

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK - ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

SESİN KOD UYARMALI DOĞRUSAL ÖNGÖRÜMLÜ KODLANMASININ
İNCELENMESİ

83307

Elektronik Müh. Salim KAHVECİ

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde

“Elektronik Yüksek Mühendisi”

Ünvanı Verilmesi için Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 03.06.1999

Tezin Savunma Tarihi : 06.07.1999

Tez Danışmanı : Yrd. Doç. Dr. Ali GANGAL

Jüri Üyesi : Yrd. Doç. Dr. İsmail KAYA

Jüri Üyesi : Yrd. Doç. Dr. Cemal KÖSE

Enstitü Müdürü : Prof. Dr. Asım KADIOĞLU

A. Kadioğlu.

TC. YÜKSEKÖĞRETİM KURULU
DOKÜMANİZASYON MERKEZİ

Trabzon 1999

ÖNSÖZ

Bu çalışmada, ses kodlama teknikleri üzerinde ayrıntılı olarak inceleme yapılmıştır.

Yüksek lisans tezi danışmanlığımı üstlenerek, ileride daha çok üzerinde araştırma yapmak istediğim bir konu üzerinde çalışmamı sağlayan, çalışmalarım sırasında bilgisini ve yardımlarını esirgemeyen sayın hocam Yrd. Doç. Dr. ALİ GANGAL'a teşekkür etmeyi bir borç bilirim.

Ayrıca yüksek lisans çalışmalarım sırasında bana gerek maddi gerekse manevi desteğini esirgemeyen aileme ve bütün arkadaşlarıma teşekkür ederim.

Trabzon, Haziran 1999

Salim KAHVECİ

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	II
İÇİNDEKİLER	III
ÖZET	V
SUMMARY	VI
ŞEKİLLER DİZİNİ.....	VII
TABLolar DİZİNİ.....	IX
SEMBOLLER DİZİNİ	X
1. GENEL BİLGİLER.....	1
1.1. Giriş.....	1
1.2. Sayısal Ses İşleme	6
1.2.1. Ses İşleme İçin Zaman-Domeni Yöntemleri	8
1.2.2. Kısa Zaman Fourier Analizi.....	10
1.3. Paralel İşlem Yaklaşımını Kullanarak Perde Periyodunun Kestirimi.....	13
1.4. Ses Dalgasının Sayısal Gösterimi	18
1.4.1. Seviyelendirme.....	19
1.5. Formant Kestirimi	26
1.6. Sesin Doğrusal Öngörümlü Kodlanması	28
1.6.1. Doğrusal Öngörüm Analizin Ana Prensipleri	28
1.6.2. Özilişki Yöntemi	31
1.6.3. Kovaryans Yöntemi	33
1.7. Değişik Ses Parametreleri Arasındaki İlişkiler	34
1.8. Doğrusal Öngörümlü Parametrelerden Sesin Sentezi	36
2. YAPILAN ÇALIŞMALAR.....	38
2.1. Sabit Kestirimli Dalgabiçimli Kodlama [15]	39
2.2. Perde Uyarmalı Doğrusal Öngörümlü Kodlayıcı [15].....	44
2.3. Adaptif Öngörümlü Dalgabiçimli Kodlama [15].....	47
2.4. Uyarma Modeli [15].....	49

2.5.	Hata Ağırlıklandırma.....	50
2.6.	Analiz-Sentez İşlemi	51
2.7.	Uzun Dönem Kestiricisi	52
2.8.	Çoklu Darbe Uyarmalı DÖK Uygulaması [16].....	53
2.9.	Düzenli Darbe Uyarmalı Doğrusal Öngörümlü Kodlayıcı.....	55
2.10.	Kod Uyarmalı Doğrusal Öngörümlü Kodlayıcı Uygulaması.....	56
2.10.1.	KUDÖK Alıcısı.....	57
2.10.2.	KUDÖK Vericisi [20]	59
2.10.2.1.	Doğrusal Öngörümlü Analizi	59
2.10.2.2.	Adaptif Kod Tablosu Arama.....	60
2.10.2.3.	Stokastik Kod Tablosu Arama.....	62
2.10.3.	Hata Düzeltme	62
2.10.4.	Kod Tablosu Arama Yöntemleri.....	62
3.	BULGULAR.....	68
4.	İRDELEME	72
5.	SONUÇLAR	73
6.	ÖNERİLER.....	74
7.	KAYNAKLAR.....	75
8.	ÖZGEÇMİŞ	77

ÖZET

Sayısal kodlama teknikleri kullanarak ses işaretinin düşük kapasiteli kanallar üzerinden iletimi ve az bellek harcayarak depolanması mümkün olmaktadır. Bu tezde, ses kodlama sistemlerinin genel yapısıyla birlikte Kod Uyarmalı Doğrusal Öngörümlü Ses Kodlayıcılar incelenmiştir.

Ses işaretlerinin kodlanması için en genel anlamda analiz-sentez işlemi yapılır. Analiz işlemi verici tarafta yapılırken sentez işlemi alıcı tarafta gerçekleştirilir. Analiz ve sentez işlemlerinde Doğrusal Öngörümlü Kodlama tekniği kullanılır. Analiz işleminde ilgili ses işareti için stokastik ve adaptif kod tablolarına ilişkin indisler ve kazançlar arama algoritmasıyla belirlenerek alıcıya gönderilir. Alıcı tarafta bu indisler ve kazançlar kullanılarak aynı stokastik ve adaptif kod tablolarından gönderilen ses işareti sentezlenir.

Yapılan çalışmada, ses kodlama performansı üzerinde önemli payı olan kestirici tasarımı, pencere fonksiyonlarının ve perde periyodunun belirlenmesi gibi konular ayrıntılı olarak incelenmiştir. Değişik kodlama teknikleri işlem miktarı, veri hızı, işaret/gürültü oranı ve işitsel kalite bakımından karşılaştırılmıştır.

Elde edilen sonuçlardan, kod uyarmalı doğrusal öngörümlü ses kodlayıcıların üstün performansa sahip olduğu görülmüştür.

Anahtar Kelimeler: Ses İşaretleri, Ses Kodlama, Doğrusal Öngörüm Katsayıları, Perde Periyodu, Doğrusal Kestirim.

SUMMARY

Studying of the Code-Excited Linear Prediction of Speech Coding

It is possible that transmission of speech signal over the channels which have low capacity by using digital coding techniques and its storage by using a low memory. In this thesis, it was examined a general structure of the speech coding systems with code-excited linear predictive coders.

Analysis-synthesis process is generally done for coding of speech signals. While analysis process is implemented at the transmitter, synthesis process is implemented at the receiver. Linear prediction coding technique is used in the analysis-synthesis process. At analysis process, indexes and gains of stochastic and adaptive code tables for the related speech signal that are found with search algorithms. Synthesis of the speech signals sent from same stochastic and adaptive code tables is done by using these indexes and gains at the receiver.

In this working subjects as predictor conception, determination of window functions and pitch period, all of which effective on performing of the speech coding, were investigated with details. Various coding techniques were compared from the point of view amount of process, data rate, signal/noise ration and audial quality.

Obtained results show that code-excited linear prediction speech coders have an excellent performance.

Key Words: Speech Signals, Speech Coding, Linear Predictive Coefficients, Pitch Period, Linear Prediction.

ŞEKİLLER DİZİNİ

Sayfa No

Şekil 1. Ses kodlayıcı blok diyagramı	1
Şekil 2. Ses kodlama sisteminin elemanları	3
Şekil 3. Ses işaretinin gösterimi	6
Şekil 4. Sayısal ses işleme teknikleri	7
Şekil 5. Kısa-zaman enerji gösterimi	9
Şekil 6. Kısa-zaman analiz prensibinin genel gösterimi	10
Şekil 7. Lineer süzgeçnin kısa zaman analiz yorumu, a) Kompleks yapı, b) Reel çalışması	12
Şekil 8. Zaman-domeni perde kestiricisi paralel işlemin blok diyagramı	15
Şekil 9. Sinüzoidal giriş ile pik ve çukurlardan üretilen ilişki darbe dizileri	17
Şekil 10. Ara ve ikinci harmonikli giriş ile pik ve çukurlardan üretilen ilişki dizini	17
Şekil 11. Zaman-domeni perde kestiricisinin her perde periyodunun temel işlemi	18
Şekil 12. Sayısal ses dalgasının genel blok diyagramı	19
Şekil 13. Seviyelendirme ve kodlama işlemi, a) Kodlayıcı, b) Kod çözücü	20
Şekil 14. Feed-forward adaptif PCM'nin blok diyagramı	23
Şekil 15. Feedback adaptasyon PCM'nin blok diyagramı	23
Şekil 16. Ses üretimi için sayısal model	26
Şekil 17. Ses üretim modelinin blok diyagramı	29
Şekil 18. Doğrusal öngörümlü sentezleyicinin blok diyagramı	37
Şekil 19. Ses kodlama çeşitleri [15]	38
Şekil 20. Analiz-sentez doğrusal öngörümlü kodlayıcıda kullanılan analiz işleminin blok diyagramı	39
Şekil 21. Fark PCM'in blok diyagramı	40
Şekil 22. Delta modülatörünün blok diyagramı	42
Şekil 23. Perde uyarmalı DÖK vericisi	45
Şekil 24. Perde uyarmalı DÖK alıcısı	45
Şekil 25. Artık uyarmalı LPC'nin blok diyagramı	49

