

# Spektral çıkarma tabanlı kalman filtresi ile ses sinyallerinin iyileştirilmesi

**Cem KUTLU\***, **Muhammet Ali ARSERİM**

*Dicle Üniversitesi, Elektrik Elektronik Mühendisliği Bölümü, 21280, Diyarbakır*

## Özet

*Kalman filtresinin, özellikle GPS ve Navigasyon uygulamalarında sunduğu üstün tahmin yeteneği, son yıllarda ses sinyallerinin işlenmesinde de kullanılmaya çalışılmıştır. Kalman filtresi en etkili ses iyileştirme yöntemlerinden biridir 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, farklı tipteki gürültülerle bozulmuş sinyallere, Spektral çıkarma, Wiener filtresi ve Kalman filtresi ayrı ayrı uygulanmıştır. Kalman filtresi için gerekli olan katsayılar temiz sinyal kullanılarak hesaplanmıştır. Daha sonra Spektral çıkarma ile birleştirilmiş Kalman filtresi uygulanmıştır. Kalman filtresi için gerekli olan parametreler Spektral çıkarma yöntemi ile iyileştirilmiş sinyal kullanılarak belirlenmiştir. 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 bastıracağı da gözlemlenmiştir.*

**Anahtar Kelimeler:** *Ses iyileştirme, Kalman filtresi, Spektral çıkarma*

\*Yazışmaların yapılacağı yazar: Cem KUTLU. ckutlu@dicle.edu.tr; Tel: (412) 248 82 17 (3561)

## Speech enhancement with spectral subtraction based kalman filter

### Extended abstract

Speech enhancement techniques aim 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 technique 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 autoregressive (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, the 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 all three methods mentioned above 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.

The results achieved in this study has shown that, if the AR coefficients of the original signal and noise variance is known Kalman filter is sufficient alone for the enhancement of noisy speech signals. In addition, musical noise which occurs with the methods based on the noise spectrum estimation, does not occur with the Kalman filtering. However, the assumption that these parameters are known is not workable in practice. Therefore, the AR coefficients are calculated by using the signal enhanced with Spectral subtraction and also noise variance is calculated by subtracting the enhanced signal from the noisy signal. In the last stage, Kalman filter was applied to the noisy signal using the parameters determined with Spectral subtraction.

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. When the SNR values are taken into account, it is seen that the Kalman filter alone provided a better SNR improvement than combined Kalman filter. This is because of using the original signal while calculating the AR parameters for the Kalman filter alone.

**Keywords:** Speech enhancement, Kalman filter, Spectral subtraction

## **Giriş**

Ses iyileştirme sistemleri, konuşmanın kalitesini ve anlaşılabilirliğini geliştirmeyi amaçlayan sistemlerdir. 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 sistemlerinde, sinyal ve gürültü süreçlerinin zamansal-spektral yapıları kullanılarak gürültünün duyulabilirliği bastırılmaya çalışılır. (Vaseghi, 2008)

## **Konuşma haberleşmesinde gürültü ve bozulmanın etkileri**

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 karışmalarını ve kayıtlardaki arka plan gürültüsünü kapsadığı gibi, yankı etkisi veya yükselteçlerden kaynaklanan bozulmaları da kapsamaktadır.

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 ise genelde bir sinyaldeki gönderme kanalının ideal olmayan karakteristiğinden, yankı, eko ve eksik örneklemeden kaynaklanan, sistematik ve istenmeyen değişimdir.

Gürültü ve bozulmalar haberleşme ve ölçüm sistemlerinin en önemli sınırlayıcı etkenleridir. Bu nedenle, gürültünün ve bozulma etkilerinin modellenmesi ve yok edilmesi haberleşme ve sinyal işlemenin çekirdeğini oluşturmaktadır. (Cecelioğlu, 2010)

## **Sinyal iyileştirme**

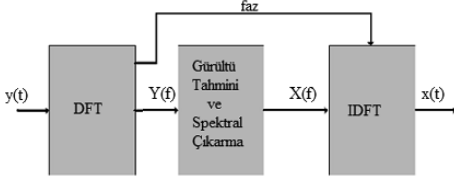
İnsanların en temel haberleşme şekli olan konuşma ile iletişimde, sokakların kalabalığı, arabalar, fabrika makineleri gibi çevresel faktörlerden kaynaklanan bir çok gürültü ses sinyallerinde bozulmaya yol açar. Bu durum da konuşmanın anlaşılabilirliğini dolayısıyla iletişimi olumsuz yönde etkiler.

