

T.C
DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KALMAN FİLTRESİ İLE SES SİNYALLERİNİN İYİLEŞTİRİLMESİ

Cem KUTLU

YÜKSEK LİSANS TEZİ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

DIYARBAKIR

Haziran-2012

T.C
DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KALMAN FİLTRESİ İLE SES SİNYALLERİNİN İYİLEŞTİRİLMESİ

Cem KUTLU

YÜKSEK LİSANS TEZİ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

DİYARBAKIR

Haziran-2012

T.C. DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ
DİYARBAKIR

Cem KUTLU tarafından yapılan “Kalman Filtresi ile Ses Sinyallerinin İyileştirilmesi” konulu bu çalışma, jürimiz tarafından Elektrik-Elektronik Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

Jüri Üyeleri

Başkan : Yard.Doç.Dr. M.Ali ARESRİM

Üye : Yard.Doç.Dr. Abdulnasır YILDIZ

Üye : Yard.Doç.Dr. Sezai ASUBAY

Tez Savunma Sınavı Tarihi: 06/06/2012

Yukarıdaki bilgilerin doğruluğunu onaylarım.

.../...../.....

Prof. Dr. Hamdi TEMEL

Enstitü Müdürü

TEŞEKKÜR

Bu tez çalışması süresince yardımlarını esirgemeyen danışman hocam Yrd. Doç. Dr. M. Ali ARSERİM'e, çalışmalarım boyunca her konuda destek veren hocalarım Arş Gör. Muhittin BAYRAM, Arş. Gör. Hüseyin ERDOĞAN, Arş. Gör. Hüseyin ACAR, Arş. Gör. S.Berat EFE ve Dicle Üniversitesi Müh. Fak. Elektrik Elektronik Müh. Böl. öğretim elemanlarına teşekkür ederim.

Tezin hazırlanması sırasında her konuda fedakârlık gösteren, maddi ve manevi yardımlarını esirgemeyen aileme sonsuz şükranlarımı sunarım.

İÇİNDEKİLER

	Sayfa
TEŞEKKÜR.....	I
İÇİNDEKİLER.....	II
ÖZET.....	IV
ABSTRACT.....	V
ÇİZELGE LİSTESİ.....	VI
ŞEKİL LİSTESİ.....	VII
KISALTMA VE SİMGELER.....	X
1. GİRİŞ.....	1
1.1. Konuşma Haberleşmesinde Gürültü ve Bozulmanın Etkileri.....	1
1.2. Tezin İçeriği.....	2
2. KAYNAK ÖZETLERİ.....	3
3. MATERYAL VE METOT.....	7
3.1. Ses.....	7
3.1.1. Gürültü Çeşitleri.....	9
3.2. İstatistiksel Kavramlar.....	10
3.2.1. Beklenen Değer (Ortalama).....	10
3.2.2. Varyans.....	11
3.2.3. Korelasyon (İlinti).....	12
3.2.4. Kovaryans.....	13
3.2.5. Spektral Güç Yoğunluğu.....	14
3.3. Doğrusal Öngörülü Kodlama (Linear Predictive Coding).....	14
3.4. Pencereleme.....	17
3.5. Ayrık Fourier Dönüşümü.....	19
3.6. Sinyal İyileştirme.....	21
3.6.1. Spektral Çıkarma.....	21
3.6.1.1. Spektral Güç Çıkarma.....	25
3.6.1.2. Spektral Genlik Çıkarma.....	26
3.6.1.3. Spektral Çıkarma Yönteminin Dezavantajları.....	26
- Müzikal gürültü.....	26
- Distorsiyon.....	27
3.6.2. Gürültü Tahmini.....	28

3.6.2.1. VAD (Voice Activity Detection).....	28
3.6.2.2. Spektral Güç Yoğunluğu ile Gürültü Tahmini.....	30
3.6.3. Wiener Filtresi.....	30
3.6.3.1. Frekans Domeninde Wiener Filtresi.....	33
3.6.3.2. Wiener Filtresi ile Spektral Çıkarma Arasındaki İlişki.....	34
3.6.3.3. Wiener Filtresi ile Ses Sinyallerinin İyileştirilmesi.....	36
3.6.4. Kalman Filtresi.....	38
3.6.4.1. Filtrenin Sayısal Kökenleri.....	41
- Güncelleme Adımı.....	42
- Filtre Kazancı.....	44
- Kalman Filtresi ile Örnek Bir Uygulama.....	47
3.6.4.2. Kalman Filtresinin Parametre Tahminine Uygulanması.....	49
3.6.4.3. Kalman Filtresi ile Ses İyileştirme.....	54
4. ARAŞTIRMA BULGULARI.....	59
4.1. Kullanılan Veritabanı.....	59
4.2. Sinyal Gürültü Oranı (SNR).....	59
4.3. Uygulanan Yöntem.....	60
4.4. Benzetim Sonuçları.....	62
5. TARTIŞMA VE SONUÇ.....	83
6. KAYNAKLAR.....	85

ÖZET

KALMAN FİLTRESİ İLE SES SİNYALLERİNİN İYİLEŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Cem KUTLU

DİCLE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

2012

Ses sinyallerinin iyileştirilmesi ile, çeşitli algoritmalar kullanılarak gürültülü sinyallerin kalitesinin veya anlaşılabilirliğinin artırılması amaçlanmaktadır. Ses iyileştirme yöntemleri hem zaman hem de frekans domeninde gerçekleştirilebilir. Frekans domeninde gerçekleştirilen ve pratikte uygulanabilirliği en fazla olan Spektral çıkarma yöntemi, etkili bir iyileştirme tekniği olmasına karşın, iyileştirdiği sinyalde meydana gelen müzikal gürültü oldukça büyük bir dezavantajdır.

Ses iyileştirme için kullanılan bir başka yöntem ise Spektral Çıkarmanın filtre olarak uygulandığı Wiener filtresidir. Wiener filtresinin dezavantajı ise tüm frekanslarda, filtrenin sabit bir frekans cevabının olması ve temiz sinyal ile gürültünün spektral güç yoğunluklarının tahminini gerektirmesidir.

Kalman filtresi ile ses iyileştirme, en etkili yöntemlerden biridir. Özellikle GPS ve Navigasyon uygulamalarında sunduğu üstün tahmin yeteneği, son yıllarda ses sinyallerinin işlenmesinde de kullanılmaya çalışılmıştır. Ancak Kalman filtresi ile ses sinyallerini iyileştirebilmek için, bir takım parametrelerin bilinmesi gerekmektedir. Temiz sinyale ait AR (Autoregressive) katsayıları ve gürültüye ait kovaryans matrisi, Kalman filtresinin başarısını çok büyük ölçüde etkileyen ve bilinmesi gereken parametrelerdir. Pratikte mevcut olan sadece gürültülü sinyal olduğu için bu parametrelerin tahmin edilmesi oldukça zordur ve hala bu konu üzerinde çalışmalar devam etmektedir.

Bu çalışmada, Kalman filtresi için gerekli olan parametreler Spektral Çıkarma yöntemi ile iyileştirilmiş sinyal kullanılarak belirlenmiştir. Öncelikle Spektral Çıkarma, Wiener Filtresi ve Kalman filtresi ayrı ayrı incelenmiştir. Daha sonra, bir ses sinyaline ait AR katsayıları, hem en sık kullanılan yöntemlerden biri olan Doğrusal Öngörülü Kodlama (LPC) ile hem de Kalman filtresi ile hesaplanmıştır.

Farklı tipteki gürültülerle bozulmuş sinyallere, yukarıda belirtilen her üç yöntem ve son olarak bu çalışmada önerilen Spektral Çıkarma ile birleştirilmiş Kalman Filtresi uygulanmıştır. Uygulama sonuçları, objektif bir ölçüm olan SNR değerleri baz alınarak karşılaştırılmıştır.

Elde edilen sonuçlar; birleştirilmiş Kalman filtresinin Wiener filtresine ve Spektral Çıkarmaya oranla daha iyi bir SNR artışı sağladığını göstermiştir. Ayrıca birleştirilmiş Kalman filtresinin Spektral Çıkarmadan kaynaklanan müzikal gürültüyü bastırdığı da gözlemlenmiştir.

Anahtar Kelimeler : Ses İyileştirme, Kalman Filtresi, Spektral Çıkarma

ABSTRACT

SPEECH ENHANCEMENT WITH KALMAN FILTER

MsC THESIS

Cem KUTLU

**DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING
INSTITUTE OF NATURAL AND APPLIED SCIENCES
UNIVERSITY OF DICLE**

2012

Speech enhancement techniques aims to improve the quality or intelligibility of speech signals contaminated with background noise and can be implemented both in time and frequency domains. Spectral Subtraction, one of the most feasible methods in practice, is an effective way to enhance the noisy speech signals. However, a residual noise called musical noise occurs with the estimated speech signal and this is the major inconvenience of Spectral Subtraction.

Wiener Filter is an alternative approach for speech enhancement in the manner of Spectral Subtraction filter. The drawback of the Wiener Filter is the fixed frequency response at all frequencies and the requirement to estimate the power spectral density of the clean signal and the noise prior to filtering.

Kalman filtering is also one of the most effective methods in speech enhancement. In recent years, due to its magnificent accurate estimation characteristics especially in the research field of navigation and GPS, researchers tried to manipulate its advantages for useful purposes in signal processing.

However, to improve the speech signals with the Kalman Filter, some parameters such as the AR coefficients of the clean signal and the noise covariance matrix must be known. Determining the AR coefficients of clean speech signal plays a crucial role for the success of the Kalman Filter while the only noisy observations are available. In such condition it is very difficult to estimate these parameters and today researches on this issue are ongoing.

In this study, these parameters necessary to implement the Kalman Filter is determined using Spectral Subtraction. First of all, Spectral Subtraction, Wiener Filter and Kalman Filter is analyzed respectively. Then the AR coefficients of a speech signal is calculated using both Kalman Filter and the method of Linear Predictive Coding (LPC) that is frequently used in the literature.

All three methods mentioned above for speech enhancement are carried out for speech signals corrupted with different types of noise. Finally, Kalman Filter combined with Spectral Subtraction proposed in this study is applied to those signals and all results are compared based on output SNR values as an objective measurement for the enhancement performance.

Considering the obtained results, combined Kalman filter provided a better SNR improvement compared to the Wiener filter and Spectral Subtraction. Also combined Kalman filter suppressed the musical noise that occurred owing to Spectral Subtraction

Key Words: Speech enhancement, Kalman Filter, Spectral Subtraction

ÇİZELGE LİSTESİ

<u>Çizelge No</u>		<u>Sayfa</u>
Çizelge 3.1.	Kalman Filtresi Denklemleri	45
Çizelge 4.1.	SpAER veritabanından alınan ses dosyaları	57
Çizelge 4.2.	Noisex-92 veritabanından alınan gürültü örnekleri	57
Çizelge 4.3.	Farklı gürültü tipleri için Cümle_1'e uygulanan yöntemlerin sağladığı SNR değerleri	79
Çizelge 4.4.	Farklı gürültü tipleri için Cümle_2'ye uygulanan yöntemlerin sağladığı SNR değerleri	79

ŞEKİL LİSTESİ

<u>Sekil No</u>	<u>Sayfa</u>
Şekil 3.1. Ses yolunun yapısı	6
Şekil 3.2. (a)Pembe gürültünün, (b)Kahverengi gürültünün zaman-genlik grafiği	8
Şekil 3.3. x ve y değişkenlerinin farklı korelasyon değerleri için dağılım şemaları	11
Şekil 3.4. N=32 için Hamming Pencere fonksiyonu	17
Şekil 3.5. N=32 için Hanning Pencere fonksiyonu	18
Şekil 3.6. Zaman ve frekans domenlerinde temiz ve gürültülü sinyal örnekleri	19
Şekil 3.7. Spektral Çıkarma yönteminin genel yapısı	21
Şekil 3.8. Gürültünün sinyal üzerindeki etkisi:a)temiz sinyal b)gürültülü sinyal c)temiz sinyal spektrumu d)gürültülü sinyal spektrumu	22
Şekil 3.9. Sırasıyla a) temiz b) gürültülü c) iyileştirilmiş sinyallerin spektrogramları	26
Şekil 3.10. Konuşmanın olduğu kesitlerin 1, olmadığı kesitlerin 0 olarak belirlendiği bir VAD uygulaması	28
Şekil 3.11. Kalman Filtresinin her iterasyon için gerçekleştirdiği işlemler	40
Şekil 3.12. Kalman filtresi blok diyagramı	42
Şekil 3.13. Gürültülü bir AR sinyaline Kalman filtresi uygulanarak elde edilen iyileştirilmiş sinyal	47
Şekil 3.14. Kalman Filtresi ile elde edilen 1.katsayı	49
Şekil 3.15. Kalman Filtresi ile elde edilen 2.katsayı	50
Şekil 3.16. Kalman Filtresi ile elde edilen 3.katsayı	50
Şekil 3.17. Kalman Filtresi ile elde edilen 4.katsayı	51
Şekil 3.18. Kalman Filtresi ile tahmin edilen katsayı değerleri	52
Şekil 3.19. Kalman Filtresinin tahmin ettiği katsayılar ile ses sinyallerinin tekrar oluşturulmasına ait simülasyon sonuçları	53
Şekil 3.20. Ses sinyalinin pencereleme işlemi ile çerçevelere bölünmesi	55
Şekil 4.1. Birleştirilmiş Kalman filtresi ile ses iyileştirme yöntemine ait blok diyagramı	59
Şekil 4.2. Fabrika gürültüsünde -5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	61
Şekil 4.3. Fabrika gürültüsünde 0dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	62

Şekil 4.4. Fabrika gürültüsünde 5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	63
Şekil 4.5. Beyaz gürültüde -5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	64
Şekil 4.6. Beyaz gürültüde 0dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	65
Şekil 4.7. Beyaz gürültüde 5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	66
Şekil 4.8. F16 gürültüsünde -5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	67
Şekil 4.9. F16 gürültüsünde 0dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	68
Şekil 4.10. F16 gürültüsünde 5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	69
Şekil 4.11. Fabrika gürültüsünde -5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	70
Şekil 4.12. Fabrika gürültüsünde 0dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları	71

- Şekil 4.13.** Fabrika gürültüsünde 5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 72
- Şekil 4.14.** Beyaz gürültüde -5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 73
- Şekil 4.15.** Beyaz gürültüde 0dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 74
- Şekil 4.16.** Beyaz gürültüde 5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 75
- Şekil 4.17.** F16 gürültüsünde -5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 76
- Şekil 4.18.** F16 gürültüsünde 0dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 77
- Şekil 4.19.** F16 gürültüsünde 5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz
(b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma
ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve
spektrogramları 78

KISALTMA VE SİMGELER

AR	: Otoregresif (Autoregressive)
a_i	: Otoregresif katsayıları
c_{xy}	: Kovaryans
c_{xx}	: Otokovaryans
$C(n)$	: Gözlem vektörü
$D(n)$	: Giriş vektörü
DFT	: Ayırık Fourier Dönüşümü (Discrete Fourier Transform)
EM	: Estimate-Maximize
$E[.]$	: Beklenen Değer
f	: Frekans değişkeni
$F(n)$	: Durum geçiş matrisi
FIR	: Sonlu Tepki Cevabı (Finite Impulse Response)
IIR	: Sonsuz Tepki Cevabı (Infinite Impulse Response)
$K(n)$	: Kalman kazancı
LPC	: Doğrusal Öngörülü Kodlama (Linear Predictive Coding)
r_{xy}	: Korelasyon
r_{xx}	: Otokorelasyon
σ	: Standart sapma
σ^2	: Varyans
μ	: Ortalama
PSD	: Spektral Güç Yoğunluğu (Power Spectral Density)
P_{xx}	: Güç Spektrumu
$Q(n)$	: Süreç gürültüsünün kovaryansı
$R(n)$	: Ölçüm gürültüsünün kovaryansı
SNR	: Sinyal Gürültü oranı (Signal to Noise Ratio)
VAD	: Voice Activity Detection
VoIP	: İnternet Protokolü üzerinden Ses (Voice over Internet Protocol)
wifi	: Kablosuz bağlantı (wireless fidelity)
ZCR	: Sıfır Geçiş oranı (Zero Crossing Rate)

1. GİRİŞ

Ses iyileştirme sistemleri, konuşmanın kalitesini ve anlaşılabilirliğini geliştirmeyi amaçlayan sistemlerdir. Bununla birlikte, ortam gürültüsü, iletişim esnasında meydana gelen veri kaybı, sınırlı bant genişliği gibi bir çok etken ses sinyallerinin kalitesini veya anlaşılabilirliğini olumsuz etkileyebilir. Dolayısıyla sinyal iyileştirme; wifi, mobil telefonlar, VoIP, hands-free kullanılan iletişim araçları, uzaktan iletişimli toplantılar (telekonferans), taşıt (kabin) haberleşmesi, işitme cihazları gibi bir çok uygulamada oldukça fayda sağlamaktadır.

Günlük hayatta, yukarıda da değinilen bir çok uygulamada tek girişli sistemler (single microphone) kullanıldığından yapılan sinyal iyileştirme çalışmaları da genellikle tek kanallı sistemler üzerine gerçekleştirilmektedir. Tek kanallı gürültü azaltma sistemleri, sinyal ve gürültü süreçlerinin zamansal-spektral yapılarını kullanarak gürültünün duyulabilirliğini bastırmaya çalışır. Diğer yandan çok kanallı sistemler ise gürültüyü, gürültülü sinyalden ayırmayı amaçlar (Vaseghi 2008).

1.1. Konuşma Haberleşmesinde Gürültü ve Bozulmanın Etkileri

Ses haberleşmesi bağlamında elektronik sistemler, konuşma sinyalinin iletim, kayıt, tekrar çalma, analiz veya sentez işlemlerini gerçekleştirmektedir. Bu belirtilen amaçlardan herhangi biri için bir sistem tasarlanırken, gürültünün etkileri de mutlaka dikkatli bir şekilde hesaplanmalıdır. Değişik tiplerde gürültü ve bozulma tanımlanabilmektedir. Bazı sinyal işleme kavramları bu gürültü ve bozulmaların etkilerinin azaltılmasında kullanılabilmekte ve böylece konuşma sinyallerinin kalitesi veya anlaşılabilirliği artırılabilir. Sayısal sinyal işleme, olaya bağlı olarak çeşitli güçlü araçlar sunmaktadır.

Konuşma sinyalinin alınması, gönderilmesi veya oluşturulması süreçlerini içeren her sistem, konuşma sinyalinin kalitesinde düşüşe neden olabilecek geniş çaplı etkilere maruz kalmaktadır. Bu etkiler, dış sinyal karışımlarını ve kayıtlardaki arka plan gürültüsünü kapsadığı gibi, yankı etkisi veya analog elektro akustik cihazlar veya yükselteçlerden kaynaklanan bozulmaları da içerir.

Bu olumsuz etkilerin bir kısmı, örnekleme hızı ve kelime uzunluğu gibi sistem parametreleri ile kontrol edilebilirken, diğerlerinin ortadan kaldırılması veya en aza indirilebilmesi sayısal sinyal işleme teknikleri ile mümkün olmaktadır.

Gürültü, bir sinyal haberleşmesine karışan istenilmeyen başka bir sinyaldir. Gürültü kaynakları çok çeşitlidir. Bunlar, hareket eden titreşen veya çarpışan cisimlerden; örneğin, işleyen makineler, hareket eden araçlar, bilgisayar fanları, klavye tuş sesleri, rüzgar, yağmur gibi audio frekansındaki akustik gürültülerden, sesin gönderilmesi ve alınmasına etki eden radyo frekansındaki elektromanyetik gürültülere kadar çeşitlilik göstermektedir.

Sinyal bozulması genelde bir sinyaldeki gönderme kanalının ideal olmayan karakteristiğinden, yankı, eko ve eksik örneklemeden kaynaklanan, sistematik ve istenmeyen değişimdir (Cecelioğlu 2010).

Gürültü ve bozulmalar haberleşme ve ölçüm sistemlerinin en önemli sınırlayıcı etkenleridir. Bu nedenle, gürültünün ve bozulmaların etkilerinin modellenmesi ve yok edilmesi haberleşme ve sinyal işlemenin çekirdeğini oluşturmaktadır.

1.2. Tezin İçeriği

Arka plan gürültüsüne maruz kalmış ses sinyallerinin Kalman Filtresi ile iyileştirilmesini amaçlayan bu çalışma aşağıda anlatıldığı şekilde organize edilmiştir,

Bu tezin ikinci bölümünde, Kalman Filtresi ile ses sinyallerinin iyileştirilmesine yönelik günümüze dek yapılan çalışmaların literatür taraması sunulmuştur.

Üçüncü bölümde, ses iyileştirme yöntemlerinden Spektral Çıkarma, Wiener Filtresi ve Kalman Filtresi ayrı ayrı incelenmiştir. Ayrıca Kalman Filtresi ile parametre tahmini de gerçekleştirilmiştir.

Dördüncü bölümde, Spektral Çıkarma ile birleştirilmiş bir Kalman Filtresi uygulanmıştır. Uygulanan tüm yöntemlerin simülasyon sonuçları karşılaştırılıp, elde edilen SNR iyileştirmeleri tablolar halinde sunulmuştur.

Bu tezin son bölümünde ise, elde edilen sonuçlar irdelenerek katkılar vurgulanmıştır. Ayrıca ileride yapılabilecek benzeri çalışmalar ve uygulama alanları için öneriler tartışılmıştır.

2. KAYNAK ÖZETLERİ

Ses sinyallerinin işlenmesinde, ses kalitesi ve anlaşılabilirliği, arka plan gürültülerinin etkisiyle önemli ölçüde bozulabilir. Özellikle temiz konuşma sinyallerinin uygulandığı kodlayıcılar veya otomatik konuşma tanıma gibi bir çok sistemin etkisiz hale gelmesine neden olabilir. Bu gibi nedenlerden dolayı konuşma iyileştirme algoritmaları son yıllarda oldukça ilgi çeken bir çalışma alanı olmuştur. İyileştirme algoritmaları arasında Wiener (1949) ve Kalman (1960) tarafından yapılan çalışmalara dayanan çok sayıda algoritma vardır.

Yapılan çalışmaların genelinde ses sinyallerinin modellenmesi “Autoregressive” (AR) yöntem baz alınarak yapılmıştır. AR modellemede, genellikle Doğrusal Öngörü (Linear Predictive) katsayıları olarak adlandırılan AR parametrelerinin bilinmesi gerekir. Bu parametrelerin hesaplanmasında da bir çok yöntem kullanılırken en sık kullanılan yöntemlerden biri Yule-Walker denklemlerinin çözümü olarak öne çıkmaktadır.

Kalman filtresini oluşturan denklemlerin gürültü etkisiyle bozulmuş ses sinyallerine uygulanabilmesi, hem ses sinyali hem de gürültü parametrelerinin ayrı ayrı bilinmesine bağlıdır. Ancak pratikte çoğu zaman mevcut olan sinyal zaten bozuk olduğu için, diğer bir ifadeyle temiz sinyal ile gürültü karışık olduğu için temiz sinyalin ve gürültünün ayrı ayrı parametrelerini bilebilmek pek de mümkün değildir. Dolayısıyla bu durumda filtrelemeden önce parametre tahmini yapabilmek gibi bir problem ortaya çıkmaktadır.

Kalman filtresini ses sinyallerinin iyileştirilmesinde uygulayan ilk araştırmacılardan olan Paliwal ve Basu (1987), deneysel çalışmalarında Kalman filtresinin Wiener filtresine göre belirgin bir farkla daha avantajlı olduğunu ortaya koymuşlardır. Ancak, Paliwal ve Basu çalışmalarında ses sinyaline ait bilinmeyen parametreleri gürültülü konuşma sinyalinden değil, temiz konuşma sinyalinden elde etmişlerdir, ki bu durum pratikte karşılaşılmaması pek mümkün olmayan bir durumdur. Problemin kendisi gürültülü sinyalden temiz sinyali elde etmek olduğu için modellemede ve filtrelemede kullanılacak sinyal parametrelerinin de doğrudan gürültülü sinyalden tahmin etmek gerekmektedir.

2. KAYNAK ÖZETLERİ

Ses sinyalinin gürültüden arındırma problemi Widrow (1975) tarafından Expectation-Maximization (EM) algoritmasının frekans domeninde kullanılmasıyla çözülmeye çalışılmıştır.

Sinyallerin ve bu sinyallere ait parametrelerin bir arada tahmininde kullanılan EM algoritması Weinstein (1987) tarafından kapsamlı çalışmalarla sunulmuştur.

Weinstein (1990), konuşma sinyalinin iyileştirilmesi problemini zaman domeninde ele almıştır. Bu yaklaşım sinyali lineer dinamik durum denklemleri ile modelleme ve EM algoritmasının uygulanmasını içermektedir. Sonuçta uygulanan algoritma, Lim ve Oppenheim algoritmasıyla benzer bir yapıdadır.

Gibson (1991), farklı gürültü çeşitleri için iyileştirme performansını arttırmak adına renkli gürültü kullanarak Kalman filtresinin kullanım alanını genişletmiştir. Gibson'ın önerdiği algoritma, gürültülü konuşma sinyallerine uygulanan Kalman filtresi ile iyileştirilen sinyalden elde edilen parametre tahmini arasında dolaşır. Çalışmada Yule-Walker denklemlerinin çözümüyle AR parametrelerinin kestirimi yapılmıştır.

Lee (1992), daha güçlü bir Kalman filtresi önermiştir. Önceki çalışmalara benzer olarak sinyalin iyileştirilmesi ile parametre tahmini birlikte yürütülmüştür. Bu çalışmayı diğerlerinden ayıran yenilik ise kullanılan non-lineer kestirim yöntemleridir.

Shen ve Deng (1999), ses sinyali iyileştirmesine H_∞ filtreleme diye adlandırılan bir yaklaşım kazandırmışlardır. Bu yaklaşımın geleneksel Kalman filtresinden ayrıldığı nokta, filtre hata kriterinin tanımlanmasından oluşmaktadır.

Fujimoto ve Ariki (2000), AR modeli kullanmadan Kalman filtresini frekans domeninde uygulamışlardır. Çalışmalarında klasik “Spektral Çıkarma” yöntemini kullanmışlardır.

İyileştirme veya geliştirme; bir şeyin değerini ya da kalitesini artırma anlamına gelir. Ses sinyallerinin iyileştirilmesinden ise sesin anlaşılabilirliğinin veya kalitesinin artırılması anlaşılabilir.

Ses sinyallerinin iyileştirilmesi iki sebepten ötürü oldukça zor bir problemdir. Bunlardan birincisi, sese bozucu etki yapan gürültü sinyallerinin doğası ve karakteristiği, zamanla ve uygulamadan uygulamaya çarpıcı bir biçimde değişebilir.

Bu yüzden pratikte, değişen şartlarda doğru sonuçlar verebilecek çok yönlü algoritmalar geliştirmek gerekir.

Diğeri ise, performans ölçümünün veya kriterlerinin her bir uygulama için ayrı ayrı tanımlanabilmesidir. Genel olarak performans ölçümü, kalite ve anlaşılabilirlik olmak üzere iki tane algısal kriter üzerine yapılır.

Gürültü azaltma teknikleri üzerine araştırmalar yaklaşık 40 yıldan fazla bir süre önce Schroeder tarafından Bell laboratuvarlarında yapılmıştır. Schroeder çalışmalarında Spektral Çıkarma yönteminin analog uygulamaları üzerinde durmuştur.

Boll (1979), yaklaşık 15 yıl sonra bu uygulamayı tekrar ele alarak sayısal domende çalışmıştır. O zamandan sonra bu yöntem üzerinden bir çok çalışma geliştirilerek yapılmıştır.

Sinyal alt uzay ayrışımına dayanan bir diğer önemli çalışma ise Ephraim ve Van Trees (1995) tarafından sunuldu. Bu çalışmada önerilen yöntem, gürültülü sinyalin vektör uzayının, sinyal-artı-gürültü alt uzayı ve gürültü alt uzayı olmak üzere ortogonal iki alt uzaya ayrıştırılmasıydı.

Yukarıda bahsedilen her iki yöntem de parametrik olmayan (non-parametrik) yöntemlerdir. Ses sinyallerinin iyileştirilmesinde kullanılan diğer algoritmalar ise parametrik yöntemlerden oluşur. Bu yöntemlerde ses sinyali, normal dağılımlı (Gaussian) gürültüye gömülmüş AR metoduyla modellenir.

Parametrik yöntemle uygulanmak istenen ses iyileştirme algoritmaları iki adımdan oluşur:

- AR katsayılarının ve gürültü varyanslarının tahmini
- Tahmin edilen parametreleri kullanarak Kalman filtresinin uygulanmasıyla temiz ses sinyalinin tahmin edilmesi

Günümüzde gürültü azaltmada uygulanan bir çok yöntemde tek mikrofon (single channel / single microphone) kullanılmaktadır. Daha öncede belirtildiği gibi aynı anda hem kaliteyi hem de anlaşılabilirliği arttırmak pek mümkün değildir.

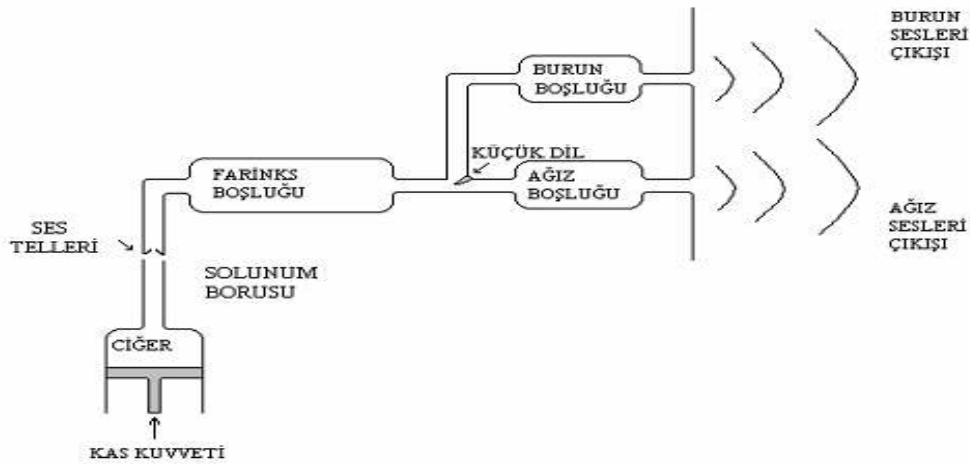
3. MATERYAL VE METOT

3.1. Ses

Ses dalgası, ses üretim sistemini meydana getiren anatomik yapıların istemli hareketleri sonucunda oluşan, akustik bir basınç dalgasıdır. Bu sistemin ana bölümleri ciğerler, nefes borusu, gırtlak, boğaz, ağız boşluğu ve burun boşluğudur.

Teknik terim olarak boğaz ve ağız boşluğu ‘ses yolu’ olarak tanımlanır. Dolayısıyla ses yolu, gırtlak çıkışından başlayıp, dudaklarda sona erer. Burun yolu ise damaktan başlar burun deliklerinde sona erer. Ses üretimi için kritik olan anatomik yapılar, ses telleri, damak, dil, dişler ve dudaklardır. Ağız iyice açıldığı zaman ağız boşluğunun arka tarafında duran damağın yumuşak uzantısına ‘küçük dil’ denir. Ses yolunu oluşturan bu anatomik yapılar, farklı pozisyonlar alarak değişik sesleri oluştururlar. Çene de ses yolunun şekil değişimini etkilediği için bu yapıların arasında yer alabilir.

Ses yolu uzunluğu, erişkin bir kadında yaklaşık ortalama 14 cm, erkekte 17cm civarındadır. Ses yolunu oluşturan yapıların, değişik konumlar alarak sesi oluşturmada sırasında, ses yolunun kesit alanı; 0-20 cm² arasında değişir. Ses yolunun temel yapısı Şekil 3.1.’de görülmektedir. Burun boşluğu, ses iletiminde işlevi olan bir yapı olup, uzunluğu erkeklerde yaklaşık 12cm’dir. Burun boşluğunun ses üretiminde ve iletimindeki kontrolü, damak, dolayısıyla küçük dil tarafından gerçekleştirilir. Bu sebeple burun yolu, ağızdan yayılan ses dalgası sıklık karakteristiğini etkiler.



