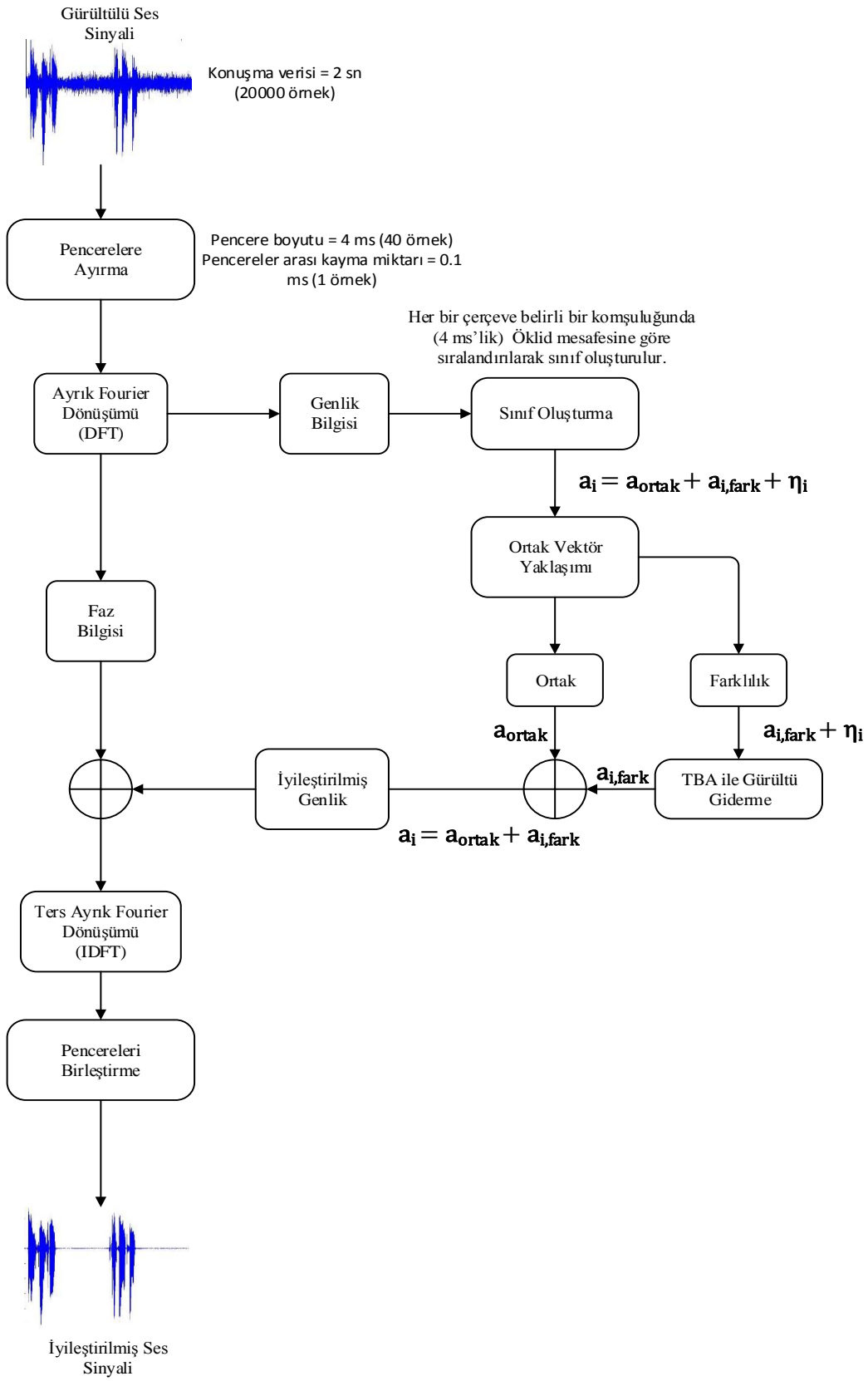
İkinci bölümde bugüne kadar olan yapılmış önemli gürültü giderme yöntemleri ve bu yöntemlerin dezavantajlarından bahsedilmişti. Şüphesiz yapılan çalışmalarda en sık karşılaşılan problemlerden biri müzikal gürültüdür. Bu problemin olmadığı yöntemler daha umut verici olarak görülmektedir.

Gerçek hayatta karşılaşılabileceğimiz gürültüler her frekans spektrumunda düzgün olarak dağılmaz ve değişkendir. Bu nedenle insan kulağından ilhamla her frekans spektrumu küçük parçalar halinde ele alınırsa ideale yakın bir gürültü giderme gerçekleştirilmiş olur.

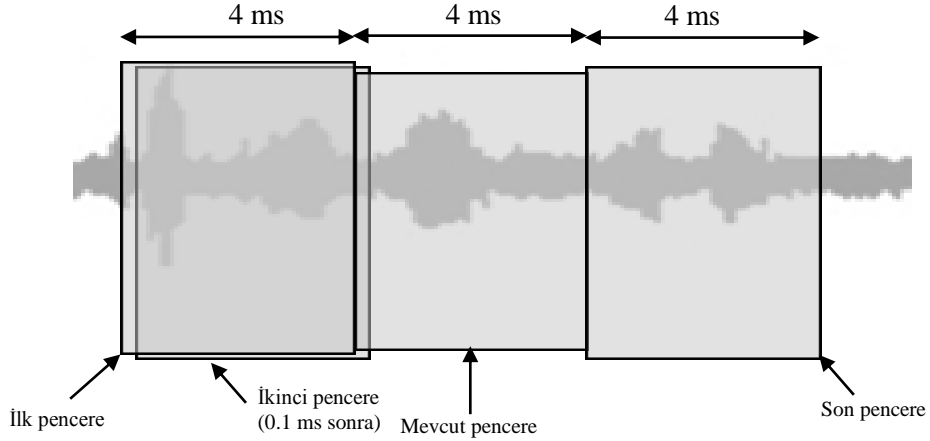
Yukarıda anlatılanlar bu tez kapsamında yapılan çalışmanın çıkış noktası olmuştur. Konuşma verisinde giderilecek gürültünün günlük hayatta karşılaşılabilecek gürültüler olması esas alınmış ve gürültülü sesin her frekans spektrumunda gürültü olabileceği düşünülüp en küçük parçalar halinde işlenip müzikal gürültünün oluşmaması amaçlanmıştır.

## 4.2. Önerilen Yöntem

Gürültülü konuşma verisini  $x_\eta(n) = x(n) + \eta(n)$  olarak tanımlayalım. Burada  $x$  gürültüsüz konuşma verisini,  $\eta$  ise beyaz veya renkli gürültüyü temsil etmektedir. Gürültülü konuşma verisi her bir frekans spektrumunda gürültü barındırabileceği için yapılacak işlem öncesinde küçük örnekler halinde pencerelere ayrıştırılır, pencereler  $a_\eta(n) = a(n) + \eta(n)$  olarak gösterilebilir. Burada  $a_\eta(n)$  pencerelere ayrılmış gürültülü ses sinyali,  $a(n)$  ise gürültüsüz ses sinyalidir. Her bir pencere Ayrık Fourier Dönüşümü (DFT) ile zaman alanından frekans alanına dönüştürülür. Frekans alanında bu pencereler gerçek ve sanal olarak  $a = a_r + ia_i$  şeklinde ifade edilir. Bu pencerelerde genlik  $\sqrt{(a_r^2 + a_i^2)}$ , faz ise  $\arctan(a_i/a_r)$  olarak hesaplanır. Hesaplanan bu bilgilerden genlik verisinde gürültü olduğu kabul edilir. Faz bilgisi genel olarak gürültü içermediği veya içerse bile az miktarda içerdiği için saklanır. Örneklere ayrılmış her bir genlik bilgisinin belli bir komşuluğunda Öklid uzaklığına göre sıralama yapılarak en çok benzeyen  $n$  adedi ile bir sınıf oluşturulur. Bu sınıfta ortak vektör yaklaşımı uygulanarak ortak ( $\mathbf{a}_{ortak}$ ) ve farklılık ( $\mathbf{a}_{fark} + \boldsymbol{\eta}_i$ ) kısımları oluşturulur. Yöntem kapsamında ortak vektör yaklaşımı uygulanırken yetersiz veri ( $m \leq n$ ) durumuna göre gerçekleştirme yapılmıştır. Ortak vektör ve farklılık vektörü bu durum göz önüne alınarak bulunmaktadır. Bu çalışmada altuzay tekniğiyle ortak vektör ve farklılık vektörü bulunmuştur. Bu kısımlarla beraber örneklenmiş genlik bilgisi  $\mathbf{a}_i = \mathbf{a}_{ortak} + \mathbf{a}_{fark} + \boldsymbol{\eta}_i$  olarak gösterilir. Gürültünün farklılık vektöründe olduğu kabul edilerek Temel Bileşen Analiziyle (TBA) veya başka herhangi bir yöntemle gürültü, farklılık vektöründen ayrılır. Gürültüsü giderilmiş genlik bilgisi  $\mathbf{a}_i = \mathbf{a}_{ortak} + \mathbf{a}_{fark}$  olarak tanımlanır ve saklanan faz bilgisiyle birleştirilir. Ters Fourier dönüşümüyle zaman alanına dönüştürülen örneklenmiş konuşma verisi pencereleri örtüşen kısımlarıyla birlikte tekrardan birleştirilerek gürültüsü giderilmiş konuşma verisi elde edilmiş olur. Yöntemin akış şeması Şekil 4.1. de verilmiştir.

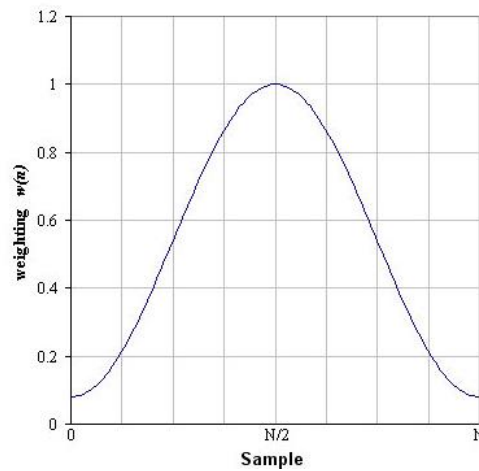


Şekil 4.1. Önerilen yöntemin akış şeması



Şekil 4.2. Belirli bir komşulukta sınıf oluşturma işlemi

Yöntemin uygulanmasındaki en önemli aşamalardan biri sınıf oluşturma aşamasıdır. Burada önemli olan oluşturulan sınıfta ses bilgisinin olmasıdır. Konuşma verisinden ayrılan pencere uzunlukları, pencereler arasındaki örtüşme oranı ve sınıf oluşturulurken arama yapılacak komşuluk oranı, Öklid mesafesine ( $\sqrt{(p_i - q_i)^2}$ ) göre sıralanan elemanlardan hangilerinin kullanılacağı önemli parametrelerimizdir. Şekil 4.2'den de görüleceği üzere belli bir komşulukta arama yapılarak benzer elemanlardan sınıf oluşturulur. Yapılan çalışmada 4 ms'lik pencerelerde 0.1 ms'lik örtüşmeyle 4 ms'lik komşuluğunda Öklid mesafesine göre belirlenen 80 vektörün en küçük 40'ı kullanılarak sınıf oluşturulmuştur. Böylelikle ortak vektör yaklaşımında yetersiz veri durumunun olması sağlanmıştır.



Şekil 4.3. Hamming penceresi ağırlıkları

Pencere oluşturulurken kullandığımız pencere fonksiyonu Şekil 4.3.'de gösterimi bulunan Hamming penceresidir. Örnek (sample) sayısı  $n$  olmak üzere bu pencerenin matematiksel ifadesi Eşitlik 4.1'de verilmiştir. Hamming pencereleme ile oluşturulan pencerelerin merkezi vurgulanmaktadır. Diğer gürültü giderme yöntemlerinde de sıklıkla kullanılan bir pencere olduğu için tercih edilmiştir.

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right); \quad n = 0 \dots N-1 \quad (4.1)$$

Yöntemin test aşamasında 20000 örnekli bir konuşma verisi için pencere boyutu 40 örnek (4 ms), örtüşme miktarı 1 örnek (0.1 ms), arama yapılan komşuluk ise  $\pm 40$  örnek (4 ms) olarak seçildiğinde en iyi sonuçlar elde edilmiştir. Bu parametreler oluşturulan sınıf içerisindeki elemanların seçimini etkilediğinden yöntemin başarısını doğrudan etkilemektedir. Örtüşme miktarı pencere sayısını artırdığından işlem yükünde artışa sebep olmaktadır. Fakat en fazla örtüşme miktarıyla benzer vektörlerin atlanmaması hedeflenmiştir.



## 5. GERÇEKLEŐTİRME VE PERFORMANS ÖLÇÜMÜ

Bu bölümde testlerin yapıldığı veri tabanı özellikleri detaylı bir şekilde anlatılacaktır. Önerilen yöntemin uygulanmasına yönelik yapılan testler hakkında bilgi verilecek ve parametrelerin test sonuçlarına etkileri incelenecektir. Nihai olarak belirlenen parametrelere göre sonuçlar zaman ve frekans alanında gösterilecektir. İyileştirilmiş ses sinyalinin kalite ve anlaşılabilirliğinin ölçümü için literatürde bulunan objektif kriterler incelenecek ve bu kriterlere göre literatürdeki önemli yöntemlerle önerilen yöntemin kıyaslamaları yapılacaktır.

### 5.1. Veri Tabanı

Bu çalışmada kullanılan ses ve gürültü dosyaları, NOIZEUS veri tabanından alınmıştır. 8 kHz ile örneklenmiş ve 16 bit olarak kaydedilmiş, yaklaşık 2 sn uzunluğunda, 3 erkek ve 3 kadın konuşmacı tarafından seslendirilmiş 30 farklı İngilizce cümle kullanılmıştır. Bu cümlelerin her biri 8 farklı gürültü çeşidi (havaalanı, kalabalık, araba, sergi salonu, restoran, tren istasyonu, sokak, tren) ve 4 farklı SNR gürültü seviyesi (0 dB, 5 dB, 10 dB, 15 dB) ile veri tabanında bulunmaktadır. Bunlara ek olarak bu cümlelere 4 farklı gürültü seviyesiyle beyaz gürültü de eklenerek genişletilmiştir. Bu veri tabanı alanında sıklıkla kullanılan en çok ses dosyasına ve gürültü çeşidine sahip olan bir veri tabanıdır. Veri tabanında bulunan cümleler ve konuşmacılar Çizelge 5.1.'de verilmiştir. Kullanılan gürültü çeşitleri gerçek hayatta karşılaşılabilecek çoğu ortamı modellemiştir.