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. Bu sistemler günümüzde de hala çalışılmaya devam edilen bir araştırma alanı olarak önemini korumaktadır. (Commins, 2005)

## **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.

Şekil 1’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 1. Spektral çıkarmanın genel yapısı

Zaman domeninde verilen gürültülü sinyal Eşitlik (1)’de verildiği gibidir. (Boll, 1979)

$$y(m) = x(m) + n(m) \quad (1)$$

$y(m)$ : gürültülü sinyal  
 $x(m)$ : temiz sinyal  
 $n(m)$ : gürültü

Eşitlik(1)’in frekans domenindeki gösterimi Eşitlik(2)’deki gibi olur.

$$Y(f) = X(f) + N(f) \quad (2)$$

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

Pencerelemiş sinyal Eşitlik(3)’teki gibi gösterilebilir;

$$\begin{aligned} y_w(m) &= w(m)y(m) \\ &= w(m)[x(m) + n(m)] \\ &= x_w(m) + n_w(m) \end{aligned} \quad (3)$$

Yine aynı şekilde pencereleme işlemi frekans domeninde de gösterilebilir;

$$\begin{aligned} Y(f) &= W(f) * Y(f) \\ &= X_w(f) + N_w(f) \end{aligned} \quad (4)$$

Eşitlik (4)’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)

$$|X(f)|^b = |Y(f)|^b - \alpha |N(f)|^b \quad (5)$$

Burada;  $|X(f)|^b$  orijinal sinyal spektrumunun tahmin edilen değerini,  $|N(f)|^b$  ise zaman-ortalama gürültü spektrumunu ifade etmektedir. Genlik spektral çıkarma için  $b=1$ , güç spektral çıkarma için  $b=2$  seçilir ve  $\alpha$  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)

$$|N(f)|^b = \frac{1}{K} \sum_{i=0}^{K-1} |N_i(f)|^b \quad (6)$$

Eşitlik (6)’da geçen  $|N_i(f)|^b$ ,  $i$ . gürültü bölümünün spektrumunu ifade etmektedir.  $K$  değeri ise sadece gürültünün mevcut olduğu bir periyotta mevcut olan bölüm sayısını göstermektedir. Ortalama gürültü spektrumu birinci dereceden alçak geçiren dijital bir filtrenin çıkışı olarak kabul edilebilir,

$$|N_i(f)|^b = p |N_{i-1}(f)|^b + (1-p) |N_i(f)|^b \quad (7)$$

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  $\theta_r(k)$  ile birleştirilerek, zaman domenine ters ayrık Fourier Transformu ile eşitlik (8)’deki gibi dönüştürülür.

$$x(m) = \sum_{k=0}^{N-1} |x(k)| e^{j\theta_s(k)} e^{-j\frac{2\pi}{N}km} \quad (8)$$

### Wiener filtresi

En küçük karesel hata (Least Square 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. (Vaseghi, 2008)

LSE filtreleri ile amaçlanan, çıkışta elde edilen sinyal ile istenen sinyal arasındaki ortalama uzaklığın karesinin minimize edilmesidir. Wiener filtresinin çözümü (filtre katsayılarının hesaplanabilmesi) için, girişe uygulanan sinyalin otokorelasyon fonksiyonunun ve bu sinyal ile çıkışta elde etmek istediğimiz sinyalin çapraz korelasyon fonksiyonunun tahmin edilmesi gerekmektedir.

Wiener filtresinin giriş-çıkış ilişkisini gösteren eşitlik (9)'da verilmiştir. (Diniz, 1997)

$$x(m) = \sum_{k=0}^{P-1} w_k y(m-k) \quad (9)$$

$m$  : zaman indeksini,

$y^T = [y(m), \dots, y(m-P-1)]$ : giriş sinyalini,

$x(m)$  : filtre çıkışını,

$w^T = [w_0, w_1, \dots, w_{P-1}]$ : Wiener filtresinin katsayı vektörünü ifade etmektedir.