Şekil 3.1. Ses yolunun yapısı

3. MATERYAL ve METOT

Ses, insan kulağını etkileyerek işitme duyusu oluşturan hava molekülleri titreşimleri, ya da bunların neden olduğu ufak hava basınç değişimleri gibi, ya da bu fiziksel olayın neden olduğu işitsel izlenim gibi tanımlanır.

Ses fizyolojisi ile ilgili bazı önemli kavramları şöyle açıklanmaktadır:

Ses Dalgası : Ses sıvı, katı, gaz ortamlarında 20 Hz ile 20 KHz arasındaki insan kulağının algılayabileceği basınç değişiklikleri olarak tanımlanmaktadır. Bu frekans aralığındaki mekanik dalgalar işitme duyumuzu uyardıklarından, bizim için özellikle önemli olan ses dalgalarını oluştururlar (İkiz 2006).

Frekans : İki tepe arasındaki uzaklık dalga boyudur. Bir saniyede gözlenen dalga tepesi sayısına frekans denir. Frekans sesin tizliğini belirler. Saniyedeki çevrim (Cycle Per Second-CPS) veya Hertz (Hz) ile gösterilir. Düşük frekanslar bas sesler, yüksek frekanslar ise yüksek seviyeli seslerdir (İkiz 2006).

Genlik : Sesin diğer bir karakteristik özelliği genliğidir. Sesler yumuşak veya yüksek olurlar. Bu havada, havayı sıkıştırmak için kullanılan güce bağlı olan bir tür küçük veya büyük basınca karşılık gelir. Ses gücü veya seviyesi için desibel (dB) birimi kullanılmaktadır. Kulağın algılama özelliği ile ilgili yapılan çalışmalarda ses gücünün artması ile hissedilen ses artışının doğrusal olmadığı ve logaritmik bir ses şiddeti ile duyma olduğu anlaşılmıştır. Bu nedenle algılanan ses logaritmik bir büyüklüktür. Haberleşme sistem ve cihazlarının yapısı ve ölçü birimleri de bu sebepten logaritmik olarak düzenlenmiştir (İkiz 2006).

Gürültü : Periyodik olmayan titreşimlerdir. Kulağın, teknik duyumu bakımından sınırları zorlayan ve psikolojik rahatsızlık doğuran seslere genel olarak gürültü denmektedir (İkiz 2006).

Sesin oluşturulmasında ses yolunun fiziksel olarak değişmesi, insan sesinin dalga şeklinin de zamana bağlı olarak değişmesine neden olur. Fakat insanların ses üretme ve sesi algılama sistemlerindeki organların yapılarından kaynaklanan bazı sınırlamalardan dolayı, ses dalga şeklinde değişime sebep olan yapıların pozisyonlarının kısa zaman aralıklarında değişmediği farz edilebilir. Bu durum konuşma işleme algoritmaları için çok kullanışlı bir özellik ortaya koymaktadır.

Özellikle frekans bölgesi konuşma işleme algoritmalarında konuşma sinyalinin, benzer özellikler taşıyan kısa süreli konuşma parçalarına ayrılarak incelenmesi yöntemine sıklıkla başvurulmaktadır.

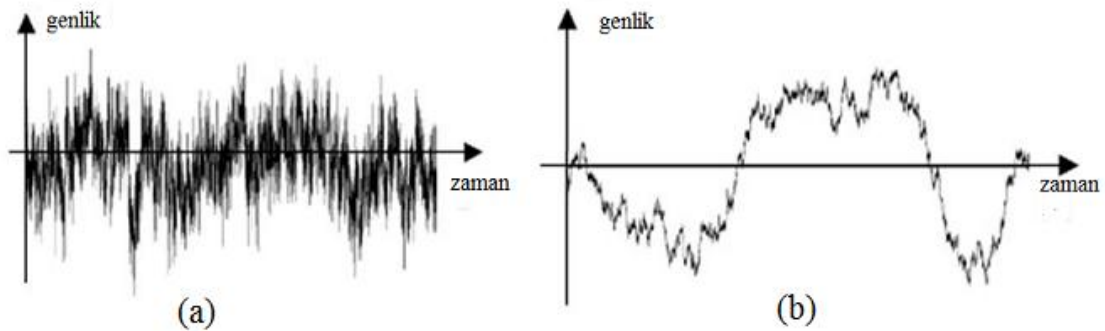
Zaman bölgesinde ses sinyallerinin dalga şekilleri incelendiğinde, sinyal periyotları, yoğunlukları, süreleri ve her bir ses parçasının sınırları tespit edilebilir. Ancak, “birleşik telaffuz” (coarticulation) olarak adlandırılan, ardı ardına gelen seslerin birbirlerini etkilemesi durumunda bu sınırları belirlemek mümkün olmayabilir (Cecelioğlu 2010).

3.1.1. Gürültü Çeşitleri

Gürültülerin karakterlerini tespit edip onları tanımak, gürültüden arındırma işine kalkışmadan önce yapılması gereken ilk çalışmadır. Gürültü yapısı, uygulanacak olan sinyal iyileştirme yöntemini belirlemek adına önemli bir faktördür. Bu nedenle doğru modellenmiş gürültü, geliştirilen algoritmanın iyi bir performans göstermesine önemli bir katkı sağlar. Ayrıca sinyal iyileştirme algoritmasının veya modelinin farklı gürültüler üzerinde etkilerinin bilinmesi bu algoritmanın veya modelin kullanılabilirliği açısından önem taşımaktadır

Yapı ve özelliklerine göre gürültüler aşağıdaki gibi gruplandırılabilirler;

Renkli Arka Plan Gürültüsü: Genellikle bozucu etki yarattığı sinyal ile ilintisizdir(uncorrelated). Araba, sokak, fan, makine, kokpit, helikopter sesleri gibi gürültüler günlük hayatta karşılaşılan gerçek gürültü tipleridir ve zamanla değişen karakteristiğe sahiptirler. Renkli gürültünün Pembe ve Kahverengi isimli iki klasik çeşidi Şekil 3.2.’de görülmektedir.



Şekil 3.2. (a)Pembe gürültünün, (b)Kahverengi gürültünün zaman-genlik grafiği

3. MATERYAL ve METOT

Birbirine karışan konuşmalar: Bu tip gürültüler, özellikle, çok sayıda konuşmacının aynı anda konuşması durumunda, örneğin kafeterya, gürültülü bir sınıf gibi yerlerde ortaya çıkan ve ilgilenilen konuşma sinyalinin alınmasında sinyal kalitesini önemli ölçüde etkileyen, günlük hayatta sıklıkla karşılaşılan gürültü tipidir.

Dürtü gürültüsü: Hızlıca çarpılmış kapı gürültüsü, arşivlenmiş gramafon kaydında bulunan gürültü, anahtarlama gürültüsü şeklinde, kısa süreli “açma/kapama” gürültü darbelerinden oluşur.

Toplamsal olmayan gürültü: Bu tip gürültüler, mikrofonların doğrusal olmayan kısımlarından, konuşmacılardan ve kanal bozulmalarından (iletim hatlarındaki konuşma sinyali) kaynaklanır.

Sinyalle ilişkili gürültü: Yankı ve eko, sinyalle ilişkili gürültü tiplerine örnek olarak gösterilebilirler.

Beyaz gürültü: Beyaz gürültü bütün frekanslarda eşit güce sahip, sinyalle ilişkili olmayan bir gürültü tipidir.

Genel olarak durağan olmayan ve hakkında ön bilgiye sahip olunmayan gürültülerle uğraşılması daha zor olmaktadır. Bu tarz gürültüler zamanla değiştiği için, önceki zaman aralıklarındaki gürültü karakteristiğine göre tahmin yapılarak gürültünün yapısının tespit edilmesi pek mümkün olmamaktadır. Bununla birlikte ilgilenilen ses ile benzer özelliklerdeki rakip seslerin oluşturduğu gürültülerden arındırmak da oldukça güçtür (Şahin 2007).

3.2. İstatistiksel Kavramlar

3.2.1. Beklenen Değer (Ortalama)

Beklenen değer ifadesi sinyal işlemede oldukça önemli bir rol oynar. Rastgele değişen süreçlerin olasılık modelleri genellikle beklenen değerlerin bir fonksiyonu olarak ifade edilir. Sinyal işleme uygulamalarında, uygun bir istatistiksel model seçebilmek için genellikle önceki deneyimler veya eldeki mevcut veriler ve hatta sezgisel tahminler (örneğin Gaussian dağılım) kullanılır. Modeli tanımlamak için, sürecin ortalama değeri ve kovaryansı gibi genellikle beklenen değerleri olan model parametrelerinin belirlenmesi gerekir.

Gürültü azaltma için tasarlanan filtrelerde veya doğrusal kestirim (linear prediction) gibi bir çok uygulamada sürecin ortalama değeri veya korelasyon fonksiyonu gibi değerlerin tahmin edilmesi gerekmektedir.

X gibi bir değişkenin dağılımının en önemli değeri dağılımın merkez değeridir. Bu merkez değeri genellikle *ortalama* ya da *beklenen değer* olarak adlandırılır ve Eşitlik (3.1)'deki gibi ifade edilir (Vaseghi 2008).

$$E[X] = \mu_X = \frac{1}{N} \sum_{i=0}^{N-1} X_i \quad (3.1)$$

3.2.2. Varyans

Ortalama değer, bir rastgele değişkenin, merkezi çevresindeki yayılımı hakkında bilgi vermez. Bu yayılmayı ölçmek için kullanılan parametre *varyans* olarak adlandırılır ve Eşitlik (3.2)'de gösterildiği gibi ifade edilir (Vaseghi 2008).

$$\text{Var}(X) = \sigma_X^2 = E\left[(X - E[X])^2\right] \quad (3.2)$$

Varyansın karekökünün pozitif değeri *standart sapma* olarak adlandırılır.

$$\sigma_X = \sqrt{E\left[(X - E[X])^2\right]} \quad (3.3)$$

X bir rastgele değişken, $c \in \mathbb{R}$ ve k bir doğal sayı olmak üzere:

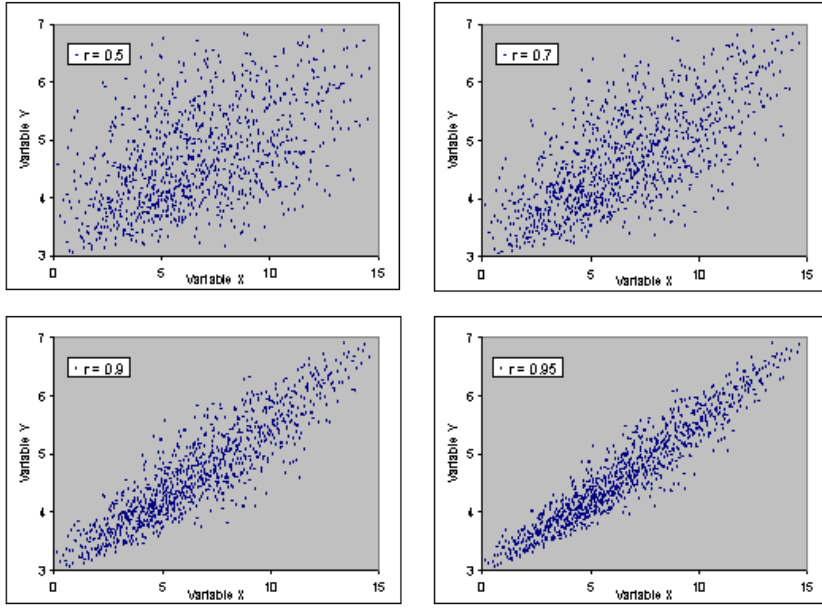
- a) $E\left[(X - c)^k\right]$ değerine X 'in c ye göre k 'inci momenti,
- b) $E[X^k]$ değerine X 'in k 'inci momenti,
- c) $E[X]$ değerine X 'in beklenen değeri,
- d) $E\left[(X - E[X])^2\right]$ değerine X 'in varyansı denir.

Sinyal işleme uygulamalarında en sık kullanılan *beklenen değerler* birinci dereceden moment yani *ortalama* ve ikinci dereceden momentler yani *korelasyon*, *kovaryans* ve *spektral güç yoğunluğu (psd)* fonksiyonlarıdır.

3.2.3. Korelasyon (İlinti)

Korelasyon, iki sinyalin birbirleriyle olan benzerliklerini veya birbirlerine olan bağımlılıklarını ifade eder ve ikinci dereceden momente karşılık gelir.

Şekil 3.3., x ve y gibi iki değişkenin farklı korelasyon değerleri için çizilmiş dağılım şemalarını göstermektedir. Şekilden de anlaşılabacağı gibi iki değişkenin korelasyon değerleri arttıkça dağılım şeması da doğrusal bir çizgiye yaklaşmaktadır.



Şekil 3.3. x ve y değişkenlerinin farklı korelasyon değerleri için dağılım şemaları

Bir sinyal ile yine bu sinyalin zamanda geciktirilmiş versiyonunun korelasyonu *otokorelasyon* olarak adlandırılır. $x(n)$ gibi bir sinyalin otokorelasyon fonksiyonu $r_{xx}(n_1, n_2)$ olarak gösterilip, aşağıdaki gibi tanımlanır (Vaseghi 2008).

$$r_{xx}(n_1, n_2) = E[x(n_1)x(n_2)] \quad (3.4)$$

Rastgele değişen bir sürecin n_1 anındaki sonucuyla n_2 anındaki sonucu birbirinden bağımsız ise, $x(n_1)$ ve $x(n_2)$ 'in ilintisiz ve dolayısıyla $r_{xx}(n_1, n_2) = 0$ olduğunu söyleyebiliriz. Örneğin, beyaz gürültü ilintisiz (uncorrelated) sinyale bir örnektir.

Otokorelasyon fonksiyonunun bazı özellikleri aşağıda verilmiştir ,

$k = n_1 - n_2$ zaman gecikmesi olmak üzere;

- $r_{xx}(n_1 + \tau, n_2 + \tau) = r_{xx}(n_1, n_2) = r_{xx}(k)$
- $r_{xx}(-k) = r_{xx}(k)$
- $r_{xx}(k) \leq r_{xx}(0)$
- $r_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x(n)x(n+k)$
- Sıfır-ortalamalı sinyallerde $r_{xx}(0)$ değeri sinyalin gücünü verir.

3.2.4. Kovaryans

Kovaryans, x ve y gibi iki rastgele değişkenin ortak değişimlerini, diğer bir deyişle ortak varyanslarını ifade eder ve aşağıdaki gibi gösterilir,

$$\begin{aligned} c_{xy} &= E[(x - E[x])(y - E[y])] \\ &= E[xy] - E[x]E[y] \end{aligned} \quad (3.5)$$

x ve y değişkenlerinin birbirinden bağımsız olduğu durumlarda kovaryans sıfıra eşit olur. Korelasyon fonksiyonu, kovaryans cinsinden Eşitlik (3.6)'daki gibi ifade edilir

$$r_{xy} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}} = \frac{c_{xy}}{\sigma_x \sigma_y} \quad (3.6)$$

Otokovaryans ise, $x(n)$ gibi bir sinyalin ortalama değeri civarındaki dağılımını ölçen bir fonksiyondur ve Eşitlik (3.7)'deki gibi gösterilir.

$$\begin{aligned} c_{xx}(n_1, n_2) &= E[(x(n_1) - \mu_x(n_1))(x(n_2) - \mu_x(n_2))] \\ &= r_{xx}(n_1, n_2) - \mu_x(n_1)\mu_x(n_2) \end{aligned} \quad (3.7)$$

Eşitlik (3.7)'den de anlaşılacağı gibi sıfır-ortalamalı sinyallerde otokorelasyon ve otokovaryans fonksiyonları birbirine eşittir. Ayrıca $c_{xx}(n_1, n_2)$ ifadesi sinyalin varyansını da verir.

3.2.5. Spektral Güç Yoğunluğu

Spektral Güç Yoğunluğu (Power Spectral Density) olarak bilinen ve kısaca *psd* olarak adlandırılan fonksiyon, bir sinyale ait güç spektrumunun, frekans eksenini boyunca farklı frekanslardaki dağılımını verir.

$X(n)$ gibi durağan bir sürecin *psd* fonksiyonu Eşitlik (3.8)'de gösterildiği gibi otokorelasyon fonksiyonunun Fourier dönüşümü alınarak elde edilir (Vaseghi 2008).

$$\begin{aligned} P_{xx}(f) &= E[X(f)X^*(f)] \\ &= \sum_{k=-\infty}^{\infty} r_{xx}(k)e^{-j2\pi fk} \end{aligned} \quad (3.8)$$

Dolayısıyla spektral güç yoğunluğu bilinen bir sürecin otokorelasyonu, *psd* fonksiyonuna ters fourier dönüşümü uygulanarak elde edilebilir.

$$r_{xx}(k) = \int_{-\infty}^{\infty} P_{xx}(f)e^{j2\pi fk} df \quad (3.9)$$

3.3. Doğrusal Öngörülü Kodlama (Linear Predictive Coding)

Doğrusal öngörü, isminden de anlaşılacağı üzere, işarete ait bir örneğin, önceki örneklerin doğrusal bir fonksiyonu olarak öngörülebilmesi esasına dayanan bir yöntemdir. Rastgele bir sürecin örnek değerlerinden oluşan bir küme verildiğinde, bu durağan ayırık zamanlı rastgele sürecin ileri veya geri yönde bir sonraki veya önceki değerinin kestirilmeye çalışılmasıdır.

Doğrusal Öngörülü Kodlama (Linear Predictive Coding – LPC) yöntemi ses dalgasının durağan kabul edilebileceği kısa bir aralıkta n 'inci örneğin kendinden önce gelen p adet ses örneğinin lineer bir birleşimi olarak yazılabileceği varsayımından hareketle ortaya konmuştur. Bu varsayımın altında, konuşma sırasında ses oluşumuna ilişkin organların ve ses yolunun kısa süreler boyunca akustik model çerçevesinde durağan bir yapıda kaldığı gerçeği yatmaktadır. Böylelikle kısa süreler boyunca ses yolunun akustik yapısı parametrik olarak modellenmiş olacaktır. Bu varsayım Eşitlik (3.10) ile formülize edilmiştir.

$$\hat{s}(n) = \sum_{k=1}^p a_k s(n-k) \quad (3.10)$$

$\{ u(n) \ u(n-1) \ u(n-2) \ \dots \ u(n-M) \}$ durağan ayrık zamanlı rastgele bir sürecin n anına kadar ve n anından önceki toplam $M+1$ tane değerini göstermek üzere ; $\{ u(n-1) \ u(n-2) \ \dots \ u(n-M) \}$ değerlerini kullanarak $u(n)$ değerini elde etmeye ($u(n-1)$ anına göre) ileri yönlü tek adımlı doğrusal öngörü, $\{ u(n) \ u(n-1) \ u(n-2) \ \dots \ u(n-M+1) \}$ değerlerini kullanarak $u(n-M)$ değerini elde etmeye ($u(n-M+1)$ anına göre) geri yönlü tek adımlı doğrusal öngörü adı verilir (Kırçiçek 2007).

Tek adımlı ileri yönlü doğrusal öngörücü kullanarak konuşma işaretlerinin sentez ve analizinde kullanılan doğrusal öngörü katsayılarının (LPC katsayılarının) bulunmasını inceleyelim;

$$\hat{s} \approx a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p) = \sum_{k=1}^p a_k s(n-k) \quad (3.11)$$

Eşitlik (3.11) ifadesine göre öyle a_k katsayıları tespit edilebilir ki, $s(n)$ örneği, önceki p adet örneğin doğrusal birleşimi olarak tanımlanabilir. a_k katsayılarının işlenen çerçeve boyunca sabit olması kabulü ile (sürecin durağan olması) sinyelimiz bir AR rastgele süreç olarak modellenenebilir.

Herhangi bir n anı için ses yolunun akustik modelinin değişmediği kısa aralığın, ses sinyalinin Eşitlik (3.12) ile verilen pencereleme fonksiyonundan geçirilmesiyle elde edilmiş olan $s_n(m)$ ile ifade edildiği durumda, gerçek değer ile tahmin değeri arasındaki farktan dolayı Eşitlik (3.13) ile verilen tahmin hatası oluşur.

$$s_n(m) = \begin{cases} s(n+m)w(m) & 0 \leq m \leq N-1 \\ 0 & \text{diğer} \end{cases} \quad (3.12)$$

$$e_n(m) = s_n(m) - \hat{s}_n(m) \quad (3.13)$$

Tahmin katsayılarına ilişkin en iyi kestirim Eşitlik (3.14) ile verilen hata kareleri toplamının a_k katsayılarına göre minimize edilmesi ile Eşitlik (3.15) elde edilebilir.(Uslu, 2007)

$$E_n = [e_n(m)]^2$$

$$= \sum_{m=0}^{N-1} \left[s_n(m) - \sum_{k=1}^p a_k s_n(m-k) \right]^2 \quad (3.14)$$

$$\frac{\partial E_n}{\partial a_k} = 0 \quad k = 1, \dots, p \quad (3.15)$$

Eşitlik (3.15) ile verilen koşul tüm a_k katsayıları için uygulandığında çözümün Eşitlik (3.16)'yı sağlayan değerler olduğu görülür (Rabiner ve Juang 1993). $\phi_n(i, k)$ terimi kovaryans fonksiyonu olarak isimlendirilir.

$$\sum_{m=0}^{N-1} s_n(m-i) s_n(m) = \sum_{k=1}^p a_k \underbrace{\sum_{m=0}^{N-1} s_n(m-i) s_n(m-k)}_{\phi_n(i, k)} \quad (3.16)$$

Eşitlik (3.16) ile verilen koşulu sağlayan a_k katsayılarının bulunması için iki yöntem kullanılır. Bu yöntemlerden ilki otokorelasyon katsayıları yöntemi diğeri ise kovaryans matrisi yöntemidir (Uslu 2007).

$\phi_n(i, k)$ terimi için değişken dönüşümü yapılırsa kovaryans fonksiyonu otokorelasyon katsayıları cinsinden Eşitlik (3.17) ile verilen şekilde yazılabilir.

$$\phi_n(i, k) = r_n(i-k) = \sum_{m=0}^{N-1-(i-k)} s_n(m) s_n(m-i-k) \quad \begin{matrix} 1 \leq i \leq p \\ 0 \leq k \leq p \end{matrix} \quad (3.17)$$

Otokorelasyon katsayıları için $r(x) = r(-x)$ simetriklik koşulu da dikkate alınarak otokorelasyon katsayıları yöntemi Eşitlik (3.18) ile ifade edilebilir.

$$\begin{bmatrix} r_n(0) & r_n(1) & r_n(2) & \cdots & r_n(p-1) \\ r_n(1) & r_n(0) & r_n(1) & \cdots & r_n(p-2) \\ r_n(2) & r_n(1) & r_n(0) & \cdots & r_n(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_n(p-1) & r_n(p-2) & r_n(p-3) & \cdots & r_n(0) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} \begin{bmatrix} r_n(1) \\ r_n(2) \\ r_n(3) \\ \vdots \\ r_n(p) \end{bmatrix} \quad (3.18)$$

$m = m' + i$ değişken dönüşümü yapılarak $\phi_n(i, k)$ terimi m' cinsinden yazılırsa kovaryans matrisi yöntemi Eşitlik (3.19) ile verilen şekliyle elde edilir.

$$\begin{bmatrix} \phi_n(1,1) & \phi_n(1,2) & \phi_n(1,3) & \cdots & \phi_n(1,p) \\ \phi_n(2,1) & \phi_n(2,2) & \phi_n(2,3) & \cdots & \phi_n(2,p) \\ \phi_n(3,1) & \phi_n(3,2) & \phi_n(3,3) & \cdots & \phi_n(3,p) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_n(p,1) & \phi_n(p,2) & \phi_n(p,3) & \cdots & \phi_n(p,p) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} \begin{bmatrix} \phi_n(1,0) \\ \phi_n(2,0) \\ \phi_n(3,0) \\ \vdots \\ \phi_n(p,0) \end{bmatrix} \quad (3.19)$$

Doğrusal öngörü analiz sistemini gerçeklemek için öncelikle elde edilen doğrusal eşitliklerin efektif bir yöntemle çözülmesi gerekmektedir. p bilinmeyenli p doğrusal eşitliği çözmek için kullanılan farklı etkinlikte pek çok yöntem mevcuttur.

İster ses kodlamada olsun, ister ses özelliklerini ortaya çıkarmakta kullanılsın, LPC ile asıl amaçlanan ses işaretinin daha az sayısal değerle, a_k katsayılarıyla, temsil edebilmektir. İncelenen konuşma işaretine ilişkin a_k katsayıları ve gerekli uyarım işareti elde olduğu sürece, modelin elverdiği ölçüde, işaret yeniden oluşturulabilir (Uslu 2007).

3.4. Pencereleme

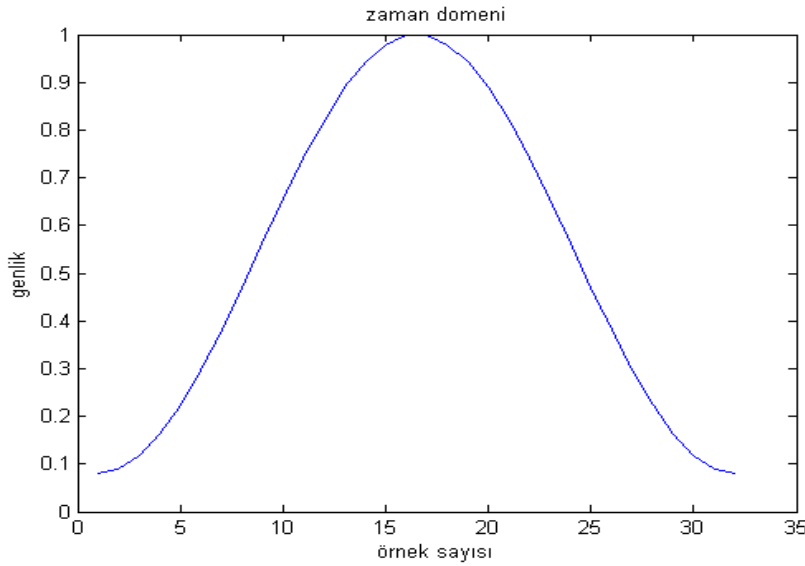
Pencereleme işlemi ile ses verisinin kısa zamanlı frekans spektrumunun elde edilmesi sağlanabilir. Pencereleme işleminde kullanılan fonksiyon ses sinyalinin durağan özellik gösterdiği varsayılan süre boyutunda tanımlanarak, ses sinyali üzerinde gezdirilmek suretiyle ses sinyali ile pencereleme fonksiyonu belirli aralıklarla çarpılır.

Pencereleme işlemlerinde çeşitli fonksiyonlar uygulanabilir. Bunlara örnek olarak dikdörtgen, Gauss, Hamming, Hann, Bartlett, üçgen, Bartlett-Hann, Blackman, Kaiser, Nuttal, Blackman- Harris ve Blackman-Nuttal fonksiyonları verilebilir.

Ses sinyallerinin işlenmesinde en çok kullanılan pencereleme fonksiyonları Hamming ve Hanning pencere fonksiyonlarıdır. Hamming ve Hanning pencere fonksiyonları uygulanarak pencere uçlarındaki süreksizlikler giderilebilmekte fakat uç noktaların özellikleri kaybedilmektedir. Bu sebeple pencereleme birbiri ardınca değil pencere boyutundan daha kısa boyutta kaydırılarak uygulanır. Kaydırma genel olarak pencere boyutunun yarısı kadar uygulanır (Cecelioğlu 2010).

Hamming pencere fonksiyonu Eşitlik (3.20) ile verilmiş olup zaman boyutundaki yapısı Şekil 3.4.'de gösterilmiştir.

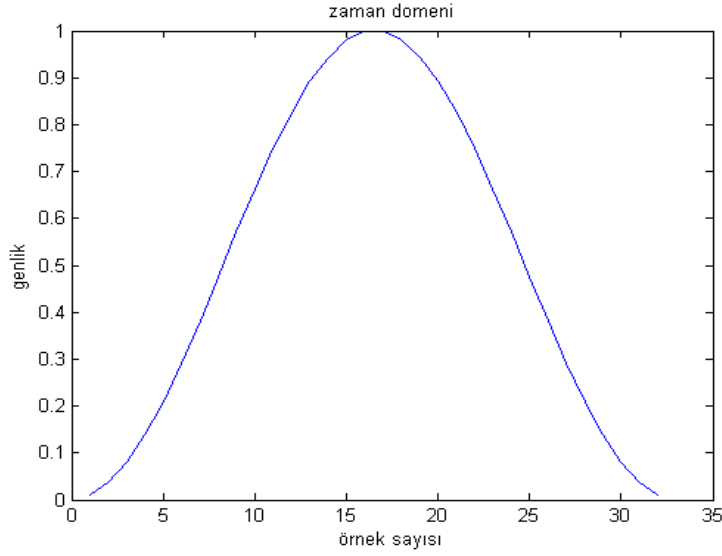
$$w(n) = \begin{cases} 0.54 + 0.46 \cos\left(\frac{2\pi n}{N-1}\right), & |n| \leq \frac{N-1}{2} \\ 0 & \text{diğer} \end{cases} \quad (3.20)$$



Şekil 3.4. N=32 için Hamming Pencere fonksiyonu

Hanning pencere fonksiyonu Eşitlik (3.21) ile verilmiş olup zaman boyutundaki yapısı Şekil 3.5.'de gösterilmiştir..

$$w(n) = \begin{cases} 0.5 + 0.5 \cos\left(\frac{2\pi n}{N-1}\right), & |n| \leq \frac{N-1}{2} \\ 0 & \text{diğer} \end{cases} \quad (3.21)$$

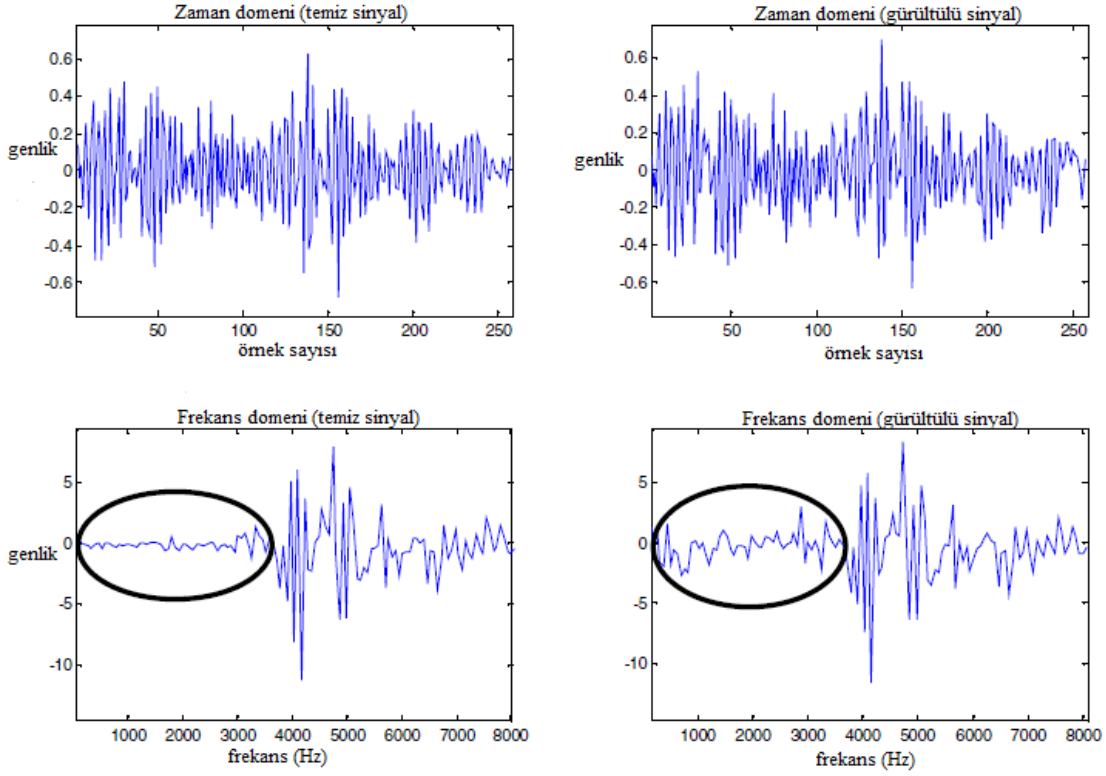


Şekil 3.5 N=32 için Hanning Pencere fonksiyonu

Konuşma tanıma açısından pencerelemenin önemi, kısa zamanlı ayrık Fourier dönüşümü yapılacağında, pencereleme sonucu sinyalin başlangıç ve bitiş değerlerini birbirine yaklaştırması ile pencere uçlarında oluşabilecek süreksizlikleri önlemesindedir. Bu sayede elde edilecek kısa zamanlı frekans spektrumu süreksizliklerden kaynaklanacak gereksiz spektrum verisinden ayıklanmış olacaktır (Uslu 2007).