Çizelge 5.1. NOIZEUS veri tabanı içerisindeki cümle metinleri ve özellikleri

Dosya Adı	Konuşmacı	Cinsiyet	Cümle Metni
sp01.wav	CH	E	The birch canoe slid on the smooth planks.
sp02.wav	CH	E	He knew the skill of the great young actress.
sp03.wav	CH	E	Her purse was full of useless trash.
sp04.wav	CH	E	Read verse out loud for pleasure.
sp05.wav	CH	E	Wipe the grease off his dirty face.
sp06.wav	DE	E	Men strive but seldom get rich.
sp07.wav	DE	E	We find joy in the simplest things.
sp08.wav	DE	E	Hedge apples may stain your hands green.
sp09.wav	DE	E	Hurdle the pit with the aid of a long pole.
sp10.wav	DE	E	The sky that morning was clear and bright blue.
sp11.wav	JE	K	He wrote down a long list of items.
sp12.wav	JE	K	The drip of the rain made a pleasant sound.
sp13.wav	JE	K	Smoke poured out of every crack.
sp14.wav	JE	K	Hats are worn to tea and not to dinner.
sp15.wav	JE	K	The clothes dried on a thin wooden rack.
sp16.wav	KI	K	The stray cat gave birth to kittens.
sp17.wav	KI	K	The lazy cow lay in the cool grass.
sp18.wav	KI	K	The friendly gang left the drug store.
sp19.wav	KI	K	We talked of the sideshow in the circus.
sp20.wav	KI	K	The set of china hit the floor with a crash.
sp21.wav	SI	E	Clams are small, round, soft and tasty.
sp22.wav	SI	E	The line where the edges join was clean.
sp23.wav	SI	E	Stop whistling and watch the boys march.
sp24.wav	SI	E	A cruise in warm waters in a sleek yacht is fun.
sp25.wav	SI	E	A good book informs of what we ought to know.
sp26.wav	TI	K	She has a smart way of wearing clothes.
sp27.wav	TI	K	Bring your best compass to the third class.
sp28.wav	TI	K	The club rented the rink for the fifth night.
sp29.wav	TI	K	The flint sputtered and lit a pine torch.
sp30.wav	TI	K	Let's all join as we sing the last chorus.

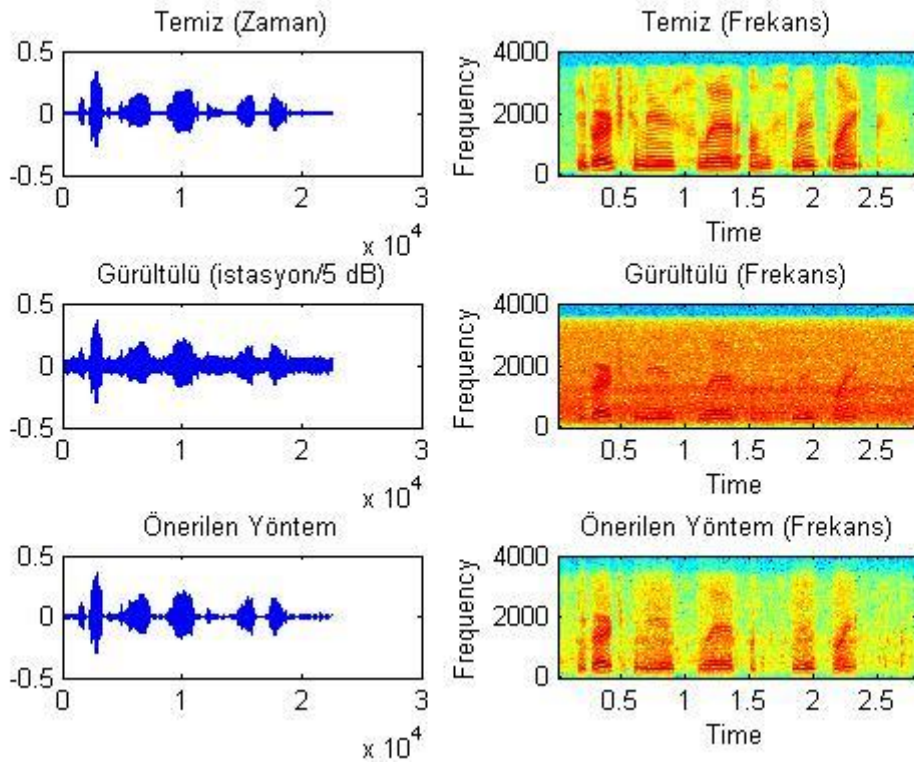
## 5.2. Gerçekleştirme

Bu çalışma kapsamında tüm yöntemlerin uygulanması, değerlendirmesi ve kıyaslama aşamaları, matematik hesaplamaları ve görselleştirme MATLAB paket programı ortamında yapılmıştır.

Önerilen yöntemin uygulanmasında kullanılan veri tabanından dolayı 9 ayrı gürültü 4 farklı gürültü seviyesi kullanılmıştır. Yapılan testlerin önemli bir kısmı yöntemin her ortam, gürültü çeşidine göre cinsiyet ayrımı gözetmeden en iyi sonuçları veren tek bir parametre bulmasına yöneliktir. Bu sebeple seçilen parametrelerin etkileri bu bölümde incelenecektir.

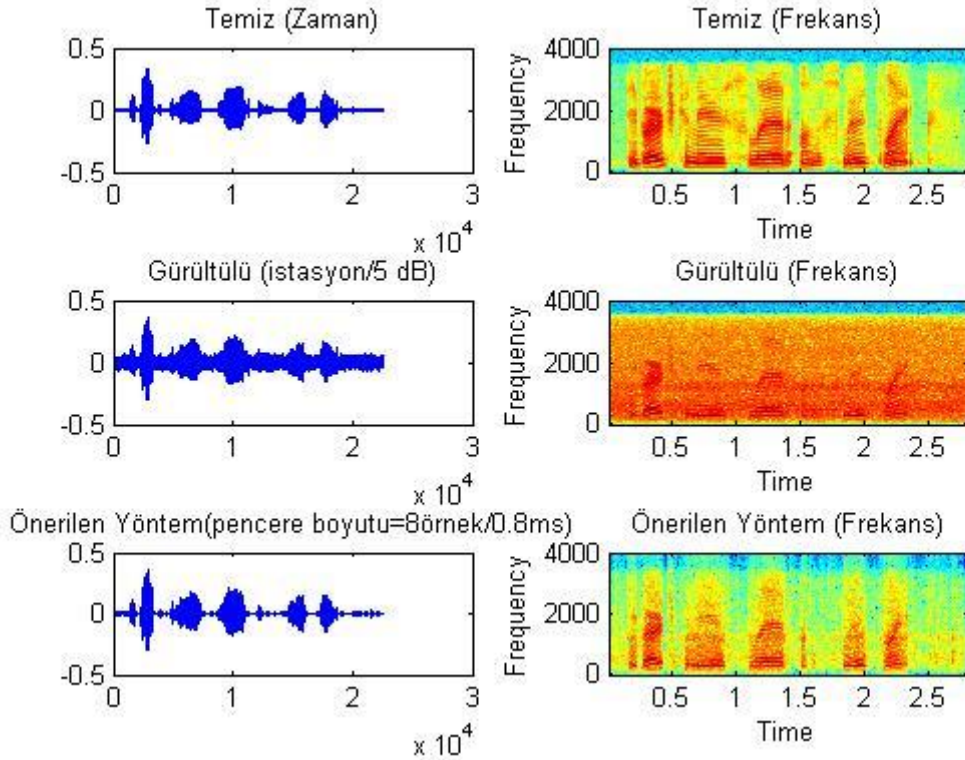
### 5.2.1. Pencere boyutunun gürültü gidermeye etkisi

Yöntemin uygulanmasında konuşma verisi pencerelere ayrılır ve bu pencere içinde belli bir komşulukta Öklid mesafesine göre sıralama yapılarak eğitim seti sınıfları oluşturulur. Bu sebeple pencere boyutu sınıf oluşturulmasında etkili bir parametredir. Pencere sayısının fazla olması kullanılabilir veri sayısını arttırmaktadır. Bu nedenle boyutun küçük ama yeterli sayıda eleman içermesi gerekmektedir. Yapılan çalışmada konuşma uzunluğu yaklaşık 2 sn olan bir veri tabanında çalışıldığı için pencere boyutu 40 örnek (4 ms) olarak belirlenmiştir. Bu bölümde yaptığımız uygulamaların hepsinde kullandığımız veri 'Noizeus' veri tabanındaki 'sp01.wav' dosyasına istasyon gürültüsü eklenmiş konuşma verisidir. Belirlediğimiz parametrelere göre gürültü giderme işlemi yapılmış ve bu işlem için SNR değeri 9.26 dB olarak bulunmuştur. Sırasıyla temiz, gürültülü ve önerilen yöntemle iyileştirilmiş ses sinyali zaman alanı ve frekans alanı olmak üzere Şekil 5.1. de gösterilmiştir.



Şekil 5.1. Pencere boyutu 40 örnek (4 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları

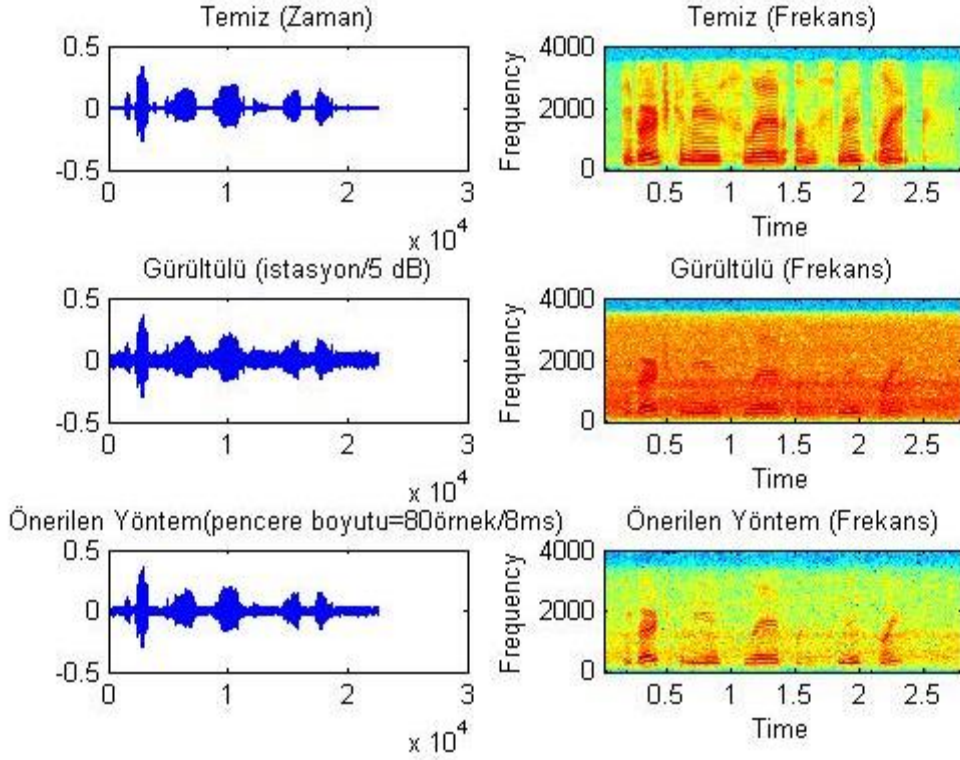
Bu değer testler sonucunda bulunmuş ve hem gürültü giderme oranı hem de işlem yükü olarak makul bir seviyede olması açısından ideale yakındır. Pencere boyutu küçültmek toplam pencere sayısını artıracığından kullanılacak veri sayısını ve işlem yükünü de artırmaktadır. Pencere boyutunu 8 örneğe düşürdüğümüz zaman SNR değerimiz 8.99 dB'ye düşmüştür. Yapılan uygulamanın sonuçları Şekil 5.2.'de verilmiştir.



Şekil 5.2. Pencere boyutu 8 örnek (0.8 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları

Detaylı olarak incelendiğinde frekans gösteriminde gürültünün olduğu bileşenler bir önceki pencere boyutu (40 örnek) ile yapılan çalışmaya göre daha az giderilmiştir. Dolayısıyla SNR değeri de düşmüştür.

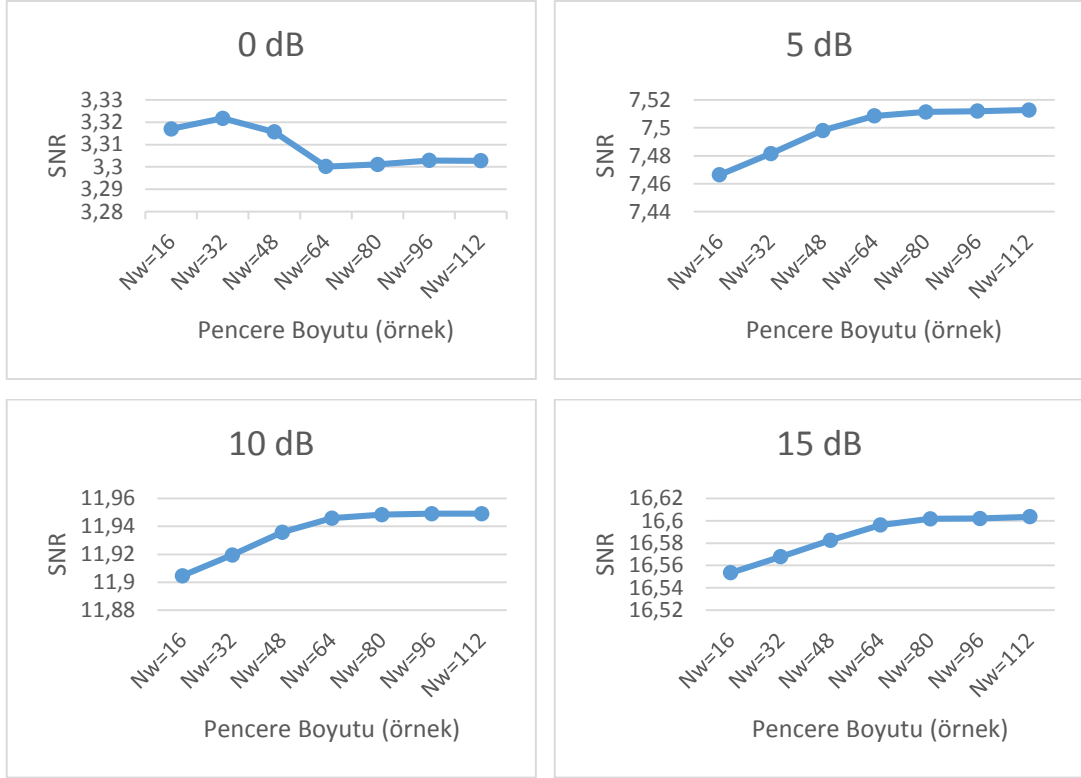
Pencere boyutunu artırdığımızda ise her pencerede konuşmanın gerekli kısmından daha fazlası olacağından Öklid mesafesi çok değişken olur ve sınıfı oluşturan eleman seçimi doğru olmaz. Eğer pencere boyutu çok fazla arttırılırsa gürültü bastırılmayacağından iyileştirme gerçekleşmez.