Şekil 26. Kaynak analiz modelin blok diyagramı	51
Şekil 27. Uzun dönem kestiricili bir DÖK sentezleyicinin blok diyagramı	53
Şekil 28. Doğrusal öngörümlü kodlayıcı, a) Verici, b) Alıcı	54
Şekil 29. Federal standart 1016 kod uyarmalı kodlayıcı alıcısı [20].	58
Şekil 30. Federal standart 1016 KUDÖK vericisi [20].	59
Şekil 31. KUDÖK sentezleyicisi	64
Şekil 32. KUDÖK analizi	64
Şekil 33. Kod bloğu arama	66
Şekil 34. KUDÖK giriş çıkış ses dosyaları, a) “ia.wav” giriş ses dosyası, b) “oa.wav” çıkış ses dosyası	68
Şekil 34. c) “iceloffma.wav” giriş ses dosyası, d) “oceloffma.wav” çıkış ses dosyası ...	69



TABLULAR DİZİNİ

	<u>Sayfa No</u>
Tablo 1. Ses kodlama tekniklerinin özellikleri [1].	4
Tablo 2. Formant kodlayıcılar için hesaplanan bazı değerler [9].	27
Tablo 3. ADM'nin parametreleri	43
Tablo 4. Çoklu darbe uyarmalı kodlayıcı analiz-sentezleyicisinin parametreleri.....	54
Tablo 5. Düzenli darbe uyarmalı kodlayıcı analiz sentezleyicisinin parametreleri.....	55
Tablo 6. Kod uyarmalı kodlayıcının bazı karakteristikleri.....	62
Tablo 7. KUDÖK analiz ve sentezleyicinin tipik parametreleri [15].....	67
Tablo 8. Ses kodlama tekniklerinin performanslarının karşılaştırılması [21].	67
Tablo 9. Çoklu darbe uyarmalı kodlayıcı için İGO ve SİGO değerleri	69
Tablo 10. Düzenli darbe uyarmalı kodlayıcı için İGO ve SİGO değerleri.....	70
Tablo 11. Kod uyarmalı kodlayıcı için İGO ve SİGO değerleri	70

SEMBOLLER DİZİNİ

A/D, D/A	: Analog dijital, dijital analog dönüştürücü.
ADM	: Adaptive delta modülasyon.
APC	: Adaptive predictive coding.
E	: Toplam enerji.
X(n)	: Giriş kaynağı.
T[]	: Doğrusal veya doğrusal olmayan dönüşüm.
W(n)	: Pencere fonksiyonu.
X _a	: Giriş örnek işareti.
X'	: Sonlu örnekli kaynak.
C(n)	: Kod kelimesi.
Δ	: Kuvantalama adım sayısı.
V(z)	: Sesli ses bölgesi transfer fonksiyonu.
F _k	: Formant frekanslar.
FEC	: Forward error correction.
S(z)	: Filtre çıkışı.
U(z)	: Filtre girişi.
G	: Sistem kazancı.
a _k	: Doğrusal öngörüm katsayıları.
p	: Kestirici dercesi.
A(z)	: Kısa-dönem kestirici filtresi.
B(z)	: Uzun-dönem kestirici filtresi.
f(n)	: Çoklu darbe uyarma kaynağı.
δ(n)	: Birim vuruş işareti.
β _k	: Ağırlık katsayıları.
H(n)	: Filtrenin birim vuruş cevabı.
W _f	: Ağırlıklı filtre transfer fonksiyonu.
i _a , i _s	: Adaptif ve stokastik indeksler.

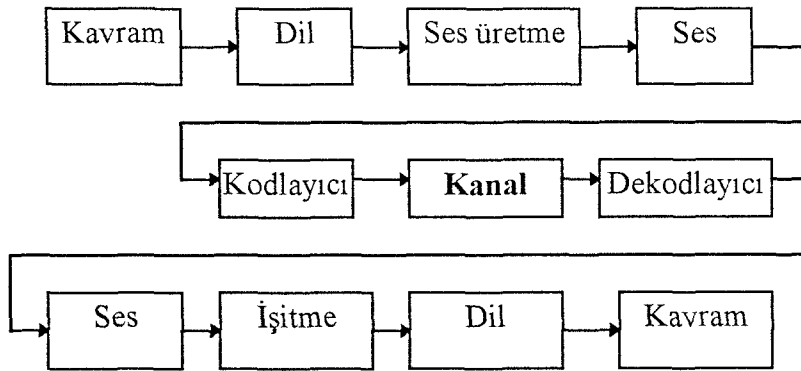
g_a, g_s	: Adaptif ve stokastik kod kelimeleri kazancı.
R	: Normalizeli çapraz ilişki.
KDK	: Kısa dönem kestirici.
UDK	: Uzun dönem kestirici.
ÇSÇ	: Çizgi spektrum çiftleri.
PCM	: Pulse code modülasyon.
LPC	: Linear predictive coding.
DPCM	: Differential pulse code modulation.
MPELPC	: Multipulse-excited linear prediction coding.
RPELPC	: Regular pulse-excited linear prediction coding.
CELPC	: Code-excited linear predictive coding.
KUDÖK	: Kod uyarmalı doğrusal öngörümlü kodlama.
MSPE	: Mean square predictive error.
d.d.	: Diğer durumlar.
AGS	: Alçak geçirgen süzgeç.
DÖK	: Doğrusal öngörümlü süzgeç.
İGO	: İşaret gürültü oranı.
VQ	: Vektör kuvantalama.
i, j, k	: Alt ve üst indisler.
n, m	: Örnek noktası.
N	: Alınan toplam örnek sayısı.

1. GENEL BİLGİLER

1.1. Giriş

Ses iletim sistemlerinde, ses işareti pek çok yolla işlenebilir, iletilebilir ve depolanabilmektedir. Genelde ses iletim sistemlerinde iki büyük kavram vardır. Biri ses işaretindeki bilginin aynen korunabilmesidir. Diğeri ise ses işaretinin gösterildiği formun, sesteki bilgiyi önemli sayılabilecek derecede bozmadan orijinal sesi ifade etmesidir. Ses işaretlerinin depolanması, kodlanması, iletilmesi, tanınması, analizi, sentezi ve daha birçok uygulamalarda işaret işleme yöntemleri esas rol oynamaktadır.

Şekil 1, ses kodlama sistemini kullanan ses iletim işleminin basit bir blok diyagramını göstermektedir. Bu genel sistemin amacı bir kavramı bir yerdeki konuşmacıdan başka bir konuşmacıya iletim kanalıyla bozulmadan göndermektir. Bu işlemde, kavram veya fikir öncelikle cümleye çevrilir. Ses kodlama sisteminin girişindeki ses, mikrofon yardımıyla hava basıncının oluşturduğu akustik dalga olarak alınır. Ses kodlayıcı, ses işaretini sayısallaştırır ve sesi sayısal bir bit dizisi halinde gösterir. Ancak bit dizisinin bit hızı iletim kanalının iletim kapasitesiyle uygun olmak zorundadır. Alıcıda, ses dekodlayıcısı sayısal bit dizisini alır ve mümkün olduğu kadar orijinal ses işaretine benzeyen yeni ses işaretini üretir. Bu işaret daha sonra dinleyicinin kulağına hoparlör yardımıyla akustik olarak gönderilir.