3.5. Ayrık Fourier Dönüşümü

Sinyaller zaman domeninden frekans domenine DFT (Discrete Fourier Transform) ile dönüştürülür. Gürültünün nispeten dar bir frekans aralığı ile sınırlı olduğu göz önünde bulundurulursa, DFT dönüşümü ile bir sinyalde hangi frekansların gürültü içerdiği ve hangilerinin içermediği belirlenebilir. Böylece gürültüye ait frekans bileşenleri çıkarılabilir veya azaltılabilir



Şekil 3.6. Zaman ve frekans domenlerinde temiz ve gürültülü sinyal örnekleri

DFT'nin önemi Şekil 3.6.'dan da açık bir şekilde anlaşılabilir. Zaman domeninde temiz sinyal ile gürültü arasında net bir ayrım yapılamazken, frekans domeninde fark daha açıktır.

Fourier dönüşümü Eşitlik (3.22)'de verildiği gibi ifade edilir;

$$x(e^{jw}) = \sum_{t=0}^{L-1} x(t)e^{-jwt} \quad (3.22)$$

Burada L pencere uzunluğu olmak üzere, ters Fourier dönüşümü ise Eşitlik (3.23)'de verilmiştir.

$$x(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{jw})e^{jwk} dw \quad (3.23)$$

3.6. Sinyal İyileştirme

İnsanların en temel haberleşme şekli olan konuşma haberleşmesinde, sokakların kalabalığı, arabalar, fabrika makineleri gibi çevresel faktörlerden kaynaklanan bir çok gürültü ses sinyallerinde bozulmaya yol açar. Bu durum da konuşmanın anlaşılabilirliğini dolayısıyla iletişimi olumsuz yönde etkiler.

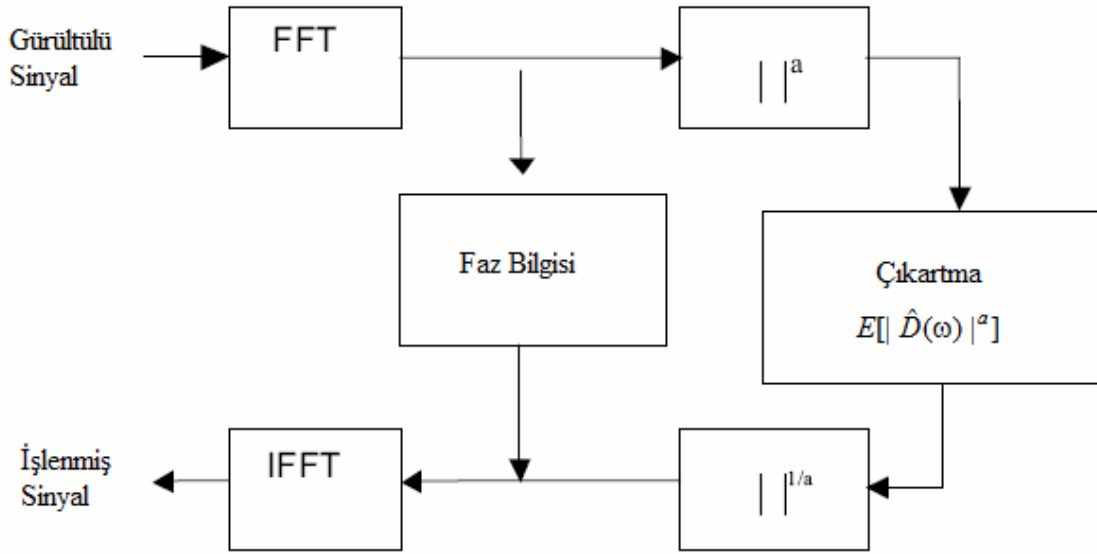
Sinyal iyileştirme sistemlerinin amacı, bu istenmeyen gürültüleri temizleyerek veya bastırarak olumsuz etkilerini minimuma indirmektir. Haberleşme, otomatik konuşma/konuşmacı tanıma sistemleri, işitme aygıtları ve daha bir çok alanda oldukça önemli olan ses iyileştirme ile ilgili ilk çalışma Boll (1979) tarafından gerçekleştirilmiştir. Günümüzde de hala çalışılmaya devam edilen bir araştırma alanı olarak önemini korumaktadır (Commins 2005).

3.6.1. Spektral Çıkarma

Spektral çıkarma, gürültülü sinyal spektrumundan tahmin edilen gürültü spektrumunun çıkarılması olarak tanımlanan bir *sinyal iyileştirme* yöntemidir. Literatürde, ses sinyali iyileştirmesine yönelik uygulanan ilk yöntemlerden biri olmakla beraber en popüler olan uygulamalardan da biridir.

Bu yöntemde gürültü spektrumu, genellikle konuşma sinyalinin olmadığı yani sadece gürültünün mevcut olduğu periyotlardan yararlanılarak tahmin edilmeye çalışılır. Ancak bu yöntemin gerçekleştirilebilmesi, göz önünde bulundurulması gereken bir kabule bağlıdır ki o da; gürültünün durağan veya çok yavaş değişim gösteren bir süreç olduğu varsayımdır.

Şekil 3.7.'de spektral çıkarma yönteminin temel yapısı özetlenmektedir. Buna göre önce Fourier dönüşümü ile frekans alanına aktarılan sinyalin, faz bilgisi alındıktan sonra genlik veya güç spektrumu elde edilir. Daha sonra belirlenen çıkarma işlemi uygulanır ve daha önce uygulanan işlemlerin tümünün tersi uygulanarak iyileştirilmiş sinyal tekrar elde edilir.

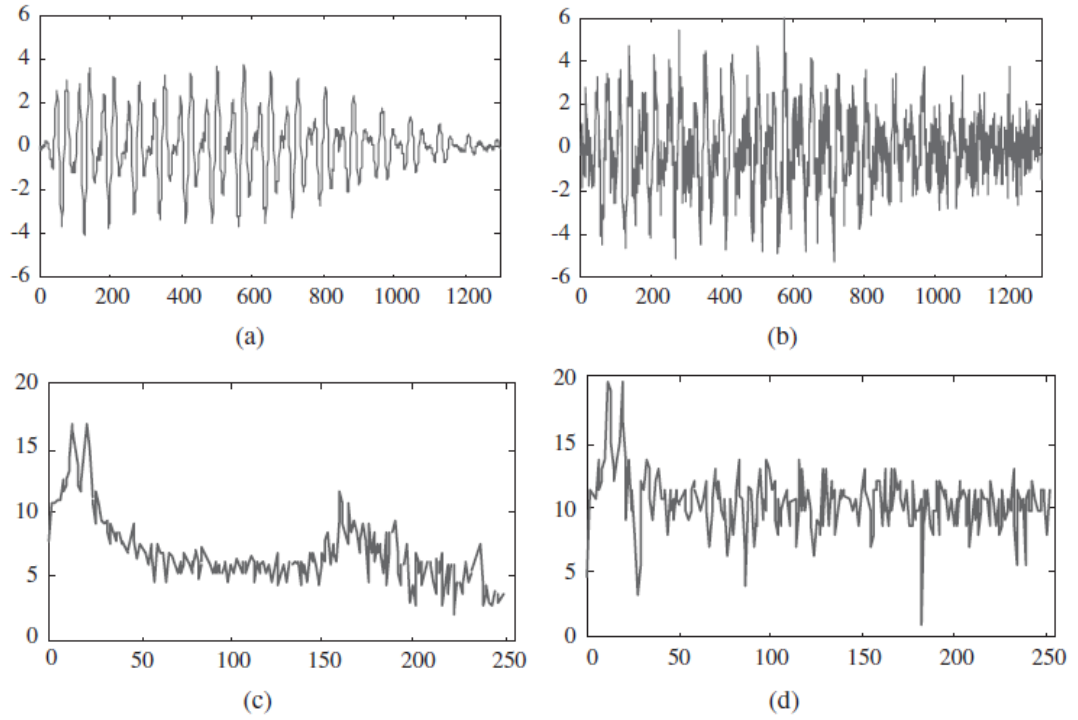


Şekil 3.7. Spektral Çıkarma yönteminin genel yapısı (Martin 1994)

Spektral çıkarma yönteminde, gürültünün rastgele değişiminden dolayı, kısa süreli genlik veya güç spektrumu tahmininde bazen negatif değerler elde edilebilir. Genlik veya güç spektrumunun negatif olmaması gerektiğinden bu değerlerin negatif olmayan değerlere dönüştürülmesi gerekmektedir. Bu dönüştürme veya düzeltme işlemi sinyalde birtakım bozulmalara neden olur. Ayrıca çıkarma işleminden sonra tekrar zaman domenine dönüştürülen iyileştirilmiş sinyalde, *müzikal gürültü* diye adlandırılan bir artık gürültü meydana gelir. Tüm bu dezavantajlarına rağmen matematiksel olarak pratikte gerçekleştirilebilir en uygun yöntemlerden biri olarak gösterilebilir.

İki kanallı sistemlerde gürültülü sinyal bir kanaldan, gürültü ayrı bir kanaldan alındığında gürültülü sinyali temizlemek oldukça rahat iken, pratikte özellikle de haberleşme sistemlerinde tek kanallı sistemler kullanıldığı için elimizde mevcut olan sadece gürültülü sinyal olacaktır. Bu tür durumlarda (rastgele) değişen gürültüyü tamamen temizlemek mümkün değildir. Ancak, gürültünün sinyal spektrumundaki ortalama etkisini azaltmak mümkündür.

Şekil 3.8.'de gösterildiği gibi gürültü, sinyalin genlik spektrumunda ortalamanın ve varyansın artmasına sebep olur. Gürültüdeki geliş güzel değişimler sinyal spektrumunun varyansının artmasına neden olur ve bu etkiyi ortadan kaldırmak pek de mümkün değildir. Sinyal spektrumunun ortalamasındaki artış ise, tahmin edilen gürültü spektrumunun, gürültülü sinyal spektrumundan çıkarılmasıyla engellenebilir.



Şekil 3.8. Gürültünün sinyal üzerindeki etkisi:a)temiz sinyal b)gürültülü sinyal c)temiz sinyal spektrumu d)gürültülü sinyal spektrumu

Zaman domeninde verilen gürültülü sinyal aşağıdaki gibidir (Boll 1979).

$$y(m) = x(m) + n(m) \quad (3.24)$$

$y(m)$ → gürültülü sinyal

$x(m)$ → temiz sinyal

$n(m)$ → gürültü

Eşitlik (3.24)'ün frekans domenindeki gösterimi aşağıda verildiği gibi olur.

$$Y(f) = X(f) + N(f) \quad (3.25)$$

Spektral çıkarmada izlenen yöntemi adım adım açıklayacak olursak;

Gelen sinyal, her biri N örnek uzunluğundan oluşan bölütlere ayrılır. Bu bölütler Hanning veya Hamming pencerelerinden bir tanesi kullanılarak oluşturulur ve her bölüt, ayırık Fourier Transformu ile N adet spektral örnek içeren bölütlere dönüştürülür.

Pencerelemiş sinyal aşağıdaki gibi gösterilebilir;

$$\begin{aligned} y_w(m) &= w(m)y(m) \\ &= w(m)[x(m) + n(m)] \\ &= x_w(m) + n_w(m) \end{aligned} \quad (3.26)$$

Yine aynı şekilde pencereleme işlemi frekans domeninde de gösterebilir;

$$\begin{aligned} Y(f) &= W(f) * Y(f) \\ &= X_w(f) + N_w(f) \end{aligned} \quad (3.27)$$

Eşitlik (3.27)'deki * operatörü konvolüsyon işlemi göstermekte olup, kolaylık sağlaması açısından pencereleme işlemi ifade eden w indisi düşürülerek, “spektral çıkarma”yı tanımlayan denklem aşağıdaki gibi ifade edilebilir (Berouti 1979).

$$|X(f)|^b = |Y(f)|^b - \alpha |\overline{N(f)}|^b \quad (3.28)$$

Burada; $|X(f)|^b$ orijinal sinyal spektrumunun tahmin edilen değerini, $|\overline{N(f)}|^b$ ise zaman-ortalamalı gürültü spektrumunu ifade etmektedir. Genlik spektral çıkarma için $b=1$, güç spektral çıkarma için $b=2$, seçilir ve α parametresi gürültülü sinyalden çıkarılan gürültü miktarını kontrol etmek amacıyla kullanılır. Tam çıkarma için $\alpha = 1$, aşırı çıkarma için $\alpha > 1$ olarak seçilir. Gürültü spektrumu, orijinal sinyalin mevcut olmadığı sadece gürültünün olduğu periyotlardan elde edilir ve aşağıdaki gibi gösterilir (Hoeldrich 1998).

$$|\overline{N(f)}|^b = \frac{1}{K} \sum_{i=0}^{K-1} |N(f)|^b \quad (3.29)$$

Eşitlik (3.29)'da geçen $|N(f)|^b$, i . gürültü bölütünün spektrumunu ifade etmektedir. K değeri ise sadece gürültünün mevcut olduğu bir periyotta mevcut olan bölüt sayısını göstermektedir. Ortalama gürültü spektrumu, birinci dereceden alçak geçiren dijital bir filtrenin çıkışı olarak kabul edilebilir.

$$|\overline{N_i(f)}|^b = p |\overline{N_{i-1}(f)}|^b + (1-p) |N_i(f)|^b \quad (3.30)$$

Burada, alçak geçiren filtrenin katsayısı olarak verilen p değeri genellikle 0.85 ile 0.99 arasında bir değer olarak seçilir (Vaseghi 2008).

Tahmin edilen ortalama gürültü spektrumu, gürültülü sinyal spektrumundan çıkarıldıktan sonra, genlik spektrumunun tahmini $|X(f)|^b$ gürültülü sinyalin fazı $\theta_Y(k)$ ile birleştirilerek, zaman domenine ters ayırık Fourier Transformu ile dönüştürülür.

$$x(m) = \sum_{k=0}^{N-1} |x(k)| e^{j\theta_Y(k)} e^{-j\frac{2\pi}{N}km} \quad (3.31)$$

3.6.1.1. Spektral Güç Çıkarma

Spektral Güç Çıkarma ifadesi, Eşitlik (3.28)'de $b=2$ alınarak elde edilir.

$$|X(f)|^2 = |Y(f)|^2 - |\overline{N(f)}|^2 \quad (3.32)$$

Yine bu ifadede çıkarma faktörü olarak adlandırılan kontrol parametresi $\alpha = 1$ olarak seçilir. Burada güç spektrumunu $E = [|X(f)|^2]$, zaman-ortalamalı güç spektrumunu $|\overline{X(f)}|^2$ ve anlık güç spektrumunu $|X(f)|^2$ olarak gösterebiliriz. Gürültülü sinyalin anlık güç spektrumu $|Y(f)|^2$ genişletilerek, Eşitlik (3.32) aşağıda verildiği gibi tekrar yazılabilir.

$$|X(f)|^2 = |X(f)|^2 + \left(|N(f)|^2 - |\overline{N(f)}|^2 \right) + X^*(f)N(f) + X(f)N^*(f) \quad (3.33)$$

(3.33)'de verilen eşitliğin her iki tarafının da beklenen değeri ($E[.]$) alınır, gürültünün ve temiz sinyalin birbirleriyle ilişkisiz olduğu varsayımı da hesaba katıldığında Eşitlik (3.34) elde edilir.

$$E[|X(f)|^2] = E[|X(f)|^2] \quad (3.34)$$

Bu ifadeden de anlaşılacağı gibi, anlık güç spektrumunun tahmin edilen değeri, temiz sinyalin güç spektrumuna yakınsamaktadır.

3.6.1.2. Spektral Genlik Çıkarma

Spektral Genlik Çıkarma, Eşitlik (3.28)'de verilen ifadede $b=1$ ve tam çıkarma için kontrol parametresi $\alpha = 1$ seçilerek elde edilir.

$$|X(f)| = |Y(f)| - |\overline{N(f)}| \quad (3.35)$$

Eşitlik (3.35)'de verilen $|\overline{N(f)}|$ ifadesi gürültünün zaman-ortalamalı genlik spektrumunu göstermektedir ve bu denklemde eşitliğin her iki tarafının beklenen değeri alınırsa Eşitlik (3.36) elde edilir.

$$\begin{aligned} E[|X(f)|] &= E[|Y(f)|] - E[|\overline{N(f)}|] \\ &= E[|X(f)| + |N(f)|] - E[|\overline{N(f)}|] \\ &= E[|X(f)|] \end{aligned} \quad (3.36)$$

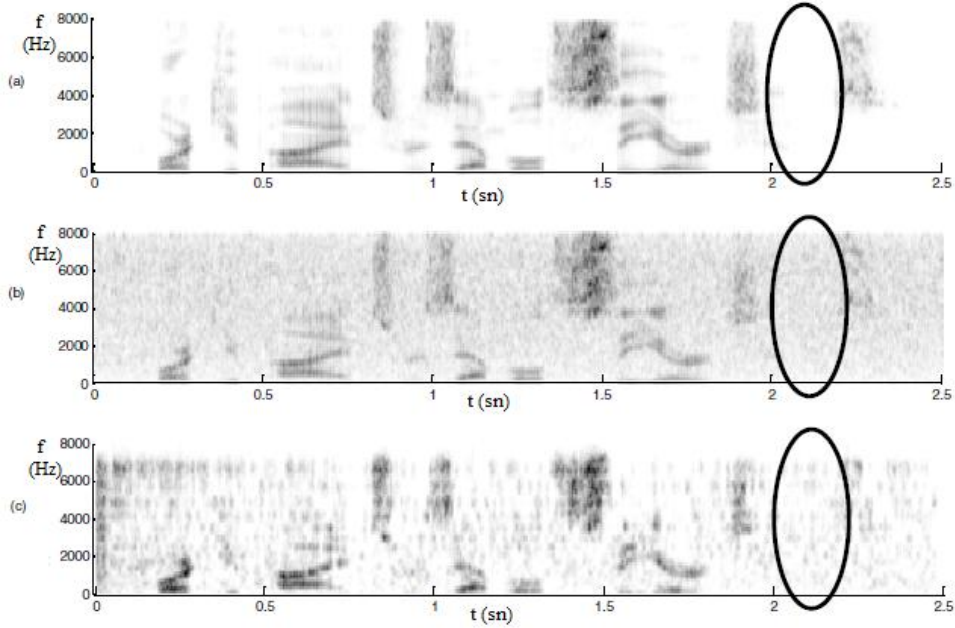
3.6.1.3. Spektral Çıkarma Yönteminin Dezavantajları

Spektral çıkarma, gürültü azaltmada kullanılan en popüler ve en etkili yöntemlerden biri olsa da bir takım dezavantajları bulunmaktadır. Spektral çıkarma ile gerçekleştirilen sinyal iyileştirmenin performansı büyük ölçüde tahmin edilen gürültü spektrumuna bağlıdır. Değişen çevresel koşullarda gürültüyü doğru bir şekilde tahmin edebilmek bu yöntemin en zor aşamasıdır. Dolayısıyla gürültü tam anlamıyla doğru tahmin edilemediğinde iki tane ciddi problem meydana gelir ki bunlar müzikal gürültü ve distorsiyondur.

- Müzikal gürültü

Gürültü tahminindeki hata, kısa süreli rasgele sinüzoidler oluşturur ve bu sinüzoidler ses sinyallerinde ton şeklinde bir müzikal gürültü meydana getirir. Yapısındaki rasgelelikten ve ton şeklindeki sesinden dolayı dinleyici açısından oldukça rahatsız edici bir gürültü tipidir. Yapılan çalışmalarda, müzikal gürültünün orijinal gürültüye oranla daha fazla rahatsız edici olduğu bir çok dinleyici tarafından belirtilmiştir.

Ses iyileştirmede amacın, bozuk sesi daha kaliteli ve anlaşılabilir bir konuşmaya dönüştürmek olduğu düşünülürse, bu durum büyük bir dezavantaj oluşturmaktadır.



Şekil 3.9. Sırasıyla a) temiz b) gürültülü c) iyileştirilmiş sinyallerin spektrogramları (Commins 2005)

Spektral çıkarma gürültülü sinyali iyileştirmesine karşın, Şekil 3.9’da görüldüğü üzere gerçek sinyalde bulunmayan bazı frekans bileşenleri, iyileştirilmiş sinyalde rastgele dağılmış bir biçimde bulunmaktadır.

- Distorsiyon

Spektral Çıkarma ile gerçekleştirilen iyileştirmenin meydana getirdiği ikinci problem ise ses sinyalinde oluşan distorsiyon yani bozulmadır. Bu bozulmaya yol açan sebep, müzikal gürültüyü de oluşturan gürültü tahminindeki hatadır.

Gürültülü sinyal spektrumundan gürültü spektrumunu çıkarırken, gürültüyle aynı frekansta olan ses sinyalinin bazı frekans bileşenleri de çıkarılmış olur. Bu durum gerçek ses sinyalinde bir miktar bozulmaya yol açar.

Meydana gelen müzikal gürültü ve distorsiyon problemlerinin her ikisini de aynı anda gidermek mümkün değildir. Bu nedenle konuyla ilgili çalışma yapan araştırmacılar müzikal gürültüyü kabul edilebilir bir eşik seviyesinin altında tutarak, bozulmayı engellemeye çalışmışlardır (Commins 2005).

3.6.2. Gürültü Tahmini

Önceki bölümlerde belirtildiği gibi gürültü karakteristiğinin doğru belirlenebilmesi, ses iyileştirme performansı için önemli bir faktördür. Gürültü tahmini için uygulanan farklı yöntemler mevcuttur. Bunlardan ‘Ses Aktivitesi Algılama’ (Voice Activity Detection) olarak adlandırılan yöntem sıklıkla tercih edilmektedir. Bunun dışında, sinyalin ilk anlarında konuşmanın olmadığı varsayımına dayanarak, ilk birkaç çerçevenin gürültü olduğu düşünülebilir ve böylece ortalama bir gürültü hesaplanabilir. Ancak gürültü ani değişimler gösteren bir yapıya sahip ise bu yöntemle çok da doğru bir gürültü tahmini yapılamayabilir.

3.6.2.1. VAD (Voice Activity Detection)

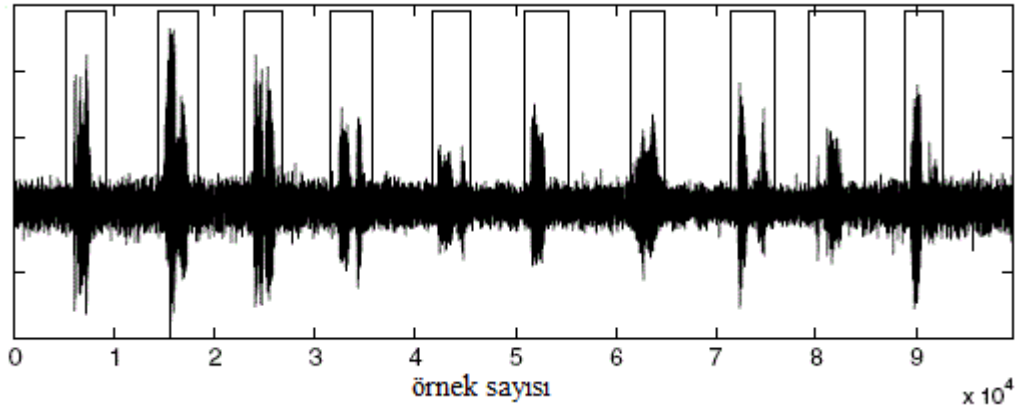
Bölütlere ayrılmış gürültülü bir sinyalde, hangi kesitlerin konuşma içerdiğini ve hangilerinin içermediğini belirlemek için çeşitli VAD algoritmaları kullanılabilir. Böylece konuşmanın olmadığı her bir kesit gürültü içeriyor demektir ve bu kesitlerden gürültüye ait istatistiksel bilgiler edinilebilir.

Literatürde *Zero Crossing Rate* (ZCR) olarak adlandırılan bir VAD uygulamasında, sinyalin genliğinin bir kesit içerisinde kaç kez x eksenini kestiği hesaplanır. Konuşmanın olmadığı kesitlerde, konuşma içeren kesitlere oranla daha düşük bir ortalama ZCR değeri gözlenir. Dolayısıyla her bir kesitte, ZCR değeri belli bir eşik seviyesinin altında ise bu kesit sadece gürültü, üstünde ise sadece gürültülü sinyal olarak belirlenir (Commins 2005).

Rabiner ve Sambur (1979) yaptıkları çalışmada ZCR değerinin yanı sıra, gürültülü sinyalin kısa-süreli enerji seviyesini de hesaplamışlardır. Örneğin, m . kesitteki gürültülü sinyalin enerjisi $Y(m)$, tahmin edilen ortalama gürültü enerjisi $N(m)$ ’den büyük ise o kesiti sadece gürültülü konuşma, aksi halde sadece gürültü olarak değerlendirmişlerdir.

$$VAD = \begin{cases} 1 & ZCR(m) > \alpha \text{ ve } Y(m) > N(m) \\ 0 & \text{diğer} \end{cases} \quad (3.37)$$

α eşik değeri olmak üzere VAD; 1 ise kesitin gürültülü konuşma, 0 ise gürültü içerdiğini ifade etmektedir.



Şekil 3.10. Konuşmanın olduğu kesitlerin 1, olmadığı kesitlerin 0 olarak belirlendiği bir VAD uygulaması

Literatürde uygulanan farklı VAD yöntemleri olmasına karşın tümünün ortak bir takım dezavantajları bulunmaktadır. Konuşma duraklamaları arasında çok kısa bir zaman dilimi olduğunda, tasarlanan VAD algoritması bu anları tespit edemeyebilir ve bu zaman diliminde gürültü karakteristiğinde ani değişimler meydana gelmiş olabilir. Bu gibi durumlar daha önce de değinilen müzikal gürültü ve distorsiyon gibi sorunlar oluşturmaktadır. Ayrıca bir çok VAD uygulaması düşük Sinyal/Gürültü Oranı (SNR) seviyelerinde gürültü ile konuşmayı ayıramayabilir. Bu nedenle iyileştirme esnasında gürültü ile birlikte ses sinyalinin bazı bileşenleri de yok olur veya bastırılır.

3.6.2.2. Spektral Güç Yoğunluğu ile Gürültü Tahmini

Literatürde VAD tekniği kullanmadan gürültü tahmini üzerine yapılan çalışmalar mevcuttur. Bu konu üzerinde duran ilk araştırmacılardan biri olan Doblinger'in (1995) yöntemine kısaca değinecek olursak;

Gürültülü sinyal alt bantlarına ayrılır ve k . alt bandın m . kesitinde gürültülü sinyalin gücü $Y(m)$ olarak tanımlanır. Gürültülü sinyal gücünün tahmini, Eşitlik (3.38)'deki gibi gösterilir.

$$Y_k(m) = \alpha Y_k(m-1) + (1-\alpha)Y_k(m) \quad (3.38)$$

Burada $Y_k(m)$ gürültülü sinyalin ortalama gücünü, α ise düzeltme katsayısını ifade etmektedir. m . kesitte gürültü gücünün tahmini Eşitlik (3.39)'da verildiği gibidir.

$$W_k(m) = \begin{cases} \lambda W_k(m-1) + \frac{1-\lambda}{1-\beta} (Y_k(m) - \beta Y_k(m-1)) & W_k(m-1) < Y_k(m) \\ W_k(m-1) & \text{diğer} \end{cases} \quad (3.39)$$

$W_k(m)$, gürültü gücünün tahmin edilen değerini ifade etmektedir. α , β ve λ parametreleri yapılan deneyler sonucunda, $0.7 \leq \alpha \leq 0.9$, $\beta = 0.96$ ve $\lambda = 0.998$ olarak seçilmiştir.

Martin (2001), gürültülü sinyalin spektral güç yoğunluğu tahminine dayalı bir yöntem önermiştir. Dobliger'in (1995) algoritmasından farklı olarak, sabit bir α yerine zamana bağlı güncellenen bir $\alpha_k(m)$ düzeltme katsayısı kullanmıştır.

$$\alpha_k(m) = \frac{|Y_k(m-1)|^2}{|W_k(m)|^2} \quad (3.40)$$

$|Y_k(m-1)|^2$, $(m-1)$. kesitteki gürültülü sinyal gücünün tahmin edilen değerini ve $|W_k(m)|^2$ ise m . kesitteki gürültü gücünün tahmin edilen değerini ifade etmektedir.

3.6.3. Wiener Filtresi

En küçük karesel hata (Least Squared Error) filtresinin teorisi ilk olarak Andrei Kolmogorov (1941) ve Norbert Wiener (1949) tarafından, birbirlerinden bağımsız olarak geliştirilmiştir. Kolmogorov'un geliştirdiği yöntem zaman domeninde gerçekleştirilen bir analiz iken, Wiener'in çalışması ise frekans domeninde gerçekleştirilen bir analizdir.

LSE filtreleri, çıkışta elde edilen sinyal ile istenen sinyal arasındaki ortalama uzaklığın karesini minimize etmeyi amaçlar. Wiener filtresinin çözümü için yani filtre katsayılarının hesaplanabilmesi için, girişe uygulanan sinyalin otokorelasyon fonksiyonunun ve bu sinyalle çıkışta elde etmek istediğimiz sinyalin korelasyon fonksiyonunun tahmin edilmesi gerekmektedir (Vaseghi 2008).

Wiener filtresi genellikle, gürültüyle karışık olarak gözlenen sinyallerin gerçek değerlerinin tahmin edilmesini gerektiren durumlarda kullanılır. Ayrıca Wiener filtresi de Kalman filtresi gibi roket, uydu vb. yörüngelerinin kestiriminde kullanılmaktadır. Nitekim Norbert Wiener, ikinci dünya savaşında bu problem üzerinde çalışmıştır.

Wiener filtresi, sonsuz dürtü yanıtı (IIR) veya sonlu dürtü yanıtı (FIR) filtre olarak tasarlanabilir. Ancak genellikle, hesaplanması basit ve pratiğe daha uygun olmasından dolayı FIR Wiener filtresi tercih edilir.

Wiener filtresinin giriş-çıkış ilişkisini gösteren eşitlik (3.41)'de verilmiştir (Diniz 1997).

$$x(m) = \sum_{k=0}^{P-1} w_k y(m-k) = \mathbf{w}^T \mathbf{y} \quad (3.41)$$

Bu eşitlikte; m zaman indeksini, $\mathbf{y}^T = [y(m), y(m-1), \dots, y(m-P+1)]$ vektörü giriş sinyalini, $x(m)$ filtre çıkışını ve $\mathbf{w}^T = [w_0, w_1, \dots, w_{P-1}]$ parametre vektörü ise Wiener filtresinin katsayı vektörünü ifade etmektedir.

Wiener filtresine ait hata sinyalini $e(m)$ olarak gösterirsek bu hata sinyalini, istenen sinyal ile çıkışta elde edilen sinyal arasındaki fark olarak ifade edebiliriz.

$$\begin{aligned} e(m) &= x(m) - \hat{x}(m) \\ &= x(m) - \mathbf{w}^T \mathbf{y} \end{aligned} \quad (3.42)$$

Eşitlik (3.42), hata sinyalinin filtre katsayı vektörü $\mathbf{w}$ 'ye bağlı olduğunu göstermektedir. Wiener filtresi, ortalama karesel hatayı minimize etme anlamında en uygun filtredir.