Şekil 5.3. Pencere boyutu 80 örnek (8 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları

Yapılan çalışmada pencere uzunluğu 80 örneğe çıkartılmıştır. Şekil 5.3.'te görüldüğü gibi gürültünün büyük kısmı giderilmemiş ve SNR değeri 7.77 dB ye düşmüştür.

Parametre değişimi işleminin tek bir ses dosyasında belirlenmesi yanlış sonuçlara sebebiyet verebilir. Bu sebeple tüm ses dosyaları için en uygun pencere boyutunu bulmak gerekmektedir. Tez kapsamında veri tabanında bulunan 30 ses dosyası ve 9 farklı gürültü çeşidinde parametre belirlemek için SNR ölçütüne göre testler yapılmıştır. Bu test sonuçlarının ortalaması alınarak en uygun parametre belirlenmiştir.



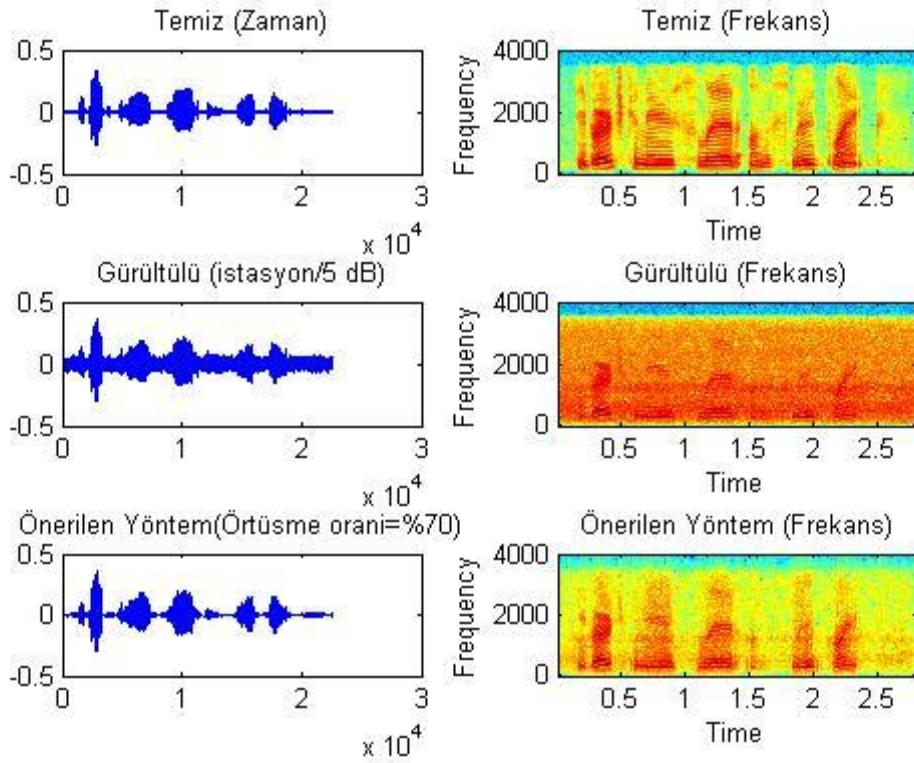
Şekil 5.4. Farklı gürültü seviyelerinde pencere boyutu - sinyal / gürültü oranı (SNR) grafiği

Şekil 5.4.'te görüldüğü gibi yapılan çalışmalarda 5 dB, 10 dB, 5 dB gürültü seviyelerinde gürültü giderme oranı bakımından pencere boyutu 48 ve 64 örnek olduğunda ortalama olarak en iyi sonuçları vermektedir. 0 dB gürültü oranında ise pencere boyutu 32 örnek olduğunda en iyi sonuçları vermektedir. Pencere boyutu, gürültü giderme oranını ve işlem yükünü doğrudan etkilediği için bu değerın uygun bir seviyede tutmak önemlidir. Tez kapsamında yapılan çalışmalarda pencere uzunluğu 40 örnek (4 ms) olarak belirlenmiştir ve tüm gerçekleştirme sonucundaki tüm iyileştirmeler bu değere göre yapılmıştır.

### 5.2.2. Örtüşme oranının gürültü gidermeye etkisi

Yapılan çalışmada konuşma verisinde sınıf oluştururken belli bir örnek kadar kayarak pencereler oluşturulur. Bu kayma miktarı her pencerenin bir önceki pencereyle örtüşecek kadar olmasını sağlamalıdır. Ne kadar fazla örtüşme olursa oluşturulacak olan seçilen sınıf elemanları da o kadar gürültü gidermeye yönelik olur. Çünkü spektrumu benzer pencerelerin komşuluklarda olması beklenir. Örtüşme oranını en fazla seçerek

oluşturacağımız eğitim setinin en uygun olması sağlanır. Bu tez kapsamında yapılan çalışmalarda kayma miktarı 1 örnek (0.1 ms) seçilerek, örtüşme oranı maksimum seviyeye çıkarıldı. Böylelikle eğitim setindeki sınıf sayısı artırıldı. Yapılan bu işlem eğitim setinin değişmez kısımlarını oluşturan ortak vektörünün ve gürültüsü giderilecek olan farklılık vektörünün doğru olarak hesaplanmasını sağlamıştır. Şekil 5.1.'de kayma miktarı 1 örnek (örtüşme oranı=%97.5) olarak belirlenen parametreler sonucundaki sonuçları göstermektedir. Kayma miktarını artırdığımızda (örtüşme oranını azaltığımızda) oluşturulan pencere sayısı azalacağından işlem yükü azalmaktadır. Bununla birlikte eğitim setindeki sınıflar içinde benzer vektör sayısı azalacağından gürültü giderme miktarı da azalacaktır. Bu durum Şekil 5.3.'te net olarak görülmektedir.

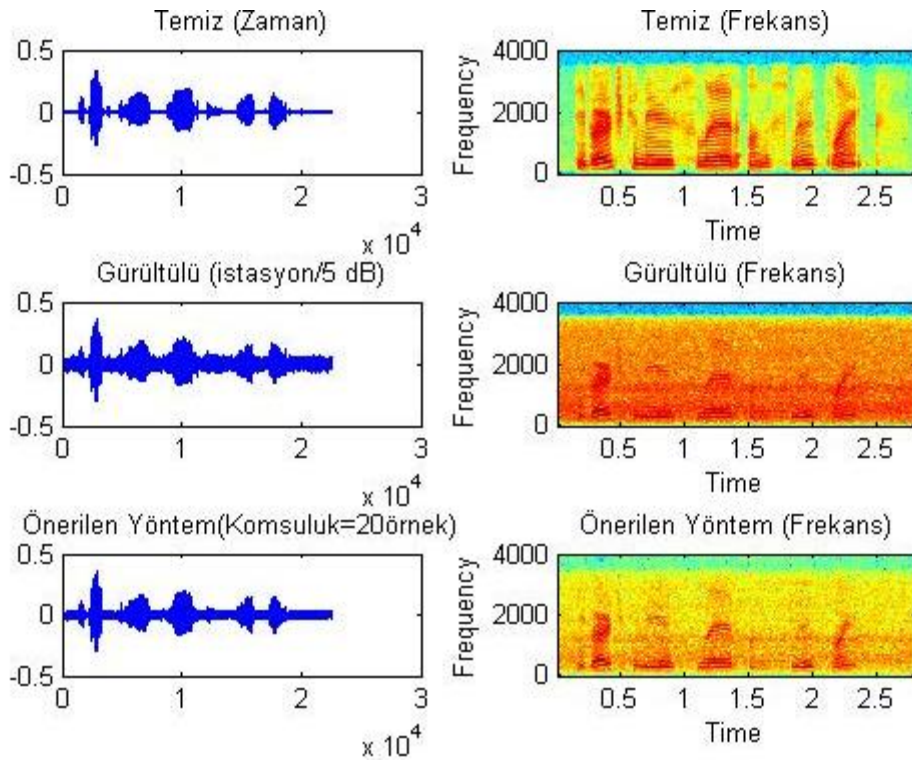


Şekil 5.5. Örtüşme oranı %70 olarak seçildiğinde gürültü giderme işlemi sonuçları

Örtüşme oranının %70'e düşürülmesi sonucunda SNR değeri de yaklaşık 0.5 dB değerinde düşmüştür. Tez kapsamında yapılan testlerde kayma miktarını 1 örnek seçerek örtüşme oranı maksimum yapılmış ve sonuçları azami seviyede artırıldığı görülmüştür.

### 5.2.3. Arama yapılan komşuluğun gürültü gidermeye etkisi

Sınıf oluştururken birbirine en yakın elemanlar Öklid mesafesine göre aynı sınıf içinde toplanır. Öklid mesafesine göre yakın olmayan elemanlar sınıfa dâhil edilmez. Bu elemanlar belli bir komşulukta aranır. Bu komşuluk az olursa yeterli sayıda eleman aranmaz, fazla olursa da işlem yükü artar ve aranan elemana yakın fakat yanlış elemanlar seçilmiş olur. Bu sebeple bu parametrenin iyi seçilmesi yöntemin gürültü gidermesi için önemlidir. Şekil 5.4.'te görüleceği gibi tez kapsamında yapılan testler sonucunda pencere boyutuyla beraber komşuluk 40 örnek olarak seçilmiştir. Arama yapılan komşuluğu azalttığımızda Şekil 5.6.'te görüldüğü gibi gürültü giderme işlemi istenen seviyelerde olmaz.



Şekil 5.6. Komşuluk 20 örnek (0.2 ms) olarak seçildiğinde gürültü giderme işlemi sonuçları

Arama yapılan komşuluğun azaltılması Şekil 5.5.'ten de görüldüğü gibi gürültünün büyük kısmının giderilmemesine neden olmuştur. SNR değeri 7.6 dB ye kadar düşmüştür. Arama yapılan komşuluğun artırılması ise gereksiz olarak işlem yükünü artırmakta ve sınıf içi elemanların yanlış seçilmesini sağlar. Bu nedenle yöntemin işlevliliğini yitirmesine neden olur.



### 5.3. Performans ölçümü

Belirlenen parametrelere göre Noizeus veri tabanında bulunan 30 ses dosyasına 9 farklı gürültü çeşidi, 4 farklı gürültü seviyesinde eklenmiş ve gürültü giderme işlemi bu ses dosyalarında yapılmıştır. Bu bölümde yapılan çalışmaya örnek olarak 'sp01.wav' ses dosyasına 5 dB seviyesinde 9 farklı gürültülü ortamda ortak vektör yaklaşımı ve literatürde bulunan 5 yöntemle gürültü giderilmiştir. Belirlenen 5 yöntem Bölüm 2'de anlatılmış olan kategorilerden seçilmiştir. Altuzay tabanlı olarak Jabloun (2003) tarafından önerilen algısal motive edilmiş altuzay tabanlı yöntem yöntem 1, istatistiksel model tabanlı olarak Hu ve Loizou (2004) tarafından önerilen Psiko-akustik motive edilmiş yöntem yöntem 2 olarak seçilmiştir. Wiener filtresi yöntemlerinden yine Hu ve Loizou tarafından önerilen dalgacık dönüşümüne dayalı wiener filtresi yöntemi yöntem 3, gürültü tahmini yöntemlerinden Rangachari ve Loizou (2006) tarafından önerilen minimum değişken kontrollü yinelemeli yöntem yöntem 4 ve Doblinger (1995) tarafından önerilen spektral takibe dayalı yöntem yöntem 5 olarak seçilmiştir.

### 5.4. Değerlendirme Kriterleri

Konuşma verisinde gürültü giderme işlemi yaptıktan sonra iyileştirmenin ölçüsü her ne kadar insan kulağıyla fark edilse de anlaşılabilirliği ve sinyal kalitesi olarak değerlendirilmesi için kriterlere ihtiyaç duyulmaktadır. Yapılan çalışmaları subjektif ve objektif ölçütlerle değerlendirmek mümkündür. Öznel ölçütler uygulama açısından çok zaman almaktadır. Bu sebeple nesnel kriterler tercih edilmektedir. Yayınlarda birçok nesnel değerlendirme yöntemi bulunmaktadır. Bu değerlendirme yöntemlerinden tez kapsamında kullanılanlar kısaca açıklamaları ile aşağıda mevcuttur.

#### 5.4.1. Sinyal / gürültü oranı (SNR)

Sinyal kalitesini ölçmek için kullanılan en yaygın yöntem şüphesiz sinyal / gürültü oranıdır. Sinyal gücünün gürültü gücüne oranının desibel cinsinden hesaplanmasıyla bulunur. Eşitlik (5.1)'deki gibi hesaplanır.

$$SNR_{dB} = 10 \log \left( \frac{\sum_n s^2(n)}{\sum_n [s(n) - \hat{s}(n)]^2} \right) \quad (5.1)$$

$s(n)$  : Temiz sinyal

$\hat{s}(n)$  : İyileştirilmiş Sinyal

#### 5.4.2. Frekans ağırlıklı parçalı SNR (fwSNRseg)

Sinyal / Gürültü oranı kadar sıklıkla kullanılan bir diğer yöntem ise frekans ağırlıklı parçalı SNR yöntemidir. Frekans alanında değerlendirme yapan bu yöntem farklı frekans bantlarında farklı ağırlıklarla hesaplamalarını yapar. (Tribolet vd., 1978)

$$fwSNRseg = \frac{10}{m} \sum_{m=0}^{M-1} \frac{\sum_{j=1}^K B_j \log[F^2(m,j)/(F(m,j) - \hat{F}(m,j))^2]}{\sum_{j=1}^K B_j} \quad (5.2)$$

Yukarıdaki eşitlikte  $B_j$ ,  $j$ . frekans bandındaki ağırlık olmak üzere,  $K$  bant sayısını,  $M$  ayrılan toplam parça sayısını,  $F(m,j)$  temiz sinyalin  $m$ . parça,  $j$ . frekans bandındaki filtrelenmiş genlik değerini,  $\hat{F}(m,j)$  ise iyileştirilmiş sinyalin aynı parça ve frekans bandındaki genlik değerini temsil etmektedir.

#### 5.4.3. Algısal konuşma kalitesinin değerlendirilmesi (PESQ)

Algısal konuşma kalitesinin değerlendirilmesi (PESQ) yöntemi ITU-T (2000) tarafından standartlaştırılmış bir değerlendirme yöntemidir. Dünya çapında telekomünikasyon alanında sıklıkla kullanılan bir yöntemdir. Referans ve test edilecek sinyalleri zaman uyumundan sonra örnekler halinde analiz etmeye dayalı bir yöntemdir.

$$PESQ = a_0 - a_1 \cdot D_{ind} - a_2 \cdot A_{ind} \quad (5.3)$$

Eşitlik (5.3)'te  $A_{ind}$  ortalama asimetric karışıklık değeri,  $D_{ind}$  ortalama asimetric karışıklık değeri olmak üzere  $a_0$  değeri 4.5,  $a_1$  değeri 0.1,  $a_2$  değeri 0.0309 olarak ifade edilmektedir (Hu ve Loizou, 2008).