Şekil 1. Ses kodlayıcı blok diyagramı

Ses kodlayıcısının amacı, mümkün olan en düşük bit hızında mümkün olan en yüksek ses kalitesini üretmektir. Mümkün olan bit hızının alt sınırı sesteki şekilsel bilgiyle belirlenir. Sesteki mümkün olan bit hızının alt sınırı yaklaşık olarak saniyede 50 bit olarak kestirilen şekilsel bilgi hızıyla bulunur. Ses işaretindeki toplam bilgi hızı, saniyede yaklaşık olarak 400 bit olarak kestirilir. Diğer taraftan, doğrusal A/D bir sistem kullanan yüksek kaliteli sesin sayısal gösterimi için sn.'de 100000 bitten daha fazlasına ihtiyaç duyulmaktadır [1]. Eğer iyi kaliteye yakın bir 400 bitlik ideal bir ses kodlayıcı elde edilebiliyorsa, sesteki enerji, sıfır geçişleri, formant frekansları ve perde periyodu gibi ayrıntılı bilgileri elde etmek mümkün olabilmektedir.

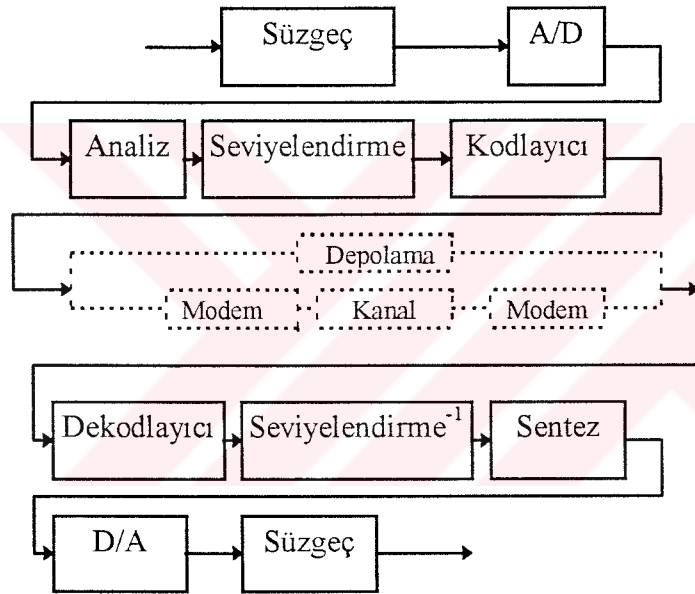
Perde uyarmalı kodlayıcı ve dalgabiçimli uyarmalı kodlayıcı modelleri modern ses kodlayıcılarında kullanılmıştır [1]. Her iki model vokal tırak süzgece sahip olup bu süzgeç yavaşça değişen parametrik olarak kontrol edilebilen doğrusal süzgeçtir. Genişçe kullanılan parametrik süzgeç tüm-kutup yapısına sahip olup, parametreleri doğrusal kestirim analizi kullanılarak elde edilmektedir. Kodlayıcıların bu sınıfı doğrusal öngörümü kodlayıcı olarak adlandırılır. Vokal tırak süzgeç kullanan bütün ses kodlayıcılarında, yapılan işlem alıcıda süzgeci kontrol eden parametreleri çıkarmak için orijinal işareti analiz etmektir. Bu parametreler daha sonra kodlanır ve alıcıya gönderilir. Bunlar, alıcıda çıkış ses işaretini üretmek için kullanılan vokal tırak süzgeç modelini kontrol etmek için kullanılır.

Perde uyarmalı ses modeli uyarma işaretini, ya sesli sesler için yarı periyodik darbeler dizini yada sessiz sesler için beyaz gürültü işareti olarak gösterir. Bu uyarma modelinin parametreleri, sesli sesler için perde periyodu, sesli/sessiz anahtarlama konumu ve uyarma kazancıdır. Perde esaslı uyarma modeli genel olarak uyarma parametreleri için 1000 bitin altındaki bit hızlarında çalışır [2].

Perde uyarmalı modele alternatif bir model dalgabiçimli uyarmalı kodlayıcıdır [1]. Uyarma işareti, kodlanmış bir uyarma işaretiyle uyarılan perde kestiricinin çıkışı olarak alınır. Bu modelde perde, perde kestirici tarafından modellenir. Uyarma işareti perde kestirici tarafından modellenmez. Çoğu ses kodlayıcılar bu genel modele uygundur. Bunlar adaptif öngörümü kodlayıcılar, artık uyarmalı kodlayıcılar, kod uyarmalı kodlayıcılar, çoklu darbe uyarmalı kodlayıcılar ve kendi kendine uyarmalı kodlayıcılarıdır[1].

Genelde bu sınıf kodlayıcılar uyarma işaretini göstermek için çok fazla bit gereksinimi duymaktadırlar. Ancak parametrik model, perde uyarmalı modelden daha az sınırlandırılır. İlave olarak bu kodlayıcılar kesin olarak sesli/sessiz karar vermeye ihtiyaç duymaz ve perde hatalarına daha az duyarlıdır. Dalgabiçimli uyarmalı kodlayıcılar düşük bit hızları için iyi sistemler olup yüksek kalitede ses kodlayıcılarıdır.

Adaptif öngörümlü kodlayıcılar, artık uyarmalı kodlayıcılar, çoklu darbe uyarmalı ve kod uyarmalı kodlayıcılar frekans domeni kodlayıcılarıdır. Frekans domeni kodlayıcılar kesin olarak kodlanmış gürültüyü sınırlandırır. Sayısal ses kodlama sisteminin bütün elemanları şekil 2’de verilmektedir.



Şekil 2. Ses kodlama sisteminin elemanları

Sistemin girişi sürekli bir $s(t)$ ses dalgasıdır. Bu işaret alçak geçiren bir süzgeçten geçirilir ve A/D dönüştürücü ile örneklenerek $s[n]$ sayısal ses işareti elde edilir. Bu ses kodlayıcının girişidir. Ses kodlayıcısı genel olarak üç elemandan oluşur. Ses analizi, parametrelerin seviyelendirilmesi ve parametrelerin kodlanmasıdır. Analiz katının girişi sayısal ses işaretidir. Bu işaret çıkışta yeni sesi oluşturmak üzere seviyelendirilir ve kodlanır. Analiz katının çıkışı ses işaretinin nasıl modelleneceğine bağlıdır. Örneğin darbe kodlu modülasyon sistemi için analiz olmayacaktır. Çünkü, çıkış sayısal ses işaretidir. Dalgabiçimli

kodlayıcılar için çıkış girişin işlenmiş bir hali olacaktır. Parametrik kodlayıcılar için çıkış ses modelinin parametreleri olacaktır. Böylece perde uyarmalı doğrusal uyarımlı kodlayıcılarda, analiz katının çıkışı vokal tırak süzgecin doğrusal öngörüm parametreleri olacaktır. Bunlar sesli/sessiz anahtarlama noktası, perde periyodu ve uyarma kazancıdır.

Analizden sonra parametreler gerekli olan bit miktarını azaltmak için seviyelendirilir. Seviyelendiricinin çıkışının, analiz işleminin yapıldığı bloğun çıkışının gürültü olarak gösterilmiş olduğu düşünülebilir. Seviyelendiricinin çıkışı kodlayıcıyı izlemektedir. Kodlayıcı her seviyelendirilmiş duruma bir tek ikili kod kelimesi tayin eder. Bu ikili kodlar etkin iletim veya depolama için paketlenir.

Sayısal kodlanmış ses, hem iletişim uygulamalarında hem de depolama ile playback uygulamalarında sık sık kullanılır. İletişim uygulamalarında, sistem kodlama gecikmesini minimize etmelidir. Ayrıca sistemin maliyeti ile güç sınırlandırması düşünülmelidir. Sonuç olarak, gerçek iletişim sistemleri ses kodlayıcı tarafından adreslenmesi gereken bit hatalarını içermektedir. Bit hatalarını önlemek için elde edilebilir bit hızı kullanılmalıdır.

Ses depolama uygulamalarında, kodlama gecikmesi söz konusu değildir ve gerçek zamanda gerçekleşmesine ihtiyaç duymaz. İlave olarak bit hatasıyla yüz yüze gelmezler. Böylece ses depolama uygulamaları, ses iletimi uygulamalarıyla karşılaştırıldığında aynı bit hızında daha yüksek kalitede olduğu görülmektedir.

Ses dekodlayıcısı, ses kodlayıcısının yaptığı işlemlerin tersini yapar. Sayısal bit dizini dekodlandıktan sonra ses parametrelerinin seviyelendirilmiş halleri ters seviyelendiriciye gönderilir. Bu parametreler sonradan $s'(n)$ kodlanmış ses işaretini sentezlemek için kullanılır. Sentetik sayısal ses işareti $s'(n)$, D/A dönüştürücüyle analoga dönüştürülür ve anti-aliasing yapılı süzgeçten geçirilerek analog sentetik $s'(t)$ ses işareti üretilir.