Sinyalin N örnek sayısına sahip bir bölütü göz önüne alındığında, istenen sinyal vektörü $[x(0), x(1), \dots, x(N-1)]$ ile gözlem vektörü $[y(0), y(1), \dots, y(N-1)]$ kullanılarak, Eşitlik (3.42) daha açık bir ifadeyle Eşitlik (3.43)'teki gibi yazılabilir.

$$\begin{pmatrix} e(0) \\ e(1) \\ e(2) \\ \vdots \\ e(N-1) \end{pmatrix} = \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{pmatrix} - \begin{pmatrix} y(0) & y(-1) & y(-2) & \cdots & y(1-P) \\ y(1) & y(0) & y(-1) & \cdots & y(2-P) \\ y(2) & y(1) & y(0) & \cdots & y(3-P) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & y(N-3) & \cdots & y(N-P) \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{P-1} \end{pmatrix} \quad (3.43)$$

Eşitlik (3.43)'ü vektör notasyonu ile yeniden yazabiliriz;

$$\mathbf{e} = \mathbf{x} - \mathbf{Y}\mathbf{w} \quad (3.44)$$

Eşitlik (3.44)'de $\mathbf{e}$ hata vektörünü, $\mathbf{x}$ istenen sinyal vektörünü, $\mathbf{Y}$ girişe verilen sinyal matrisini ve $\mathbf{Y}\mathbf{w} = \mathbf{x}$ Wiener filtresinin çıkış sinyal vektörünü ifade etmektedir. Giriş sinyal matrisinde P adet $[y(-1), \dots, y(-P-1)]$ giriş değerinin genellikle bilindiği kabul edilir veya sıfıra eşitlenir.

Ortalama karesel hatayı $E[e^2(m)]$ ile ifade edersek, Wiener filtresinin katsayıları bu ifadenin minimize edilmesi ile bulunur.

$$\begin{aligned} E[e^2(m)] &= E\left[\left(x(m) - \mathbf{w}^T \mathbf{y}\right)^2\right] \\ &= E\left[x^2(m)\right] - 2\mathbf{w}^T E[\mathbf{y}x(m)] + \mathbf{w}^T E[\mathbf{y}\mathbf{y}^T] \mathbf{w} \\ &= r_{xx}(0) - 2\mathbf{w}^T r_{yx} + \mathbf{w}^T R_{yy} \mathbf{w} \end{aligned} \quad (3.45)$$

Eşitlik (3.45)'de geçen $R_{yy} = E[\mathbf{y}(m)\mathbf{y}^T(m)]$ giriş sinyalinin otokorelasyonunu ve $r_{yx} = E[x(m)y(m)]$ ise giriş sinyali ile istenen sinyalin korelasyonunu ifade etmektedir. Eşitlik (3.45)'in daha geniş bir formu Eşitlik (3.46)'da ki gibidir.

$$E[e^2(m)] = r_{xx}(0) - 2 \sum_{k=0}^{P-1} w_k r_{yx}(k) + \sum_{k=0}^{P-1} w_k \sum_{j=0}^{P-1} w_j r_{yy}(k-j) \quad (3.46)$$

Ortalama karesel hatanın, filtre katsayılarına göre türevi Eşitlik (3.47)'de verildiği gibi olur.

$$\begin{aligned}\frac{\partial}{\partial w} E[e^2(m)] &= -2E[x(m)y(m)] + 2w^T E[y(m)y^T(m)] \\ &= -2r_{yx} + 2w^T R_{yy}\end{aligned}\quad (3.47)$$

Burada eğim vektörü $\frac{\partial}{\partial w}$ aşağıdaki aşağıda ki gibidir.

$$\frac{\partial}{\partial w} = \left[\frac{\partial}{\partial w_0}, \frac{\partial}{\partial w_1}, \dots, \frac{\partial}{\partial w_{P-1}} \right]^T \quad (3.48)$$

En küçük karesel hataya sahip Wiener filtresi, Eşitlik (3.47)'nin sıfıra eşitlenmesi ile elde edilir.

$$R_{yy}w = r_{yx} \quad (3.49)$$

ya da

$$w = R_{yy}^{-1}r_{yx} \quad (3.50)$$

Eşitlik (3.50)'de verilen Wiener filtresinin çözümü genişletilmiş bir şekilde yeniden yazılabilir.

$$\begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{P-1} \end{pmatrix} = \begin{pmatrix} r_{yy}(0) & r_{yy}(1) & r_{yy}(2) & \cdots & r_{yy}(P-1) \\ r_{yy}(1) & r_{yy}(0) & r_{yy}(1) & \cdots & r_{yy}(P-1) \\ r_{yy}(2) & r_{yy}(1) & r_{yy}(0) & \cdots & r_{yy}(P-1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{yy}(N-1) & r_{yy}(N-2) & r_{yy}(N-3) & \cdots & r_{yy}(0) \end{pmatrix} \begin{pmatrix} r_{yx}(0) \\ r_{yx}(1) \\ r_{yx}(2) \\ \vdots \\ r_{yx}(P-1) \end{pmatrix} \quad (3.51)$$

3.6.3.1. Frekans Domeninde Wiener Filtresi

Frekans domeninde Wiener Filtresinin çıkışı $X(f)$, giriş sinyali $Y(f)$ ile filtrenin frekans cevabı $W(f)$ 'in çarpımıdır.

$$X(f) = W(f)Y(f) \quad (3.52)$$

Hata sinyali, istenen sinyal $X(f)$ ile filtrenin çıkışı $X(f)$ arasındaki fark olarak tanımlanır.

$$\begin{aligned} e(f) &= X(f) - W(f)Y(f) \\ &= X(f) - W(f)Y(f) \end{aligned} \quad (3.53)$$

Dolayısıyla f frekansındaki ortalama karesel hata Eşitlik (3.54)'de verildiği gibi olur.

$$E\left[|e(f)|^2\right] = E\left[(X(f) - W(f)Y(f))^* (X(f) - W(f)Y(f))\right] \quad (3.54)$$

Eşitlik (3.54)'deki “*” sembolü karmaşık eşlenik ifadesidir ve ortalama karesel hatanın zaman domenindeki değeri ile frekans domenindeki değeri arasında bulunan ilişki Eşitlik (3.55)'de gösterildiği gibidir.

$$\sum_{m=0}^{N-1} e^2(m) = \int_{-0.5}^{0.5} |e(f)|^2 df \quad (3.55)$$

Karesel hatanın minimum olması için Eşitlik (3.54)'ün $W(f)$ filtresine göre türevi sıfıra eşitlenir.

$$\frac{\partial E\left[|e(f)|^2\right]}{\partial W(f)} = 2W(f)P_{YY}(f) - 2P_{XY}(f) = 0 \quad (3.56)$$

Burada;

$$P_{yy} = E\left[Y(f)Y^*(f)\right], \quad Y(f) \text{ 'in güç spektrumunu, } P_{XY}(f) = E\left[X(f)Y^*(f)\right]$$

ise $X(f)$ ile $Y(f)$ 'in çapraz güç spektrumunu ifade eder.

Eşitlik (3.56) kullanılarak frekans domenindeki Wiener filtresi elde edilir.

$$W(f) = \frac{P_{XY}(f)}{P_{YY}(f)} \quad (3.57)$$

3.6.3.2. Wiener filtresi ile Spektral Çıkarma arasındaki ilişki

Spektral çıkarma eşitliği, gürültülü sinyal spektrumu ile spektral çıkarma filtresine ait frekans cevabının çarpımı olarak ifade edilebilir.

$$\begin{aligned}
|X(f)|^2 &= |Y(f)|^2 - |\overline{N(f)}|^2 \\
&= H(f)|Y(f)|^2
\end{aligned} \tag{3.58}$$

Spektral Çıkarma filtresinin frekans cevabı olan $H(f)$, Eşitlik (3.59)'daki gibi tanımlanabilir.

$$H(f) = 1 - \frac{|\overline{N(f)}|^2}{|Y(f)|^2} = \frac{|Y(f)|^2 - |\overline{N(f)}|^2}{|Y(f)|^2} \tag{3.59}$$

Wiener filtresi gürültülü sinyallere uygulanırken filtrenin frekans cevabı Eşitlik (3.60)'daki gibi tanımlanır.

$$W(f) = \frac{E[|Y(f)|^2] - E[|N(f)|^2]}{E[|Y(f)|^2]} \tag{3.60}$$

Eşitlik (3.59) ve Eşitlik (3.60)'a bakarak $H(f)$ ile $W(f)$ 'yi karşılaştırsak; Wiener filtresinin, sinyalin ve gürültünün ortalama spektrumuna bağlı olduğunu, Spektral çıkarma filtresinin ise gürültülü sinyalin anlık spektrumuna ve gürültünün zaman-ortalamalı spektrumuna bağlı olduğu görülür.

Ancak, filtreleme işleminden önce sinyal pencerelenerek 20-30ms uzunluğunda bölütlere ayrılır ve bu nedenle filtrelemeye tabi tutulan her bir sinyal kesiti durağan kabul edilir. Dolayısıyla gürültülü sinyalin, spektral çıkarma denklemindeki anlık spektrumu yerine zaman-ortalamalı spektrumunu koyarsak;

$$H(f) = \frac{|\overline{Y(f)}|^2 - |\overline{N(f)}|^2}{|\overline{Y(f)}|^2} \tag{3.61}$$

Spektral Çıkarma filtresinin frekans cevabı $H(f)$, Wiener filtresinin frekans cevabı $W(f)$ 'ye yaklaşmış olur.

3.6.3.3. Wiener filtresi ile Ses Sinyallerinin İyileştirilmesi

Ses iyileştirme için Wiener Filtresi, Spektral Çıkarmayı daha verimli bir hale getirebilmek adına Lim ve Oppenheim (1979) tarafından önerilmiştir. Burada doğrudan çıkarma yerine, bir Wiener kazanç fonksiyonu hesaplanıp gürültülü spektrum ile çarpılarak gürültüye ait frekans bileşenleri zayıflatılmıştır.

$n(m)$ gibi bir gürültüyle karışmış $s(m)$ gibi bir ses sinyali ele alınırsa, gürültülü sinyal

$$x(m) = s(m) + n(m) \quad (3.62)$$

olarak gösterilebilir. Burada sinyal ile gürültünün ilintisiz (uncorrelated) olduğu kabul edilir ($r_{sn}(m) = 0$) ve gürültülü sinyalin otokorelasyon matrisi, temiz sinyalin ve gürültünün otokorelasyon matrisleri toplamına eşit olur.

$$R_{xx} = R_{ss} + R_{nn} \quad (3.63)$$

Ayrıca;

$$r_{sx} = r_{ss} \quad (3.64)$$

Eşitlik (3.50)'de verilen Wiener filtresi, Eşitlik (3.63) ve Eşitlik (3.64) kullanılarak yeniden yazılabilir.

$$w = (R_{ss} + R_{nn})^{-1} r_{ss} \quad (3.65)$$

Eşitlik (3.65), gürültü azaltma için en uygun lineer filtredir. Frekans domenindeki gösterimi ise Eşitlik (3.66)'da verilmiştir.

$$W(f) = \frac{P_{ss}(f)}{P_{ss}(f) + P_{nn}(f)} \quad (3.66)$$

$P_{ss}(f)$ ve $P_{nn}(f)$, sırasıyla sinyalin ve gürültünün güç spektrumların ifade etmektedir.

Ses sinyalinin iyileştirilmesi için, önce gürültülü sinyal Hamming veya Hanning fonksiyonlarından biri ile pencerelenir. Pencerelenmiş gürültülü sinyal w alt indisi ile $y_w(m)$ olarak gösterilebilir. Sinyalin durağan olabilmesi için her bir pencere uzunluğu genellikle 10-30ms arasında ve art arda gelen bu kesitlerin örtüşme oranı da %50 ile %70 arasında seçilir.

Orijinal spektrum $S_w(f)$ 'in kısa süreli tahmini $S(f)$ 'yi elde etmek için, pencerelenmiş olan gürültülü sinyal $x_w(m)$ 'e, frekans domeninde Wiener filtresi uygulanır.

$$S_w(f) = W(f)X_w(f) \quad (3.67)$$

Filtre çıkışında elde edilen sinyale ters Fourier dönüşümü uygulanarak daha önce pencerelenmiş sinyal birleştirilerek tekrar oluşturulur.

Eşitlik (3.66)'da verilen Wiener filtresini uygulayabilmek için ses sinyalinin ve gürültünün spektral güç yoğunluklarının (psd) tahmin edilmesi gerekmektedir. Ses sinyalinin spektral güç yoğunluğu yerine $|S_w(f)|^2$ ifadesini, gürültünün spektral güç yoğunluğu yerine de $|\overline{N(f)}|^2$ olarak tanımlanan zaman-ortalamalı spektrumu kullanılabilir.

Bu koşullar altında Wiener filtresi Eşitlik (3.68)'de verildiği gibi yeniden düzenlenebilir.

$$W(f) = \frac{|S_w(f)|^2}{|S_w(f)|^2 + |\overline{N(f)}|^2} \quad (3.68)$$

Eşitlik (3.67) ve Eşitlik (3.68) birleştirilip, sinyalin spektral güç yoğunluğu ifadesi yerine tahmin edilen değeri koyularak aşağıdaki eşitlik elde edilir.

$$S_w(f) = \frac{|S_w(f)|^2}{|S_w(f)|^2 + |\overline{N(f)}|^2} X_w(f) \quad (3.69)$$

Eşitlik (3.69)'un birinci çözümü sıfır iken, ikinci çözümü Eşitlik (3.70)'de verildiği gibi olur.

$$|S_w(f)| = \sqrt{|X_w(f)|^2 - |N(f)|^2} \quad (3.70)$$

3.6.4. Kalman Filtresi

Kalman filtresi, adını aldığı Rudolph E. Kalman tarafından ayrık verilerin lineer filtreleme probleminin çözümüne yönelik yayımladığı makaleyle (Kalman 1960), istatistiksel tahmin teorisine yeni bir bakış açısı kazandırmıştır. Kalman filtresi, gürültülü ve eksik veri serisinden, dinamik bir sistemin durumunu kestiren, verimli bir tekrarlamalı filtredir (Pınarcı 2007).

Teorik olarak Kalman Filtresi, beyaz gürültüye maruz kalmış lineer bir dinamik sistemin anlık durumunu tahmin eden bir filtre olarak tanımlanabilir. Pratikte ise istatistiksel tahmin teorisi tarihindeki en büyük buluşlardan biri olarak gösterilmektedir. Özellikle uçak, gemi, uzay aracı üretimi gibi süreçlerde oldukça zor olan kompleks dinamik sistemlerin kontrolü, Kalman filtresinin sunduğu büyük avantajlarla sağlanabilmektedir.

Dinamik bir sistemi kontrol edebilmek için öncelikle bu sistemin ne yaptığını, nasıl işlediğini bilmek gerekir. Ancak pratik uygulamalarda kontrol etmek istediğimiz tüm değişkenleri ölçmek her zaman mümkün olan veya istenen bir durum değildir. Böyle bir durumda Kalman Filtresi eksik veya gürültülü ölçümlerden gerçek durumu veya bilgiyi elde edebilecek yeteneğe sahiptir (Grewal 2001).

Durum uzay modeli ile gösterilen bir dinamik sistemde, modelin önceki bilgileriyle birlikte giriş ve çıkış bilgilerinden sistemin durumu tahmin edilebilir. Gözlemlene teorisi, karar verilen bir bakış açısı temelinde, sistemin durum tahmini için izlenecek bir yoldur. Eğer sistemin stokastik ve rasgele gürültülü olduğu durum hesaba katılırsa minimum varyans tahmini ya da Kalman filtresi uygun olmaktadır. Kalman filtresi, geleneksel tahmin yöntemlerinde olduğu gibi filtreleme özelliğine rağmen, sistemin ölçülemeyen durumlarını tahmin etmede de çok güçlü ve yeteneklidir.

Kalman Filtresi, ortalama kestirim hatasının karesini minimize etmeye yönelik bir hesaplama mantığı ile duruma bakar (Ankışhan 2007).

Kalman filtresi, bir işlemi kestirirken, bir çeşit geri-besleme kontrol mekanizması kullanır. Filtre, bir zaman adımı için işlemin kestirim değerini hesaplar ve daha sonra gürültülü ölçümden geri-besleme alır. Bu yüzden Kalman filtresi denklemleri, zaman güncelleme denklemleri ve ölçüm güncelleme denklemleri olarak iki gruba ayrılır.

Basit bir anlatımla, zaman güncelleme denklemleri geleceği tahmin eden denklemler, ölçüm güncelleme denklemleri ise düzeltici denklemlerdir.

Kalman filtresinin uygulanabilmesi için başlangıç anındaki sistem durum vektörü ve hata kovaryans matrisinin bilinmesi gerekmektedir. Genellikle bu değerlerin gerçek değerleri bilinmediği için kullanıcı tarafından tahmin edilmeleri gerekmektedir. Sistem durumuna ait yapılan tahminin doğruluğu ile orantılı olarak hata kovaryans matrisinin değerinin küçük olması gerekmektedir. Hata kovaryans matrisinin gerçek değerine göre aşırı küçük veya büyük olması durumunda filtre ıraksamaları görülebilmektedir. Bunun dışındaki durumlarda, başlangıç anında filtre hızlı bir şekilde doğru değere ulaşmakta ve sonrasında filtrenin ilerlemesi ilk tahminden bağımsız olarak gerçekleşmektedir.

Durum-uzay modeli, sistemin durumunu gösteren ancak gözlenemeyen $\{x_n, n=0,1,2,\dots\}$ stokastik süreci ile ilgili bir durum eşitliği ve gözlenebilen $\{y_n, n=0,1,2,\dots\}$ stokastik süreci ile ilgili bir ölçüm (gözlem) eşitliğinden oluşan

$$x(n+1) = F(n+1,n)x(n) + D(n)v(n) \quad (3.71)$$

$$y(n) = C(n)x(n) + w(n) \quad (3.72)$$

şeklinde bir modeldir (Babacan 2009). Burada;

- **x(n)**, n anındaki durum vektörünü
- **F(n)**, durum geçiş matrisini
- **v(n)**, sıfır ortalamalı olan beyaz Gaussian süreç gürültüsünü
- **y(n)**, gürültülü gözlemi
- **w(n)**, sıfır ortalamalı olan ölçüm gürültüsünü

ifade etmektedir.

Süreç gürültüsü $v(n)$ ve ölçüm gürültüsü $w(n)$ 'in ortalamaları sıfır olduğundan beklenen değerleri sıfıra eşittir. Ayrıca bu iki gürültü birbiriyle tamamen bağımsız olup süreç gürültüsünün kovaryans matrisi $Q(n)$, ölçüm gürültüsünün kovaryansı $R(n)$ olarak tanımlanır (Paliwal ve Basu 1987).

$$E[v(n)] = 0, E[w(n)] = 0 \quad (3.73)$$

$$E[v(n)w^T(k)] = E[v(n)]E[w^T(k)] = 0 \quad \forall n, k \quad (3.74)$$

$$E[v(n)v^T(n)] = Q(n) \quad (3.75)$$

$$E[w(n)w^T(n)] = R(n) \quad (3.76)$$

Buradaki amaç; herhangi bir t anında mevcut olan bilgilerle durum vektörü $x(n)$ 'i tahmin edebilmektir. Dolayısıyla göz önünde bulundurulması gereken üç durum vardır (Najim 2008).

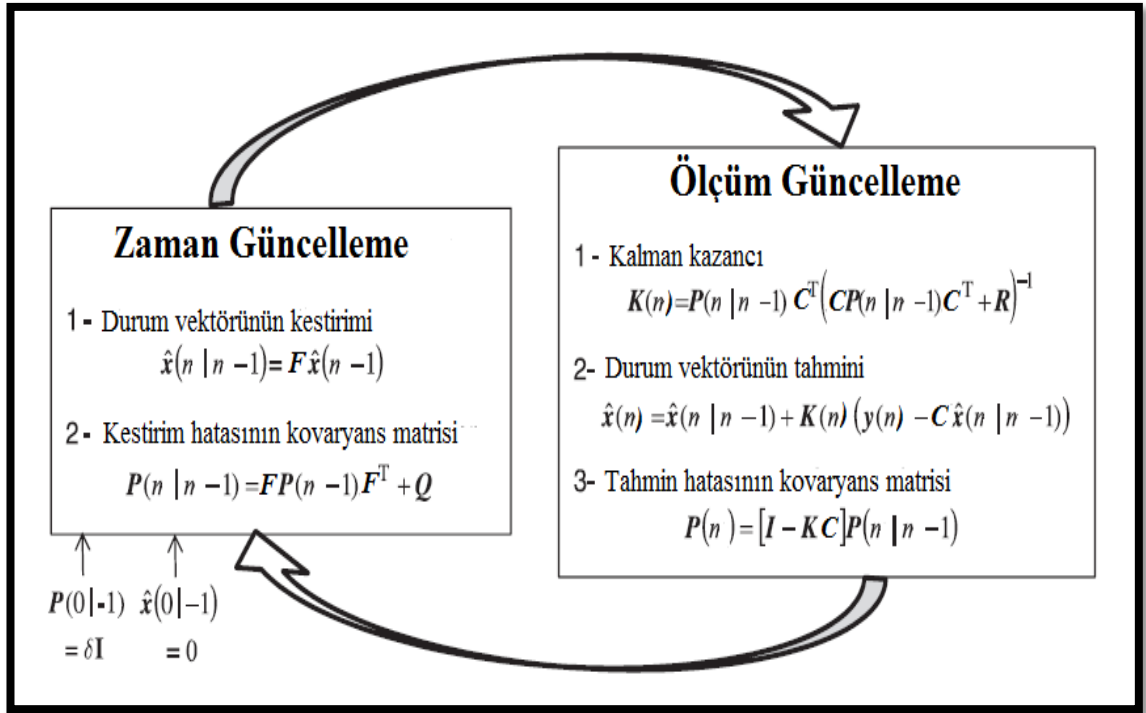
* $n=t$: Durum vektörü $x(n)$, n anına kadar elde edilmiş tüm gözlem değerleri dikkate alınarak tahmin edilmelidir ki bu durumda yapılan işlem *filtreleme*'dir.

* $n < t$: Elde edilmiş tüm gözlem değerlerinin sadece bir kısmı alınır ve bu durumda yapılan işlem *düzeltilme (smoothing)* olarak adlandırılır.

* $n > t$: Bu koşulda yapılan işlem *kestirim (prediction)* olarak adlandırılır.

Yukarıda değinilen üç durumun tümü için $x(n/t)$ notasyonu, t anına kadar elde edilmiş gözlem bilgileri dahilinde, $x(n)$ 'in n anında tahmin edilen değerini ifade etmektedir.

Kalman Filtresi en basit açıklamayla; $n+1$ anında elde edilen ölçüm değerini ve bir önceki anda yani n anında yapılan tahmini kullanarak yeni durumu tahmin eder. Kalman filtresinin her adım için gerçekleştirdiği işlemler Şekil 3.11.'de verilmiştir.



Şekil 3.11 Kalman Filtresinin her iterasyon için gerçekleştirdiği işlemler

3.6.4.1. Filtrenin sayısal kökenleri

Tahmin edilen durum vektörü $x(n+1/n)$ 'i

$$x(n+1/n) = E[x(n+1)|y(1), \dots, y(n)] \quad (3.77)$$

olarak tanımlanıp Eşitlik (3.77)'de yerine konulursa;

$$x(n+1) = F(n+1, n) E[x(n)|y(1), \dots, y(n)] + D(n) E[v(n)|y(1), \dots, y(n)] \quad (3.78)$$

olur. Ayrıca,

$$E[v(n)|y(1), \dots, y(n)] = 0 \quad (3.79)$$

olduğu için

$$x(n+1/n) = F(n+1, n) x(n/n) \quad (3.80)$$

olarak sadeleşir. $\tilde{e}(n/n)$ ifadesi, durum vektörü $x(n)$ 'in n anında yapılan tahminindeki hata olarak tanımlanırsa;

$$\tilde{e}(n/n) = x(n) - x(n/n) \quad (3.81)$$

n anındaki hatanın (*posteriori*) kovaryans matrisi

$$P(n/n) = E \left\{ \left[x(n) - x(n/n) \right] \left[x(n) - x(n/n) \right]^T \right\} = E \left[\tilde{e} \tilde{e}^T \right] \quad (3.82)$$

olur ve Eşitlik (3.71)'den Eşitlik (3.80)'i çıkararak;

$$x(n+1) - x(n+1/n) = F(n+1, n) \left[x(n) - x(n/n) \right] + D(n)v(n) \quad (3.83)$$

Eşitlik (3.83) elde edilir. Bu durumda hatanın (*priori*) kovaryans matrisi

$$P(n+1/n) = E \left\{ \left[x(n+1) - x(n+1/n) \right] \left[x(n+1) - x(n+1/n) \right]^T \right\} \quad (3.84)$$

olur ve Eşitlik (3.83), Eşitlik (3.84)'de yerine konulursa;

$$\begin{aligned} P(n+1/n) &= F(n+1, n) E \left\{ \left[x(n) - x(n/n) \right] \left[x(n) - x(n/n) \right]^T \right\} F^T(n+1, n) \\ &\quad + D(n) E \left[v(n)v^T(n) \right] D^T(n) \end{aligned} \quad (3.85)$$

olarak elde edilir. Yine burada Eşitlik (3.82) ve (3.85)'den yararlanılarak;

$$P(n+1/n) = F(n+1, n) P(n/n) F^T(n+1/n) + D(n) Q(n) D^T \quad (3.86)$$

eşitliği oluşturulabilir.

- Güncelleme adımı

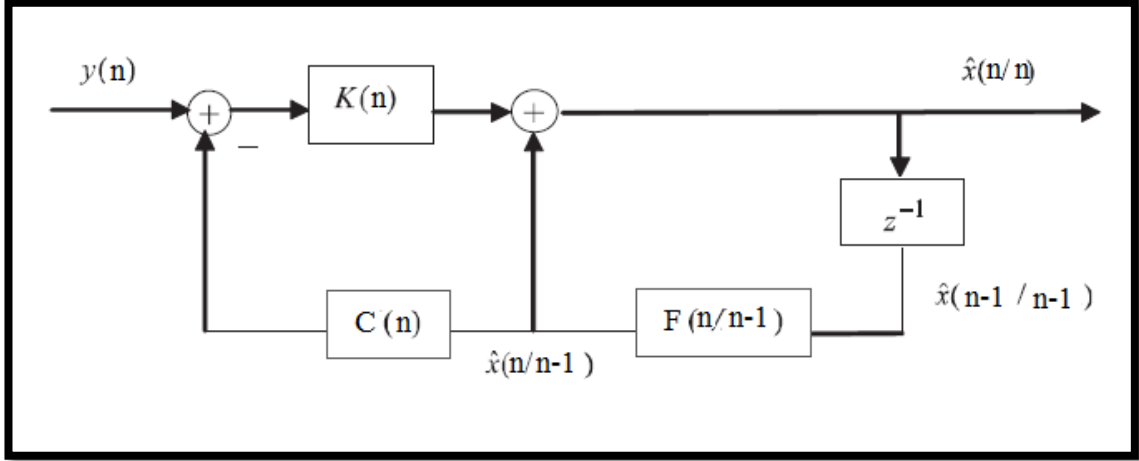
Daha önce Kalman Filtresi'ni oluşturan denklemler gösterilirken, durum vektörünü tahmin etmede kullanılan denklem,

$$x(n/n) = x(n/n-1) + K(n) \left[y(n) - C(n)x(n/n-1) \right] \quad (3.87)$$

olarak ifade edilmişti.

Eşitlik (3.87)'de geçen $K(n)$ ifadesi, *filtre kazancı* ya da *Kalman kazancı* olarak adlandırılır. Burada, n anındaki durum vektörünün tahmini (estimation) olan $x(n/n)$, bir önceki ($n-1$) anın ölçümlerine dayanarak kestirilen (prediction) durum vektörü $x(n/n-1)$ 'in, bir nevi düzeltici diye adlandırabilinen bir değerle güncellenmesiyle elde

edilir. “*düzeltilici*” olarak adlandırılan ifade, ölçülen gerçek değer ile kestirilen değer arasındaki farktır. Bu fark sıfıra ne kadar yakın olursa, tahminde yapılan hatanın da o kadar az olduğu sonucuna varılabilir. Kalman filtresi ile yapılan güncellemeye ait blok diyagramı Şekil 3.12.’de verilmiştir



Şekil 3.12 Kalman filtresi blok diyagramı

Durum vektörünün tahminindeki hatayı ifade eden Eşitlik (3.81)’de, $x(n/n)$ yerine Eşitlik (3.87)’deki ifade konulursa, hata ifadesi;

$$\begin{aligned}
 \tilde{e}(n/n) &= x(n) - \hat{x}(n/n) \\
 &= x(n) - \hat{x}(n/n-1) - K(n) \left[y(n) - C(n)\hat{x}(n/n-1) \right] \\
 &= x(n) - \hat{x}(n/n-1) - K(n) \left[C(n)x(n) + w(n) - C(n)\hat{x}(n/n-1) \right] \\
 &= [I - K(n)C(n)] \left[x(n) - \hat{x}(n/n-1) \right] - K(n)w(n)
 \end{aligned} \tag{3.88}$$

olarak elde edilir. Eşitlik (3.88)’den faydalanarak $P(n/n)$, $P(n/n-1)$ ’in bir fonksiyonu cinsinden ifade edilebilir.

$$\begin{aligned}
 P(n/n) &= E \left[\tilde{e}(n/n) \tilde{e}^T(n/n) \right] \\
 &= E \left\{ \left[(I - K(n)C(n))\tilde{e}(n/n-1) - K(n)w(n) \right] \right. \\
 &\quad \left. \left[(I - K(n)C(n))\tilde{e}(n/n-1) - K(n)w(n) \right]^T \right\}
 \end{aligned} \tag{3.89}$$

Daha sade bir ifadeyle,

$$P(n/n) = [I - K(n)C(n)]P(n/n-1)[I - K^T(n)C^T(n)] + K(n)R(n)K^T(n) \quad (3.90)$$

olur.

- Filtre Kazancı

Filtrenin en uygun sonucu verebilmesi için, $K(n)$ kazancının, durum vektörünün tahminindeki ortalama karesel hatayı (*mean square error*) minimum yapabilecek şekilde seçilmesi gereklidir.

Bunun için gerekli koşul, daha önce tahmin hatasının kovaryans matrisi olarak tanımlanan $P(n/n)$ matrisi izinin (trace), $K(n)$ 'e göre türevinin sıfıra eşit olmasıdır.

$$\frac{\partial \text{trace} \left\{ [I - K(n)C(n)]P(n/n-1)[I - K^T(n)C^T(n)] + K(n)R(n)K^T(n) \right\}}{\partial K(n)} = 0 \quad (3.91)$$

Eşitlik (3.91)'in çözümü, Eşitlik (3.90)'ın daha açık bir şekilde yazılıp $K(n)$ 'e göre türevinin alınmasıyla,

$$P(n/n-1)C^T(n) = K(n)C(n)P(n/n-1)C^T(n) + K(n)R(n) \quad (3.92)$$

olarak bulunur. Buradan da $K(n)$ 'i çekilirse filtrenin kazancı,

$$K(n) = \frac{P(n/n-1)C^T(n)}{C(n)P(n/n-1)C^T(n) + R(n)} \quad (3.93)$$

olarak ifade edilebilir (Najim 2008).

$P(n/n)$ ile $P(n/n-1)$ arasındaki bağıntıyı veren Eşitlik (3.90) yeniden düzenlenirse,

$$\begin{aligned} P(n/n) &= P(n/n-1) - K(n)C(n)P(n/n-1) - P(n/n-1)C^T(n)K^T(n) \\ &\quad + K(n)C(n)P(n/n-1)C^T(n)K^T(n) + K(n)R(n)K^T(n) \\ &= P(n/n-1) - K(n)C(n)P(n/n-1) - P(n/n-1)C^T(n)K^T(n) \\ &\quad + K(n)[C(n)P(n/n-1)C^T(n) + R(n)]K^T(n) \end{aligned} \quad (3.94)$$

olur ve Eşitlik (3.94)'teki köşeli parantez içerisinde bulunan ifade yerine, denklem(3.93)'deki karşılığı konulabilir ve

$$\begin{aligned} P(n/n) &= P(n/n-1) - K(n)C(n)P(n/n-1) - P(n/n-1)C^T(n)K^T(n) \\ &\quad + P(n/n-1)C^T(n)K^T(n) \end{aligned} \quad (3.95)$$

olur. Eşitlik (3.95) yeniden düzenlenirse,

$$P(n/n) = (I - K(n)C(n))P(n/n-1) \quad (3.96)$$

olarak elde edilir. Kalman filtresini oluşturan tüm bu eşitliklerin genel bir gösterimi Çizelge 3.1.'de verilmiştir.