#### 5.4.4. Logaritmik benzerlik oranı (LLR)

İstatistiksel olarak referans ve iyileştirilmiş ses sinyalinin karşılaştırılmasına dayanan bir yöntemdir. Quackenbush (1988) bu oranı aşağıdaki gibi ifade etmiştir.

$$d_{LLR}(\mathbf{a}_p, \mathbf{a}_c) = \log \left( \frac{\mathbf{a}_p \mathbf{R}_c \mathbf{a}_p^T}{\mathbf{a}_c \mathbf{R}_c \mathbf{a}_c^T} \right) \quad (5.4)$$

$\mathbf{a}_c$  : Temiz sinyal için doğrusal öngörü vektörü

$\mathbf{a}_p$  : İyileştirilmiş sinyal için doğrusal öngörü vektörü

$\mathbf{R}_c$  : Temiz sinyal için otokorelasyon matrisi

#### 5.4.5. Ağırlıklı spektral eğim mesafesi ölçümü (WSSD)

Ağırlıklı spektral eğim mesafesi, ses sinyalinin frekans alanında alt bantlara ayrılıp her bant için spektral eğim hesaplanmasına dayalı bir değerlendirme yöntemidir (Quackenbush, 1988; Klatt, 1982).

$$WSSD = \sum_{k=1}^N w(k) [V_x(k) - V_y(k)]^2 \quad (5.5)$$

Eşitlik (5.5)'te  $w(k)$  ağırlık katsayısını,  $k$  bant indeksini,  $V_x(k)$  iyileştirilmiş ses sinyalinin spektral eğimini,  $V_y(k)$  ise referans sinyalin spektral eğimini temsil etmektedir. Bu mesafenin 0'a yakın olmasıyla iyileştirme yönteminin kalitesi orantılıdır.

#### 5.4.6. Itakura-Saito mesafe ölçümü (IS)

IS mesafe ölçümü temiz ve iyileştirilmiş ses sinyallerinin doğrusal öngörü kod (LPC) kazançlarına göre mesafe ölçmeye dayalı bir yöntemdir (Quackenbush, 1988)

$$d_{IS}(\mathbf{a}_p, \mathbf{a}_c) = \frac{\sigma_c^2}{\sigma_p^2} \left( \frac{\mathbf{a}_p \mathbf{R}_c \mathbf{a}_p^T}{\mathbf{a}_c \mathbf{R}_c \mathbf{a}_c^T} \right) + \log \left( \frac{\sigma_c^2}{\sigma_p^2} \right) - 1 \quad (5.6)$$

$\mathbf{a}_c$  : Temiz sinyal için doğrusal öngörü vektörü

$\mathbf{a}_p$  : İyileştirilmiş sinyal için doğrusal öngörü vektörü

$R_c$  : Temiz sinyal için otokorelasyon matrisi

$\sigma_c^2$  : Temiz sinyalin LPC kazancı

$\sigma_p^2$  : İyileştirilmiş sinyalin LPC kazancı

#### 5.4.7. Cepstrum mesafe ölçümü (CEP)

CEP mesafe ölçümü LPC katsayılarına dayalı bir mesafe ölçüm yöntemidir (Kitawaki, 1988).

$$d_{cep}(c_x, c_{\hat{x}}) = \frac{10}{\log_e 10} \sqrt{2 \sum_{k=1}^p [c_x(k) - c_{\hat{x}}(k)]^2} \quad (5.7)$$

$c_x(k)$  : Temiz sinyalin cepstrum katsayıları

$c_{\hat{x}}(k)$  : İyileştirilmiş sinyalin cepstrum katsayıları

#### 5.4.8. Bileşik ölçümler

Bileşik ölçümler objektif değerlendirme yöntemleriyle subjektif değerlendirme yöntemlerine yakın sonuçlar almak için ilişki katsayılarının artırılmasına dayanan ölçüm yöntemidir. Çoklu doğrusal regresyon analiziyle ortaya çıkan bu ölçümler aşağıdaki gibi ifade edilir (Hu ve Loizou, 2006).

$$\begin{aligned} C_{sig} &= 3.093 - 1.029LLR + 0.603PESQ - 0.009WSS \\ C_{bak} &= 1.634 - 0.478PESQ + 0.007WSS - 0.063segSNR \\ C_{ovl} &= 1.594 + 0.805PESQ - 0.512LLR - 0.007WSS \end{aligned} \quad (5.8)$$

Yukarıdaki eşitlikte  $C_{sig}$  sinyal bozulması,  $C_{bak}$  gürültü bozulması ve toplam kalite ölçütü  $C_{ovl}$  olarak ifade edilmektedir. Bu ölçümler 1 ve 5 arasında değer almaktadır. Sinyal bozulması için 1= hiç doğal değil, 5= çok doğal, gürültü bozulması için 1= çok gürültü karışmış, 5= fark edilmeyecek kadar az gürültü karışmış, toplam kalite ölçütü için 1=kötü

5= mükemmel şeklinde ifade edilmektedir. Yani bütün bu değerlerin 5'e yakın olması istenir.

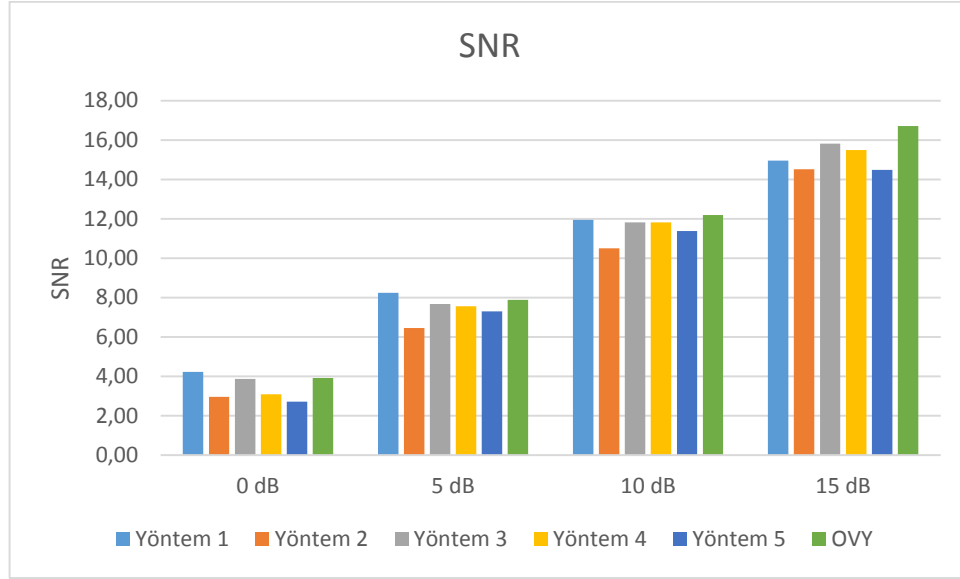
### **5.5. Elde Edilen Sonuçların Değerlendirilmesi**

Tez kapsamında yapılan ortak vektör yaklaşımıyla önerilen yöntem (OVY) bahsedilen 5 farklı yöntemle yukarıda anlatılan 8 objektif kriterde kıyaslanmıştır. Her bir gürültü çeşidinde ve seviyesinde 30 farklı ses dosyasında yapılan çalışmaların ortalaması alınarak çizelgeler oluşturulmuştur. Sonuçları yorumlarken dikkat edilecek bir husus mesafe ölçüm kriterlerinde ve logaritmik benzerlik oranında (LLR) ideal durumun sıfır (veya sıfıra yakın) olduğudur. Kalın (bold) olarak belirtilen ortalamalar o gürültü çeşidinde ve seviyesindeki en iyi sonucu göstermektedir. Kıyaslama sonuçları aşağıda mevcuttur.

Çizelge 5.2. SNR kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	<b>3,81</b>	2,81	3,60	2,14	2,11	3,60
	5 dB	<b>7,83</b>	6,52	7,37	6,84	6,82	7,52
	10 dB	11,65	10,34	11,35	11,18	10,96	<b>11,86</b>
	15 dB	14,83	14,25	15,54	15,26	14,35	<b>16,44</b>
<i>Kalabalık</i>	0 dB	<b>4,17</b>	3,07	3,97	2,16	1,97	3,57
	5 dB	<b>7,81</b>	6,42	7,41	6,85	6,71	7,61
	10 dB	11,68	10,62	11,53	11,51	11,12	<b>11,99</b>
	15 dB	14,72	14,33	15,41	15,24	14,29	<b>16,32</b>
<i>Araba</i>	0 dB	<b>4,94</b>	3,65	4,03	4,73	3,94	4,86
	5 dB	<b>8,90</b>	6,63	7,76	8,64	8,13	8,60
	10 dB	12,28	10,15	11,71	12,46	11,91	<b>12,65</b>
	15 dB	15,25	14,25	15,95	15,94	14,80	<b>17,11</b>
<i>Sergi Salonu</i>	0 dB	<b>3,88</b>	2,60	3,68	2,40	2,22	3,47
	5 dB	<b>8,02</b>	6,46	7,81	6,91	6,78	7,54
	10 dB	11,79	10,69	12,02	11,37	11,06	<b>11,98</b>
	15 dB	14,88	14,67	15,92	15,26	14,34	<b>16,74</b>
<i>Restoran</i>	0 dB	2,55	2,23	<b>2,94</b>	0,99	1,20	2,39
	5 dB	<b>6,77</b>	5,91	6,74	5,93	6,03	6,60
	10 dB	11,01	10,10	11,01	10,62	10,45	<b>11,23</b>
	15 dB	14,48	14,36	15,32	14,89	14,02	<b>15,97</b>
<i>Tren İstasyonu</i>	0 dB	<b>4,10</b>	3,08	3,52	3,34	3,03	3,91
	5 dB	7,90	6,24	7,12	7,78	7,49	<b>7,90</b>
	10 dB	11,98	10,45	11,58	11,99	11,53	<b>12,26</b>
	15 dB	14,95	14,05	15,50	15,30	14,41	<b>16,63</b>
<i>Sokak</i>	0 dB	3,28	2,24	3,16	2,36	2,46	<b>3,92</b>
	5 dB	7,45	5,68	6,89	7,18	7,30	<b>7,67</b>
	10 dB	11,01	9,79	10,90	11,48	11,26	<b>12,01</b>
	15 dB	14,63	14,07	15,31	15,30	14,42	<b>16,57</b>
<i>Tren</i>	0 dB	<b>3,93</b>	1,70	3,72	2,92	2,43	3,57
	5 dB	<b>8,71</b>	5,88	7,97	7,48	7,14	7,72
	10 dB	<b>12,42</b>	10,36	12,28	11,85	11,32	11,93
	15 dB	14,99	14,90	16,02	15,47	14,43	<b>16,53</b>
<i>Beyaz Gürültü</i>	0 dB	<b>7,37</b>	5,13	6,09	6,66	5,07	5,91
	5 dB	<b>10,72</b>	8,32	9,91	10,44	9,20	9,77
	10 dB	13,65	11,96	13,97	13,85	12,71	<b>13,77</b>
	15 dB	15,86	15,69	17,39	16,69	15,22	<b>18,07</b>

Yöntem Çizelge 5.2.'de görüldüğü gibi SNR kriterinde genel olarak 10 ve 15 dB gürültü seviyelerinde en iyi yöntem, 0 ve 5 dB gürültü seviyelerinde ise en iyi ikinci yöntem olarak görülmektedir. Tren istasyonu ve sokak gürültü çeşitlerinde ise tüm gürültü seviyelerinde en yüksek SNR değerine sahiptir.



Şekil 5.7. Tüm gürültü çeşitlerinde ortalama SNR grafiği

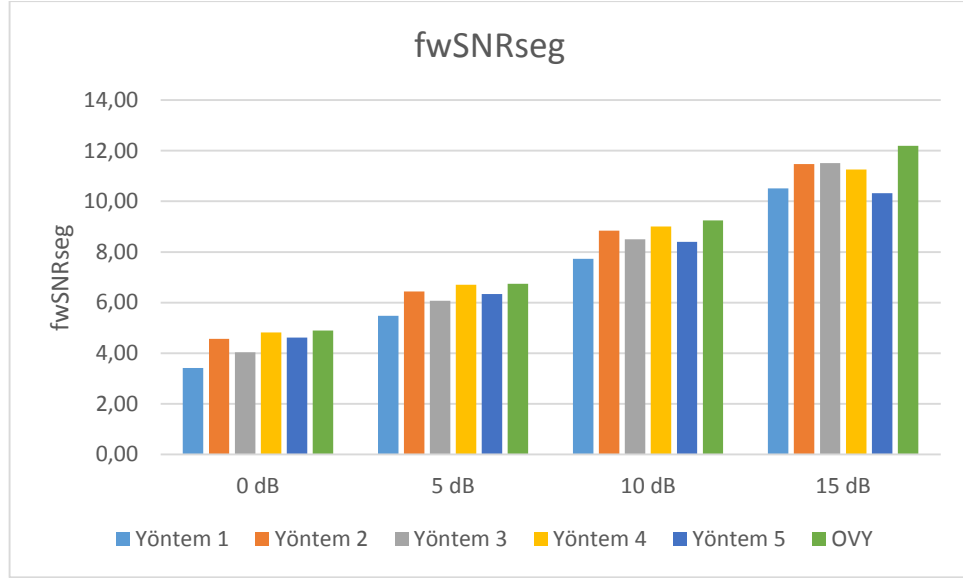
Şekil 5.7.'da görüleceği gibi gürültü çeşitlerinde 30 ses dosyasında ortalaması alındığında performans olarak 0 ve 5 dB gürültü seviyelerinde en iyi ikinci yöntem, 10 ve 15 dB gürültü seviyelerinde en iyi yöntemdir.