Wiener filtresine ait hata sinyali  $e(m)$  olarak tanımlanırsa, bu hata sinyali, istenen sinyal ile çıkışta elde edilen sinyal arasındaki fark olarak ifade edilebilir.

$$\begin{aligned} e(m) &= x(m) - x(m) \\ &= x(m) - w^T y \end{aligned} \quad (10)$$

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

$$\begin{aligned} E[e^2(m)] &= E[(x(m) - w^T y)^2] \quad (11) \\ &= E[x^2(m)] - 2w^T E[yx(m)] + w^T E[yy^T] w \\ &= r_{xx}(0) - 2w^T r_{yx} + w^T R_{yy} w \end{aligned}$$

En küçük karesel hataya sahip Wiener filtresi, Eşitlik (12)'nin sıfıra eşitlenmesi ile Eşitlik (14)'teki gibi elde edilir. (Diniz, 1997)

$$\frac{\partial}{\partial w} E[e^2(m)] = -2r_{yx} + 2w^T R_{yy} \quad (12)$$

$$R_{yy} w = r_{yx} \quad (13)$$

$$w = R_{yy}^{-1} r_{yx} \quad (14)$$

$R_{yy}$  : Giriş sinyalinin otokorelasyonu

$r_{yx}$  : Giriş sinyali ile istenen sinyalin çapraz korelasyonu

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

$$X(f) = W(f)Y(f) \quad (15)$$

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

$$E[e(f)^2] = E[(X(f) - W(f)Y(f))^* (X(f) - W(f)Y(f))] \quad (16)$$

Karesel hatanın minimum olması için Eşitlik(17) sıfıra eşitlenir.

$$\frac{\partial E[e(f)^2]}{\partial W(f)} = 2W(f)P_{YY}(f) - 2P_{XY}(f) \quad (17)$$

$$W(f) = \frac{P_{XY}(f)}{P_{YY}(f)} \quad (18)$$

$P_{YY}$  : Giriş sinyalinin güç spektrumu

$P_{XY}$ : Giriş sinyali ile istenen sinyalin çapraz güç spektrumu

### Kalman filtresi

Kalman filtresi, adını aldığı Rudolph E. Kalman tarafından, ayrı 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)

Durum uzay modeli ile gösterilen bir dinamik sistemde, modelin önceki bilgileriyle birlikte giriş ve çıkış bilgilerinden sistemin durumu tahmin edilebilir. Gözlemlenebilirlik teorisi, karar verilen bir bakış açısı temelinde, sistemin durum tahmini için izlenecek bir yoldur. Eğer sistemin stokastik ve rastgele 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 etmek için çok güçlü ve yeteneklidir. Kalman filtresi, ortalama kestirim hatasının karesini minimize etmeye yönelik bir hesaplama mantığı ile duruma bakar. (Ankışhan, 2007)

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

$$x(n+1) = F(n+1, n)x(n) + D(n)v(n) \quad (19)$$

$$y(n) = C(n)x(n) + w(n) \quad (20)$$

şeklinde bir modeldir. (Deviren, 2000)

Burada;

$\mathbf{x}(n)$ ,  $n$  anındaki durum vektörünü

$\mathbf{F}(n)$ , durum geçiş matrisini

$\mathbf{v}(n)$ , sıfır ortalamalı olan beyaz Gaussian süreç gürültüsünü

$\mathbf{y}(n)$ , gürültülü gözlemi

$\mathbf{w}(n)$ , sıfır ortalamalı olan ölçüm gürültüsünü

$\mathbf{D}(n)$ , giriş vektörünü

$\mathbf{C}(n)$ , gözlem vektörünü

ifade etmektedir.

Süreç gürültüsü  $v(n)$  ve ölçüm gürültüsü  $w(n)$  sıfır ortalamalı oldukları için beklenen değerleri sıfıra eşittir. Ayrıca bu iki gürültü birbiriyle tamamen bağımsız olup süreç gürültüsünün kovaryansı  $Q(n)$ , ölçüm gürültüsünün kovaryansı  $R(n)$  olarak tanımlanır.