En yaygın olarak kullanılan modern ses kodlama sistemlerinin bazı özellikleri Tablo 1'de verilmiştir.

Tablo 1. Ses kodlama tekniklerinin özellikleri [1].

Kodlama tekniği	Bit hızı (kbps)	Gürültü maskeleme	İşitme frekansı	İşitme fazı	Hece enerjisi	Vokal tırak model	Kısa zaman sabiti	Perde
Doğrusal PCM	80-120	Yok	Yok	Yok	Yok	Yok	Yok	Yok
Delta modülasyon	50-80	Var	Yok	Yok	Yok	Var	Yok	Yok
DPCM	40-80	Var	Yok	Yok	Yok	Var	Yok	Yok
ADM	16-40	Var	Yok	Yok	Var	Var	Yok	Belki
Altband kodlayıcı	10-32	Var	Var	Belki	Var	Yok	Yok	Belki
ATC	8-32	Var	Var	Belki	Var	Var	Var	Var
APC	8-32	Var	Yok	Yok	Var	Var	Var	Belki
MPLPC	8-16	Var	Yok	Var	Var	Var	Var	Var
CELP	4-16	Var	Yok	Var	Var	Var	Var	Var
SEV	4-16	Var	Yok	Var	Var	Var	Var	Var
LPC kod.	0.6-2.4	Var	Yok	Var	Var	Var	Var	Var

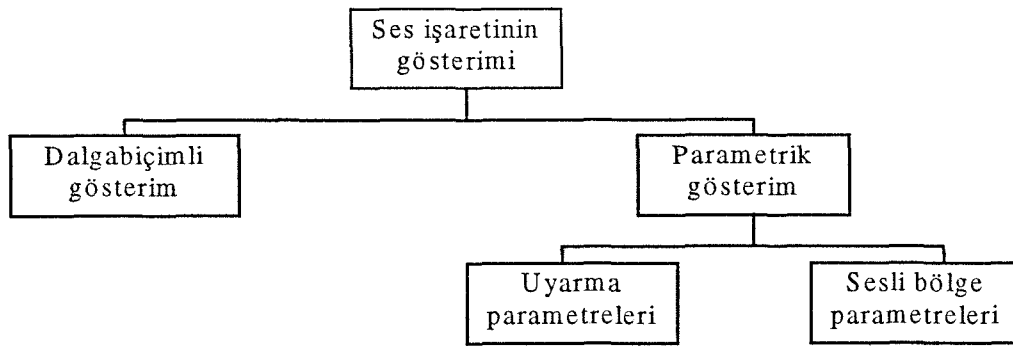
Tablo 1'deki CELP(Code-excited linear prediction) kodlama tekniği 4kb/sn. ile 16kb/sn. aralığında çalışmakta ve gürültü maskeleme özelliğine sahiptir. Ayrıca hece enerjisine sahip olup vokal tırak modelini içermektedir. Perde periyodu ve kısa zaman enerjisi bilgilerini elde etmemize imkan vermektedir. Tablo 1'de verilen bütün kodlama teknikleri hem arka plan gürültüsü ve hem de iletim hatalarına oldukça duyarlıdır.

Ses kodlama standartları kullanılan ve geliştirilecek olan ses kodlayıcılarda önemli bir rol oynamaktadır. En basit standartlar, ses kodlama standardı olarak kesin biçimde belirtilmemekle birlikte sayısal compact diskler için kullanılan formatlardır. Böyle sistemler 16 bitlik doğrusal PCM kodlamayı kullanırlar. İlk standartlardan biri olan ve halen kullanılan adaptif delta modülasyonu formudur. Bu ya 16kbps veya 32 kbps'dir. 16 kbps askeri amaçlı olanlar için uygundur. Diğer standartlar U.S. Federal standart 1015 ve 1016'dır. Diğer bir

önemli standart Uluslararası Telekomünikasyon Birimi (ITU)'dur. Bu yaygın olarak sayısal anahtarlamalı uygulamalarda 64 kbps'yi kullanır. 3.4 kHz ses işaretleri için diğer önemli telefon standartları (G.726) 16,24,32 ve 40 kbps'de çalışırlardır. Düşük bit hızlarında yapılan ses kodlama işlemlerinde sesin kalitesi ancak anlaşılabilir düzeyde olmaktadır. Değişik kodlayıcı türü olmakla birlikte ses kodlayıcılar, analiz-sentez ana bloklarından oluşmaktadır. Bu kodlayıcı türlerinden biri olan ve düşük bit hızlarında uygulanan Kod Uyarımlı Doğrusal Öngörümlü Kodlayıcıdır (KUDÖK).

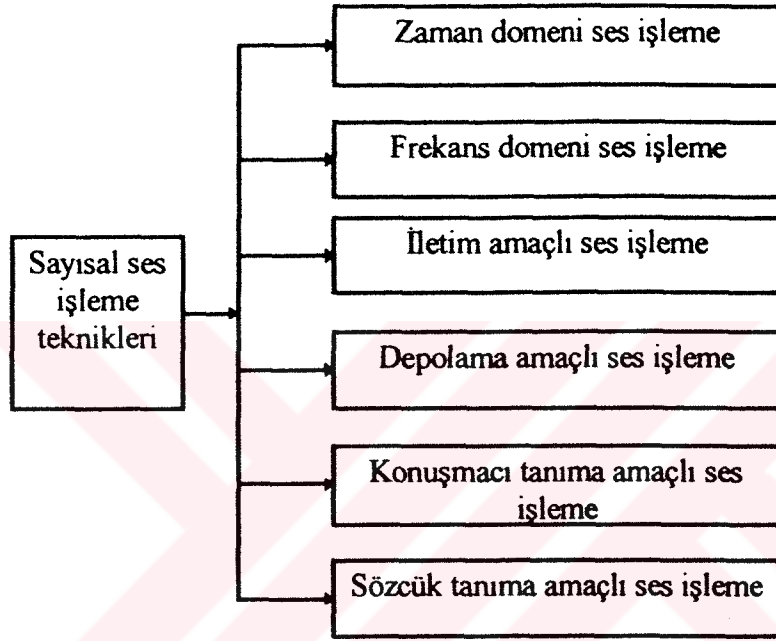
1.2. Sayısal Ses İşleme

Ses işaretlerinin sayısal formda gösterilmesi temel kavramlardan birisidir. Zaman domeninden ayrık-zamanlı domene geçilmesi için işaretin örnekleme teoremine [1] uygun olarak örneklenmesi gerekir. Böylece ses işaretinin ayrık gösterimi mümkün olmaktadır. Şekil 3'te ses işaretlerinin iki ana grupta sınıflandırıldığı gösterilmektedir. Bu gruplar dalga biçimli ve parametrik gösterimdir. Dalga biçimli gösterim ses işaretinin örneklenmesini ve seviyelendirilmesi işlemlerini içermektedir. Parametrik gösterim ise ses üretimi için kullanılan modelin çıkışı olarak gösterilmektedir. Modelin parametreleri elde edilmeden önce ses işareti örneklenir ve seviyelendirilir. Modelin parametreleri ya uyarma parametreleri ya da sesli ses bölgesi cevabı parametreleri olarak sınıflandırılır. Ses işlemede ilk çalışma 1930'lu yıllarda Homer DUDLEY tarafından gerçekleştirilen "vocoder" ya da "voice coder" lar olarak bilinmektedir [2]. Bu kodlayıcıların (vocoder) amacı ses işaretinin iletilmesi için gerekli olan band genişliğini azaltmaktır. Üstelik bu sayısal ses sistemlerinde daha düşük bit hızını mümkün kılmaktadır.



Şekil 3. Ses işaretinin gösterimi

Dalgabiçimli gösterim; analog ses işaretinin örnekleme seviyelendirilmesi işlemlerinde ses dalgasının biçimini korumaktadır [1]. Parametrik gösterim kaynak kodlamayı içermekte olup, dalgabiçimli gösterim kaynak kodlamayı içermemektedir. Çoğu uygulamalarda, ses işaretlerinin iletimindeki etkileri pek çok yolla sınırlandırılmaktadır. Böyle durumlarda ses işaretinin kalitesini artırmak için işaret işleme teknikleri kullanılmaktadır. Sayısal ses işleme teknikleri şekil 4'te verilmiştir.