3. MATERYAL ve METOT

Çizelge 3.1. Kalman Filtresi Denklemleri

Model denklemi	$x(n+1) = F(n+1, n)x(n) + D(n)v(n)$
Gözlem denklemi	$y(n) = C(n)x(n) + w(n)$
Başlangıç bilgileri	$E[v(n)] = 0 \quad \text{ve} \quad E[w(n)] = 0$ $E[v(n)w(k)] = 0 \quad \forall n, k$ $E[v(n)v^T(k)] = Q(n)$ $E[w(n)w(k)] = R(n)$ $E[x_0(n)v^T(n)] = 0 \quad \forall n \geq 0$ $E[x_0(n)w(n)] = 0 \quad \forall n \geq 0$
Filtre denklemleri	$\hat{x}(n/n-1) = F(n, n-1)\hat{x}(n-1/n-1)$ $\hat{x}(n/n) = \hat{x}(n/n-1) + K(n)[y(n) - C(n)\hat{x}(n/n-1)]$
Kazanç ifadeleri	$K(n) = P(n/n-1)C^T(n)[C(n)P(n/n-1)C^T(n) + R(n)]^{-1}$ $K(n) = P(n/n)C^T(n)R^{-1}(n)$
Tahminin hata kovaryans matrisi	$P(n/n) = [I - K(n)C(n)]P(n/n-1)$
Kestirimin hata kovaryans matrisi	$P(n/n-1) = F(n, n-1)P(n-1/n-1)F^T(n, n-1) +$ $D(n)Q(n)D^T(n)$
İlk koşullar	$\hat{x}(0/0) = E[x(0)]$ $P(0/0) = P(0) = E\{[x(0) - \hat{x}(0/0)][x(0) - \hat{x}(0/0)]^T\}$

- Kalman Filtresi ile örnek bir uygulama

Beyaz ve normal dağılımlı bir gürültüyle karışmış otoregresif (AR) bir sinyalin Kalman Filtresi ile tahmini

$x(m) \longrightarrow$ temiz sinyal

$n(m) \longrightarrow$ gürültü

olsun. Temiz sinyalin otoregresif modeli;

$$x(m) = \sum_{k=1}^p a_k x(m-k) + e(k) \quad (3.97)$$

olur ve gürültülü sinyal aşağıdaki gibi gösterilir,

$$y(m) = x(m) + n(m) \quad (3.98)$$

Eşitlik (3.97), vektör formunda Eşitlik (3.99)'da ki gibi yazılabilir

$$\begin{pmatrix} x(m) \\ x(m-1) \\ x(m-2) \\ \vdots \\ x(m-p+1) \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x(m-1) \\ x(m-2) \\ x(m-3) \\ \vdots \\ x(m-p) \end{pmatrix} + \begin{pmatrix} e(m) \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (3.99)$$

Kalman filtresi denklemlerinde geçen gözlem matrisinin $C = I$ birim matris olduğu varsayımı ile Kalman filtresi denklemleri uygulanır.

İlk koşullar:

$$P(-1) = \delta I \quad (3.100)$$

$$x(0/-1) = 0 \quad (3.101)$$

Zaman güncelleme denklemleri:

Sinyal kestirim denklemi;

$$x(m/m-1) = F x(m) \quad (3.102)$$

Kestirim hatası;

$$e(m/m-1) = x(m) - F x(m-1) \quad (3.103)$$

Kestirim hatasının kovaryans matrisi;

$$P(m/m-1) = FP(m-1)F^T + Q \quad (3.104)$$

Ölçüm güncelleme denklemleri:

Kalman kazancı;

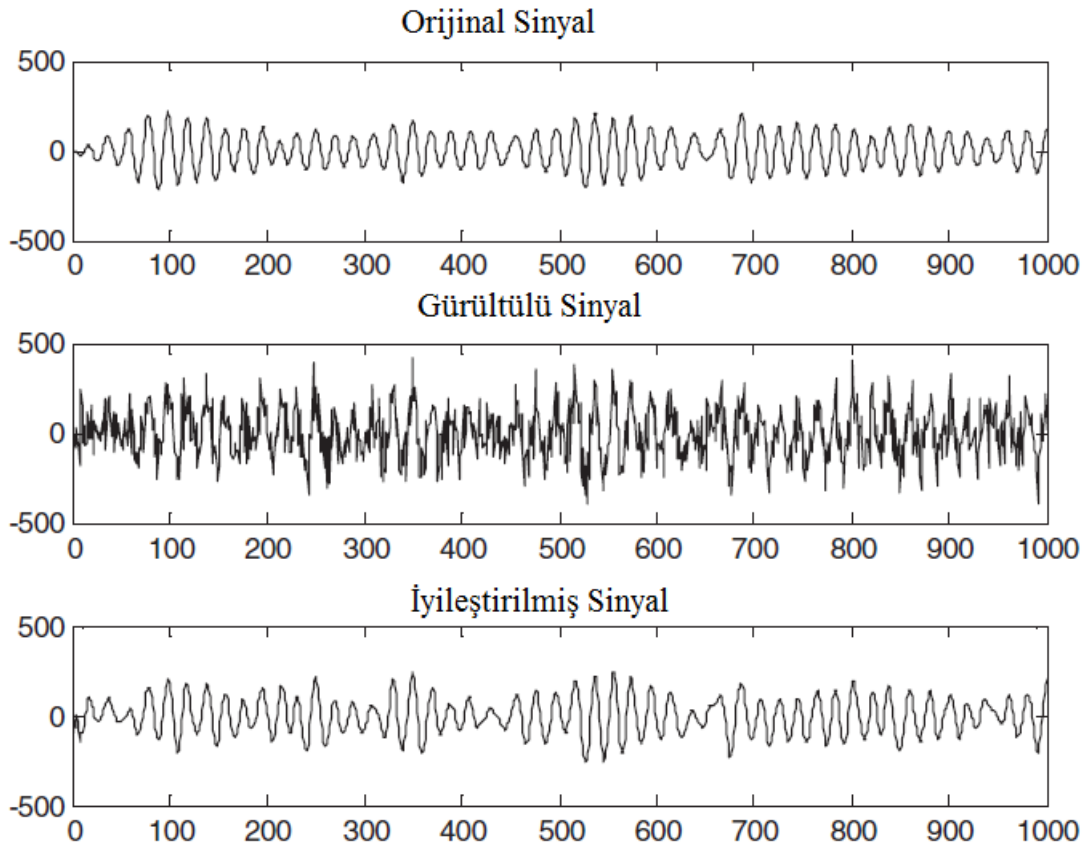
$$K(m) = P(m/m-1)(P(m/m-1) + R)^{-1} \quad (3.105)$$

Sinyal Tahmin denklemi;

$$x(m) = x(m/m-1) + K(m)(y(m) - x(m/m-1)) \quad (3.106)$$

Tahmin hatasının kovaryans matrisi;

$$P(m) = [1 - K(m)]P(m/m-1) \quad (3.107)$$



Şekil 3.13. Gürültülü bir AR sinyale Kalman filtresi uygulanarak elde edilen iyileştirilmiş sinyal

3.6.4.2. Kalman Filtresinin Parametre Tahminine Uygulanması

Önceki bölümlerde belirtildiği gibi ses sinyallerini özbağımlı veya başka bir deyişle otoregresif olarak modelleyebiliriz. Bu durumda Eşitlik (3.108)'deki gibi bir AR süreci tanımlanırsa,

$$y(k) = -\sum_{i=1}^p a_i y(k-i) + u(k) \quad (3.108)$$

$y(k)$ k anındaki sinyali, $u(k)$ ise birim varyansa sahip ve sıfır ortalamalı beyaz Gaussian gürültüyü temsil eder. Burada öngörü katsayıları veya AR katsayıları diye adlandırılan $\{a_i\}_{i \in [1, \dots, p]}$ parametrelerinin tahmininin minimum hata ile gerçekleştirilmesi gerekmektedir.

Durum uzayı gösterimi için kullanılan $\tilde{x}(k)$ durum vektörü, AR katsayılarına karşılık gelen $\{x_i(k)\}_{i \in [1, \dots, p]}$ elemanlarından oluşturulabilir.

$$\tilde{x}(k) = \begin{bmatrix} x_1(k) \\ \cdot \\ \cdot \\ \cdot \\ x_p(k) \end{bmatrix} = \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{bmatrix} \quad (3.109)$$

Diğer bir gösterimle;

$$\tilde{x}(k) = [a_1 \dots a_p]^T = \theta(k) \quad (3.110)$$

Yine durum uzayı gösterimindeki gözlem vektörü $H(k)$ gözlemin geçmiş p değerinden oluşturulabilir,

$$H(k) = [-y(k-1) \dots -y(k-p)] \quad (3.111)$$

Burada model durağan olduğu sürece $\{a_i\}_{i \in [1, \dots, p]}$ parametrelerinin de sabit olduğu kabul edilebilir. Dolayısıyla durum uzay modeli (3.112) ve (3.113)'de verilen iki denklemlerle tanımlanabilir.

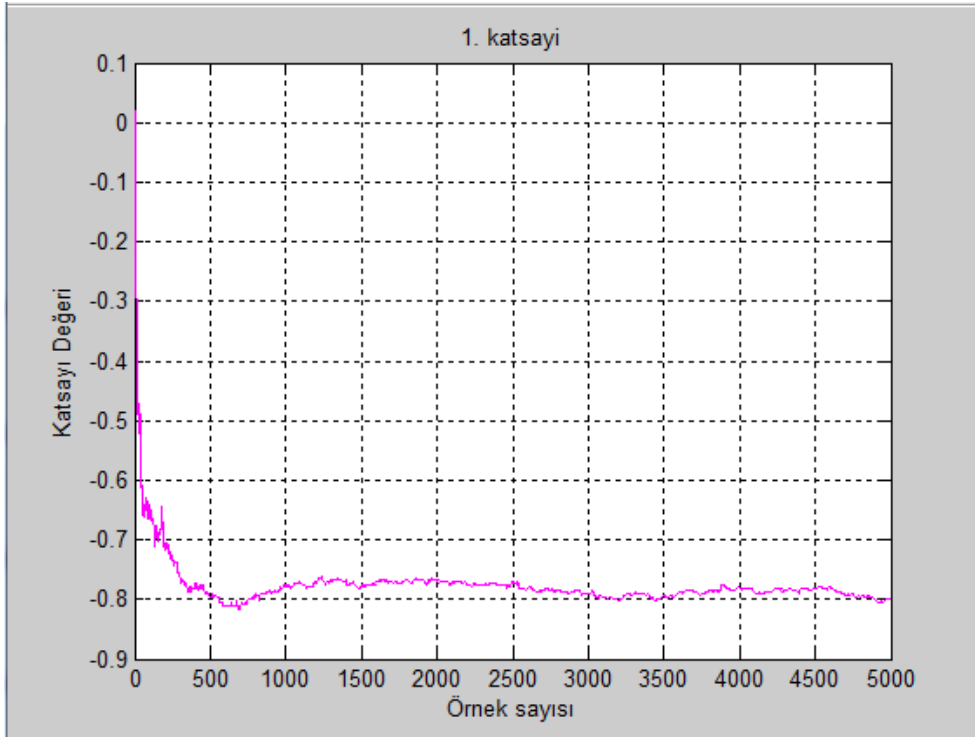
$$\tilde{x}(k+1) = \tilde{x}(k) \quad (3.112)$$

$$y(k) = H(k)\tilde{x}(k) + u(k) \quad (3.113)$$

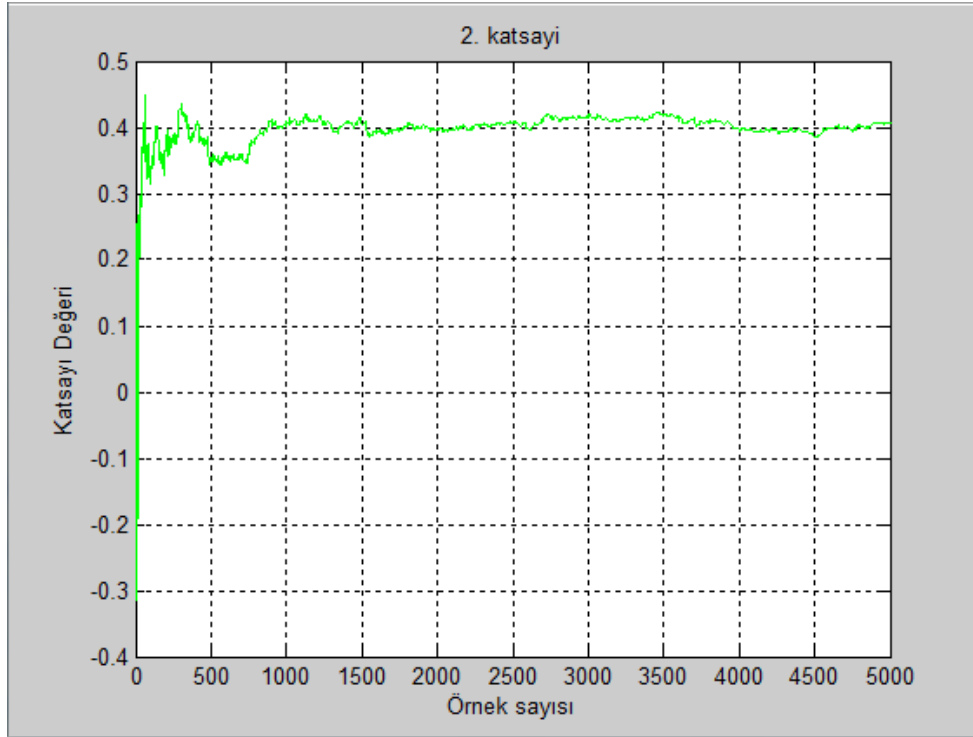
Eşitlik (3.112) ve Eşitlik (3.113)'de verilen durum uzayı gösterimi incelendiğinde süreç gürültüsü kovaryansının $Q(k)=0$ ve geçiş matrisinin $A(k+1,k)=I$ birim matris olarak seçildiği görülmektedir (Najim 2008),

Kalman filtresinin diğer parametre tahmin yöntemlerine üstünlük sağlayan özelliği, parametreleri izleyebilme, diğer bir deyişle takip edebilme yeteneğinden kaynaklanır.

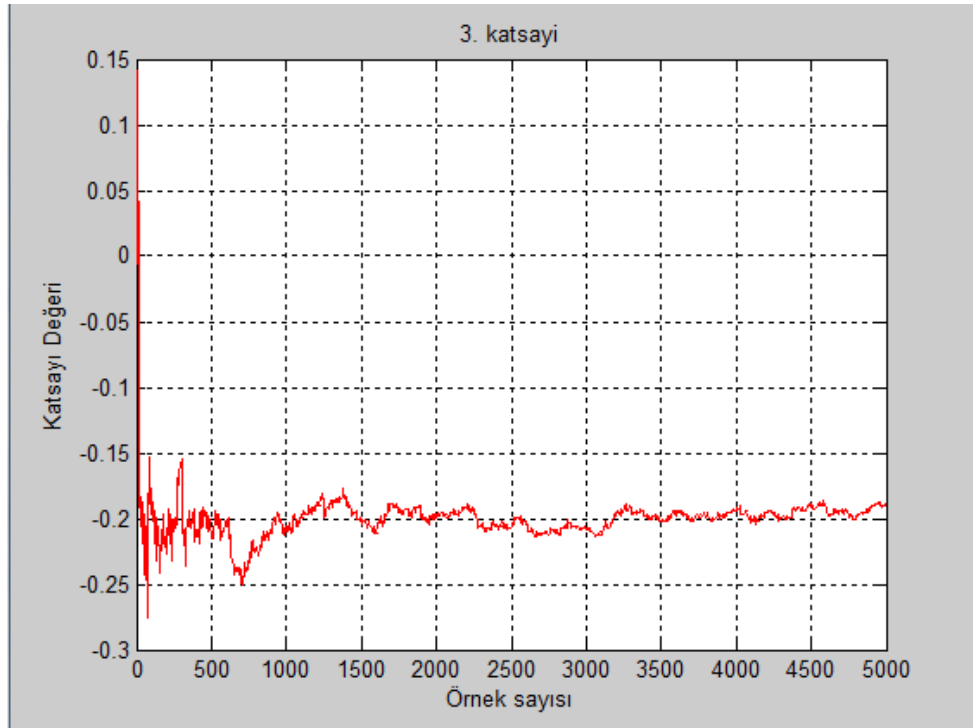
Bu durumu daha iyi kavrayabilmek adına, rastgele bir sinyal oluşturulmuş ve bu sinyal, katsayıları -0.8, 0.4, -0.2, 0.5 olan 4.dereceden bir sayısal filtreden geçirilmiştir. Neticede oluşan sinyalin belirtilen katsayılarının, kalman filtresi ile tahmin edilip edilemeyeceği test edilmiştir. Elde edilen benzetim sonuçları sırasıyla Şekil 3.14., Şekil 3.15., Şekil 3.16. ve Şekil 3.17.'de gösterilmiştir.



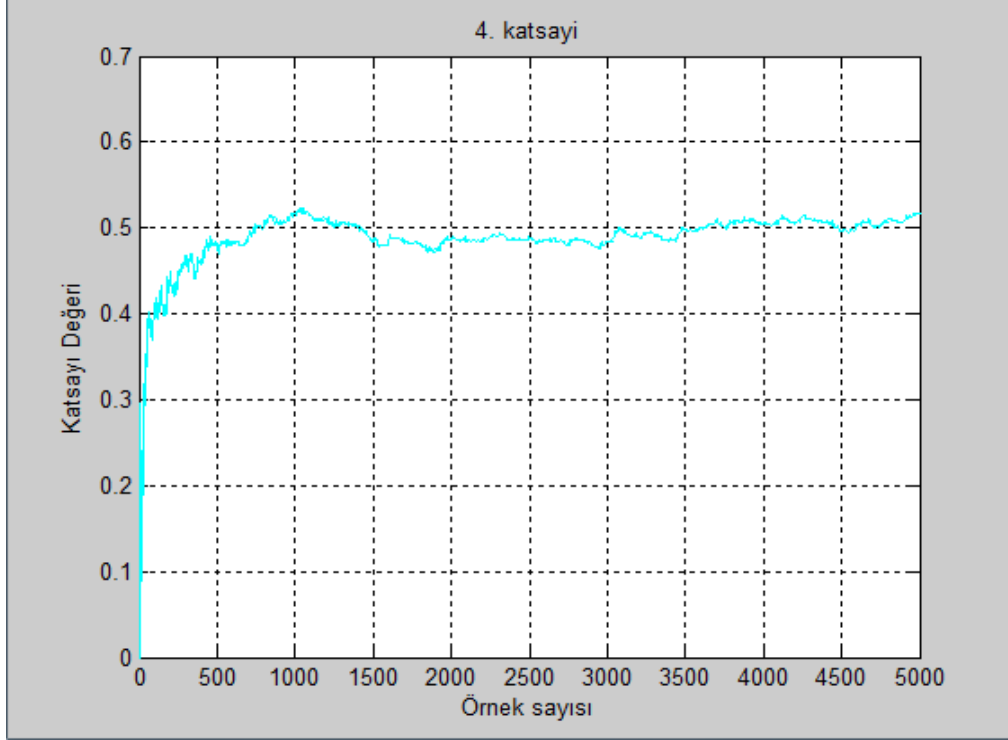
Şekil 3.14. Kalman Filtresi ile elde edilen 1.katsayı



Şekil 3.15 Kalman Filtresi ile elde edilen 2.katsayı

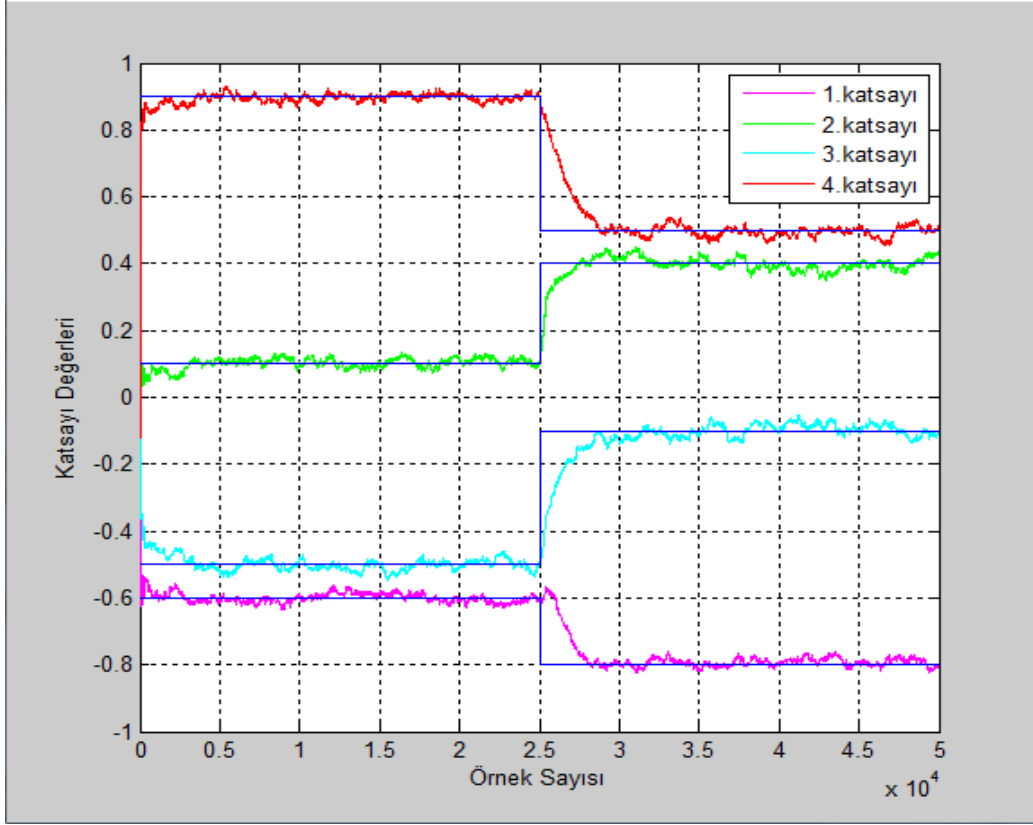


Şekil 3.16 Kalman Filtresi ile elde edilen 3.katsayı



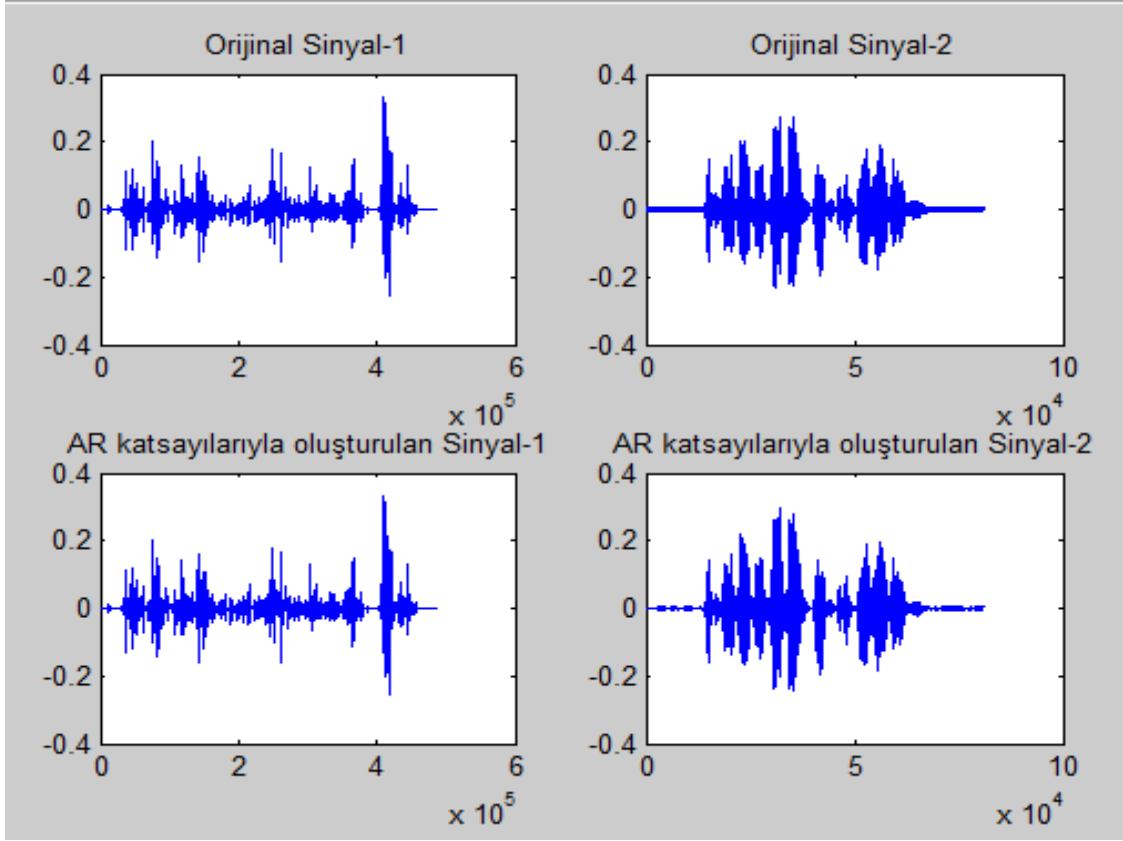
Şekil 3.17. Kalman Filtresi ile elde edilen 4.katsayı

Kalman filtresinin takip (tracking) yeteneğini daha iyi görebilmek için bir diğer testte ise, üretilen rastgele sinyalin yarısı $-0.6, 0.1, -0.5, 0.9$ katsayıları ile diğer yarısı ise $-0.8, 0.4, -0.1, 0.5$ katsayıları ile filtrelenmiştir ve oluşan sinyale Kalman filtresi uygulanarak katsayılar tahmin edilmeye çalışılmıştır. Elde edilen benzetim sonuçları Şekil 3.18.'de gösterilmiştir. Yine bu sonuçlar, Kalman filtresinin farklı zamanlarda farklı katsayıları doğru bir şekilde tahmin edebildiğini ve takip sürecine hızlı bir şekilde adapte olabildiğini göstermiştir.



Şekil 3.18. Kalman Filtresi ile tahmin edilen katsayı değerleri

Simülasyon ortamında rastgele oluşturduğumuz bir sinyale uygulanan bu algoritmanın ses sinyalleri üzerindeki etkisini veya başarısını gözlemleyebilmek için, “Oregon Graduate Institute of Science and Technology”ye ait “Speech Enhancement Assessment Resource (SpEAR) Database” veritabanından alınan ses sinyalleri Matlab ortamına aktarılarak aynı algoritma uygulanmış, elde edilen katsayılarla ses sinyali tekrar oluşturulup orijinaliyle karşılaştırılmıştır. Şekil 3.19.’da verilen simülasyon sonuçları, Kalman filtresi ile ses sinyalinin AR katsayılarının doğru bir şekilde bulunduğunu ve sinyalin bu katsayılarla başarılı bir şekilde tekrar oluşturulduğunu göstermektedir. Bu durum aynı zamanda ses sinyallerinin otoregresif modellenebildiğinin bir göstergesi olarak da değerlendirilebilir.



Şekil 3.19. Kalman Filtresinin tahmin ettiği katsayılar ile ses sinyallerinin tekrar oluşturulmasına ait simülasyon sonuçları.

3.6.4.3. Kalman Filtresi ile Ses İyileştirme

Ses sinyallerinin anlaşılabilirliğine ve kalitesine bozucu bir etki yapan gürültülerin bu etkisini azaltabilmek için ses iyileştirme yöntemi olarak Kalman Filtresi kullanılabilir. Problemi, sadece arka plan gürültüsüyle bozulmuş bir ses sinyali mevcutken ve bu sinyale ait hiçbir ön bilgi yokken ses sinyalinin kalitesini iyileştirmek olarak tanımlanabilir.

Ses sinyali $s(k)$ olarak gösterilip, p . dereceden otoregresif (AR) bir model gibi ifade edilebilir (Paliwal ve Basu 1987).

$$s(k) = -\sum_{i=1}^p a_i s(k-i) + u(k) \quad (3.114)$$

Ölçülen veya gözlemlenen sinyal ise varyansı R olan sıfır ortalamalı bir gürültü ile ses sinyalinin birleşiminden oluşur.

$$y(k) = s(k) + b(k) \quad (3.115)$$

Daha öncede değinildiği gibi Kalman Filtresi durum tahmini için özyinelemeli bir prosedür yürütür. Buradaki amaç ses sinyalinin tahminini gerçekleştirmektir.

Bunun için, ses sinyalinin $q(\geq p)$ tane önceki değeri alınarak $\underline{x}(k)$ durum vektörü oluşturulabilir (Najim 2008).

$$\underline{x}(k) = [s(k-q+1).....s(k)]^T \quad (3.116)$$

Eşitlik (3.115) ve Eşitlik (3.116)'da ki durum uzay gösterimini aşağıdaki gibi düzenleyebiliriz;

$$\underline{x}(k) = \phi \underline{x}(k-1) + Gu(k) \quad (3.117)$$

$$y(k) = H \underline{x}(k) + b(k) \quad (3.118)$$

ϕ , geçiş matrisi olarak adlandırılır ve $q \times q$ boyutundadır (Najim 2008).

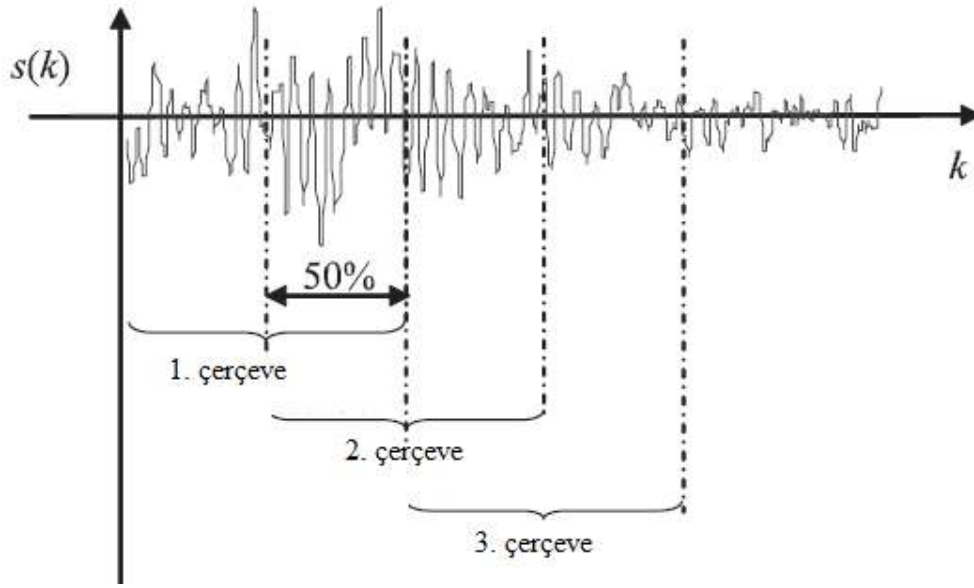
$$\phi = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & \dots & \dots & \vdots \\ \vdots & & \vdots & 0 & \ddots & & & \vdots \\ \vdots & & \vdots & \vdots & \ddots & & & \vdots \\ \vdots & & \vdots & \vdots & \ddots & & & 0 \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & 0 & \underbrace{-a_p \quad -a_{p-1} \quad \dots \quad -a_1}_p & \dots & \dots & \dots & \dots \end{bmatrix} \quad (3.119)$$

$\underbrace{\quad}_{q-p}$

G ve H sırasıyla giriş ve gözlem vektörleri olarak aşağıdaki gibi tanımlanır;

$$G = H^T = \begin{bmatrix} 0 \dots 0 1 \\ \vdots \\ 1 \end{bmatrix}^T \quad (3.120)$$

Kalman Filtresi ile ses sinyallerinin iyileştirilmesinde birçok farklı yaklaşım izlenebilir. İyileştirme süreci genellikle pencereleme işlemi ile başlar. Ses sinyalleri yaklaşık 20 – 30ms'lik çerçevelere bölünerek iyileştirme algoritması her çerçeveye ayrı ayrı uygulanır. Pencereleme işleminin amacı sinyalin durağan olmasını sağlamaktır.



Şekil 3.20. Ses sinyalinin pencereleme işlemi ile çerçevelere bölünmesi (Najim 2008)

İşlem kolaylığı sağlaması açısından tüm bölütleri eşit uzunlukta ve pencereleme fonksiyonu olarak da %50 örtüşme ile Hanning veya Hamming penceresi alınabilir.

Hanning veya Hamming Penceresi fonksiyonu aşağıdaki gibi tanımlanır;

$$w(k) = a - (1-a) \cos\left(2\pi \frac{k}{N}\right), \quad k = 0, \dots, N-1 \quad (3.121)$$

Hanning için $a=0.5$ ve Hamming için $a=0.54$ seçilir. Ayrıca N her bir bölütteki örnek sayısını ifade eder.