Çizelge 5.3. fwSNRseg kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	3,49	4,65	4,11	4,88	4,84	<b>5,05</b>
	5 dB	5,43	6,49	6,04	6,64	6,50	<b>7,03</b>
	10 dB	7,90	9,22	8,65	9,04	8,67	<b>9,71</b>
	15 dB	10,89	11,96	11,91	11,34	10,50	<b>12,69</b>
<i>Kalabalık</i>	0 dB	3,32	4,44	4,10	4,48	4,48	<b>4,94</b>
	5 dB	5,41	6,54	6,08	6,49	6,31	<b>6,99</b>
	10 dB	7,87	9,15	8,76	8,86	8,39	<b>9,61</b>
	15 dB	10,89	11,87	11,93	11,25	10,34	<b>12,73</b>
<i>Araba</i>	0 dB	3,28	4,47	3,72	<b>4,72</b>	4,46	4,66
	5 dB	5,38	6,23	5,73	<b>6,64</b>	6,23	6,45
	10 dB	7,23	8,45	8,01	<b>8,91</b>	8,27	8,87
	15 dB	10,28	11,20	11,16	11,22	10,27	<b>11,80</b>
<i>Sergi Salonu</i>	0 dB	3,70	4,92	4,29	<b>5,08</b>	4,74	4,89
	5 dB	5,88	<b>6,78</b>	6,47	6,96	6,49	6,72
	10 dB	7,78	8,96	8,77	<b>9,30</b>	8,63	9,24
	15 dB	10,29	11,32	11,38	11,49	10,47	<b>12,18</b>
<i>Restoran</i>	0 dB	3,30	4,84	4,14	4,94	4,91	<b>5,30</b>
	5 dB	5,74	6,93	6,55	6,86	6,69	<b>7,34</b>
	10 dB	8,02	9,43	8,91	9,22	8,80	<b>10,10</b>
	15 dB	11,14	12,32	12,31	11,49	10,58	<b>13,20</b>
<i>Tren İstasyonu</i>	0 dB	3,54	4,47	3,91	4,80	4,64	<b>4,85</b>
	5 dB	5,51	6,59	6,00	6,77	6,44	<b>6,81</b>
	10 dB	7,63	8,75	8,14	8,94	8,39	<b>9,23</b>
	15 dB	10,33	11,46	11,16	11,16	10,34	<b>12,27</b>
<i>Sokak</i>	0 dB	3,10	4,64	3,90	4,91	4,82	<b>5,08</b>
	5 dB	5,11	6,49	5,88	6,75	6,57	<b>6,83</b>
	10 dB	7,73	9,05	8,48	9,12	8,71	<b>9,54</b>
	15 dB	10,16	11,38	11,21	11,17	10,42	<b>12,36</b>
<i>Tren</i>	0 dB	3,44	4,92	4,55	<b>5,14</b>	4,90	5,11
	5 dB	5,57	6,36	6,24	6,82	6,48	<b>6,84</b>
	10 dB	7,91	8,71	8,75	9,04	8,37	<b>9,19</b>
	15 dB	10,55	11,42	11,68	11,31	10,38	<b>12,13</b>
<i>Beyaz Gürültü</i>	0 dB	3,60	3,71	3,65	<b>4,49</b>	3,77	4,15
	5 dB	5,24	5,58	5,63	<b>6,38</b>	5,38	5,73
	10 dB	7,47	7,83	8,02	<b>8,59</b>	7,42	7,77
	15 dB	10,08	10,28	10,84	<b>10,89</b>	9,58	10,34



Çizelge 5.3.'te görüleceği üzere frekans ağırlıklı parçalı SNR (fwSNRseg) kriterinde ise SNR'a göre daha iyi başarıma sahiptir. Bunun sebebi önerilen yöntemin frekans alanında gürültü gidermesidir. Frekans alanında gürültüye ait bileşenler bastırıldığı için bu alandaki değerlendirme daha iyi sonuçlar elde etmemizi sağlamıştır.



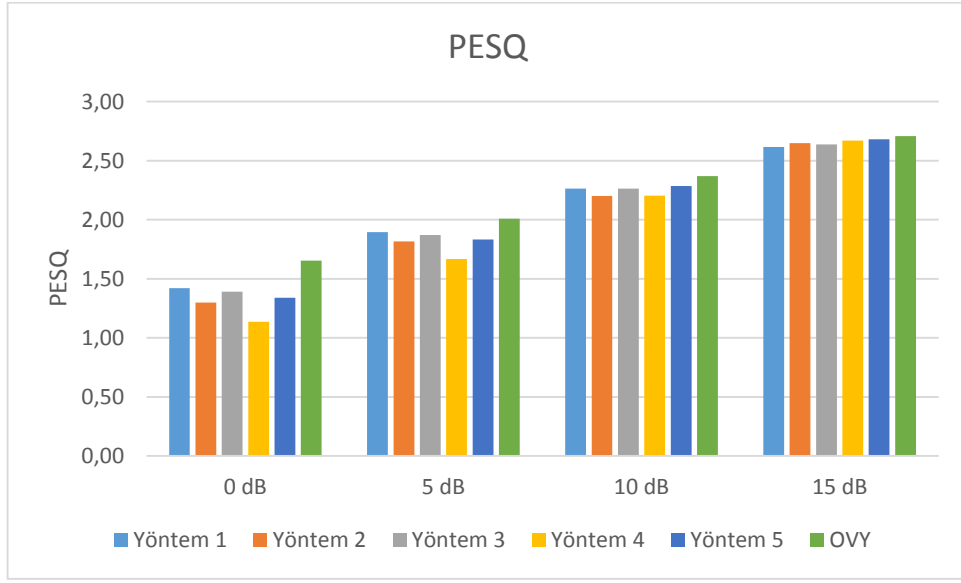
Şekil 5.8. Tüm gürültü çeşitlerinde ortalama fwSNRseg grafiği

Şekil 5.8.'de görüldüğü gibi önerdiğimiz yöntem tüm gürültü çeşitlerinde ortalama olarak diğer yöntemlerden yüksek fwSNRseg değerine sahiptir.

Çizelge 5.4. PESQ kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	1,43	1,31	1,40	1,24	1,32	<b>1,69</b>
	5 dB	1,87	1,68	1,84	1,77	1,82	<b>2,05</b>
	10 dB	2,23	2,09	2,24	2,20	2,26	<b>2,40</b>
	15 dB	2,63	2,59	2,68	2,72	2,73	<b>2,74</b>
<i>Kalabalık</i>	0 dB	1,37	1,20	1,34	1,13	1,34	<b>1,64</b>
	5 dB	1,84	1,73	1,82	1,62	1,78	<b>2,02</b>
	10 dB	2,23	2,16	2,25	2,24	2,24	<b>2,39</b>
	15 dB	2,63	2,60	2,65	2,69	2,70	<b>2,74</b>
<i>Araba</i>	0 dB	1,46	1,27	1,41	1,23	1,40	<b>1,68</b>
	5 dB	1,91	1,92	1,89	1,67	1,86	<b>1,97</b>
	10 dB	2,24	2,25	2,25	2,17	2,28	<b>2,34</b>
	15 dB	2,57	2,70	2,60	2,66	<b>2,71</b>	2,70
<i>Sergi Salonu</i>	0 dB	1,42	1,29	1,39	1,15	1,31	<b>1,62</b>
	5 dB	1,91	1,89	1,89	1,68	1,85	<b>1,98</b>
	10 dB	2,25	2,17	2,25	2,21	2,28	<b>2,32</b>
	15 dB	2,62	2,61	2,60	2,64	2,64	<b>2,67</b>
<i>Restoran</i>	0 dB	1,37	1,24	1,42	1,21	1,26	<b>1,70</b>
	5 dB	1,84	1,74	1,83	1,71	1,80	<b>2,02</b>
	10 dB	2,23	2,08	2,22	2,22	2,22	<b>2,41</b>
	15 dB	2,60	2,57	2,64	2,67	2,63	<b>2,72</b>
<i>Tren İstasyonu</i>	0 dB	1,44	1,35	1,41	1,20	1,36	<b>1,66</b>
	5 dB	1,88	1,87	1,89	1,75	1,83	<b>2,07</b>
	10 dB	2,24	2,25	2,27	2,24	2,31	<b>2,38</b>
	15 dB	2,53	2,61	2,59	2,65	2,66	<b>2,72</b>
<i>Sokak</i>	0 dB	1,32	1,18	1,24	1,04	1,24	<b>1,62</b>
	5 dB	1,79	1,68	1,78	1,62	1,76	<b>2,00</b>
	10 dB	2,23	2,13	2,21	2,19	2,21	<b>2,38</b>
	15 dB	2,57	2,59	2,57	2,60	2,56	<b>2,68</b>
<i>Tren</i>	0 dB	1,22	1,10	1,16	0,81	1,09	<b>1,57</b>
	5 dB	1,86	1,67	1,75	1,42	1,65	<b>1,92</b>
	10 dB	2,28	2,14	2,21	2,07	2,19	<b>2,29</b>
	15 dB	2,64	2,63	2,62	2,64	2,63	<b>2,65</b>
<i>Beyaz Gürültü</i>	0 dB	1,77	<b>1,76</b>	1,75	1,23	1,73	1,70
	5 dB	2,17	<b>2,17</b>	2,16	1,77	2,16	2,06
	10 dB	2,44	2,55	2,47	2,31	<b>2,57</b>	2,42
	15 dB	2,75	<b>2,95</b>	2,80	2,76	2,87	2,75

Önerdiğimiz yöntemin Çizelge 5.4.'te görüldüğü gibi PESQ ölçütünde başarımı oldukça iyidir. Sadece beyaz gürültü çeşidinde önerilen yöntem diğer yöntemlere göre düşük performans göstermiştir. Fakat yine de beyaz gürültüde dahi en iyi sonuçlara yakın sonuçlar elde edilmiştir.



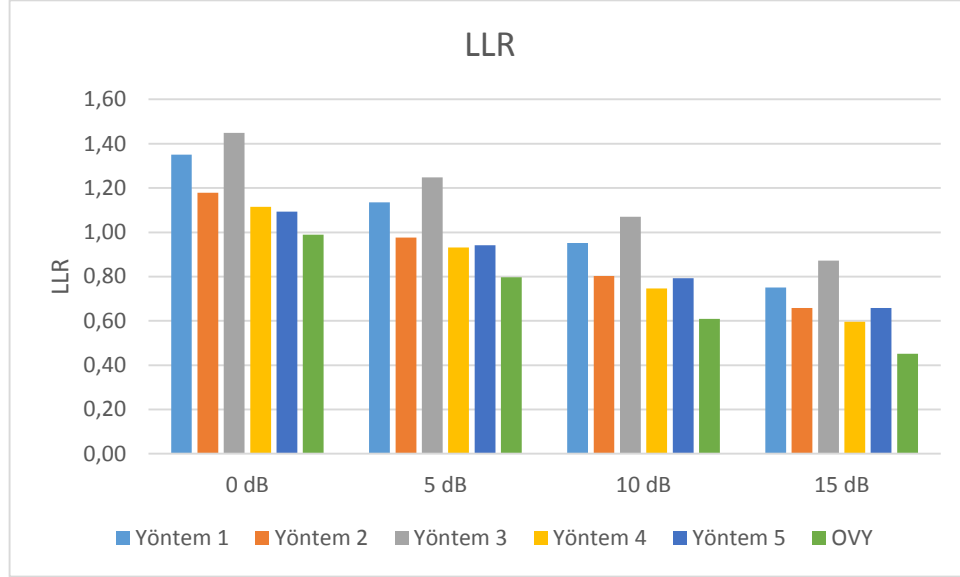
Şekil 5.9. Tüm gürültü çeşitlerinde ortalama PESQ grafiği

PESQ değerlendirme ölçütünde Şekil 5.9.'de görüldüğü gibi tüm gürültü çeşitlerinde diğer yöntemlerden daha iyi sonuçlar elde edilmektedir. Fakat 0 dB gürültü seviyesinde diğer yöntemlerle arasındaki başarı farkı oldukça fazladır.

Çizelge 5.5. LLR kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	1,29	1,09	1,37	1,03	0,96	<b>0,85</b>
	5 dB	1,10	0,92	1,16	0,86	0,83	<b>0,66</b>
	10 dB	0,86	0,70	0,94	0,67	0,69	<b>0,49</b>
	15 dB	0,66	0,56	0,72	0,54	0,56	<b>0,35</b>
<i>Kalabalık</i>	0 dB	1,30	1,14	1,35	1,08	1,03	<b>0,89</b>
	5 dB	1,10	0,92	1,17	0,90	0,88	<b>0,70</b>
	10 dB	0,88	0,71	0,93	0,69	0,72	<b>0,50</b>
	15 dB	0,67	0,56	0,72	0,55	0,59	<b>0,36</b>
<i>Araba</i>	0 dB	1,35	1,10	1,47	1,08	1,03	<b>0,94</b>
	5 dB	1,11	0,89	1,22	0,86	0,86	<b>0,73</b>
	10 dB	0,98	0,75	1,10	0,67	0,71	<b>0,54</b>
	15 dB	0,71	0,57	0,82	0,51	0,57	<b>0,37</b>
<i>Sergi Salonu</i>	0 dB	1,34	1,19	1,52	1,17	1,14	<b>1,06</b>
	5 dB	1,13	0,98	1,31	0,98	0,97	<b>0,85</b>
	10 dB	1,02	0,86	1,19	0,78	0,81	<b>0,64</b>
	15 dB	0,81	0,74	1,01	0,61	0,66	<b>0,47</b>
<i>Restoran</i>	0 dB	1,35	1,12	1,39	1,02	0,98	<b>0,85</b>
	5 dB	1,04	0,88	1,08	0,84	0,82	<b>0,66</b>
	10 dB	0,88	0,73	0,94	0,68	0,68	<b>0,49</b>
	15 dB	0,70	0,60	0,77	0,56	0,60	<b>0,36</b>
<i>Tren İstasyonu</i>	0 dB	1,28	1,10	1,39	1,06	1,00	<b>0,89</b>
	5 dB	1,08	0,87	1,21	0,84	0,83	<b>0,68</b>
	10 dB	0,89	0,72	1,02	0,66	0,69	<b>0,50</b>
	15 dB	0,72	0,59	0,83	0,52	0,56	<b>0,35</b>
<i>Sokak</i>	0 dB	1,39	1,19	1,48	1,09	1,07	<b>0,94</b>
	5 dB	1,16	0,99	1,28	0,89	0,90	<b>0,75</b>
	10 dB	0,93	0,78	1,05	0,72	0,75	<b>0,57</b>
	15 dB	0,78	0,68	0,94	0,59	0,63	<b>0,42</b>
<i>Tren</i>	0 dB	1,41	1,18	1,43	1,14	1,14	<b>1,05</b>
	5 dB	1,20	1,06	1,28	0,99	1,01	<b>0,88</b>
	10 dB	1,01	0,86	1,10	0,82	0,86	<b>0,68</b>
	15 dB	0,80	0,66	0,87	0,64	0,69	<b>0,50</b>
<i>Beyaz Gürültü</i>	0 dB	1,44	1,50	1,64	1,36	1,49	<b>1,43</b>
	5 dB	1,30	1,27	1,52	1,22	1,37	<b>1,26</b>
	10 dB	1,11	1,12	1,36	1,03	1,23	<b>1,07</b>
	15 dB	0,91	0,96	1,17	0,84	1,06	<b>0,88</b>

Çizelge 5.5.'te görüldüğü gibi logaritmik benzerlik oranında (LLR) yöntem tüm gürültü çeşitleri ve seviyelerinde en yüksek başarıya sahiptir.