Şekil 4. Sayısal ses işleme teknikleri

Ses haberleşmesinde bu tekniklerin çoğundan yararlanılmaktadır. Perde periyoduna sahip sesli sesler için sayısal işlemler frekans domeninde gerçekleştirilmektedir. İşaretin zamanla değişimi ise zaman domeni modellerinde ele alınır. Ses işleme uygulamalarında iletim ve depolama amaçlı ses işleme teknikleri çok önemli bir yere sahiptir. İletim amaçlı ses işlemede; gerekli olan band aralığını azaltmak ve kanal hatalarını düzeltmek esas amaçtır. Bunlar ayrıca uydu, mikrodalga ve optik iletim sistemlerinde de başta gelen problemlerdir. Sayısal ses depolamada kanal hataları söz konusu olmamaktadır. Depolama amaçlı ses işlemede MP3 ve MP4 gibi sıkıştırma oranları yüksek teknikler geliştirilmiştir. MPEG kodlanmış hareketli görüntüleri göstermek, ses ve bunların varyasyonları için standart geliştirmek amacıyla 1988'de gerçekleştirilmiştir. MP3, MPEG-1'in 3.ses tabakasıdır.

MPEG-1 standardında 3 farklı tabaka vardır. Tabakalar arasındaki ana fark kodlama ve dekodlamayı göstermeleridir. Yüksek tabakalar çok daha karmaşıktır ve daha çok fonksiyona sahiptir. MP4 ise geniş ve oldukça farklı ses formatlarını tanımlamak için kullanılır. MP4 terimi MP3 standardının üstünlüğünü artırmak için kullanılır. MP4 tüketici elektronik sistemleri ve MP3 için yazılım tabakalarıyla birlikte kullanılamaz. Bu, yüksek kaliteli ses verisini ve görüntüleri iletmek için güvenli bir yoldur. Genelde ses kalitesi kullanılan kodlayıcı ve dekodlayıcının karmaşıklığına bağlıdır.

Ses sentez sistemleriyle birleştirilen ses tanıma amaçlı sistemler düşük bit hızlı iletişim sistemlerini kapsamaktadırlar [3].

1.2.1. Ses İşleme İçin Zaman-Domeni Yöntemleri

Burada sayısal işaret işleme metotlarının ses işaretlerine nasıl uygulanabileceği anlatılacaktır. Ses işaretlerini işlemedeki amaç, ses sinyali tarafından taşınan bilginin çok daha uygun bir şekilde kullanılabilmesi ve gösterilebilmesidir. Ses işareti sesli, sessiz ve gürültü olmak üzere üç bileşenden oluşmaktadır. Zaman domen ölçümlerinde ses işaretinin gösterilmesi, ortalama sıfır geçiş hızı, enerji ve öz ilişki fonksiyonlarını içermektedir. Böyle bir gösterim gerekli sayısal işleme için çok önemlidir. Ayrıca bu gösterim ses işaretinin önemli özelliklerinin kestirimini daha da kolaylaştırmaktadır. Bu gösterimle amaç, zaman domen gösteriminin ses işlemede nasıl daha etkili bir şekilde kullanılabilceğidir.

Ses işaretinin özellikleri zamana bağlı olarak değişmektedir. Örneğin uyarma, sesli ile sessiz ses arasında değişir. Bu ses işaretinin pik değerinin yeterli bir derecede değişmesi anlamına gelmektedir. Sayısal ses işlemede kısa-zaman işlem tekniğinin formu şöyledir.

$$Q_n = \sum_{m=-\infty}^{\infty} T[x(m)]W(n-m) \quad (1)$$

Ses işareti bir dönüşüme tabi tutulur. Bu T[] dönüşümü ya lineer yada nonlineer olabilmektedir. Bu bazı ayarlanabilir parametre veya parametreler kümesine bağlıdır. Elde edilen kaynak n indeksli örneğe ilişkin bir pencere fonksiyonuyla çarpılır. Çarpım bütün

sıfırdan farklı değerler için toplanır. Genelde pencere fonksiyonu sonlu süreli olacaktır. Böylece Q_n değerleri $T[x(m)]$ kaynağının lokal ağırlıklı ortalama değerleri kaynağı olmaktadır. Ayırık-zamanlı bir işaretin enerjisi şöyle verilir.

$$E = \sum_{m=-\infty}^{\infty} x^2(m) \quad (2)$$

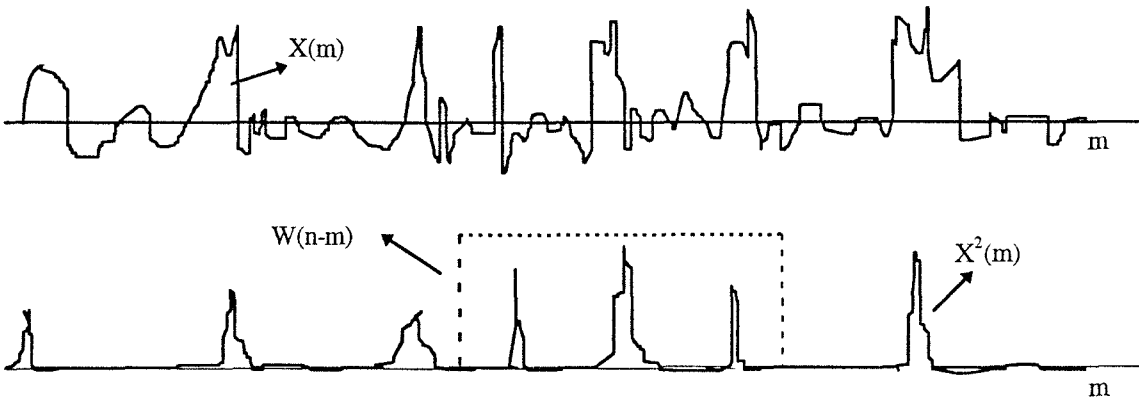
Böyle bir miktar, ses için az bir anlam taşımaktadır. Çünkü bu ses işaretinin zaman-bağımlı özellikleri hakkında az bilgi vermektedir. Kısa-zaman enerjisinin basit bir tanımı şöyledir.

$$E_n = \sum_{m=n-N+1}^n x^2(m) \quad (3)$$

Bu ifade n . örnekteki kısa-zaman enerjisi N örneğin karelerinin toplamıdır. Genel ifadedeki $T[]$ işlemi burada basit bir karesel işlemdir.

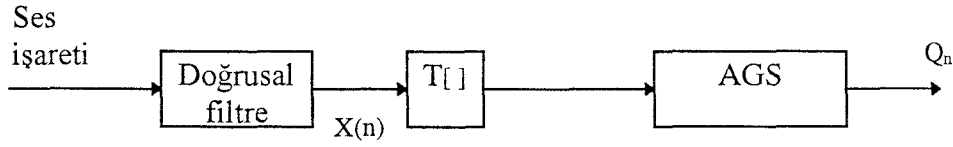
$$W(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{d.d.} \end{cases} \quad (4)$$

Genel ifade açık bir şekilde pencere fonksiyonuyla $T[x(n)]$ kaynağının ayırık bir katlaması formunda olduğunu göstermektedir. Kısa-zaman enerjisi şekil 5'te gösterilmiştir.



Şekil 5. Kısa-zaman enerji gösterimi

Şekil 5'te m örneklı herhangi bir ses işareti ele alınmıştır. Bu $X[m]$ ses işaretinin karesi alınarak $X^2[m]$ işareti elde edilir ve böylece negatif örnek değerleri pozitif hale dönüştürülmüş olunur. Daha sonra $W(n-m)$ pencere fonksiyonu şekil 5'te gösterilen $X^2[m]$ işareti üzerinde kaydırılarak enerji hesabı yapılır. Bu olay blok diyagram olarak şekil 6'da gösterilmiştir. Şekil 6'da ses işareti süzgeçten geçirilerek kare alma işlemi olan $T[]$ dönüşümüne tabi tutulur. $T[]$ dönüşümü yapıldıktan sonra gürültüyü elemine etmek için AGS'den geçirilir ve (1) nolu denklemlle ifade edilen Q_n değeri bulunur.



Şekil 6. Kısa-zaman analiz prensibinin genel gösterimi

1.2.2. Kısa Zaman Fourier Analizi

Fourier gösteriminin işaret işlemede kullanışlı olmasının iki temel sebebi vardır. İlki doğrusal sistemler için sinüzoidal veya üstel durumlara olan cevabın belirlenmesi bu gösterimle uygun olmaktadır. İkincisi ise Fourier gösteriminin orijinal işarete tam olarak görülmeyen veya belli olmayan temel özellikleri göstermesidir.