$$E[v(n)] = 0, E[w(n)] = 0 \quad (21)$$

$$E[v(n)w^T(k)] = E[v(n)]E[w^T(k)] = 0 \quad (22)$$

$$E[v(n)v^T(n)] = Q(n) \quad (23)$$

$$E[w(n)w^T(n)] = R(n) \quad (24)$$

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

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

$$s(n) = -\sum_{i=1}^p a_i s(n-i) + v(n) \quad (25)$$

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

$$y(n) = s(n) + w(n) \quad (26)$$

Kalman filtresi durum tahmini için özyinelemeli bir prosedür yürütür. Buradaki amaç orijinal ses sinyalinin tahminini gerçekleştirmektir. Bunun için, ses sinyalinin  $q(\geq p)$  tane önceki değeri alınarak  $\underline{x}(k)$  durum vektörü oluşturulabilir. (Najim, 2008)

$$\underline{x}(n) = [s(n-q+1), \dots, s(n)]^T \quad (27)$$

Eşitlik (26) ve Eşitlik (27)'de ki durum uzay gösterimi aşağıdaki gibi düzenlenebilir;

$$\underline{x}(n) = F\underline{x}(n-1) + Dv(n) \quad (28)$$

$$y(n) = C\underline{x}(n) + w(n) \quad (29)$$

$F$ , geçiş matrisi olarak adlandırılır ve  $q \times q$  boyutundadır; (Najim, 2008)

$$F = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 & 1 \\ 0 & \dots & 0 & -a_p & -a_{p-1} & \dots & \dots & -a_1 \end{bmatrix} \quad (30)$$

$\{a_i\}_{i \in \{1, \dots, p\}}$ , orijinal ses sinyaline ait AR katsayılarını ifade etmektedir.  $D$  ve  $C$  sırasıyla giriş ve gözlem vektörleri olarak aşağıdaki gibi tanımlanır. (Najim, 2008)

$$D = C^T = \begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix}^T \quad (31)$$

Kalman filtresi ile ses sinyallerini iyileştirilme süreci genellikle pencereleme işlemi ile başlar. Ses sinyalleri yaklaşık 20–30ms'lik çerçevelere bölünerek iyileştirme algoritması her çerçeveye ayrı ayrı uygulanır. Pencereleme işlemi ile amaçlanan sinyalin durağanlaştırılmasıdır.

İşlem kolaylığı sağlaması açısından tüm bölütler eşit uzunlukta ve pencereleme fonksiyonu olarak ta %50 örtüşme ile Hanning veya Hamming penceresi alınabilir. Hanning veya Hamming Penceresi fonksiyonu aşağıdaki gibi tanımlanır;

$$w(n) = a - (1-a) \cos\left(2\pi \frac{n}{N}\right), \quad n = 0, \dots, N-1 \quad (32)$$

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

Her bir bölütteki sinyalin tahmin edilebilmesi için öncelikle ölçüm gürültüsünün varyansı  $R$ 'nin ve orijinal sinyale ait AR katsayılarının bilinmesi gerekmektedir.

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

$$x(n/n-1) = Fx(n-1/n-1) \quad (33)$$

$$P(n/n-1) = FP(n-1/n-1)F^T + \sigma_v^2 \quad (34)$$

$$K(n) = P(n/n-1)C^T \left( CP(n/n-1)C^T + R \right)^{-1} \quad (35)$$

$$x(n/n) = x(n/n-1) + K(n)(y(n) - Cx(n/n-1)) \quad (36)$$

$$P(n/n) = [I - K(n)C]P(n/n-1) \quad (37)$$

Böylece iyileştirilmiş sinyal Eşitlik(38)'deki gibi elde edilir;

$$\hat{s}(n) = Cx(n/n) \quad (38)$$

$n$  zaman değişkeni olmak üzere;

$x(n/n)$ : Durum vektörünün tahminini,

$x(n/n-1)$ : Durum vektörünün kestirimini,

$P(n/n-1)$ : Kestirim hatasının kovaryansını,

$P(n/n)$ : Tahmin hatasının kovaryansını,

$K(n)$ : Kalman filtresinin kazancını

ifade etmektedir.