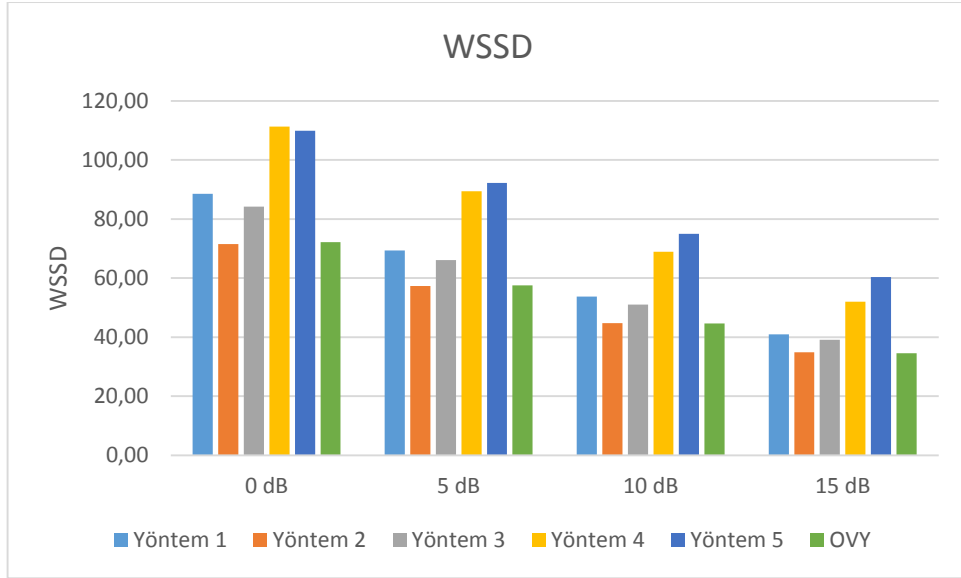
Şekil 5.10. Tüm gürültü çeşitlerinde ortalama LLR grafiği

Şekil 5.10.'da net olarak görüleceği gibi önerdiğimiz yöntemde logaritmik benzerlik oranı (LLR) ortalama olarak tüm gürültü çeşitlerinde ve tüm gürültü seviyelerinde en iyi sonuçlara sahiptir. LLR'de değerin küçük olması iyileştirilmiş ses sinyalinin gürültüsüz, orjinal sinyale benzerliğinin büyük olmasını belirtmektedir. Dolayısıyla önerdiğimiz yöntem tüm testlerde orjinal sinyale en çok benzeyen sinyal olarak görülmektedir.

Çizelge 5.6. WSSD kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	92,79	<b>73,94</b>	85,52	116,15	112,45	75,01
	5 dB	71,81	<b>58,22</b>	66,28	91,34	93,20	58,43
	10 dB	54,41	<b>44,18</b>	48,27	68,19	72,81	44,20
	15 dB	41,00	34,49	37,13	49,90	58,05	<b>34,48</b>
<i>Kalabalık</i>	0 dB	92,31	<b>73,66</b>	83,45	115,82	113,24	74,77
	5 dB	72,25	<b>58,17</b>	65,55	93,95	95,61	59,22
	10 dB	54,10	<b>44,29</b>	48,96	69,90	75,00	44,66
	15 dB	40,65	<b>33,75</b>	36,34	50,69	58,38	33,75
<i>Araba</i>	0 dB	88,06	<b>70,03</b>	86,48	108,45	107,99	70,44
	5 dB	69,10	<b>56,06</b>	66,47	85,95	90,22	56,59
	10 dB	53,80	43,50	51,99	64,38	72,75	<b>43,25</b>
	15 dB	39,29	33,53	38,53	46,71	56,88	<b>33,06</b>
<i>Sergi Salonu</i>	0 dB	89,50	<b>67,45</b>	82,53	117,81	112,25	70,62
	5 dB	72,54	<b>54,91</b>	66,07	96,93	96,10	56,78
	10 dB	56,46	<b>43,17</b>	51,47	75,19	79,34	44,51
	15 dB	43,24	34,30	40,84	56,35	63,91	<b>34,20</b>
<i>Restoran</i>	0 dB	90,62	<b>70,18</b>	81,78	109,16	107,96	71,08
	5 dB	68,65	<b>56,37</b>	62,46	87,88	90,81	56,82
	10 dB	53,23	43,17	47,77	66,31	71,29	<b>43,17</b>
	15 dB	42,00	33,89	36,47	52,02	58,33	<b>33,66</b>
<i>Tren İstasyonu</i>	0 dB	88,03	<b>72,28</b>	84,36	112,55	110,87	72,58
	5 dB	68,32	<b>56,76</b>	66,56	87,47	91,51	57,01
	10 dB	53,08	<b>44,17</b>	50,78	65,97	73,31	44,28
	15 dB	40,35	34,24	38,02	48,35	58,06	<b>33,61</b>
<i>Sokak</i>	0 dB	85,02	<b>67,18</b>	83,04	104,07	106,97	67,60
	5 dB	66,69	<b>52,98</b>	63,33	81,89	86,56	53,06
	10 dB	51,46	42,09	49,58	67,50	72,90	41,76
	15 dB	40,18	33,72	39,69	54,44	61,26	33,13
<i>Tren</i>	0 dB	81,04	<b>63,51</b>	74,51	95,47	95,38	65,34
	5 dB	62,98	<b>50,71</b>	58,45	77,52	80,04	51,87
	10 dB	49,10	<b>40,10</b>	45,16	62,47	67,60	40,64
	15 dB	37,66	<b>30,70</b>	34,21	48,38	54,76	31,02
<i>Beyaz Gürültü</i>	0 dB	89,43	85,59	96,31	122,56	121,94	<b>82,01</b>
	5 dB	71,43	71,43	80,10	102,12	106,00	<b>68,29</b>
	10 dB	58,12	57,71	65,36	80,18	89,42	<b>55,17</b>
	15 dB	44,46	45,54	50,77	61,20	73,88	<b>43,70</b>

Ağırlıklı spektral eğim mesafesi (WSSD) kriterinde ise OVY Çizelge 5.6.'dan görüleceği gibi testlerin büyük çoğunluğunda başarı olarak en iyi ikinci yöntem olarak görülmektedir.



Şekil 5.11. Tüm gürültü çeşitlerinde ortalama WSSD grafiği

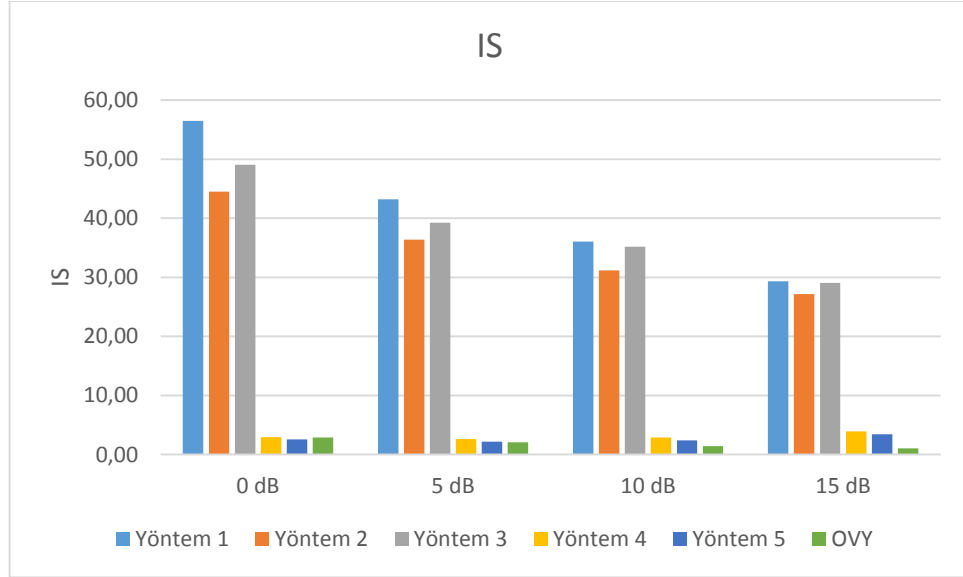
Ortalama olarak ise WSSD ölçütünde diğer tüm yöntemlere göre Şekil 5.11.'da net olarak görüleceği üzere en iyi yöntem olarak görülmektedir. LLR ölçütünde olduğu gibi WSSD ölçütünde de değerin sifıra yakınlığı yöntemin iyi olmasını gösterir.

Çizelge 5.7. IS kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	51,66	43,35	44,96	2,60	<b>2,24</b>	2,64
	5 dB	43,62	36,52	38,35	2,70	2,13	<b>1,87</b>
	10 dB	35,11	32,52	32,71	3,30	2,80	<b>1,28</b>
	15 dB	26,84	25,35	24,35	4,60	3,94	<b>0,95</b>
<i>Kalabalık</i>	0 dB	49,89	44,87	40,28	2,68	<b>2,42</b>	2,68
	5 dB	41,86	38,48	37,00	2,49	2,14	<b>1,93</b>
	10 dB	33,12	29,13	30,58	3,02	2,40	<b>1,30</b>
	15 dB	25,53	23,99	23,99	4,28	4,03	<b>0,89</b>
<i>Araba</i>	0 dB	73,24	47,04	58,31	2,70	<b>2,37</b>	2,91
	5 dB	50,86	38,29	41,47	2,40	<b>1,98</b>	2,05
	10 dB	44,60	35,92	42,25	2,88	2,07	<b>1,34</b>
	15 dB	33,89	29,60	32,19	4,26	3,25	<b>0,95</b>
<i>Sergi Salonu</i>	0 dB	54,54	43,30	49,29	2,60	<b>2,55</b>	2,98
	5 dB	37,01	33,93	35,98	2,24	<b>2,07</b>	2,18
	10 dB	36,27	31,71	38,03	2,42	2,05	<b>1,46</b>
	15 dB	32,18	31,02	34,29	3,33	2,77	<b>1,01</b>
<i>Restoran</i>	0 dB	47,48	44,74	43,75	2,86	<b>2,40</b>	2,58
	5 dB	32,97	31,86	29,64	2,54	2,12	<b>1,68</b>
	10 dB	31,89	30,33	30,18	3,32	2,75	<b>1,28</b>
	15 dB	24,88	25,34	23,88	4,33	4,04	<b>0,87</b>
<i>Tren İstasyonu</i>	0 dB	56,97	45,88	48,34	2,81	<b>2,30</b>	2,49
	5 dB	47,47	39,18	43,77	2,64	2,03	<b>1,83</b>
	10 dB	38,06	33,47	36,45	3,21	2,46	<b>1,34</b>
	15 dB	32,15	30,23	30,91	4,64	3,85	<b>1,01</b>
<i>Sokak</i>	0 dB	53,49	47,99	49,95	3,96	<b>2,69</b>	3,17
	5 dB	41,42	38,46	37,36	3,26	2,29	<b>2,08</b>
	10 dB	31,40	28,30	30,21	3,15	2,41	<b>1,47</b>
	15 dB	29,01	27,46	29,08	3,88	3,56	<b>1,11</b>
<i>Tren</i>	0 dB	55,95	45,14	47,10	3,50	<b>2,89</b>	3,13
	5 dB	41,68	36,14	36,39	2,94	2,49	<b>2,35</b>
	10 dB	33,84	27,50	31,02	2,60	2,36	<b>1,58</b>
	15 dB	25,19	21,58	24,38	3,20	2,87	<b>1,07</b>
<i>Beyaz Gürültü</i>	0 dB	65,28	38,39	59,56	<b>2,84</b>	3,31	3,36
	5 dB	51,96	34,49	53,29	<b>2,39</b>	2,69	2,58
	10 dB	40,26	31,59	45,48	2,32	2,39	<b>1,88</b>
	15 dB	34,52	29,82	38,79	3,03	2,65	<b>1,43</b>



Çizelge 5.7.'de Itakuro-Saito (IS) mesafe ölçümünde görüleceği gibi testlerin %33'ünde en yüksek yöntemdir, geri kalanında ise başarısı en yüksek 2 yöntem arasındadır.



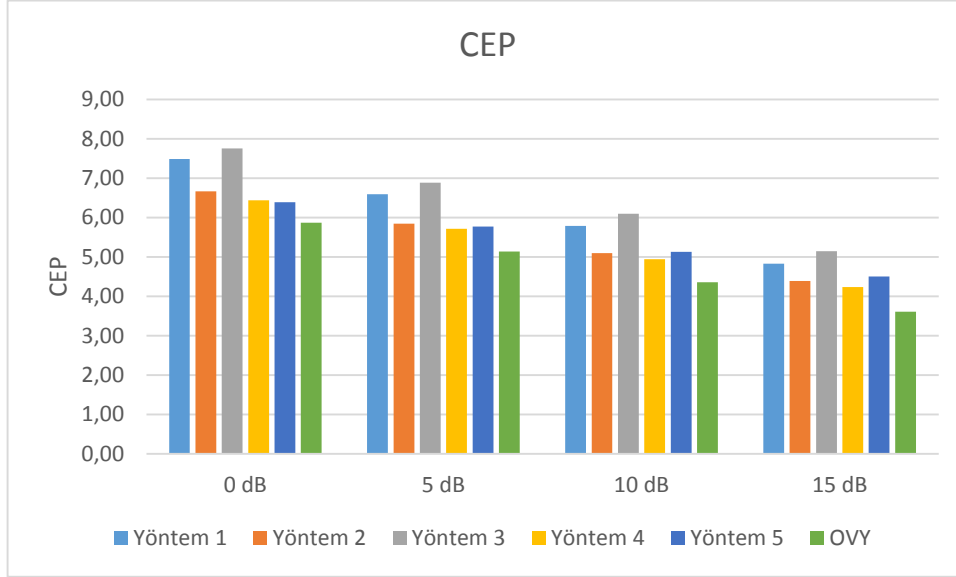
Şekil 5.12. Tüm gürültü çeşitlerinde ortalama IS grafiği

Şekil 5.12.'de görüleceği gibi ortalama sonuçlarda tüm testlerde diğer 5 yönteme göre en iyi yöntemdir. 15 dB gürültü seviyesinde ise ortalama olarak 0'a çok yakın bir değerle diğer yöntemlerden oldukça iyi sonuçlara sahiptir.