Fourier analizi yapan sistemlerin çıkış tayfi hem uyarma hemde vokal tırak frekans cevabının özelliklerini yansıtmaktadırlar. Bu yüzden standart fourier gösterim özellikleri zamanla değişen ses işaretlerine doğrudan uygulanmaz. Ses işaretlerinde enerji, sıfır eçişleri ve ilişki gibi özelliklerin 10-30 milisaniye gibi zaman aralığında sabit kaldığı düşünülür.

Ses dalgasının zamanla değişen özelliklerini yansıtan spektral gösterime ihtiyaç duyulur. Zaman bağımlı Fourier dönüşümünün kullanışlı bir tarifi şöyledir.

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(n-m)x(m)e^{-j\omega n} \quad (5)$$

Giriş işaretinin bir kısmını göstermek üzere $w(n-m)$ pencere kaynağı kullanılır. Bu dönüşüm iki değişkenin bir fonksiyonudur. Zaman indeksi olan ayrık n ve frekans değişkeni olan ω 'dır. Toplamın indeksi değiştirilirse;

$$X_n(e^{j\omega}) = e^{-j\omega} \sum_{n=-\infty}^{\infty} x(n-m)w(m)e^{j\omega m} \quad (6)$$

ifadesi elde edilir. Bu ifadeler iki ayrı biçimde yorumlanabilir. Eğer n sabit düşünülürse $X_n(e^{j\omega})$ normal bir Fourier dönüşümüdür. Diğer yorum $X_n(e^{j\omega})$ 'in sabit ω ve n zaman indeksinin bir fonksiyonudur. Bu durumda (5) ve (6) ifadeleri bir katlama biçimini alır.

Fourier Dönüşümü Yorumu : Zaman bağımlı Fourier dönüşümü bütün reel değerleri alan n zaman indeksinin bir fonksiyonudur. Ayrık zamanlı işaretlerin normal Fourier dönüşümünde olduğu gibi zaman değişimli Fourier dönüşüm, ω periyodlu olup periyodu 2π 'dir. $X_n(e^{j\omega})$, $w(n-m)x(m)$ 'in normal Fourier dönüşümü olarak şöyle yazılır.

$$w(n-m)x(m) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} X_n(e^{j\omega}) e^{j\omega m} d\omega \quad (7)$$

Bu integral 2π uzunluklu aralıkta gerçekleştirilmektedir. Eğer $w(0)$, 0'dan farklı ise $m=n$ durumu için;

$$x(n) = \frac{1}{2\pi w(0)} \int_{-\pi}^{\pi} X_n(e^{j\omega}) e^{j\omega n} d\omega \quad (8)$$

ifadesi elde edilir. Böylece $x(n)$ kaynağı $X_n(e^{j\omega})$ ' dan yeniden elde edilmektedir.

$X_n(e^{j\omega})$ 'nın önemli bir özelliği kısa zaman otokorelasyon fonksiyonuyla ilişkili olmasıdır.

$$S_n(e^{j\omega}) = |X_n(e^{j\omega})|^2 = X_n(e^{j\omega})X_n^*(e^{j\omega}) \quad (9)$$

$$R_n(k) = \sum_{m=-\infty}^{\infty} w(n-m)x(m)w(n-k-m)x(m+k)$$

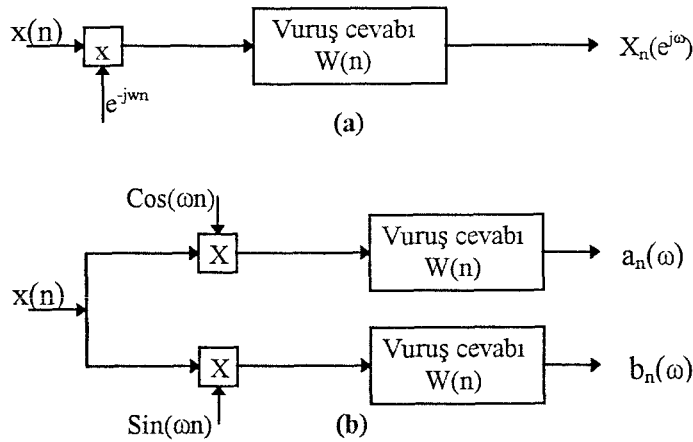
Böylece (9) denklemi kısa zaman tayfının, kısa zaman ilişkisini nasıl gösterdiğini ifade eder. $X_n(e^{j\omega})$ 'nin diğer bir gösterimi genlik ve faz terimlerini içermiştir.

$$x_n(e^{j\omega}) = |X_n(e^{j\omega})|e^{jQ_n(\omega)} \quad (10)$$

Pencere fonksiyonu biçimi zaman değişimli Fourier analizinde önemli bir etkiye sahiptir. $X(m)$ 'in $-\infty < m < +\infty$ arasındaki Fourier dönüşümü kaydırılmış pencere fonksiyonunun Fourier dönüşümünü içermektedir. Zaman bağımlı Fourier dönüşüm pencere içerisindeki işaretin Fourier dönüşümünün düzgünleştirilmiş bir halidir.

Linear Süzgeçleme Yorumu : Her ω için $X_n(e^{j\omega})$, $w(n)$ ile $x(n)e^{j\omega n}$ 'nin katlaması olduğu ifade edildi. $W(n)$ doğrusal sistemlerin birim vuruş cevabında önemli rol oynamaktadır. Şekil 7 (a)'daki doğrusal sistemin giriş ve çıkışı kompleksdir.

$$X_n(e^{j\omega}) = a_n(\omega) - jb_n(\omega) \quad (11)$$



Şekil 7. Linear süzgeçnin kısa zaman analiz yorumu, a) Kompleks yapı, b) Reel çalışması

Burada $a_n(\omega)$ ve $b_n(\omega)$ 'yı elde etmek için gerekli olan işlemler şekil 7 (b)'de gösterilmiştir. Ayrıca lineer süzgeç kullanılarak hesaplanan $X_n(e^{j\omega n})$ 'de sonsuz süreli pencerelerin kullanılması daha uygundur.

1.3. Paralel İşlem Yaklaşımını Kullanarak Perde Periyodunun Kestirimi

Perde periyodunun kestirimi ses işlemede en önemli problemlerden birisidir. Perde belirleyiciler kodlayıcılarda kullanılmaktadır [3]. Ayrıca konuşmacı belirleme ve sistem gerçekleştirilmesinde de kullanılmaktadır. Sesli sesin perde periyodunu kestirmek için farklı yaklaşımlar vardır. Bunlar frekans ve zaman domeni yaklaşımlardır. Zaman domeni yaklaşımında ses işareti perdeyi kestirmek için işlenir. Frekans domeni yaklaşımında ise ses işaretinin tayf bilgisi ve harmonik yapısı perdenin kestirilmesi için kullanılır. Zaman domeni perde periyodu kestirimi yöntemleri için, ses dalgasının merkez kırılması ve pik seçme basit algoritmaları örnek verilebilir. Bu teknikler sesin kısa zaman özilişkisine bağlıdır.

Perde periyodu hesaplama algoritmalarında göz önünde bulundurulması gereken iki ana nokta vardır. İlki kullanılan algoritmanın güvenilirliğidir. İkincisi ise kullanılacak olan algoritmadaki işlem hesabı karmaşıklığıdır.

En basit algoritmalarından biri ses işaretindeki sıfır-geçişleri arasındaki mesafeyi bulmaktır [3]. Sıfır geçişleri arasındaki mesafeyi kullanarak perdeyi kestiren algoritmada; giriş ses işareti sıfır geçiş mesafeleri hesaplanmadan önce alçak veya band geçiren süzgeçten geçirilmelidir. Çünkü yüksek frekans bileşenleri hatalara neden olmaktadır. Perde frekansı $f=1/T$ şeklinde olup, T sıfır geçişleri arasındaki mesafenin iki katı olarak alınır. Bir diğer yöntem Reetz algoritmasının geliştirilmiş halidir [4]. Burada perdeyi kestirebilmek için giriş ses işareti dört adımlık bir işleme tabi tutulmaktadır. Dört adımlık işlemin sonunda sesli ses için perde periyodu bulunur. Ancak bu yöntemle bulunan perde periyodu değeri en ilkel olanıdır. Bu adımlar şöyledir:

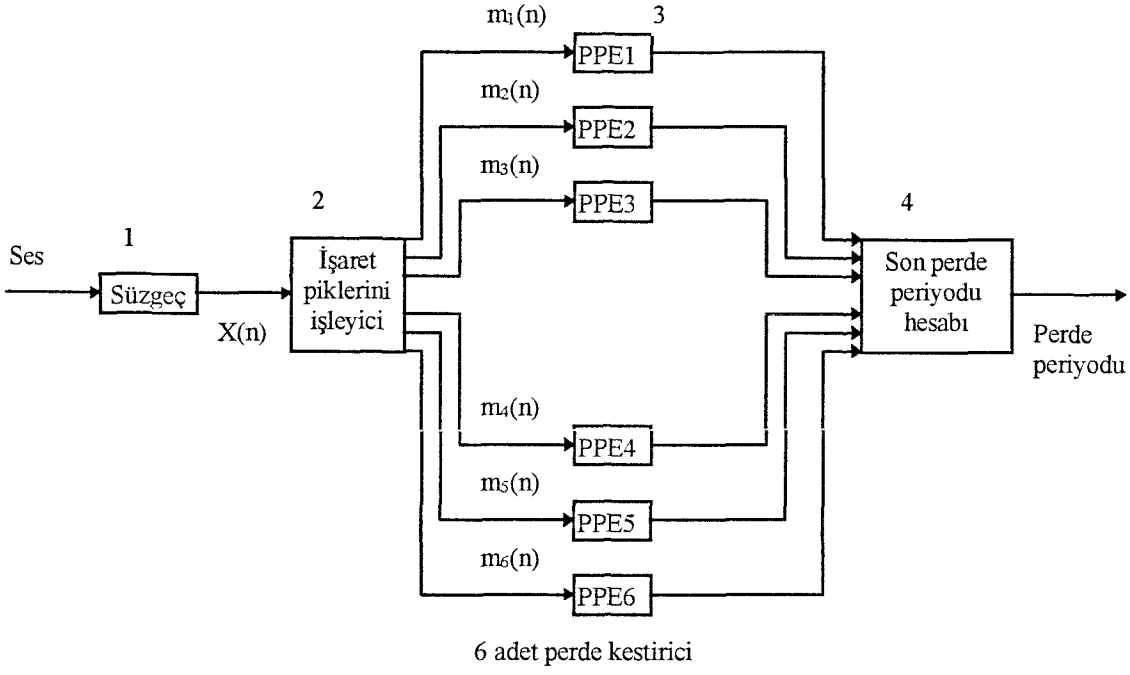
- 1) Veri(data) azaltma: Giriş işaretini belirgin olarak pozitif ve negatif piklere dönüştürmek. Ses kaydedilirken veri azaltma işlemi yapılmış olunmaktadır.

- 2) Mantıksal süzgeçleme: Çok küçük veya çok büyük genlikli pikleri yok ederek gürültüyü elemine etmektir.
- 3) Zincirleme: Bütün piklere bakılarak optimal perde için arama yapmak. Arama işleminde izlenecek yol hangi pik bloğunun tekrarlandığını tespit etmek olacaktır. Perdeyi belirliyen en çok tekrarlanan pik bloğu olacaktır.
- 4) En son işlem: Çok büyük, çok küçük veya oldukça kısa segmentleri silmektir.

Bu yöntemle perde periyodu hesaplanmasında işlem sayısı azaltılmış olunur.

Burada GOLD [4] tarafından ilk olarak önerilen kısmi perde belirleme biçimi açıklanacaktır. Perde kestiriciler geniş bir uygulama alanında başarılı olarak kullanıldı. Bu, ses işlemede paralel işlemin ana prensiplerini kullanmayı ifade eder. Bu biçim şekil 8'de gösterilmektedir. Bunun ana prensipleri şunlardır;

- 1) Ses işareti bir darbe dizisi oluşturacak şekilde işlenir. Bu vuruşlar orjinal işaretin periyodlarını izler.
- 2) Bu işlem her darbe dizinin periyodunu kestirmek için perde kestiricilerinin kullanılmasına izin vermektedir.
- 3) Bu perde kestiricilerin birkaçının değerleri ses dalgasının periyodunu göstermek üzere lojksel olarak birleştirilir.



Şekil 8. Zaman-domeni perde kestiricisi paralel işlemin blok diyagramı

Ses dalgası uygun bir biçimde örneklenir. 10kHz'lik örnekleme frekansı $T=10^{-4}$ sn içinde perdenin belirlenmesine izin verir. Ses, düzgün bir dalga üretmek için 900Hz'de alçak-geçiren süzgeçten geçirilir. Band-geçiren süzgeç 100Hz ile 900Hz arasını geçirmekte ve 60Hz'lik gürültüleri uzaklaştırabilmektedir. Bu süzgeç ya örneklemeden önce analog olarak yada örneklemeden sonra sayısal olarak gerçekleştirilebilir.

Tepeler ve çukurlar yerel max. ve yerel min. olarak adlandırılarak yerleştirilirler. Süzgeçlenmiş işaretten birkaç darbe dizini elde edilir. Herbir darbe dizini ya pik yada çukur yerlerinde meydana gelen pozitif darbelerden oluşmaktadır. Gold ve Rabiner'in önerdiği 6 durum şöyle ifade edilmektedir.

- 1) $m_1(n)$: Her pikin olduğu yerde pikin genliğine eşit bir darbe oluşturulur.
- 2) $m_2(n)$: Her pikte, pik genliği ile bu piki izleyen çukur genliği arasındaki farka eşit genlikli bir darbe oluşturulur.
- 3) $m_3(n)$: Her pikte, pik genliği ve bu piki izleyen diğer pik genliği arasındaki farka eşit genlikli bir darbe oluşturulur. Eğer bu fark negatif ise darbe sıfıra setlenir.
- 4) $m_4(n)$: Her çukurda, çukurun genliğinin negatifine eşit bir darbe oluşturulur.

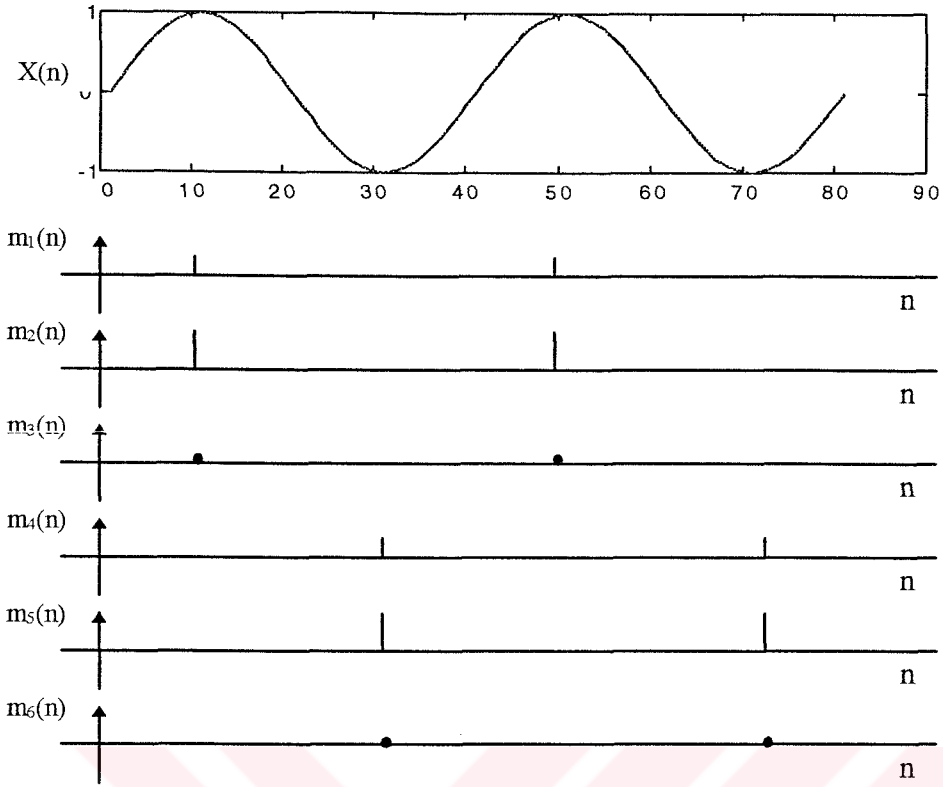
- 5) $m_5(n)$: Her çukurda, çukurun genliğinin negatifi ile bu çukuru izleyen pikin genliğinin toplamına eşit bir darbe oluşturulur.
- 6) $m_6(n)$: Her çukurda, çukurun genliğinin negatifi ile bunu izleyen yerel min. genliğinin toplamına eşit bir darbe oluşturulur. Eğer fark negatif ise darbe sıfıra setlenir.