Her bir bölütteki sinyali tahmin edebilmek için öncelikle ölçüm gürültüsünün varyansı R 'nin ve durum uzayı matrisleri $[\phi, G, H]$ 'ın bilinmesi gerekmektedir. Böylece iyileştirilmiş sinyali aşağıdaki gibi tanımlayabiliriz;

$$\hat{s}(k) = Hx(k/k) \quad (3.122)$$

Burada $x(k/k)$, k tane gözleme $\{y(1), y(2), \dots, y(k)\}$ dayanan $\underline{x}(k)$ durum vektörünün tahmin edilen değeridir. Bu durumda yapılması gereken AR katsayılarının hesaplanmasıdır. Her bir bölütün durağan olduğu göz önüne alınırsa bu katsayılar her bölüt boyunca sabit kabul edilebilir.

Gerekli parametreler belirlendikten sonra her bölüt için Kalman Filtresi'nin denklemleri uygulanır ve ses sinyali tekrar oluşturulur. Uygulanan Kalman denklemleri aşağıda verildiği gibidir (So ve Paliwal 2008).

$$e(n) = y(n) - x(n/n-1) \quad (3.123)$$

$$K(n) = P(n/n-1) \times (P(n/n-1) + R)^{-1} \quad (3.124)$$

$$x(n/n) = x(n/n-1) + K(n) \times e(n) \quad (3.125)$$

$$P(n/n) = (I - K(n)) \times P(n/n-1) \quad (3.126)$$

$$x(n+1/n) = \phi x(n/n) \quad (3.127)$$

$$P(n+1/n) = \phi P(n/n) \phi^T + \sigma_u^2 \quad (3.128)$$

4. ARAŞTIRMA BULGULARI

4.1. Kullanılan Veritabanı

Bu çalışmada kullanılan ses dosyaları, Speech Enhancement Assessment Resource (SpEAR) veritabanından alınmıştır. 16 kHz ile örneklenmiş, yaklaşık 3 sn uzunluğunda, 1 bay ve 1 bayan konuşmacı tarafından seslendirilmiş 2 cümle kullanılmıştır. Kullanılan ses dosyaları Çizelge 4.1.'de verilmiştir.

Çizelge 4.1. SpAER veritabanından alınan ses dosyaları

Konuşmacı	Dosya	Cümle
Bay	Bigtips_16.wav	“Good service should be rewarded by big tips”
Bayan	Scholars_16.wav	“Biblical scholars argue history”

Gürültü örnekleri ise Noisex-92 veritabanından, temiz konuşma sinyalleri ile aynı frekansta (16 kHz) örneklenmiş olarak alınmıştır. Üç farklı gürültü tipi her iki cümleye de, farklı Sinyal/Gürültü oranlarında (SNR) eklenerek sinyal iyileştirme algoritmaları gerçekleştirilmiştir. Kullanılan gürültü tipleri Çizelge 4.2.'de verilmiştir.

Çizelge 4.2. Noisex-92 veritabanından alınan gürültü örnekleri

Gürültü tipi	Açıklama
Fabrika	Bir otomobil fabrikasının üretim bölümünde kaydedilmiştir.
Beyaz	Yüksek kaliteli analog gürültü üretici ile elde edilip örneklenmiştir. Tüm frekanslarda aynı enerjiye sahiptir
F16	Saatte yaklaşık 925 km hızla giden bir F16 savaş uçağının kokpitinde kaydedilmiştir.

4.2. Sinyal gürültü oranı (SNR)

SNR, konuşma kalitesini sıyan objektif bir yöntemdir. Sinyalin enerjisinin gürültünün enerjisine oranı olarak tanımlanır.

$$E = \sum_{n=1}^N s^2(n) \quad (4.1)$$

(4.1) eşitliği sinyalin enerjisini göstermektedir.

4. ARAŞTIRMA BULGULARI

E : Enerji, n: örnek, N: sinyal uzunluğu, s(n) : Sinyalin genlik değerleri

$$SNR = 20\log_{10} (E_s / E_g) \quad (4.2)$$

(4.2) eşitliği SNR değerini dB cinsinden göstermektedir.

E_s : sinyalin enerjisi, E_g : gürültünün enerjisi

SNR değeri 0 dB olan gürültülü sinyal, gürültünün temiz sinyal enerjisi kadar eklendiği anlamına gelmektedir. Eşitlik (4.2) açılırsa Eşitlik (4.3), SNR değerini dB cinsinden ifade eder.

$$SNR = 10\log_{10} \left[\frac{\sum_{n=1}^N s^2[n]}{\sum_{n=1}^N (s[n] - \tilde{s}[n])^2} \right] \quad (4.3)$$

s[n]: orijinal temiz sinyal, $\tilde{s}[n]$: iyileştirilen sinyal

İyileştirilen sinyal, orijinal sinyal s[n]'e ne kadar benzerse o kadar iyi iyileştirme yapılmış olur. Bu durumda SNR artar. Orijinal temiz ses sinyali ile iyileştirilen ses sinyalinin bire bir aynı olduğu kabul edilirse bu durumda SNR değerinin sonsuz olması gerekir. İki sinyal birbirinden çok farklıysa bu durumda SNR değeri düşer (Yüksek 2010).

4.3. Uygulanan Yöntem

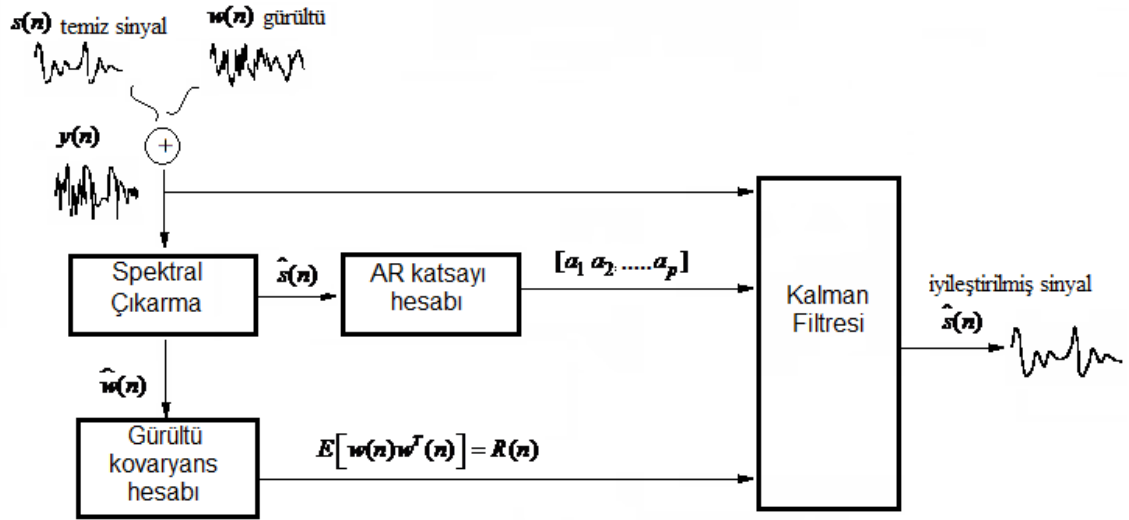
SpEAR veritabanından alınan temiz konuşma sinyalleri, Çizelge (4.2)'de verilen tüm gürültü tipleri ile ve her bir gürültü için -5dB, 0dB ve 5dB olacak şekilde üç farklı SNR değerinde bozularak toplamda 18 gürültülü konuşma sinyali elde edilmiştir. Bu sinyaller, temelleri bölüm 3'de anlatılan ses iyileştirme yöntemlerinden Spektral Çıkarma, Wiener Filtresi ve Kalman Filtresi ile iyileştirilmeye çalışılmıştır.

Spektral Çıkarma, Martin'in (1994) geliştirdiği minimum istatistiklere dayalı yöntem ile gerçekleştirilmiştir. Wiener Filtresi, Plapous ve ark. (2006) tarafından önerilen Harmonik Geri Oluşumlu Gürültü Azaltma yöntemi ile gerçekleştirilmiştir.

Kalman Filtresi bölüm 3’de verilen (3.123) – (3.128) denklemlerine göre uygulanmıştır. Filtre için gerekli olan AR katsayıları gürültü ile bozulmadan önce temiz konuşma sinyalinden elde edilmiştir.

Ancak esas problem sadece gürültülü sinyalin mevcut olmasıdır yani Kalman filtresi için gerekli olan AR katsayıları, pratikte temiz sinyal olmadığı için hesaplanamayacaktır. Ayrıca Kalman Filtresi kazancının hesaplanabilmesi için gürültüye ait kovaryans matrisinin de bilinmesi gerekmektedir.

Kalman Filtresi’nin başarısını çok büyük oranda etkileyen bu problemler, blok diyagramı Şekil 4.1.’de verilen yöntem ile aşılmaya çalışılmıştır.



Şekil 4.1. Birleştirilmiş Kalman filtresi ile ses iyileştirme yöntemine ait blok diyagramı.

Şekil 4.1.’deki Spektral Çıkarma bloğunda önce, Martin’in (2001) önerdiği minimum istatistiklere dayalı, gürültünün güç spektrum tahmini yöntemi ile gürültü karakteristiği belirlenmiş daha sonra Spektral Çıkarma ile gürültülü sinyal iyileştirilmiştir. Her ne kadar temiz sinyalin katsayılarını tam olarak vermese de iyileştirilmiş sinyal kullanılarak katsayılar hesaplanmıştır. Gürültülü sinyalden iyileştirilmiş sinyal çıkarılarak da gürültüye ait kovaryans matrisi hesaplanmıştır.

4. ARAŞTIRMA BULGULARI

Kalman Filtresi için gerekli olan parametreler belirlendikten sonra, gürültülü sinyale filtre denklemlerinin oluşturduğu algoritma uygulanmıştır ve toplamda uygulanan dört yöntem ile elde edilen simülasyon sonuçları ve SNR artışları karşılaştırılmıştır.

4.4. Benzetim sonuçları

Uygulanan yöntemlere ait simülasyon sonuçları Şekil 4.2. – Şekil 4.19.’da verilmiş olup, bu yöntemlerin sağladığı SNR değerleri Çizelge 4.3. ve Çizelge 4.4.’de gösterilmiştir.

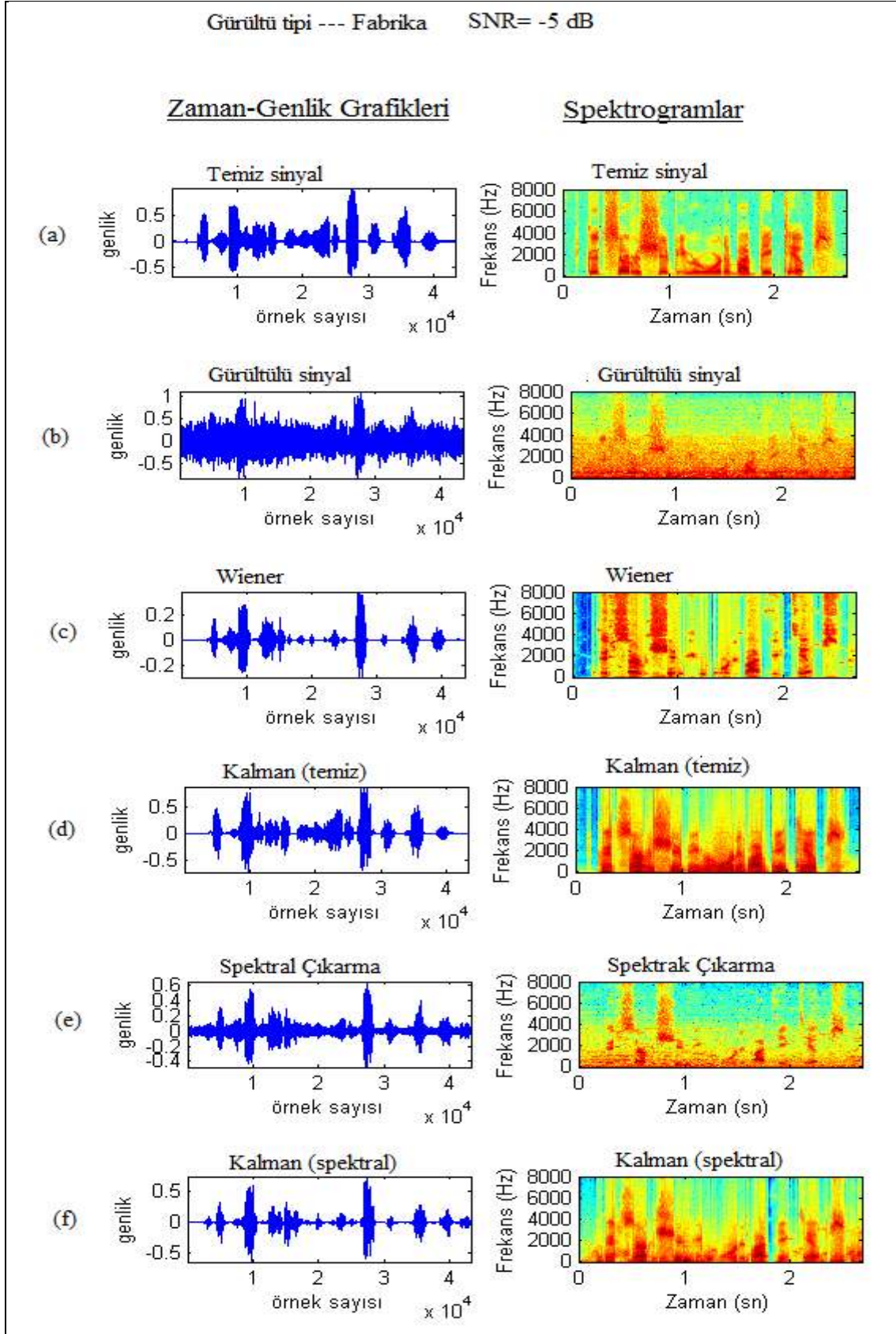
Simülasyon sonuçlarının gösterildiği şekillerde, yöntemleri ifade eden tanımlamalar şöyledir;

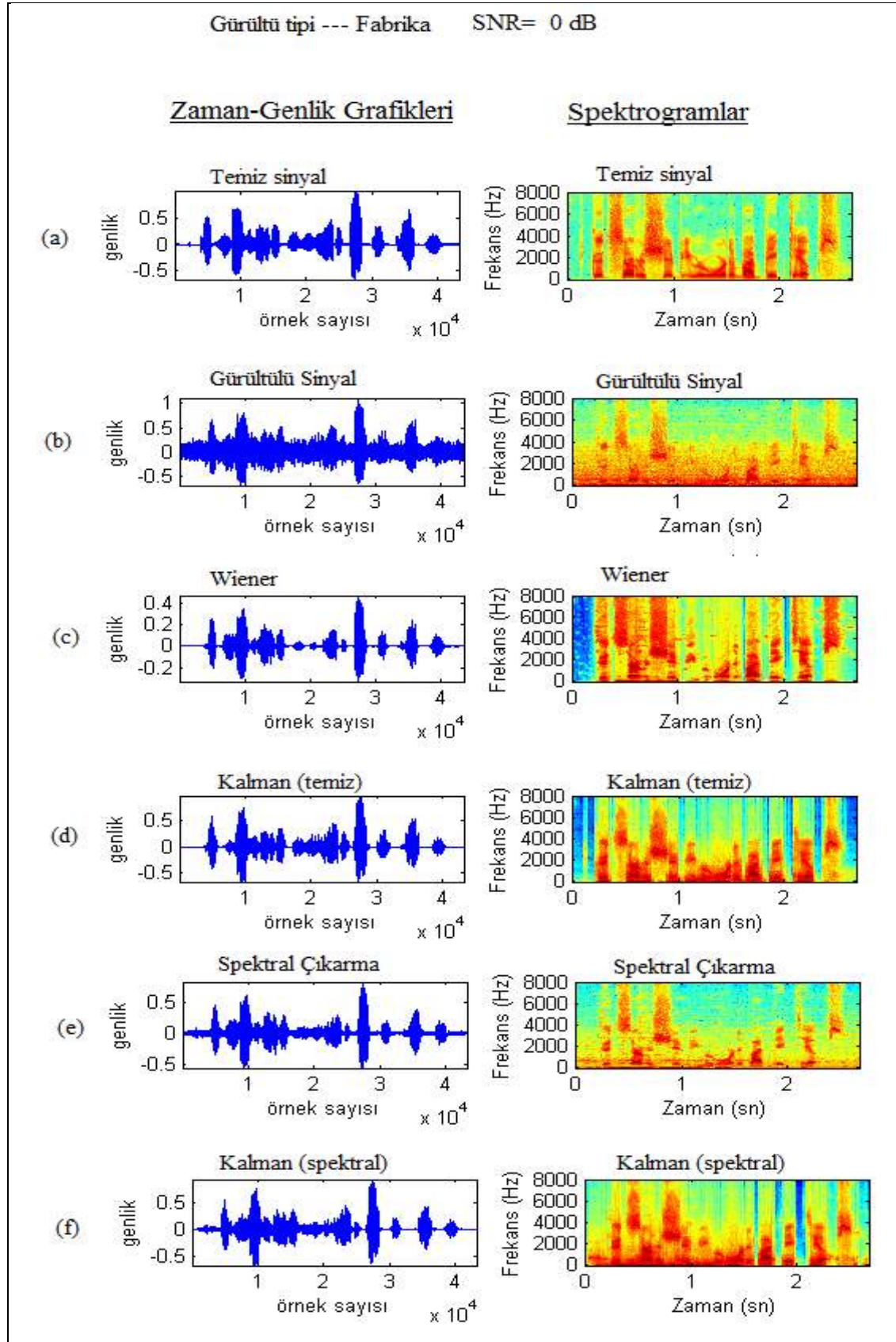
- Wiener: Plapous ve ark. (2006) tarafından önerilen yöntem.
- Spektral Çıkarma: Martin (2001) tarafından önerilen yöntem.
- Kalman (temiz): Temiz sinyalin katsayıları kullanılarak uygulanan Kalman Filtresi.
- Kalman (spektral): Bu çalışmada uygulanan birleştirilmiş Kalman Filtresi.

Ayrıca yine simülasyon sonuçlarının gösterildiği şekil açıklamalarında geçen

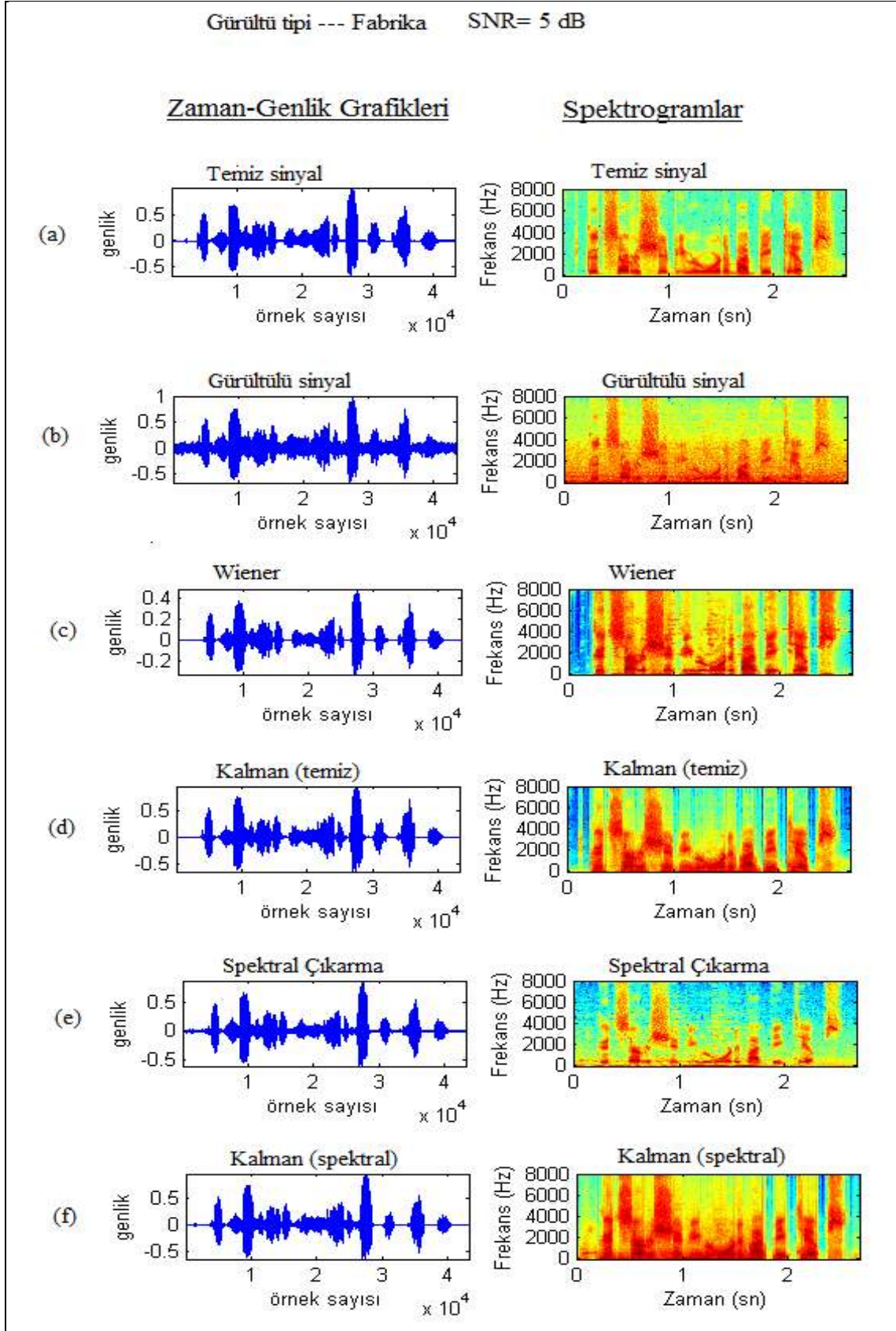
Cümle_1; “Good service should be rewarded by big tips” cümlesini ve

Cümle_2; “Biblical scholars argue history” cümlesini ifade etmektedir.

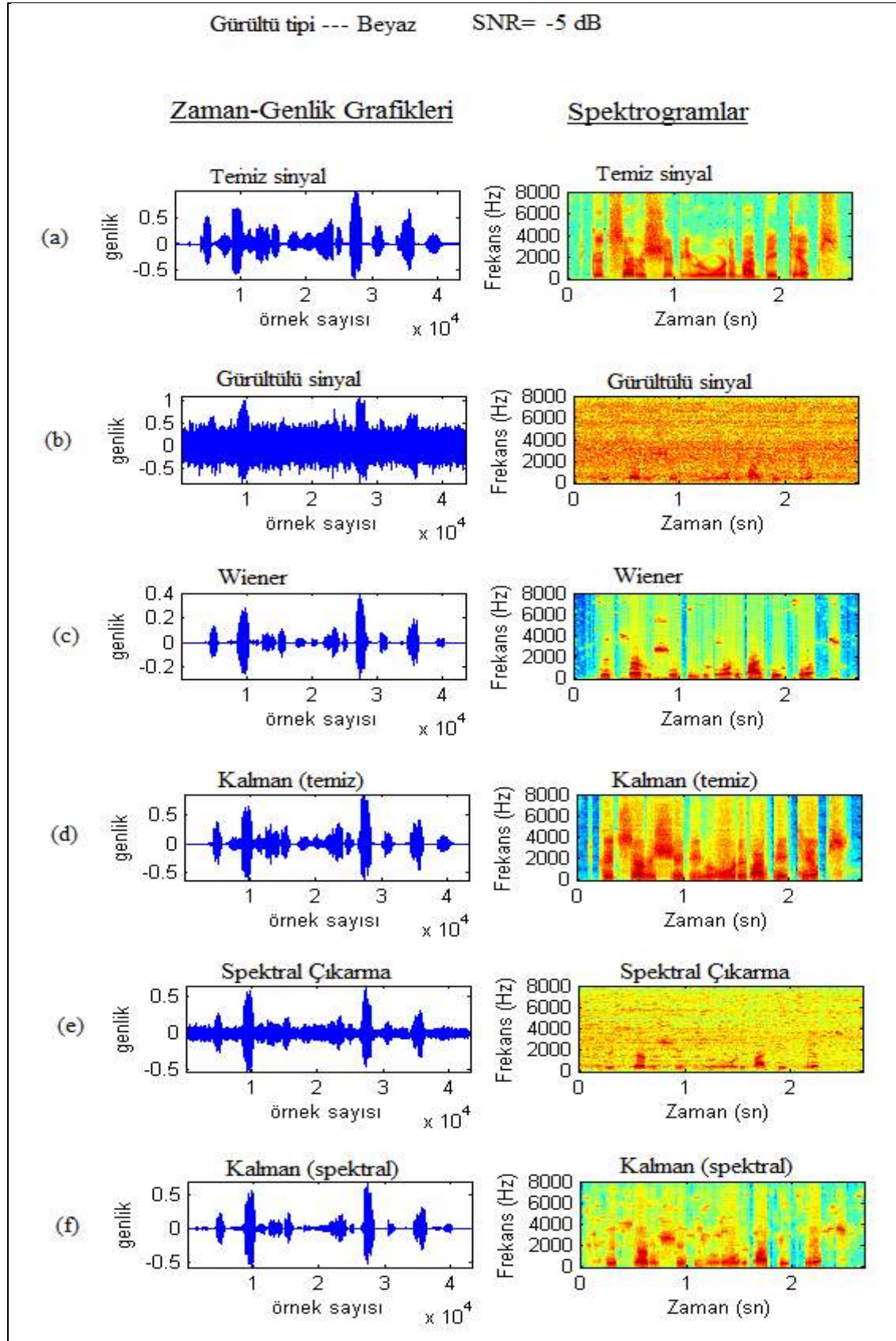




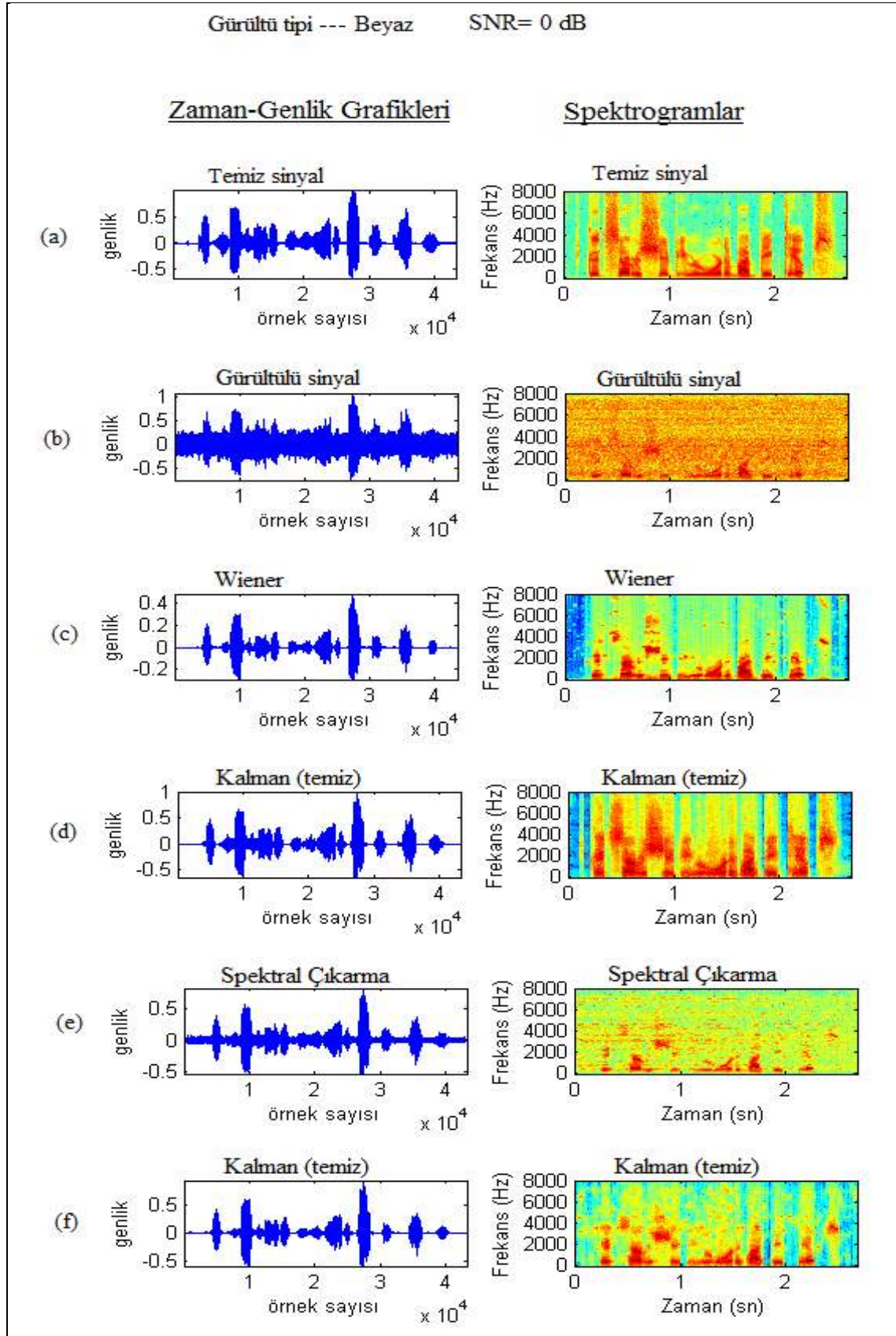
Şekil 4.3. Fabrika gürültüsünde 0dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



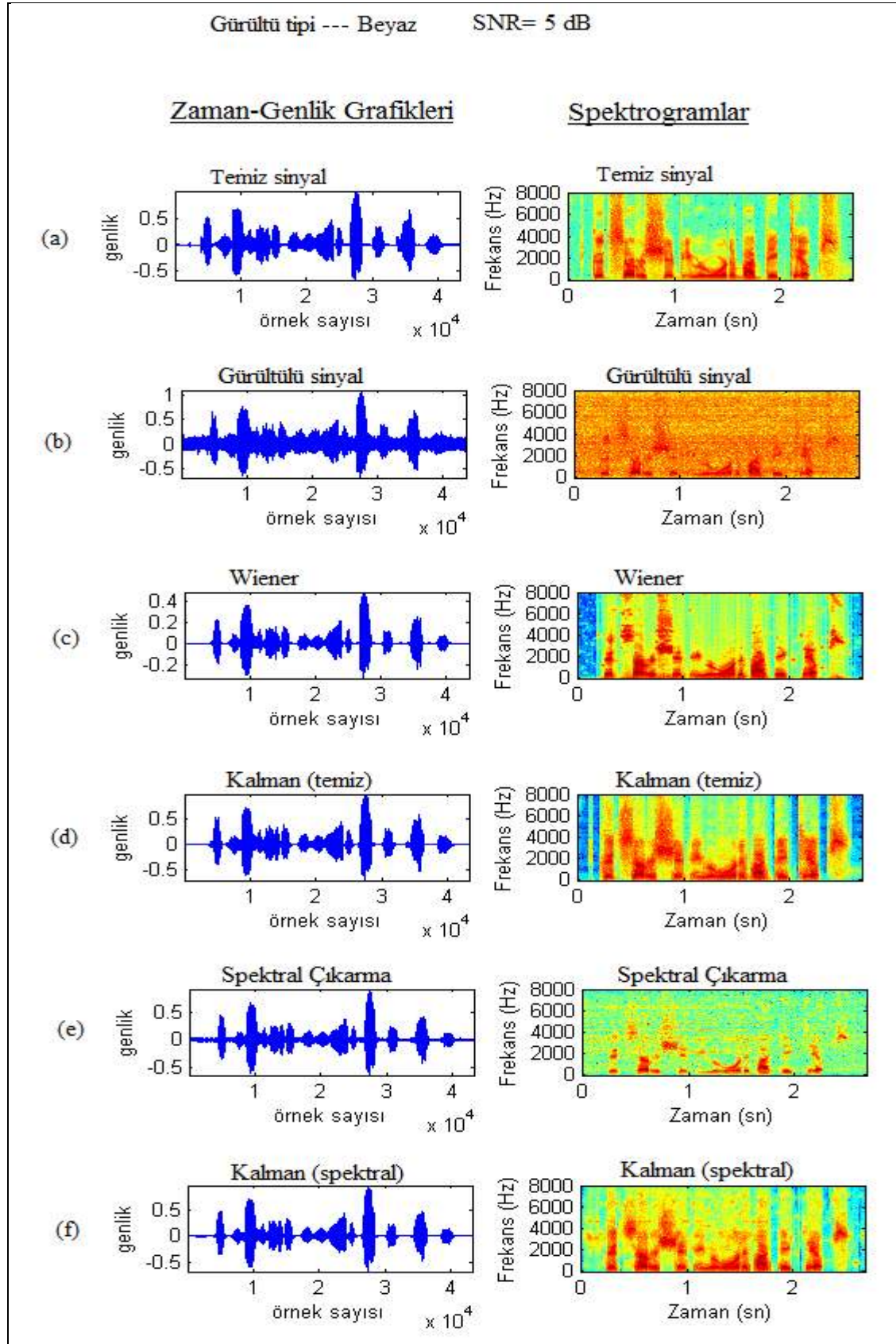
Şekil 4.4. Fabrika gürültüsünde 5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