Çizelge 5.8. CEP kriterine göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

		Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4	Yöntem 5	OVY
<i>Havaalanı</i>	0 dB	7,27	6,29	7,44	6,11	5,85	<b>5,27</b>
	5 dB	6,47	5,57	6,54	5,42	5,27	<b>4,52</b>
	10 dB	5,39	4,62	5,55	4,59	4,62	<b>3,73</b>
	15 dB	4,41	3,90	4,45	3,91	4,02	<b>3,02</b>
<i>Kalabalık</i>	0 dB	7,26	6,47	7,32	6,22	6,04	<b>5,37</b>
	5 dB	6,43	5,55	6,58	5,50	5,42	<b>4,60</b>
	10 dB	5,47	4,63	5,50	4,61	4,69	<b>3,76</b>
	15 dB	4,43	3,89	4,47	3,92	4,08	<b>3,00</b>
<i>Araba</i>	0 dB	7,46	6,23	7,77	6,15	5,96	<b>5,43</b>
	5 dB	6,49	5,40	6,74	5,30	5,31	<b>4,66</b>
	10 dB	5,92	4,76	6,19	4,50	4,62	<b>3,87</b>
	15 dB	4,63	3,90	4,88	3,74	3,94	<b>3,06</b>
<i>Sergi Salonu</i>	0 dB	7,44	6,59	8,02	6,62	6,43	<b>5,95</b>
	5 dB	6,60	5,76	7,12	5,87	5,78	<b>5,16</b>
	10 dB	6,08	5,23	6,61	5,03	5,09	<b>4,33</b>
	15 dB	5,11	4,61	5,70	4,21	4,39	<b>3,53</b>
<i>Restoran</i>	0 dB	7,50	6,45	7,58	6,07	5,89	<b>5,27</b>
	5 dB	6,19	5,42	6,21	5,29	5,18	<b>4,52</b>
	10 dB	5,46	4,74	5,55	4,58	4,57	<b>3,73</b>
	15 dB	4,57	4,06	4,66	4,05	4,16	<b>3,08</b>
<i>Tren İstasyonu</i>	0 dB	7,20	6,27	7,46	6,14	5,91	<b>5,35</b>
	5 dB	6,37	5,28	6,68	5,25	5,18	<b>4,51</b>
	10 dB	5,49	4,63	5,83	4,48	4,55	<b>3,75</b>
	15 dB	4,65	3,99	4,93	3,81	3,95	<b>3,00</b>
<i>Sokak</i>	0 dB	7,63	6,65	7,89	6,35	6,24	<b>5,62</b>
	5 dB	6,65	5,82	7,00	5,49	5,52	<b>4,84</b>
	10 dB	5,69	4,92	6,00	4,78	4,85	<b>4,08</b>
	15 dB	4,99	4,44	5,45	4,15	4,29	<b>3,37</b>
<i>Tren</i>	0 dB	7,65	6,60	7,73	6,53	6,51	<b>6,09</b>
	5 dB	6,78	6,19	7,11	5,98	6,02	<b>5,45</b>
	10 dB	5,99	5,31	6,26	5,25	5,37	<b>4,63</b>
	15 dB	5,05	4,47	5,26	4,43	4,60	<b>3,81</b>
<i>Beyaz Gürültü</i>	0 dB	8,00	8,45	8,60	<b>7,78</b>	8,66	8,45
	5 dB	7,36	7,64	8,02	<b>7,32</b>	8,29	7,97
	10 dB	<b>6,57</b>	7,01	7,36	6,69	7,79	7,36
	15 dB	<b>5,65</b>	6,25	6,48	5,91	7,09	6,61

Çizelge 5.8.'de görüldüğü gibi Cepstrum mesafe ölçümünde önerdiğimiz yöntem beyaz gürültü hariç tüm gürültü çeşitleri ve seviyelerinde en yüksek başarıya sahiptir.



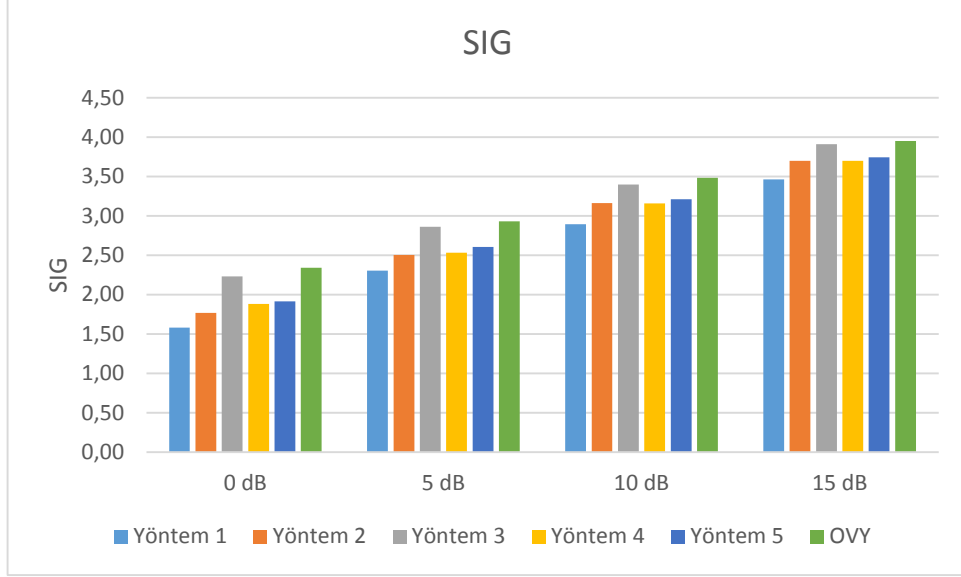
Şekil 5.13. Tüm gürültü çeşitlerinde ortalama CEP grafiği

Şekil 5.13.'de net olarak görüleceği üzere CEP ölçütünde diğer tüm yöntemlere göre ortalama olarak alınan sonuçlarda en iyi yöntem olarak görülmektedir. LLR ve WSSD ölçütlerinde olduğu gibi CEP ölçütünde de değerin sifıra yakınlığı yöntemin iyi olmasını gösterir.

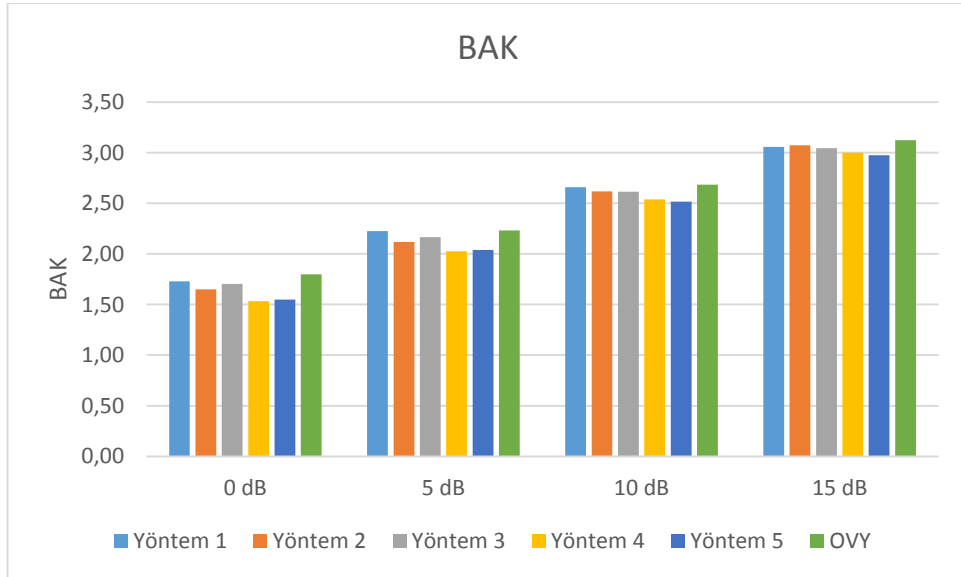
Çizelge 5.9. Bileşik ölçümlere göre önerilen yöntem ve 5 farklı yöntemin ölçüm sonuçları

	Yöntem 1			Yöntem 2			Yöntem 3			Yöntem 4			Yöntem 5			OVY		
	SIG	BAK	OVL	SIG	BAK	OVL	SIG	BAK	OVL	SIG	BAK	OVL	SIG	BAK	OVL	SIG	BAK	OVL
Havaalanı	0	1,64	1,66	1,40	1,95	1,67	1,48	2,31	1,61	1,86	1,96	1,59	2,03	1,49	1,65	2,55	1,79	1,99
	5	2,29	2,17	1,96	2,64	2,16	2,12	2,86	2,09	2,37	2,60	2,16	2,70	2,00	2,24	3,12	2,24	2,49
	10	2,97	2,61	2,52	3,28	2,62	2,69	3,45	2,56	2,88	3,24	2,52	3,36	2,52	2,80	3,64	2,70	2,97
Kalabalık	0	1,53	1,63	1,30	1,82	1,62	1,36	2,20	1,60	1,80	1,88	1,45	1,92	1,46	1,58	2,46	1,76	1,92
	5	2,30	2,14	1,94	2,55	2,08	1,99	2,81	2,07	2,33	2,52	1,94	2,10	2,62	1,95	3,05	2,21	2,44
	10	2,95	2,61	2,51	3,31	2,63	2,72	3,40	2,55	2,85	3,20	2,49	3,28	2,47	2,74	3,61	2,67	2,94
Araba	0	1,62	1,81	1,41	1,98	1,73	1,50	2,47	1,82	2,03	2,01	1,61	2,00	1,63	1,69	2,44	1,83	1,94
	5	2,36	2,27	2,02	2,64	2,14	2,07	3,08	2,25	2,54	2,67	2,08	2,22	2,71	2,07	2,26	3,00	2,23
	10	2,83	2,67	2,46	3,22	2,60	2,64	3,54	2,68	2,99	3,31	2,60	3,28	2,54	2,77	3,56	2,68	2,90
Sergi Salonu	0	1,58	1,74	1,36	1,75	1,68	1,36	2,14	1,62	1,74	1,69	1,42	1,42	1,46	1,49	2,22	1,79	1,80
	5	2,34	2,23	2,00	2,49	2,15	2,00	2,77	2,08	2,30	2,37	1,94	2,49	1,97	2,09	2,86	2,22	2,53
	10	2,81	2,64	2,44	3,11	2,63	2,61	3,27	2,54	2,77	3,03	2,46	3,09	2,45	2,62	3,42	2,66	2,82
Restoran	0	1,51	1,61	1,32	1,94	1,67	1,46	2,25	1,62	1,83	2,04	1,51	1,66	1,52	1,70	2,58	1,81	2,02
	5	2,36	2,14	1,99	2,66	2,14	2,10	2,91	2,08	2,38	2,62	1,99	2,16	2,74	2,01	2,24	3,11	2,22
	10	2,94	2,60	2,51	3,25	2,63	2,69	3,41	2,56	2,86	3,24	2,52	2,72	3,31	2,51	2,77	3,65	2,71
Tren	0	1,68	1,74	1,42	1,91	1,68	1,44	2,40	1,73	1,95	1,94	1,53	1,61	1,57	1,65	2,48	1,80	1,95
	5	2,37	2,22	2,01	2,71	2,16	2,14	3,06	2,22	2,53	2,71	2,08	2,26	2,78	2,09	2,32	3,12	2,26
	10	2,95	2,65	2,52	3,28	2,64	2,71	3,54	2,64	2,97	3,30	2,57	2,78	3,34	2,54	2,80	3,62	2,69
Sokak	0	1,51	1,68	1,34	1,70	1,62	1,29	2,17	1,69	1,80	1,96	1,58	1,62	1,98	1,58	2,43	1,83	1,91
	5	2,22	2,15	1,90	2,51	2,10	1,99	2,87	2,18	2,39	2,64	2,09	2,21	2,67	2,07	2,23	3,02	2,26
	10	2,89	2,62	2,49	3,21	2,61	2,65	3,42	2,63	2,88	3,19	2,55	2,70	3,30	2,56	2,77	3,55	2,72
Tren	0	1,47	1,68	1,27	1,63	1,51	1,16	2,24	1,80	1,87	1,96	1,64	1,64	1,98	1,65	2,26	1,80	1,81
	5	2,22	2,26	1,94	2,29	2,02	1,78	2,81	2,24	2,36	2,52	2,09	2,13	2,59	2,10	2,18	2,82	2,22
	10	2,86	2,71	2,51	3,06	2,57	2,53	3,37	2,65	2,85	3,09	2,54	2,63	3,16	2,52	2,66	3,39	2,66
Beyaz Gürültü	0	1,70	2,01	1,57	1,21	1,64	1,12	1,91	1,83	1,75	1,51	1,58	1,43	1,45	1,58	1,41	1,65	1,78
	5	2,27	2,45	2,10	2,03	2,13	1,78	2,58	2,28	2,31	2,14	2,07	1,98	2,14	2,08	2,01	2,29	2,21
	10	2,81	2,80	2,54	2,73	2,63	2,44	3,18	2,73	2,82	2,82	2,58	2,56	2,77	2,53	2,90	2,66	2,58
15	3,35	3,16	3,00	3,32	3,10	2,99	3,73	3,16	3,31	3,44	3,04	3,09	3,36	3,01	3,05	3,44	3,11	3,04

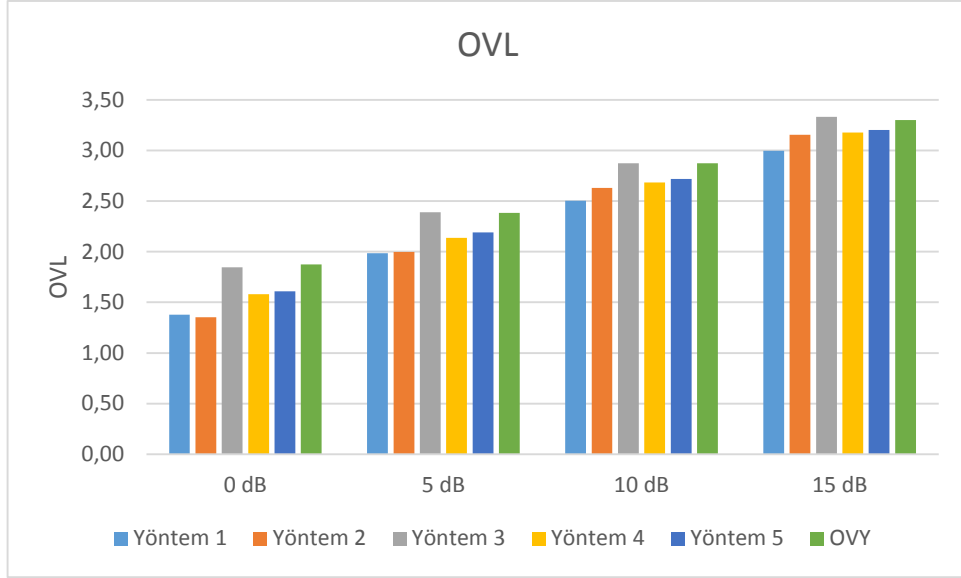
Çizelge 5.9.'da görüldüğü gibi bileşik ölçümlere göre yöntemimiz beyaz gürültü hariç diğer tüm gürültü çeşitleri ve seviyelerinde sinyal kalitesi, arka plan gürültü giderme miktarı ve toplam kalite ölçümlerinde en yüksek başarıya sahip yöntemdir.