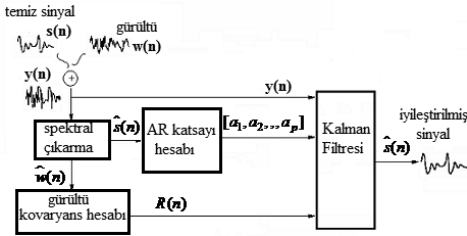
## Uygulama

Bu çalışmada kullanılan ses dosyası, Speech Enhancement Assessment Resource (SpEAR) veritabanından, gürültü örnekleri ise Noisex-92 veritabanından, temiz konuşma sinyali ile aynı frekansta (16 kHz) örneklenmiş olarak alınmıştır. Üç farklı gürültü tipi farklı Sinyal/Gürültü oranlarında (SNR) orijinal ses sinyaline eklenerek, gürültülü sinyaller 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 (33) – (37) 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 1'de verilen yöntem ile aşılmaya çalışılmıştır.



Şekil 2. Uygulanan kalman filtresine ait blok diyagram

Şekil 2'deki spektral çıkarma bloğunda önce, Martin'in(2001) önerdiği minimum istatistiklere dayalı, gürültünün güç spektrum tahmini yöntemi ile gürültü karakteristiği belirlenmiştir. Daha sonra spektral çıkarma ile gürültülü sinyal iyileştirilmiştir. Temiz sinyalin katsayılarını tam olarak vermese de iyileştirilmiş sinyal kullanılarak katsayılar hesaplanmıştır. Gürültülü sinyalden iyileştirilmiş sinyal çıkarılarak da gürültüye ait kovaryans matrisi hesaplanmıştır.

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

## Benzetim sonuçları

SpEAR veritabanından alınan ses dosyalarından biri "Good service should be rewarded by big tips" cümlesini içermektedir. Bu cümleye, fabrika gürültüsü, f16 gürültüsü ve beyaz gürültü olmak üzere üç ayrı gürültü tipi farklı SNR değerlerinde eklenmiştir. 0dB SNR ile bozulmuş cümleye uygulanan yöntemlere ait benzetim sonuçları Şekil 3 – Şekil 6'da gösterilmiştir. Şekil açıklamalarında geçen,

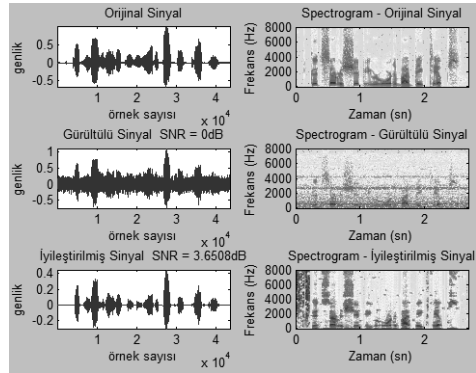
*Wiener*, Plapous ve ark. (2006) tarafından önerilen yöntemi,

*Spektral çıkarma*, Martin (2001) tarafından önerilen yöntemi,

*Kalman (temiz)*, Temiz sinyalin katsayıları kullanılarak uygulanan Kalman filtresini,

*Kalman (spektral)*, Bu çalışmada uygulanan birleştirilmiş Kalman filtresini ifade etmektedir.

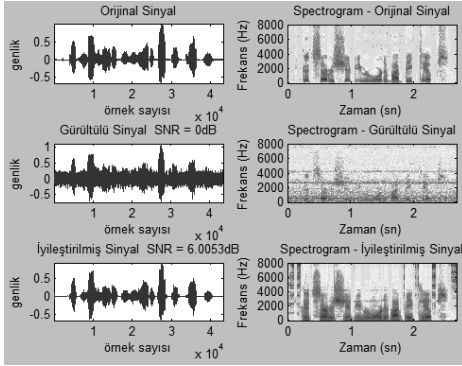
Tüm gürültü tipleri için elde edilen SNR değerleri ise Tablo 1'de verilmiştir.



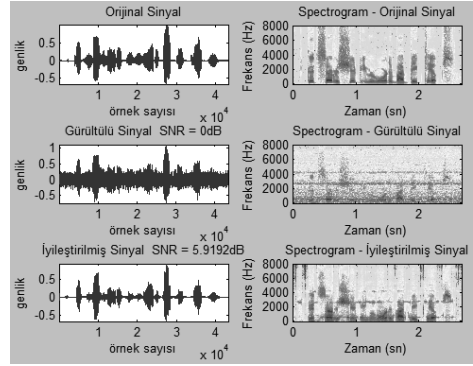
Şekil 3. F16 gürültüsü ve 0dB SNR ile bozulmuş konuşmaya ait orijinal, gürültülü ve Wiener filtresi ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