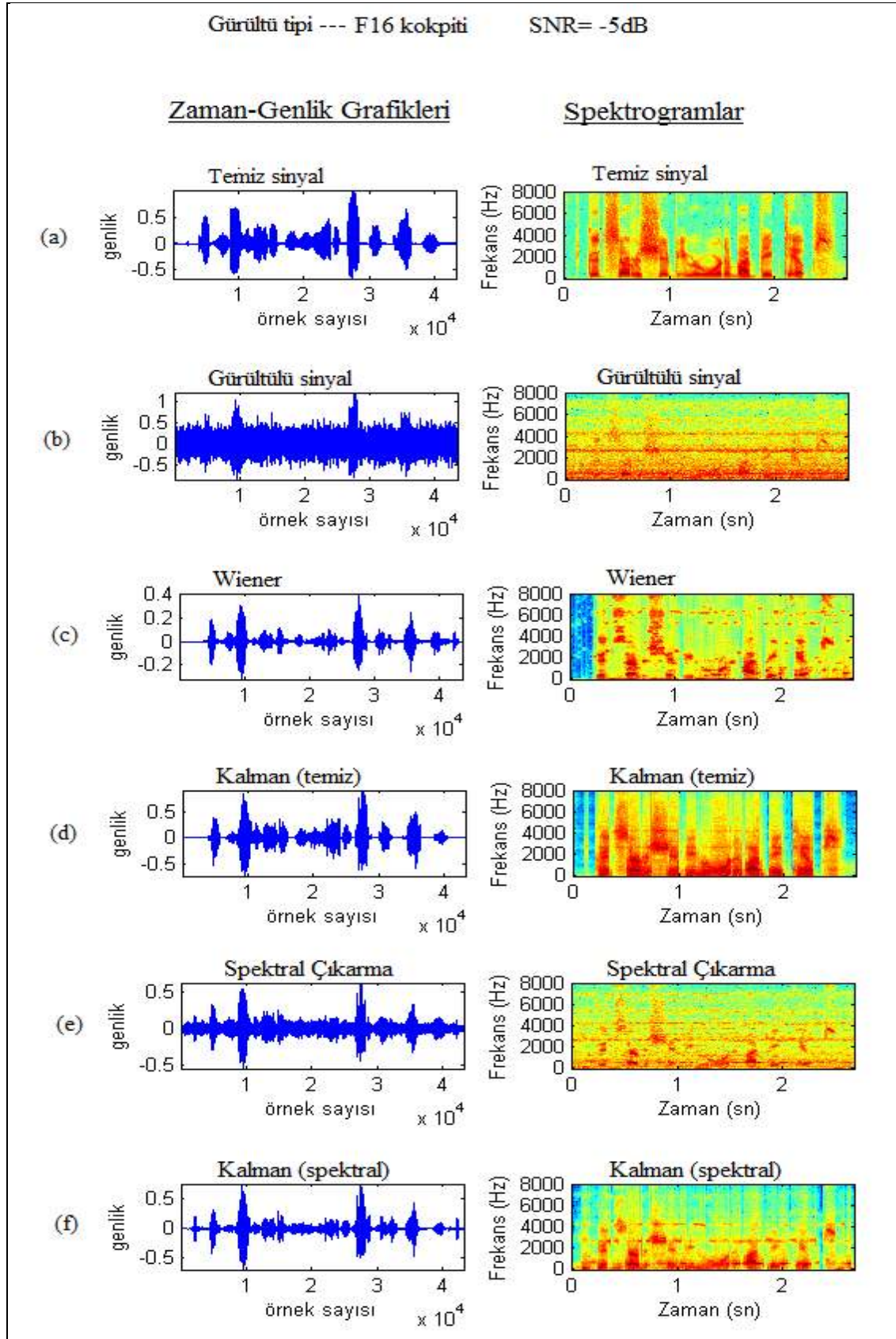
Şekil 4.5. Beyaz gürültüde -5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

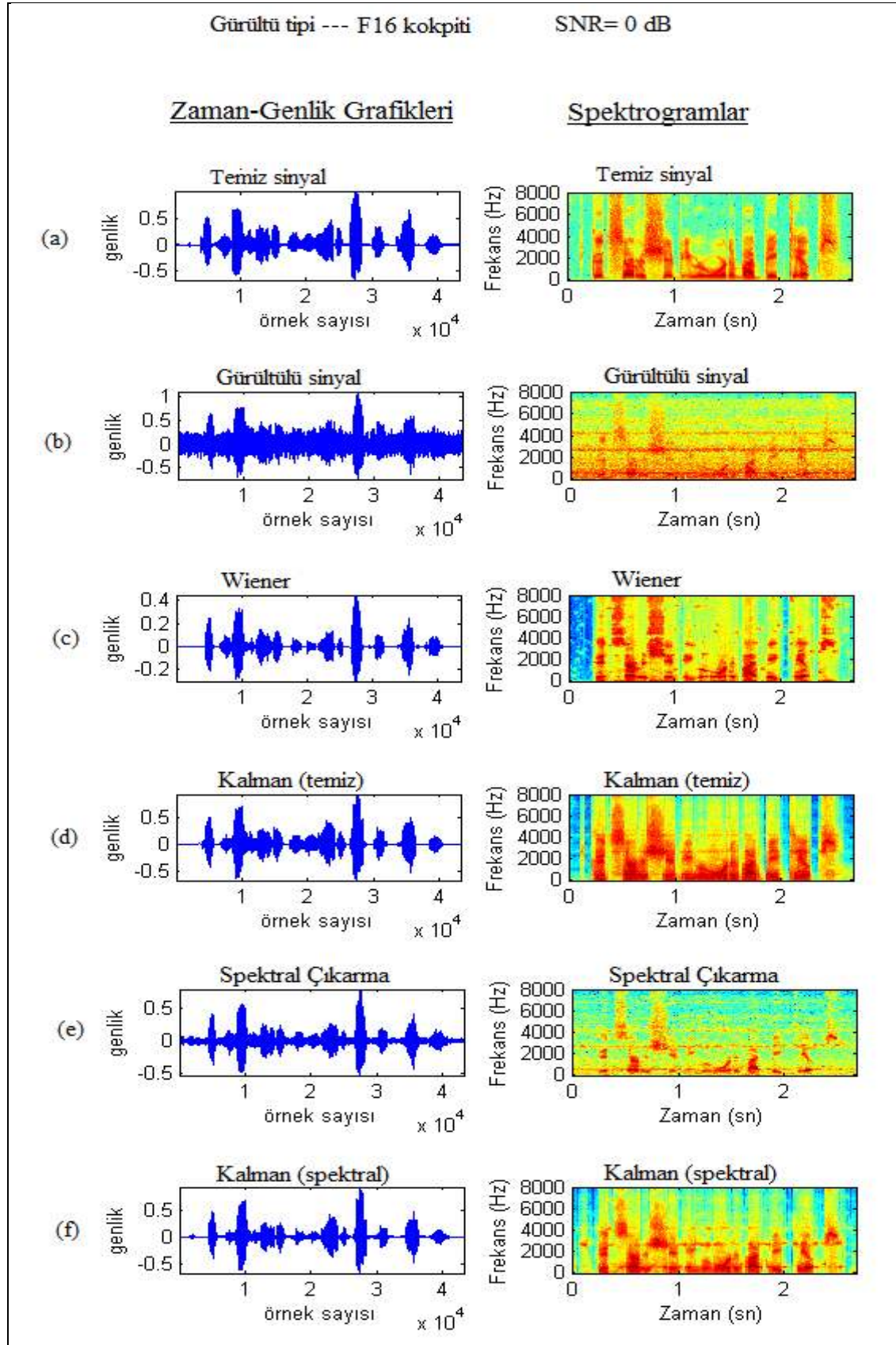


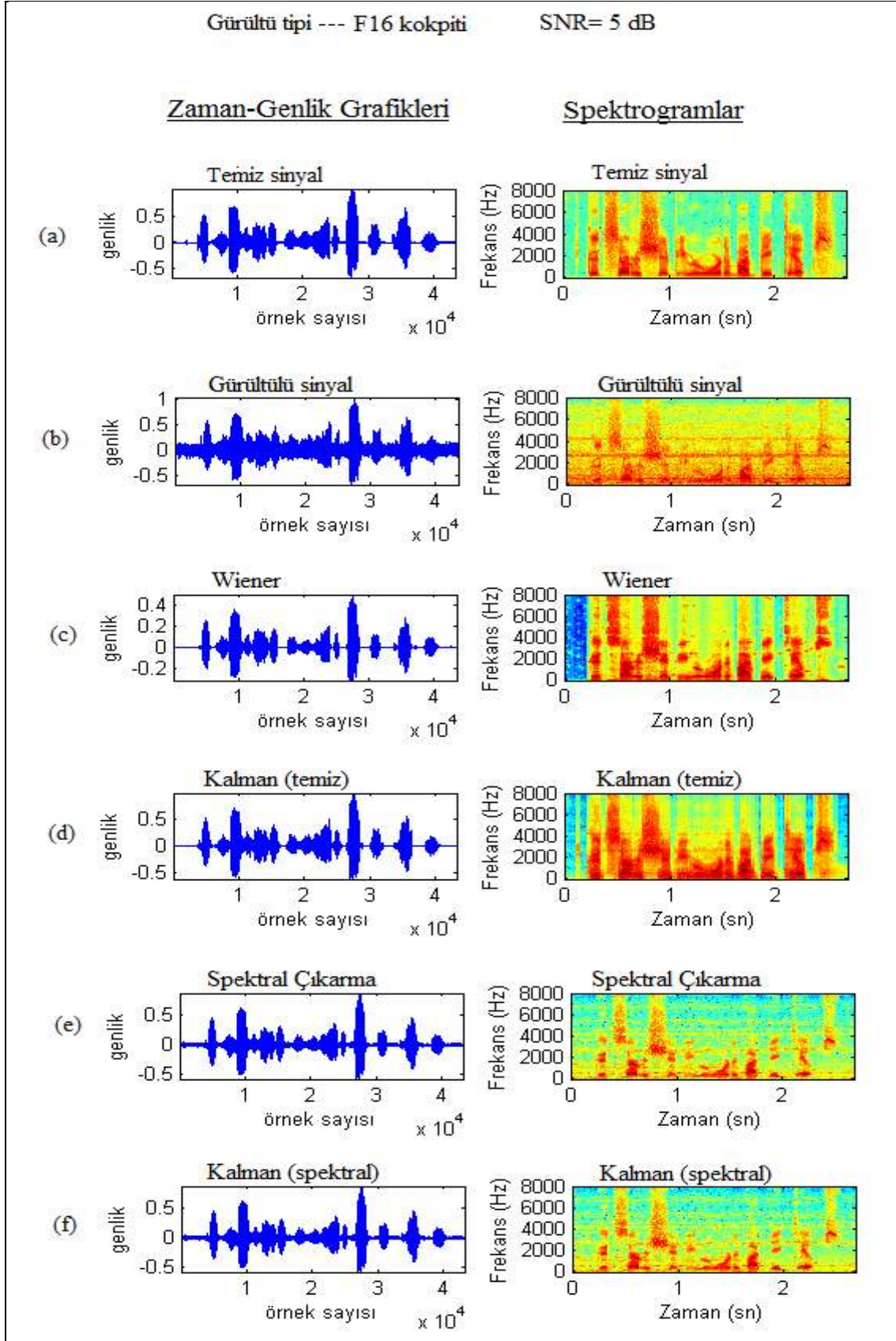
Şekil 4.6. Beyaz gürültüde 0dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

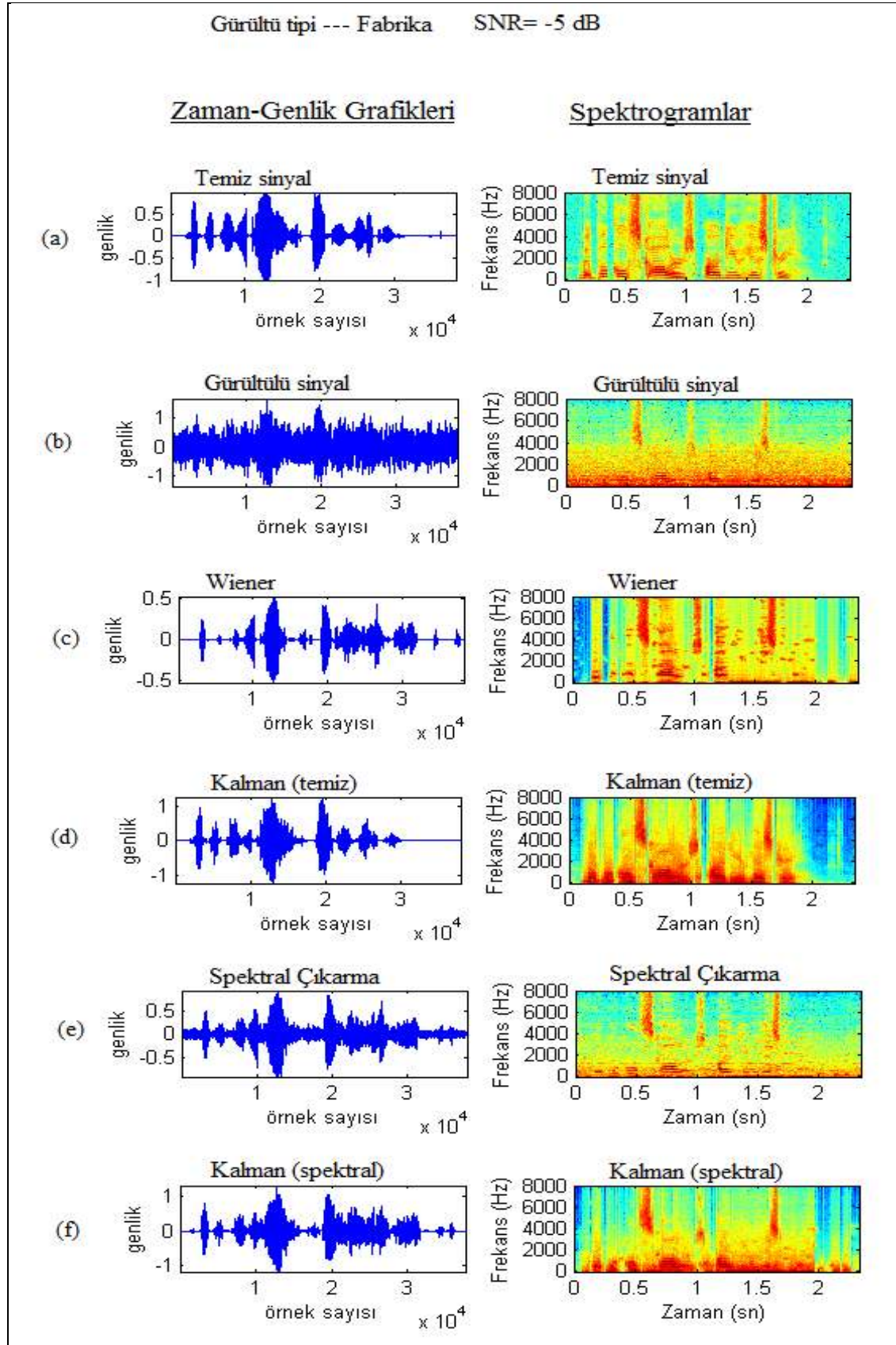


Şekil 4.7. Beyaz gürültüde 5dB SNR ile bozulmuş Cümle_1'e ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

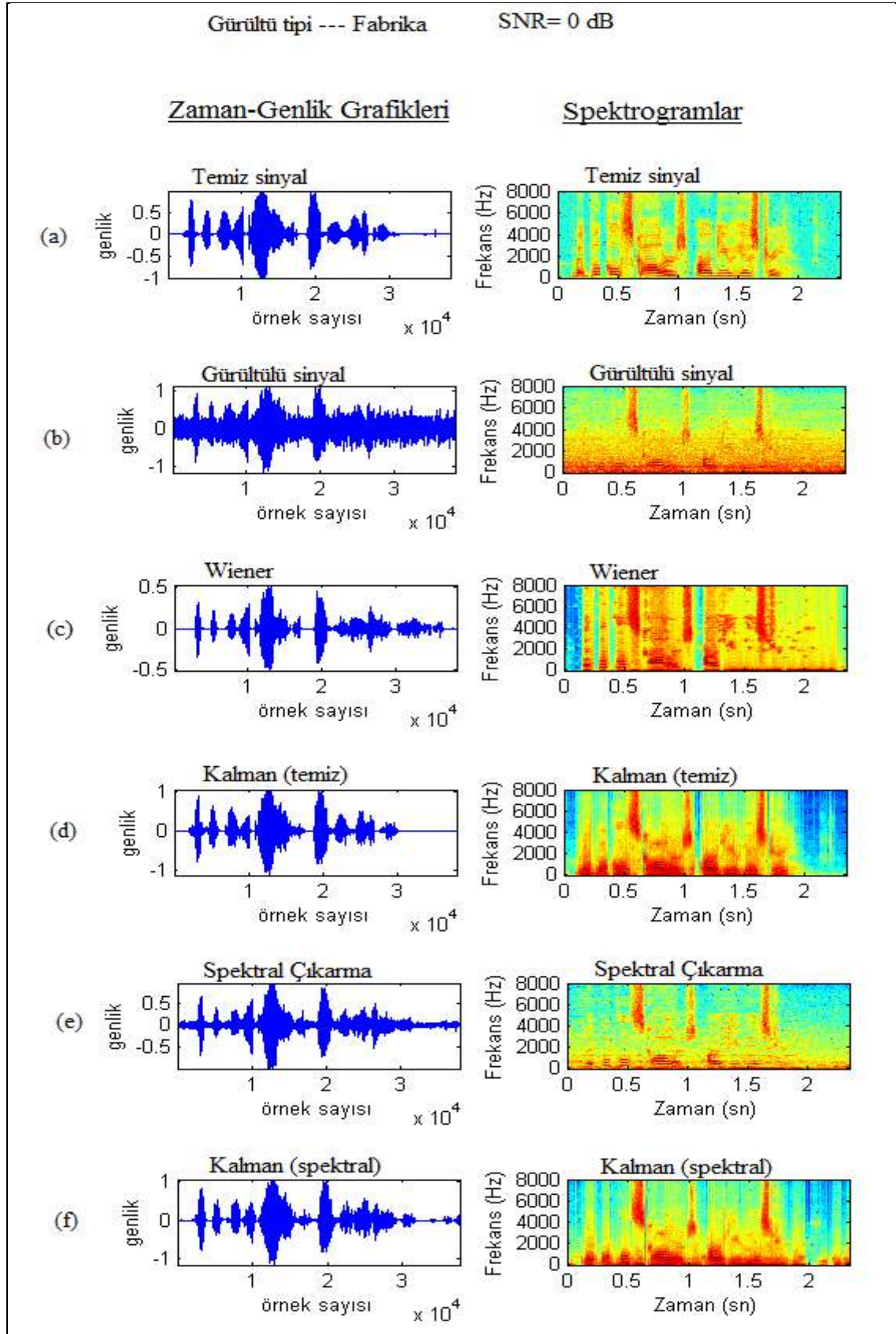


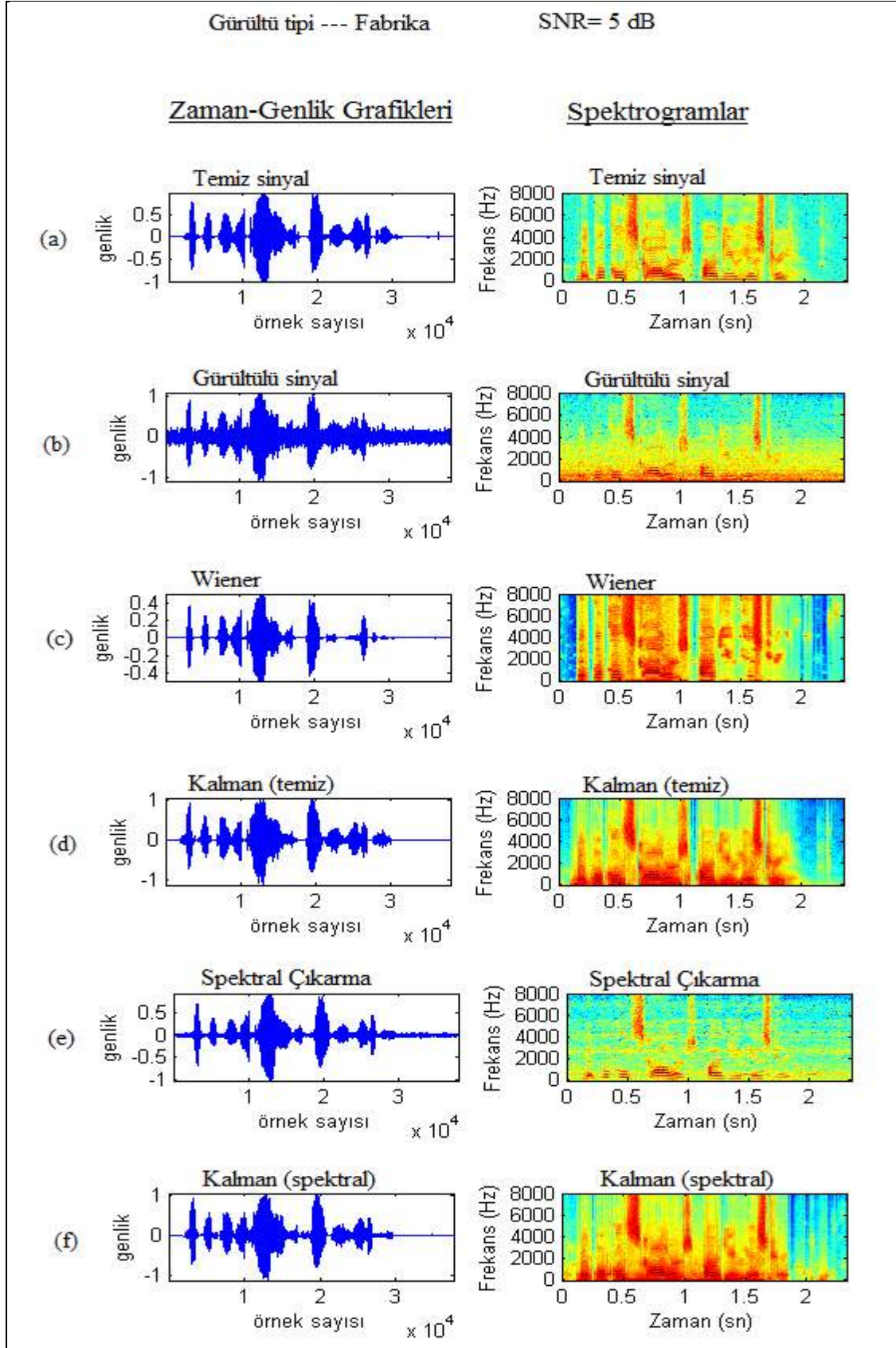




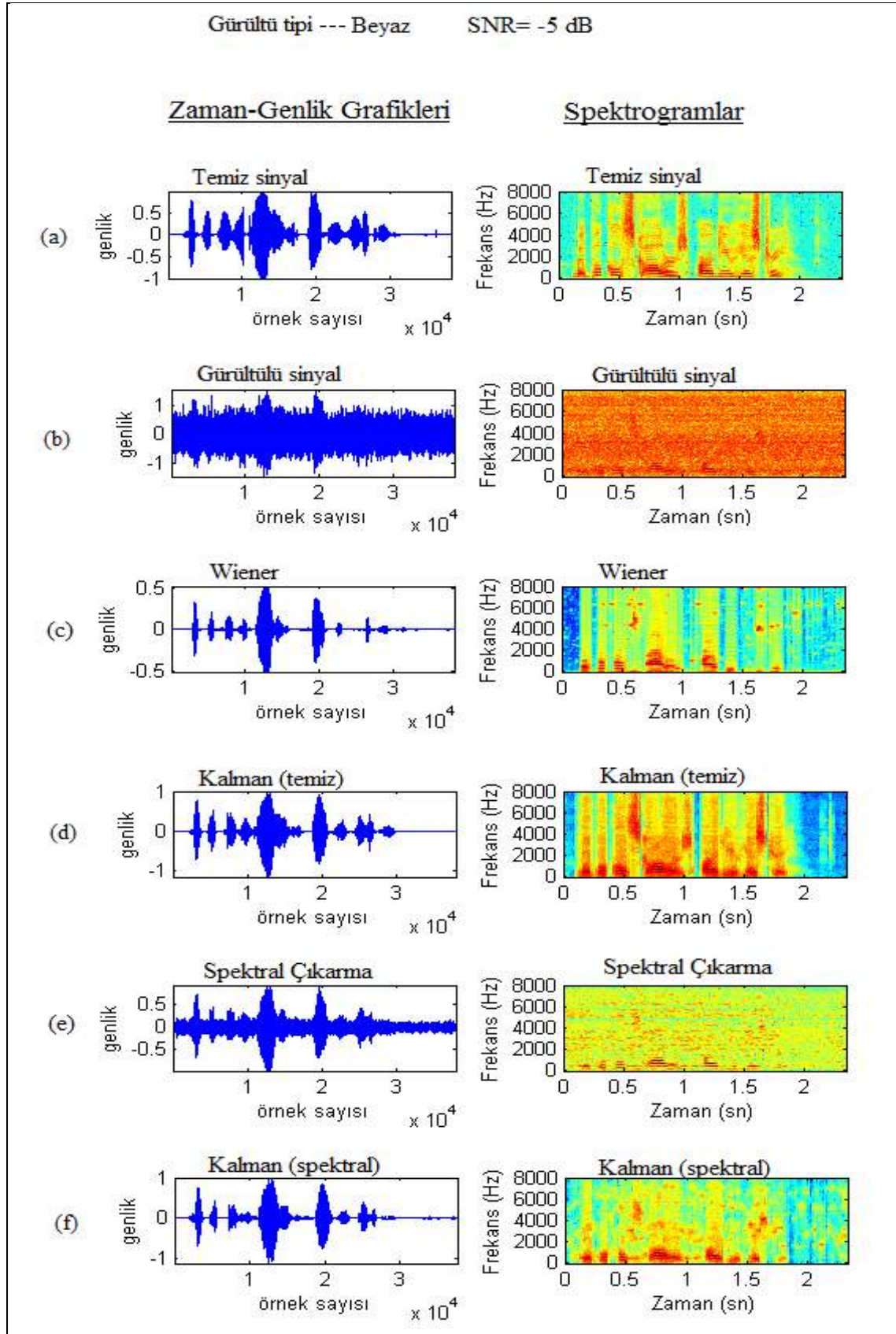


Şekil 4.11. Fabrika gürültüsünde -5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

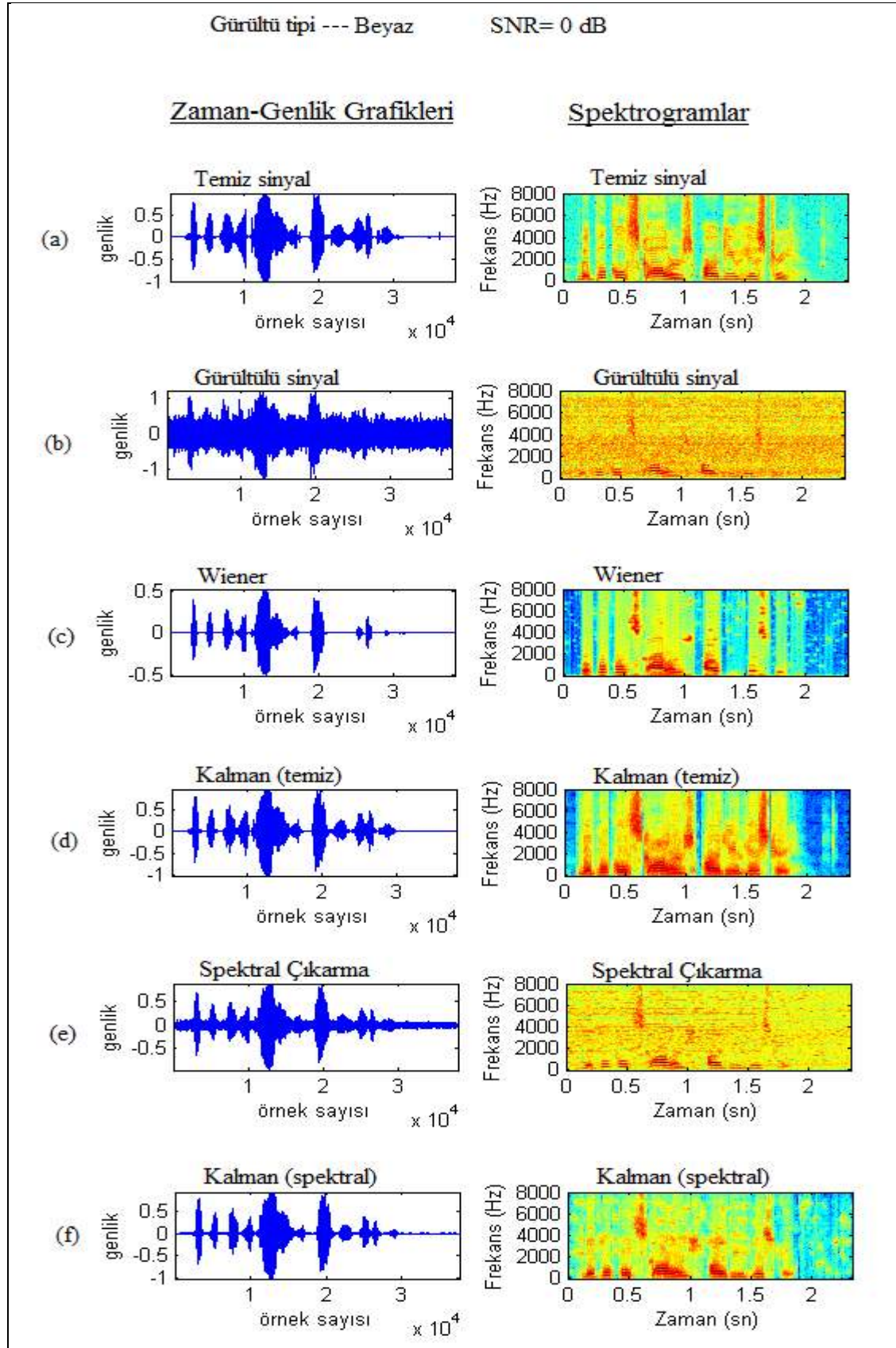




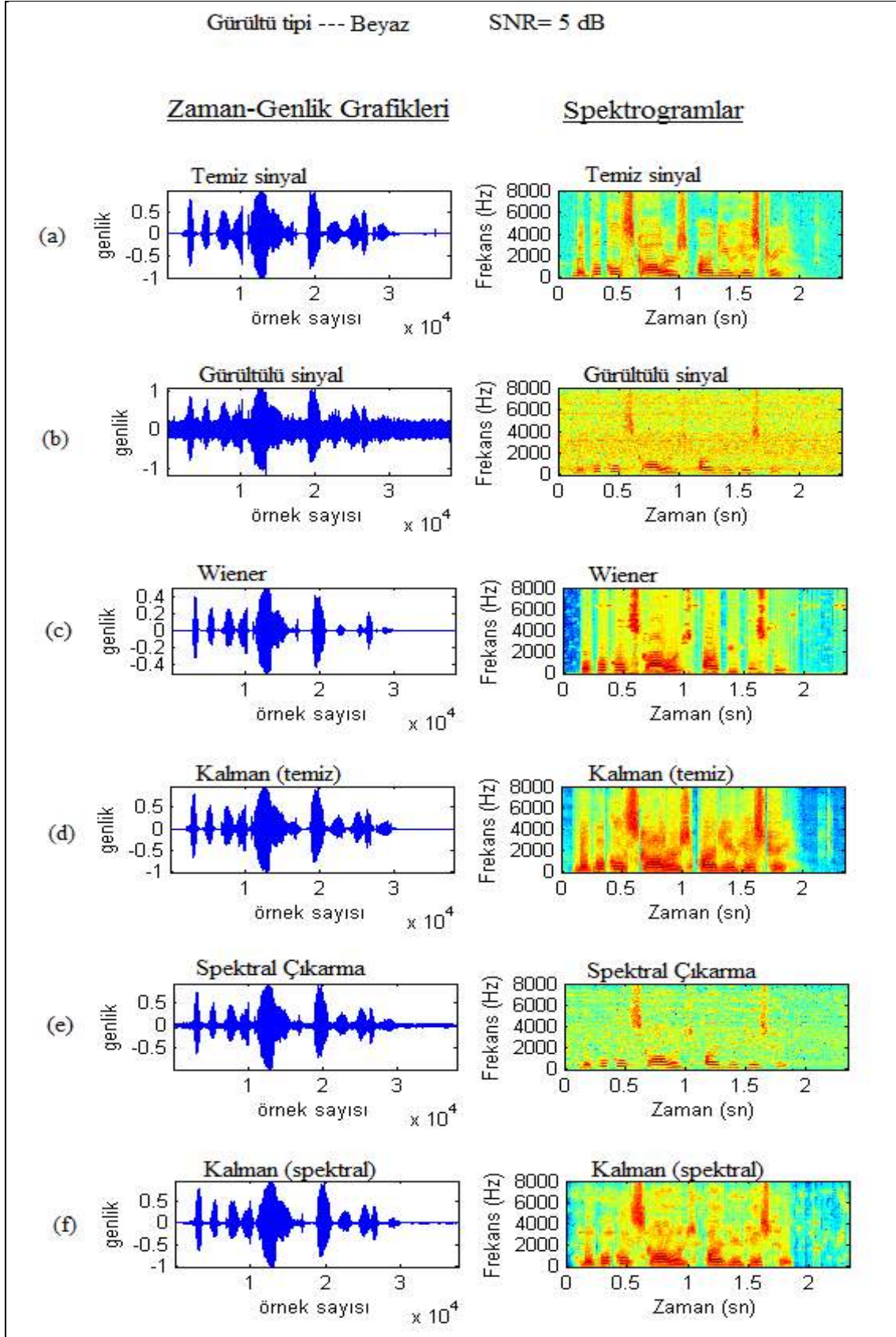
Şekil 4.13. Fabrika gürültüsünde 5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