Şekil 5.14. Tüm gürültü çeşitlerinde ortalama sinyal bozulması (SIG) grafiği



Şekil 5.15. Tüm gürültü çeşitlerinde ortalama gürültü bozulması (BAK) grafiği



Şekil 5.16. Tüm gürültü çeşitlerinde ortalama toplam kalite (OVL) grafiği

Şekil 5.14. ve Şekil 5.15.'te görüleceği gibi sinyal bozulması ve gürültü bozulması grafiklerinde ortalama olarak diğer yöntemlerden daha iyi sonuçlara sahiptir. Şekil 5.16.'te ise toplam kalite ölçümünde yöntem 3 ve önerdiğimiz yöntem yakın sonuçlar vermekte ve önerdiğimiz yöntemle az farkla daha iyi sonuçlar elde edilmektedir.

Genel olarak tüm kriterlerde bir değerlendirme yapılırsa tez kapsamında önerdiğimiz yöntemin renkli gürültüde beyaz gürültüye göre daha iyi sonuçlar elde ettiğini söylenebilir. Fakat literatürde sıklıkla kullanılan kriterlerde diğer yöntemlere karşı yüksek bir başarı oranı sağlandığı net olarak görülmektedir.

## 6. SONUÇ VE ÖNERİLER

Yüksek lisans tezi kapsamında yapılan bu çalışma gürültülü konuşma verisinde gürültü gideren, anlaşılabilirliği ve ses kalitesini yüksek tutan bir yöntem geliştirmeyi hedeflemiştir. Günlük hayatta karşılaşılma ihtimali daha yüksek olan tek kanallı sistemler göz önünde bulundurularak yöntem geliştirilmiştir. Yöntemin en önemli kısmı eğitim kümesinin oluşturulma aşamasıdır. Önemli olan eğitim kümesinde oluşturulan sınıfların doğru elemanlar içermesidir. Bu sebeple pencere uzunluğu, örtüşme oranı ve arama yapılacak komşuluk miktarı önemli parametrelerimizdir. Yapılan testlere göre tüm gürültü çeşitleri ve seviyelerine göre en iyi parametre belirlenmiştir. Parametre setinin tüm gürültü çeşitleri ve seviyeleri için bir tane olması yöntemin avantajlarından biridir.

Yapılan testlerde parametre belirlenirken ses kalitesi / anlaşılabilirliği ve yöntemin çalışma hızı göz önünde bulundurulmuştur. Altuzay tabanlı yöntemler diğer yöntemlere göre işlem yükü olarak daha ağırdır. Bu işlem yükü yöntemlerin çalışma hızını büyük oranda etkiler. Belirlediğimiz parametre ses kalitesi ve işlem yükü kriterlerinin her ikisine yönelik seçilmiştir. Seçilen parametrelerden sınıf oluşturulmasında etkili olan pencereler arası örtüşme oranının maksimum seçilmesi aynı zamanda müzikal gürültünün de büyük oranda azalmasını sağlamıştır.

Yapılan çalışmada test aşamasında Noizeus veri tabanı kullanılmıştır. Bu veri tabanı içerisinde farklı erkek ve kadın konuşmacıların olduğu 30 farklı cümlede 8 farklı renkli gürültü ve 4 farklı gürültü seviyesi ile gürültülü konuşma verileri oluşturulmuş bir veri tabanıdır. Bu veri tabanına ek olarak aynı ses dosyalarına aynı gürültü seviyelerinde beyaz gürültü de eklenmiştir. Karşılaşılabilecek gürültü çeşitliliği açısından kullanılan veri tabanının oldukça zengin olduğu söylenebilir. Genel olarak literatürde yapılan çalışmalarda bir veya birkaç ses dosyası referans alınarak o ses dosyalarında parametre belirlenmekte ve yöntem sadece seçilen ses dosyasını temel almaktadır. Tez kapsamında yaptığımız çalışma 1080 ses dosyası üzerinde çalışmıştır. Bu yönüyle kullanılabilirliği açısından diğer yöntemlerden avantajlıdır.

Yöntemlerin başarısını sadece insan kulağıyla değerlendirmek oldukça zordur ve doğruluğu düşük olabilmektedir. Bu sebeple literatürde değerlendirme yöntemleri ile ilgili çalışma bulunmaktadır. Bu değerlendirme yöntemlerinin her biri iyileştirilmiş ses sinyaline farklı bakış açısıyla bakar. Bunlardan birinden çok başarılı olabilirken diğerinden başarı oranı düşebilmektedir. Bu sebeple birkaç değerlendirme yöntemine göre başarıyı belirlemek daha doğru olmaktadır. Tez kapsamında yapılan çalışmada objektif olan 8 değerlendirme kriterine göre tüm ses dosyaları için sonuçlar elde edilmiştir. Bu sonuçların ortalamaları alınarak ortalama sonuçlar üzerinden yorumlar yapılmıştır. Erkek veya kadın ses farklılıkları ve farklı konuşmacılar arasındaki farklılıkların sonuçları yükseltmesi/alçaltması böylelikle engellenmiştir. Literatürde sıklıkla yapılan tek ses dosyası üzerinde çalışmanın yöntemin başarısını artırdığı açık şekilde görülmektedir. Yapılan değerlendirme sonuçlarında 30 ses dosyasının ortalamasını alarak bunun önüne geçilmesi istenmiştir. Değerlendirme yöntemi olarak 8 tane yöntem seçilerek geniş bir alanda değerlendirme yapılmıştır.

Kıyaslama olarak literatürde bulunan ve her biri bilinen iyi yöntemlerden olan 5 farklı yöntem kullanılmıştır. Geliştirdiğimiz yöntemin bu uygulamalar arasındaki başarısının görülmesi sağlanmıştır.

Yapılan test sonuçlarına göre önerdiğimiz ortak vektör yaklaşımıyla gürültü giderme yöntemi yüksek bir başarıya sahiptir. Diğer yöntemlere oranla değerlendirme yöntemlerinin büyük bir kısmında en yüksek gürültü giderme oranına sahiptir. Bileşik ölçümlere göre ses kalitesi olarak daha yüksek kaliteli iyileştirme gerçekleşmiştir. Beyaz gürültüye oranla renkli gürültülerde daha iyi sonuçlar elde edilmesi yöntemin günlük hayattaki tek kanallı sistemlerde kullanılabilirliğini göstermektedir.

Özetle, önerilen yöntem literatürde bulunan yöntemlerle kıyaslandığında yüksek başarıya sahiptir. Parametrenin gürültü çeşitleri ve seviyeleri için aynı olması kullanılabilirliğini arttırmaktadır. Gürültü giderme oranını artırmak için yöntem aşamalarında geliştirmeler planlanmaktadır.



## KAYNAKLAR DİZİNİ

- Berouti, M., Schwartz, R., Makhoul, 1979, J., Enhancement of speech corrupted by acoustic noise, Proceeding IEEE Transaction on Acoustics, Speech, Signal Processing, p. 208-211.
- Boll, S.F., 1979, Suppression of acoustic noise in speech using spectral subtraction, IEEE Transaction on Acoustics, Speech, Signal Processing, Vol. ASSP-27, No.2, p.113-120.
- Commins, B., 2005, Signal Subspace Speech Enhancement with Adaptive Noise Estimation, Yüksek Lisans Tezi, National University of Ireland, Galway, 107.
- Çevikalp, H., Neamtu, M., Wilkes, M. and Barkana, A., 2005, Discriminative common vectors for face recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, 27, 1, 4-13.
- Dendrinos, M., Bakamidis, S., Carayannis G., 1991, Speech enhancement from noise: A regenerative approach, Speech Communication, 10:45-47.
- Doblinger, G., 1995, Computationally Efficient Speech Enhancement By Spectral Minima Tracking in Subbands, Proc. EuroSpeech, vol. 2, p. 1513-1516.
- Edizkan, R., 1999, Gizli Markov Model ile Bilgisayarda Konuşma Tanıma: Özellik Uzayında ve Altuzayda Sınıflandırıcı Tasarımı, Doktora Tezi, Eskişehir Osmangazi Üniversitesi, Eskişehir.
- Ephraim, Y., 1992, Statistical-model based speech enhancement systems, Proc. IEEE, 80(10): 1526-1555.
- Ephraim, Y., Van Trees, H.L., 1995, A signal subspace approach for speech enhancement, IEEE Trans. Speech Audio Processing, 3:251-266.
- Ergin, S., 2004, Gürültülü Ses Parametrelerinin İyileştirilmesi ve Tanınması, Yüksek Lisans Tezi, Eskişehir Osmangazi Üniversitesi, Eskişehir.

**KAYNAKLAR DİZİNİ (devam ediyor)**

- Garg, N., Gupta, J., 2013, Review on Speech Enhancement using Signal Subspace method, International Journal of Application or Innovation in Engineering & Management, 2, 5, 2319-4847.
- Gonzalez-Rodriguez, J., Ortega-Garcia, J., Martin, C., Hernandez, L., 1996, Increasing robustness in GMM speaker recognition systems with low complexity microphone arrays, Proc. ICLSP, 1333-1336
- Gülmezoglu, M.B. and Barkana, A., 1998, Text-dependent speaker recognition by using Gram-Schmidt orthogonalization method, Proc. of IASTED Int. Conf. on Signal Processing and Applications, 438-440.
- Gülmezoglu, M.B., Dzhafarov, V. and Barkana, A., 2001, The Common Vector Approach and its relation to Principal Component Analysis, IEEE Trans. On Speech and Audio Processing, 9, 6, 655-662.
- Gülmezoglu, M.B., Dzhafarov, V., Edizkan, R. and Barkana, A., 2007, The Common Vector Approach and its comparison with other subspace methods in case of sufficient data, Computer Speech and Language, 21, 266-281.
- Gülmezoglu, M.B., Dzhafarov, V., Keskin, M., Barkana, A., 1999, A novel approach to isolated word recognition, IEEE Trans. on Acoustic Speech and Signal Processing, 7, 6, 620-628.
- Gülmezoglu, M.B., Ergin, S., 2007, An approach for bearing fault detection in electrical motors, European Transactions on Electrical Power, 17,6, 628-641.
- Günel, S., Ergin, S., Gülmezoglu, M.B. and Gerek, Ö.N., 2006, On feature extraction for spam e-mail detection, Lecture Notes in Computer Science, 4105, 635-642.
- Hansen, J., Pellom, B., An effective quality evaluation protocol for speech enhancement algorithms, Inter. Conf. on Spoken Language Processing, 7(2819), 2822.
- Hu, Y., Loizou P., 2008, Evaluation of objective quality measures for speech enhancement, IEEE Trans. Speech Audio Process., 16 (1): 229-238.

**KAYNAKLAR DİZİNİ (devam ediyor)**

- Hu, Y., Loizou, P., 2004, Incorporating a psychoacoustical model in frequency domain speech enhancement. *IEEE Signal Processing Letters*, 11(2), 270-273.
- Hu, Y., Loizou, P., 2006, Evaluation of objective measures for speech enhancement, *Proc. Interspeech*.
- Hu, Y., Loizou, P., 2004, Speech enhancement based on wavelet thresholding the multitaper spectrum, *IEEE Trans. on Speech and Audio Processing*, 12(1), 59-67.
- ITU, 2000, Perceptual evaluation of speech quality (PESQ) and objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, *ITU-T Recommendation P.862*
- Jabloun, F., Champagne, B., 2003, Incorporating the human hearing properties in the signal subspace approach for speech enhancement, *IEEE Trans. on Speech and Audio Processing*, 11(6), 700-708.
- Jensen, S. H., Hansen, P. C., Hansen, S. D., Sorensen, J.A., 1995, Reduction of broadband noise in speech by truncated qsvd, *IEEE Trans. Speech Audio Processing*, 3:439-448.
- Kamath, S. D., 2001, A Multi-Band Spectral Subtraction Method For Speech Enhancement, Master Thesis, The University of Texas, Dallas.
- Kitawaki, N., Nagabuchi, H., and Itoh, K., 1988, Objective quality evaluation for low bit-rate speech coding systems. *IEEE J. Select. Areas in Comm.*, 6(2), 262-273.
- Klatt, D., 1982, Prediction of perceived phonetic distance from critical band spectra, *Proc. IEEE Int. Conf. Acoust. , Speech, Signal Processing*, 7, 1278-1281.
- Liao, L., Gregory, M.A., 1999, "Algorithms for speech classification", *ISSPA*, p. 623-627.
- Lim, J., Oppenheim, A., 1979, Enhancement and bandwidth compression of noisy speech, *Proc. IEEE*, vol. 67, No. 12, p. 221-239.

**KAYNAKLAR DİZİNİ (devam ediyor)**

Loizou, P. C., 2007, *Speech Enhancement: Theory and Practice*, CRC Press.

Martin, R., 2001, Noise Power spectral density estimation based on optimal smoothing and minimum statistics, *IEEE Trans. on Speech and Audio Processing*, vol. 9, no. 5, p. 504-512.

McAulay, R. J., Malpass, M. L., 1980, Speech enhancement using a soft-decision noise suppression filter, *IEEE Trans. Acoust. Speech Signal Process.*, 28, 137-145.

Mittal, U., Phamdo, N., 2000, Signal/noise KLT based approach for enhancing speech degraded by colored noise, *IEEE Trans. Speech Audio Processing*, 8:159{167.

NOIZEUS, <http://ecs.utdallas.edu/loizou/speech/noizeus/>, erişim tarihi: 13.07.2015.

Özkan, K., Seke, E., 2015, Image denoising using common vector approach, *IET Image Process.*, 9, 709-715.

Quackenbush, S., Barnwell, T., and Clements, M., 1988, *Objective measures of speech quality*, NJ: Prentice-Hall, Eaglewood Cliffs.

Rabiner, L.R., Sambur, M.R., 1975, An algorithm for determining the Endpoint of Isolated Utterances, *The Bell Systems Technical journal*, Vol. 54, No.2, p.297-315.

Rangachari, S., Loizou, P.C., Hu, Y., 2004, A Noise Estimation Algorithm with Rapid Adaptation for Highly Non-Stationary Environments, *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, p. I-305-I-308.

Rangachari, S., Loizou, P., 2006, A noise estimation algorithm for highly nonstationary environments. *Speech Communication*, 28, 220-231.

Rezayee, A., Gazor, S., 2001, An adaptive KLT approach for speech enhancement, *IEEE Trans. Speech Audio Processing*, 9:87{95.

Swets, D.L. and Weng, J., 1996, Using discriminant eigenfeatures for image retrieval, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18, 8, 831-836.

**KAYNAKLAR DİZİNİ (devam ediyor)**

- Tribolet, J., Noll, P., McDermott, B., and Crochiere, R. E., 1978, A study of complexity and quality of speech waveform coders. Proc. IEEE Int. Conf. Acoust. , Speech, Signal Processing, 586-590.
- Zhang, L., Dong, W.S., Zhang, D., Shi, G.M., 2010, Two-stage image denoising by principal component analysis with local pixel grouping. Pattern Recognition 43:1531-1549.