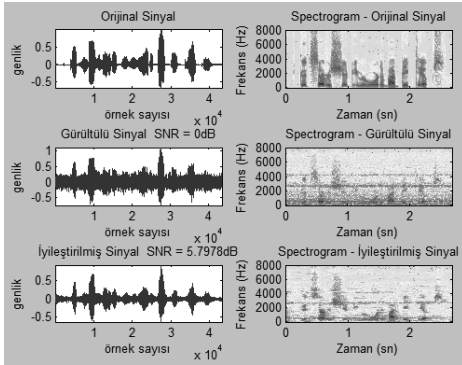
## Kalman filtresi ile ses sinyallerinin iyileştirilmesi



Şekil 4. F16 gürültüsü ve 0dB SNR ile bozulmuş konuşmaya ait orijinal, gürültülü ve Kalman (temiz) filtresi ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



Şekil 6. F16 gürültüsü ve 0dB SNR ile bozulmuş konuşmaya ait orijinal, gürültülü ve Kalman (spektral) filtresi ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları



Şekil 5. F16 gürültüsü ve 0dB SNR ile bozulmuş konuşmaya ait orijinal, gürültülü ve spektral çıkarma ile iyileştirilmiş sinyalin zaman-genlik grafikleri ve spektrogramları

## Sonuçlar

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.

Tablo 1. Farklı gürültü tipleri için 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

Ancak sadece gürültülü sinyalin mevcut olmasından ötürü, temiz sinyale ait katsayıların ve gürültü varyansının biliniyor olduğu varsayımı pratikte uygulanabilir bir yaklaşım değildir. Dolayısıyla Spektral çıkarma ile iyileştirilen sinyal kullanılarak AR katsayıları ve gürültülü sinyalden iyileştirilen sinyal çıkarılarak gürültü varyansı hesaplanmıştır. Daha sonra gürültülü sinyale, belirlenen parametreler kullanılarak Kalman filtresi uygulanmıştır.

Bu yöntemle, 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 AR katsayılarının, Kalman filtresinin performansını çok büyük bir ölçüde etkilediğinin bir göstergesi olarak görülebilir.

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 bastırıldığı gözlemlenmiştir.

## Kaynaklar

Akaike, H., (1974). A New Look at Statistical Model Identification, *IEEE Trans. Automatic Control*, AC-19, 716–723.

Ankışhan, H., (2007). Gürültülü Ses sinyali İyileştirilmesine İkili Kalman Filtre Yaklaşımı, *Doktora Tezi*, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.

Benesty, J., Makino, S. ve Chen, J., (ed.), (2005). *Speech Enhancement*, Springer-Verlag, Berlin.

Berouti, M., Schwartz, R. ve Makhoul, J., (1979). Enhancement of speech corrupted by acoustic noise in Speech Enhancement, *ICASSP-79*, Washington, 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.

Boll, S.F., (1979). Suppression of Acoustic Noise in Speech Using Spectral Subtraction, *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-27, 113-120

Cecelioğlu, S., (2010). Tek Kanallı Toplamsal Gürültülü Konuşma Sinyali İyileştirme, *Yüksek Lisans Tezi*, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.

Commins, B., (2005). Signal Subspace Speech Enhancement with Adaptive Noise Estimation, *Yüksek Lisans Tezi*, National University of Ireland, Galway.

Deviren, M., (2000). Enhancement of Noisy Speech Utilizing the Kalman Filter, *Yüksek Lisans Tezi*, Orta Doğu Teknik Üniversitesi, Ankara.

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*, 2, 1513-1516.

Ephraim, Y. ve 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, 1109-1121.

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

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

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

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

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

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

Grewal, M.S. ve Andrews, A.P., (2001). *Kalman Filtering Theory and Practice using Matlab*, 410, John Wiley&Sons, Inc..

Güler, İ., Kıymık, M.K., Akın, M. ve Alkan, A., (2001). AR spectral analysis of EEG signal by using maximum likelihood estimation, *Comput. Biol. Med.* 31, 441–450.