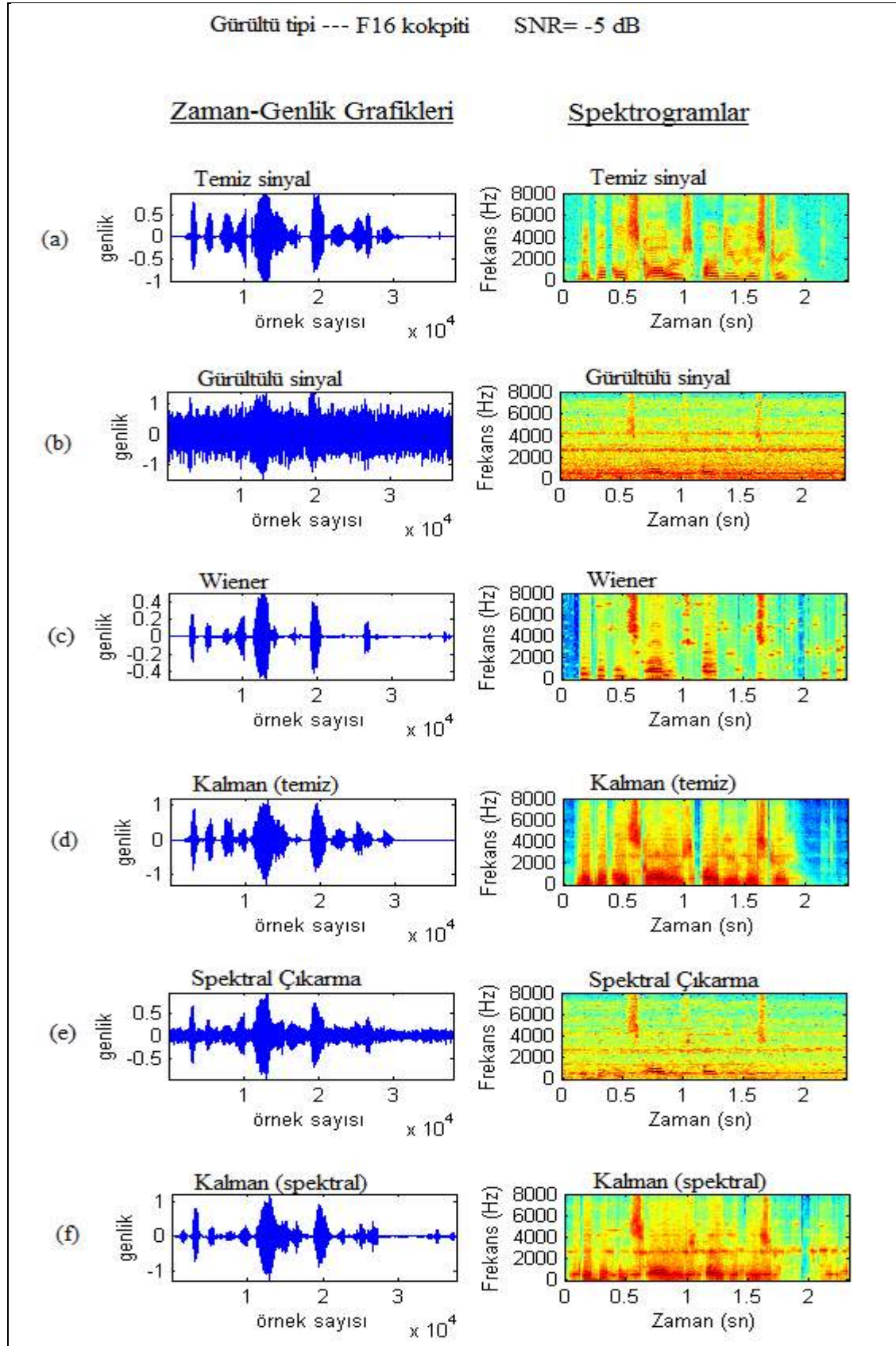
Şekil 4.14. Beyaz gürültüde -5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



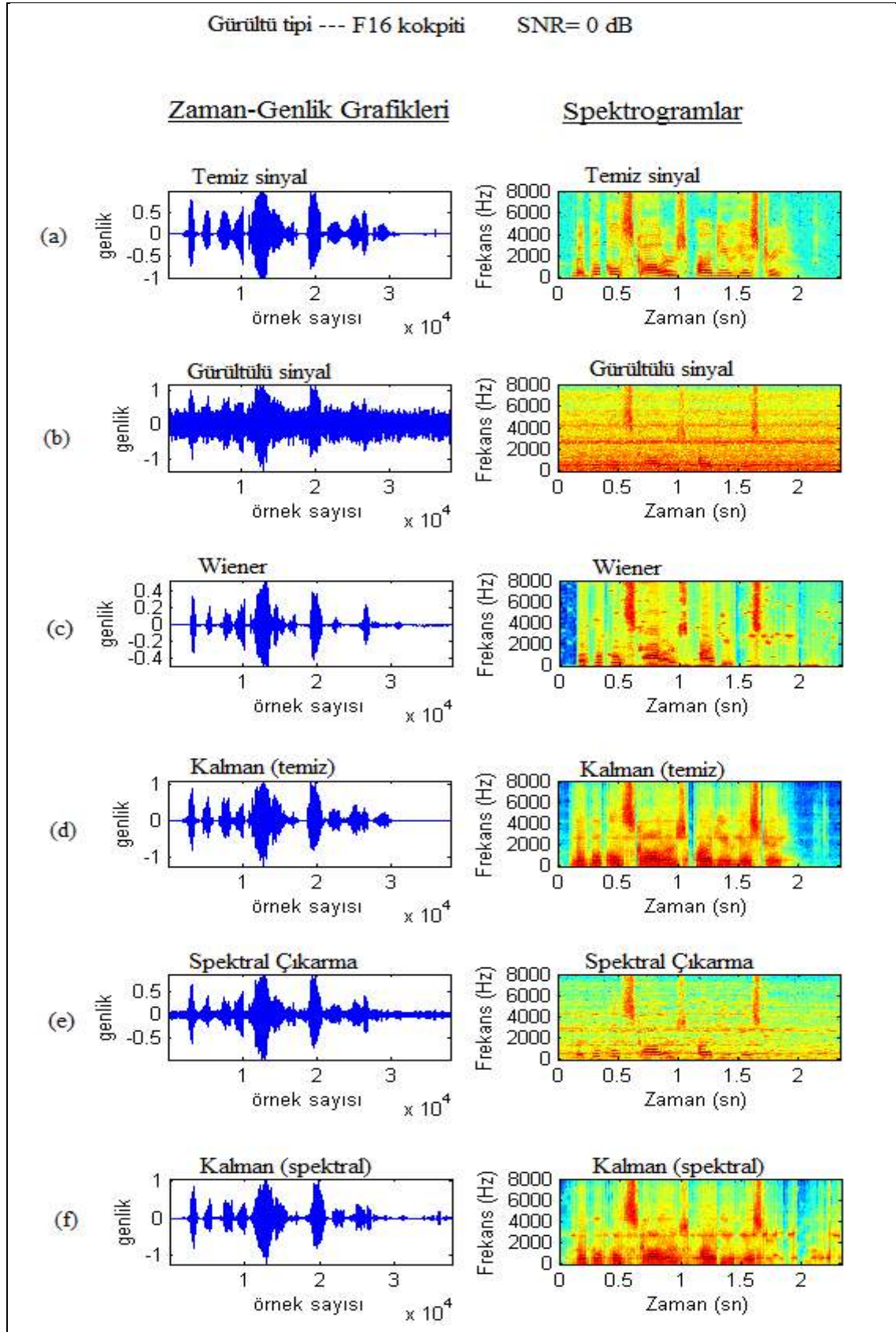
Şekil 4.15. Beyaz gürültüde 0dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

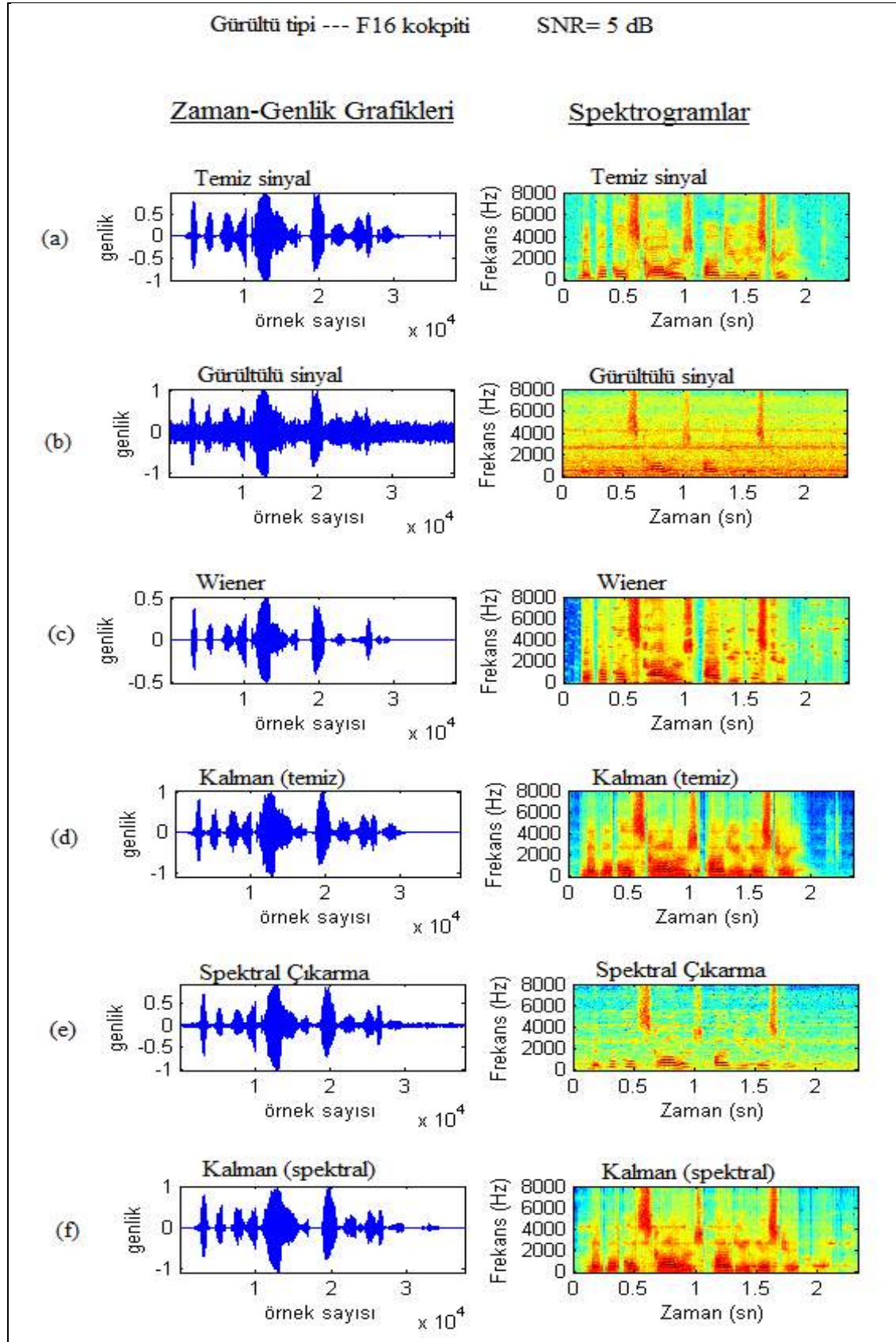


4. ARAŞTIRMA BULGULARI



Şekil 4.17. F16 gürültüsünde -5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları





Şekil 4.19. F16 gürültüsünde 5dB SNR ile bozulmuş Cümle_2'ye ait (a)Temiz (b)Gürültülü sinyalin (c)Wiener ile (d)Kalman (temiz) ile (e) Spektral Çıkarma ile (f)Kalman(spektral) ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

Çizelge 4.3. Farklı gürültü tipleri için Cümle_1'e uygulanan yöntemlerin sağladığı SNR değerleri

Gürültü Tipleri	Giriş SNR (dB)	Çıkış SNR (dB)			
		Kalman (temiz)	Wiener	Spektral Çıkarma	Kalman (spektral)
Fabrika	-5	3.57	2.58	3.02	3.18
	0	5.03	3.83	4.21	4.56
	5	8.55	5.19	7.65	8.13
Beyaz	-5	6.19	2.82	3.80	4.53
	0	9.00	4.08	6.81	7.77
	5	12.10	5.29	10.08	11.04
F16 kokpiti	-5	3.00	2.27	2.48	2.74
	0	6.00	3.65	5.79	5.91
	5	9.58	5.83	9.17	9.47

Çizelge 4.4. Farklı gürültü tipleri için Cümle_2'ye uygulanan yöntemlerin sağladığı SNR değerleri

Gürültü Tipleri	Giriş SNR (dB)	Çıkış SNR (dB)			
		Kalman (temiz)	Wiener	Spektral Çıkarma	Kalman (spektral)
Fabrika	-5	3.76	2.07	2.35	2.51
	0	6.42	3.31	4.22	4.35
	5	9.20	5.16	8.32	8.62
Beyaz	-5	7.85	3.28	4.36	6.85
	0	10.98	4.07	8.41	9.96
	5	14.22	5.45	12.00	13.49
F16 kokpiti	-5	3.92	2.89	3.68	3.74
	0	7.30	3.63	7.44	7.47
	5	10.85	5.18	11.71	10.85

5. TARTIŞMA VE SONUÇ

Ses sinyallerinin iyileştirilmesi ile amaçlanan, çeşitli algoritmalar kullanılarak gürültülü sinyallerin kalitesinin veya anlaşılabilirliğinin arttırılmasıdır. Ses iyileştirme yöntemleri zaman veya frekans domenlerinde gerçekleştirilebilir. Frekans domeninde gerçekleştirilen ve pratikte uygulanabilirliği en fazla olan Spektral Çıkarma yöntemi, etkili bir iyileştirme tekniği olmasına karşın, iyileştirdiği sinyalde meydana gelen müzikal gürültü oldukça büyük bir dezavantajdır.

Kalman filtresinin temellerini oluşturan Wiener filtresi, ses iyileştirme için kullanılan bir başka yöntemdir. Wiener filtresinin dezavantajı ise tüm frekanslarda filtrenin sabit bir frekans cevabının olması ve temiz sinyal ile gürültünün spektral güç yoğunluklarının tahminini gerektirmesidir.

Kalman filtresi ile ses iyileştirme, en etkili yöntemlerden biridir. Ancak Kalman filtresi ile ses sinyallerini iyileştirebilmek için, bir takım parametrelerin bilinmesi gerekmektedir. Temiz sinyale ait AR katsayıları ve gürültüye ait kovaryans matrisi, Kalman filtresinin başarısını çok büyük ölçüde etkileyen ve bilinmesi gereken parametrelerdir.

Bu çalışmada elde edilen sonuçlar, temiz sinyale ait AR katsayılarının ve gürültü varyansının bilindiği varsayımları altında Kalman Filtresinin tek başına yeterli olduğunu, Spektral Çıkarma ve Wiener Filtresine oranla daha iyi bir SNR iyileştirmesi sağladığını göstermektedir. Ayrıca gürültü spektrumunun tahminine dayalı iyileştirme yöntemlerinde meydana gelen ve dinleyici açısından oldukça rahatsız edici olan müzikal gürültü, Kalman Filtresi ile gerçekleştirilen iyileştirme uygulamalarında oluşmamaktadır.

Ancak temiz sinyale ait katsayıların ve gürültü varyansının biliniyor olduğu varsayımı pratikte uygulanabilir bir yaklaşım değildir. Problem, sadece gürültülü sinyalin mevcut olmasıdır ve bu parametrelerin sadece gürültülü sinyalden temin edilmesi gerekmektedir. Dolayısıyla bu çalışmada gerçekleştirilmeye çalışılan, önce gürültü karakteristiğinin olabildiğince doğru tahmin edilmesidir. Bu doğrultuda, diğer bir çok tekniğe göre daha başarılı bir tahmin gerçekleştiren Martin'in(2001) yöntemi uygulanarak gürültü spektrumu tahmin edilmiştir.

Daha sonra Spektral Çıkarma ile iyileştirilen sinyal kullanılarak AR katsayıları ve gürültülü sinyalden iyileştirilen sinyal çıkarılarak gürültü varyansı hesaplanmıştır. Son aşamada ise gürültülü sinyale, belirlenen parametreler kullanılarak Kalman Filtresi uygulanmıştır.

Elde edilen sonuçlar, birleştirilmiş Kalman filtresinin Wiener filtresine ve Spektral Çıkarmaya oranla daha iyi bir SNR artışı sağladığını göstermiştir. Ayrıca birleştirilmiş Kalman filtresinin Spektral Çıkarmadan kaynaklanan müzikal gürültüyü de bastırdığı gözlemlenmiştir.

Ancak, temiz sinyal katsayılarının kullanıldığı Kalman(temiz) filtresine oranla daha düşük bir SNR değeri elde edilmiştir. Bunun sebebi ise iyileştirilmiş sinyalden elde edilen katsayıların tam olarak gerçek katsayılar olmamasıdır. Bu durum daha önce de değinildiği gibi AR katsayılarının, Kalman Filtresinin performansını çok büyük bir ölçüde etkilediğinin bir göstergesidir.

Sonuç olarak, doğru bir gürültü tahmini ve orijinal sinyale ait AR katsayılarının belirlenmesi ile Kalman Filtresi oldukça başarılı bir iyileştirme performansı sunmaktadır. Bu bağlamda, arka plan etkisiyle bozulmuş konuşma sinyallerinden orijinal sinyale ait katsayıların belirlenmesi, çalışılması gereken bir konu olarak düşünülmektedir.

6. KAYNAKLAR

- Akaike, H. 1974. A New Look at Statistical Model Identification. IEEE Trans. Automatic Control, AC-19: S,716–723.
- Ankışhan, H. 2007. Gürültülü Ses sinyali İyileştirilmesine İkili Kalman Filtre Yaklaşımı. Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 135.
- Babacan, E.K. 2009. Kısıtlı Durum Kalman Filtresi ve Bazı Uygulamaları. Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 172.
- Babu, G.R., Rao, R. 2011. Modified Kalman Filter-based Approach in Comparison with Traditional Speech Enhancement Algorithms from Adverse Noisy Environments. International Journal on Computer Science and Engineering (IJCSE). S,744-759.
- Berouti, M., Schwartz, R., Makhoul, J. 1979. Enhancement of Speech Corrupted by Acoustic Noise. Proc. IEEE-ICASSP, cilt 4, S,208-211
- Bhatnagar, M. 2002. A Modified Spectral Subtraction Method Combined with Perceptual Weighting for Speech Enhancement. Yüksek Lisans Tezi, The University of Texas, Dallas, 110.
- Boll, S. F. 1979. Suppression of Acoustic Noise in Speech Using Spectral Subtraction. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-27. S,113-120
- Cai, H., Grivel, E., Najim, M. 2003. A Dual Kalman Filter-Based Smoother for Speech Enhancement. IEEE-ICASSP '03. S,912-915.
- Cecelioğlu, S. 2010. Tek Kanallı Toplamsal Gürültülü Konuşma Sinyali İyileştirme. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 115.
- Cohen, I. 2004. Speech Enhancement Using a Noncausal A Priori SNR Estimator. IEEE Signal Processing Letters, 11(9): S,725–728.
- Commins, B. 2005. Signal Subspace Speech Enhancement with Adaptive Noise Estimation. Yüksek Lisans Tezi, National University of Ireland, Galway, 107.
- Deviren, M. 2000. Enhancement of Noisy Speech Utilizing the Kalman Filter. Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi, Ankara, 68.
- Diniz, P.S.R. 1997. Adaptive Filtering: Algorithms and Practical Implementation, 627, Kluwer Academic Publishers.

Doblinger, G. 1995. Computationally Efficient Speech Enhancement By Spectral Minima Tracking in Subbands. Proc. EuroSpeech, Cilt 2. S,1513-1516.

Doblinger, G. 1998. Smoothing of noisy AR signals using an adaptive Kalman filter. Proceedings of the EUSIPCO. Cilt 2. S,781–784.

Ephraim, Y., Malah, D. 1984. Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. IEEE Trans. Acoust., Speech and Signal Processing, ASSP-32 (6): S,1109-1121.

Ephraim, Y. 1992. Statistical-Model Based Speech Enhancement Systems. Proc. IEEE, 80(10): S,1526–1554.

Ephraim, Y., Van Trees H. L. 1995. A signal subspace approach for speech enhancement. IEEE Trans. Speech and Audio Processing. Cilt 3. S,251-266

Feder, M., Oppenheim, A.V., Weinstein, E. 1987. Methods for Noise Cancellation Based on the EM Algorithm. Proc. IEEE ICASSP. S,201-104.

Fujimoto, M., Ariki, Y. 2000. Noisy Speech Recognition using Noise Reduction Method Based on Kalman Filter. Proc. IEEE ICASSP. S,1727-1730.

Gabrea, M., Grivel, E., Najim, M. 1999. A Single Microphone Kalman Filter-based Noise Canceller. IEEE Signal Processing Letters. Cilt 6. S,55-57.

Gabrea, M. 2001. Adaptive Kalman Filtering-Based Speech Enhancement Algorithm. Canadian Conference on Electrical and Computer Engineering '01. Cilt 1. S,521-526.

Gabrea, M. 2004. Robust Adaptive Kalman Filtering-based Speech Enhancement Algorithm. Proc. ICASSP'04. S,301–304.

Gibson, J.D., Koo, B., Gray, S.D. 1991. Filtering of Colored Noise for Speech Enhancement and Coding. IEEE Trans. on Acoustic, Speech and Signal Processing. Cilt 39. S,1732-1742.

Grivel, E., Gabrea, M., Najim, M. 2002. Speech Enhancement as a Realization Issue. Signal Processing, Cilt 82. S,1963-1978.

Grancharov, V., Samuelsson, J., Kleijn, W.B. 2005. Improved Kalman Filtering for Speech Enhancement. Proc. IEEE ICASSP, S,1109-1112.

Grewal, M.S., Andrews, A.P. 2001. Kalman Filtering Theory and Practice using Matlab. John Wiley&Sons,Inc. Sayfa:410.

- Güler, İ., Kıymık, M.K., Akın, M., Alkan, A. 2001. AR spectral analysis of EEG signal by using maximum likelihood estimation. *Comput. Biol. Med.* 31. S,441–450.
- İkiz, M. 2006. Wavelet ve Yapay Sinir Ağı Kullanarak Ses Sinyalinden Konuşmacı Tespiti. Yüksek Lisans Tezi, Dicle Üniversitesi Fen Bilimleri Enstitüsü, Diyarbakır, 170.
- Kalman, R.E. 1949. A New Approach to Linear Filtering and Prediction Problems. *Trans. of the ASME-Journal of Basic Engineering*, New York. Cilt 26. S,35-45.
- Kalman, R.E., Bucy R.S. 1961. New Results in Linear Filtering and Prediction Theory. *Trans. ASME J. Basic Eng.*, 83: S,95–108.
- Kırçiçek, Y. 2007. Doğrusal Öngörü ile Konuşma İşareti Kodlayıcısı Tasarımı. Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Fen Bil. Ens., İstanbul, 78.
- Kumar, T.K. 2008. Optimum Cascaded Design for Speech Enhancement using Kalman Filter. *World Academy of Science, Engineering and Technology* 41. S,355-359
- Lee, K.Y., Lee, B.-G., Song, I., Ann, S. 1992. Robust Estimation AR Parameters and its Application for Speech Enhancement. *Proc. IEEE ICASSP*. S,309-312.
- Lim, J.S., Oppenheim, A.V. 1978. All-pole Modelling of Degraded Speech. *IEEE Trans. on Acoustic, Speech and Signal Processing*, ASSP-26 (3): S,197-210.
- Ma, N., Bouchard, M., Goubran, R.A. 2004. Perceptual Kalman Filtering for Speech Enhancement in Colored Noise. *Proc. IEEE ICASSP*. S,717-720
- Ma, N., Bouchard, M., Goubran, R.A. 2006. Speech Enhancement Using aMasking Threshold Constrained Kalman Filter and Its Heuristic Implementations. *IEEE Trans. on Audio, Speech and Language Processing*. Cilt 14. S,19-32.
- Mack, G.A., Jain, V.K.. 1983. Speech Parameter Estimation by Time-Weighted-Error Kalman Filtering. *IEEE Trans. on Acoustics Speech and Signal Processing*, Cilt ASSP-31. S,1300-1303.
- Mai, Q., He, D., Hou, Y., Huang, Z. 2011. A Fast Adaptive Kalman Filtering Algorithm for Speech Enhancement. *IEEE Int. Con. Automation Science.and Eng.* S,327-332.
- Martin, R. 1994. Spectral Subtraction Based on Minimum Statistics. *Proc. 7th European Signal Processing Conf. EUSIPCO-94*. S,1182-1185.
- Martin, R. 2001. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Trans. Speech Audio Process.*, 9(5): S,504-512.

6. KAYNAKLAR

Mehra, R. K. 1970. On the Identification of Variances and Adaptive Kalman Filtering. IEEE Trans. on Automatic Control. Cilt AC-15. S,175-184.

Najim, M. 2008. Modelling, Estimation and Optimal Filtering. John Wiley&Sons,Inc. Sayfa:408.

Noisex-92 database, Eriřim:[http://spib.rice.edu/spib/select_noise.html], Signal Processing Information Base.

Paliwal, K. K., Basu A.1987. A Speech Enhancement Method Based on Kalman Filtering. IEEE-ICASSP '87. Dallas. S,177- 180.

Plapous C., Marro C., Scalart, P. 2005. Speech enhancement using harmonic regeneration. Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. S,157-160

Plapous C., Marro C., Pascal S. 2006. Improved signal-to-noise ratio estimation for speech enhancement. IEEE Trans. Audio, Speech, Lang. Processing, 14 (6): S,2098-2108.

Pınarcı, E. 2007. İki Boyutlu Kalman Filtresinin Yersel Lazer Tarama Verisine Uygulanması. Yüksek Lisans Tezi, Gebze Yüksek Teknoloji Enstitüsü Mühendislik ve Fen Bilimleri Enstitüsü, Gebze, 90.

Rangachari, S. 2004. Noise Estimation Algorithms for Highly Non-Stationary Environments. Yüksek Lisans Tezi, The University of Texas, Dallas, 73.

Scalart P., Filho J. V. 1996. Speech enhancement based on a priori signal to noise estimation. Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. S,629-632.

Shen, X., Deng, L. 1999. A Dynamic System Approach to Speech Enhancement using the H_{∞} Filtering Algorithm. IEEE Trans. on Speech and Audio Processing. Cilt 7. S,391-399.

So, S., Paliwal, K. 2008. A long state vector Kalman filter for speech enhancement. Proc. Int. Conf. Spoken Language Processing (INTERSPEECH-ICSLP). S,391-394.

Şahin, H. 2007. Tek kanallı sistemler için geliştirilmiş ses iyileştirme yöntemi. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 81.

Uslu, E. 2007. Gizli Markov Modeli ile Geniş Sözlüklü Sürekli Konuşma Tanıma. Yüksek Lisans Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul, 81.

Vaseghi, S.V. 2008. Advanced Digital Signal Processing and Noise Reduction. John Wiley&Sons, Ltd. ISBN: 978-0-470-75406-1, 4.baskı, Sayfa:532.

Wan E., Nelson A., Peterson, R. Speech Enhancement Assessment Resource (SpEAR) Database. Eriřim:[<http://ee.ogi.edu/NSEL/> Beta Release v1.0. CSLU], Oregon Graduate Institute of Science and Technology.

Weinstein, E., Oppenheim, A.V., Feder, M. 1990. Signal Enhancement using Single and Multi-sensor Measurements. Technical report no. 560, M.I.T., Cambridge.

Widrow, B., Glover. J.R., McCool, J.M., Kaunitz, J., Williams, C.S., Hearn, R.H., Zeider, J.R., Dong Jr., E., Goodlin, R.C. 1975. Adaptive Noise Cancelling:Principals and Applications. Proc. IEEE. Cilt 63. S,1692-1716.

Wiener, N.1949. The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. MIT Press, Cambridge.

You, C.H., Rahardja, S., Koh, S.N.2007. Autoregressive Parameter Estimation for Kalman Filtering Speech Enhancement. Proc. IEEE ICASSP. S,913-916.

ÖZGEÇMİŞ

Adı ve Soyadı: Cem KUTLU

Doğum Yeri: Diyarbakır

Doğum Tarihi: 19.07.1985

Medeni Hali: Bekar

Yabancı Dili: İngilizce

Eğitim Durumu (Kurum ve Yıl)

Lise: Milli Piyango Anadolu Lisesi / İzmir, 2003

Lisans: Dicle Üniversitesi, Elektrik Elektronik Mühendisliği / Diyarbakır, 2008

Çalıştığı Kurum/Kurumlar ve Yıl:

AnelTech A.Ş. (Anel Grup)

Elektrik-Elektronik Mühendisi (2008), Diyarbakır

Dicle Üniversitesi

Araştırma Görevlisi (2009-Halen), Diyarbakır

Ulusal ve Uluslar Arası Kongre ve Sempozyumlarda Sunulmuş Bildiriler

1- Yapay Sinir Ağları Temelli Güç Akış Analizi, Bilimde Modern Yöntemler Sempozyumu 2010, Diyarbakır

2- Yoğunlaştırıcı Güneş Enerji Santralleri ve Ilısu HES'e alternatif olarak Güneydoğu Anadolu Bölgesine Uygulanabilirliği, YEKSEM 2009, Diyarbakır

3- Neural Network Based Power Flow Analysis, EPECS-2010, Lviv, Ukraine