- Hoeldrich R. ve Lorber M., (1998). Real-Time Broadband Noise Reduction, *Proceedings of International Computer Music Conference*.
- Kalman, R.E., (1949). A New Approach to Linear Filtering and Prediction Problems, *Trans. of the ASME-Journal of Basic Engineering*, **26**, 35-45.
- Kalman, R.E. ve Bucy R.S., (1961). New Results in Linear Filtering and Prediction Theory, *Trans. ASME J. Basic Eng.*, **83**, 95-108.
- Kırççek, Y., (2007). Doğrusal Öngörü ile Konuşma İşareti Kodlayıcısı Tasarımı, *Yüksek Lisans Tezi*, Yıldız Teknik Üniversitesi Fen Bil. Ens., İstanbul.
- Lee, K.Y., Lee, B.-G., Song, I. ve Ann, S., (1992). Robust Estimation AR Parameters and its Application for Speech Enhancement, *Proc. IEEE ICASSP*, 309-312.
- Lim, J.S. ve Oppenheim, A.V., (1978). All-pole Modelling of Degraded Speech, *IEEE Trans. on Acoustic, Speech and Signal Processing*, ASSP-**26** (3), 197-210.
- Ma, N., Bouchard, M. ve Goubran, R.A., (2004). Perceptual Kalman Filtering for Speech Enhancement in Colored Noise, *Proc. IEEE ICASSP*, 717-720
- Mai, Q., He, D., Hou, Y. ve Huang, Z., (2011). A Fast Adaptive Kalman Filtering Algorithm for Speech Enhancement, *IEEE Int. Con. Automation Science and Eng*, 327-332.
- Martin, R., (1994). Spectral Subtraction Based on Minimum Statistics, *Proc. 7th European Signal Processing Conf. EUSIPCO-94*, 1182-1185.
- Martin, R., (2001). Noise power spectral density estimation based on optimal smoothing and minimum statistics, *IEEE Trans. Speech Audio Process.*, **9**(5), 504-512.
- Mehra, R. K., (1970). On the Identification of Variances and Adaptive Kalman Filtering, *IEEE Trans. on Automatic Control*, AC-**15**, 175-184.
- Najim, M., (2008). *Modelling, Estimation and Optimal Filtering*, 408, John Wiley&Sons,Inc.
- Noisex-92 database,  
[http://spib.rice.edu/spib/select\\_noise.html](http://spib.rice.edu/spib/select_noise.html)  
Signal Processing Information Base.
- Paliwal, K. K. ve Basu A., (1987). A Speech Enhancement Method Based on Kalman Filtering, *IEEE-ICASSP '87*, Dallas, 177- 180.
- Plapous C., Marro C. ve Scalart, P., (2005). Speech enhancement using harmonic regeneration. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process*, 157-160
- Plapous C., Marro C. ve Scalart, P., (2006). Improved signal-to-noise ratio estimation for speech enhancement, *IEEE Trans. Audio, Speech, Lang. Processing*, **14** (6), 2098-2108.
- Pınarcı, E., (2007). İki boyutlu Kalman filtresinin yersel lazer tarama verisine uygulanması, *Yüksek Lisans Tezi*, Gebze Yüksek Teknoloji Enstitüsü, Gebze.
- Scalart P. ve Filho J. V., (1996). Speech enhancement based on a priori signal to noise estimation, *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process*, 629-632.
- So, S. ve Paliwal, K., (2008). A long state vector Kalman filter for speech enhancement, *Proc. Int. Conf. Spoken Language Processing (INTERSPEECH-ICSLP)*, 391-394.
- Şahin, H., (2007). Tek kanallı sistemler için geliştirilmiş ses iyileştirme yöntemi, *Yüksek Lisans Tezi*, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Vaseghi, S.V., (2008). *Advanced Digital Signal Processing and Noise Reduction*, 532, John Wiley&Sons, Ltd. ISBN: 978-0-470-75406-1
- Wan E., Nelson A. ve Peterson, R. Speech Enhancement Assessment Resource (SpEAR) Database, [http://www.cslu.ogi.edu/nsel/data/SpEAR\\_database.html](http://www.cslu.ogi.edu/nsel/data/SpEAR_database.html), *Oregon Graduate Institute of Science and Technology*.
- Weinstein, E., Oppenheim, A.V. ve 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. ve Goodlin, R.C., (1975). Adaptive Noise Cancelling:Principals and Applications, *Proc. IEEE*, 1692-1716.
- Wiener, N., (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, MIT Press, Cambridge, MA; Wiley, New York; Chapman&Hall, London.