Şekil 9 ve şekil 10 buna ilişkin verilmiş iki örnektir. Açık bir biçimde darbe dizileri, orijinal giriş işaretinin ana frekansına eşit frekansa sahip oldukları görülmektedir. Bu darbe dizilerinin üretilmesinin amacı perdeyi kestirebilmeyi yapabilmektir. Perde periyodu kestirici işlemi şekil 11'de gösterilmiştir. Her darbe dizini bir zaman değişimli doğrusal olmayan sistemle işlenir. Girişte yeterli seviyede bir darbeye rastlanıldığında, çıkış o darbenin değerine setlenir ve daha sonra darbe bulunana kadar boş bir aralıkta tutulur. Bu aralığın sonunda çıkış üstel olarak azalmaya başlar. Eğer bir darbe azalan bu çıkış değerini aştığında işlem tekrarlanır. Azalma hızı ve boşluk aralığı en son kestirilen perde periyodu değerine bağlıdır.

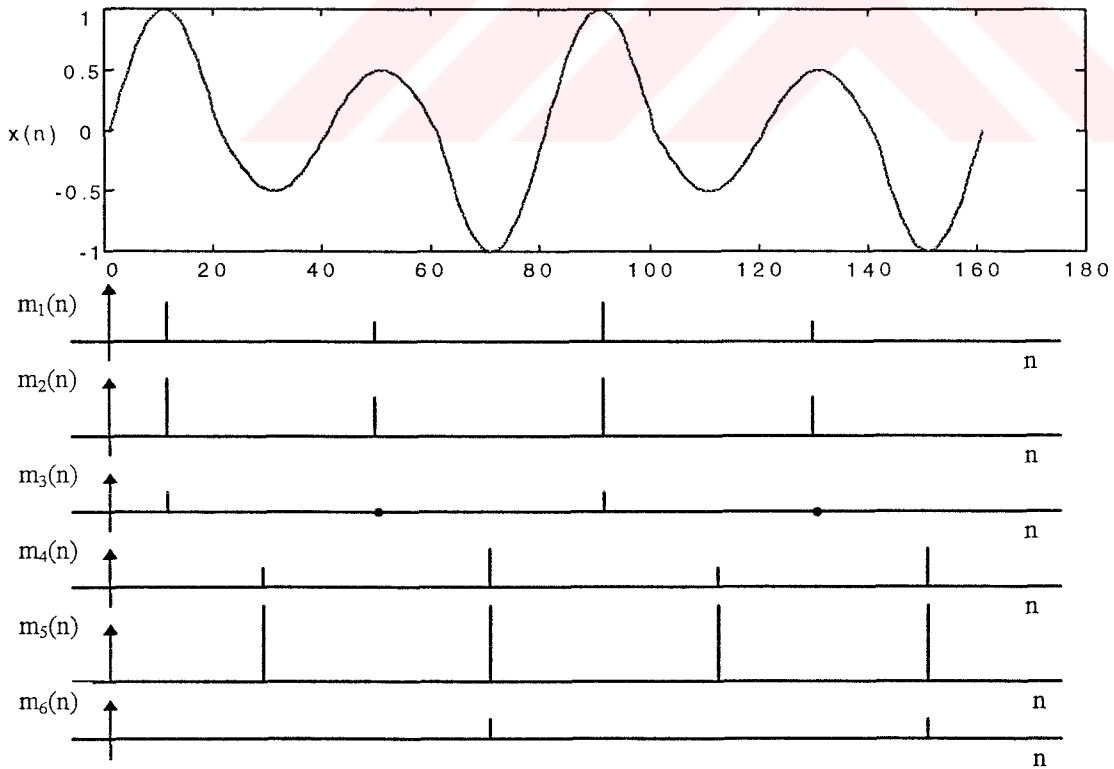
Ayrıca perde periyodunun kestirimi için değişik metotlar vardır. Bu yöntemleri üç ana sınıfta toplayabiliriz.

- 1) Frekans domeni: Eğer işaret periyodik ise F_0 ve bunun harmonikleri (F_0 'ın katları) frekans domeninde yüksek enerjili noktalar olarak ortaya çıkarlar.
- 2) Zaman domeni: Bu özilişki fonksiyonuna dayanmaktadır. $1/F_0$ ve katlarındaki piklerin aranmasıdır.
- 3) Zaman ve frekans domeni: Yukardaki iki sınıfta kapsamaktadır.

Bu yöntemler perde periyodu kestirimi için geliştirilen ilk ve basit yöntemlerdir. Kestirilen perde periyodu değeri kod uyarmalı doğrusal öngörümlü kodlayıcıda yer alan uzun dönem kestirici süzgecinde kullanılmaktadır. Perde periyodu ses işaretinin sesli ses bölgesi için geçerlidir. Sessiz ses bölgeleri için perde periyoduna gerek yoktur. Çünkü sessiz ses bölgeleri yarı periyodik bir özelliğe sahip değillerdir.



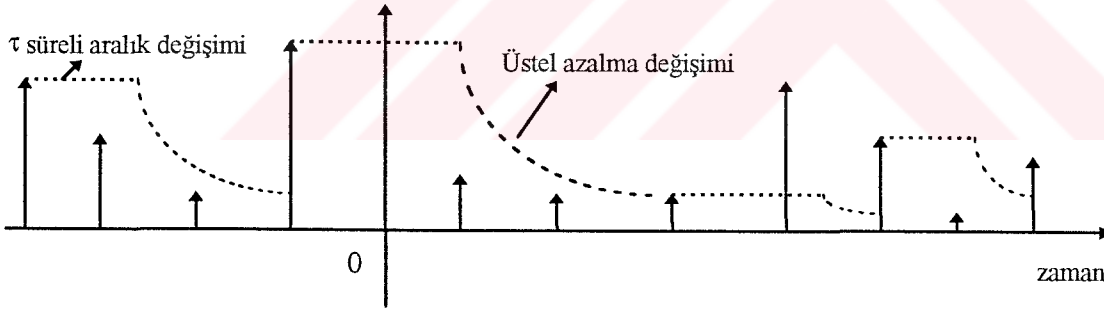
Şekil 9. Sinüzoidal giriş ile pik ve çukurlardan üretilen ilişki darbe dizileri



Şekil 10. Ara ve ikinci harmonikli giriş ile pik ve çukurlardan üretilen ilişki dizini

Sonuç düzgün bir darbe dizisidir. Üretilen yarı-periyodik darbe kaynağı şekil 11’de gösterilmiştir. Her darbenin uzunluğu perde periyodunun kestirilen değeridir. Perde periyodu periyodik olarak, örnekleme aralığında darbe sürelerinin uzunluğu ölçülerek kestirilir. Bu teknik 6 vuruş dizisinin herbirine uygulanır. Böylece 6 perde periyodu kestirilmiş olunur. Bu değerler, 6 perde kestiricisinin herbiri için enson kestirilen değerlerin ikisiyle birleştirilir. Bu değerler karşılaştırılır ve en uygun doğrulukta olanı perde periyodu olarak belirlenir. Bu işlem sesli sesin periyodunun kestirimini çok iyi bir şekilde yapmaktadır. Sessiz ses için kestirilen bu değerler arasında belli bir uyum eksikliği vardır. Bu uyum eksikliği olan ses, sessiz ses olarak sınıflandırılır. Sentetik sesin kullanılmasının amacı tamamen doğru olarak perde periyodlarının bilinmesidir. Sentetik sesin dezavantajı basit bir modelle üretilmesidir. Böylece doğal sesin özelliklerinin tamamını gösterememektedir.

İşaret işleme, ses işaretinin istenilen özelliğinin kestiriminde lojiksel karmaşıklığı artırmıştır. Çünkü lojiksel işlemler işaret işlemeden daha düşük hızda olduğu gösterilmiştir. Güvenilir perde periyodunun kestirimi için Barnwell ve arkadaşları tarafından benzer yaklaşımlar yapılmıştır [5].

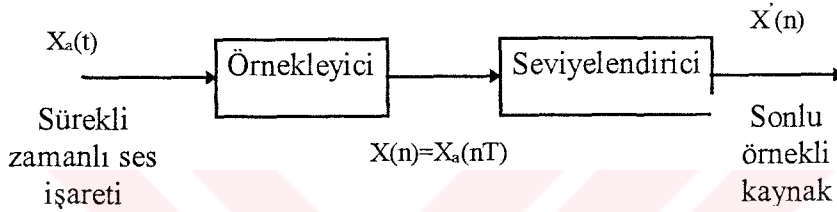


Şekil 11. Zaman-domeni perde kestiricisinin her perde periyodunun temel işlemi

1.4. Ses Dalgasının Sayısal Gösterimi

Watson; eğer hava değişimiyle şiddeti orantılı bir elektrik akımı meydana getirecek bir düzenek kurabilirsem sesin iletimini sağlayabilirim demiştir [6].

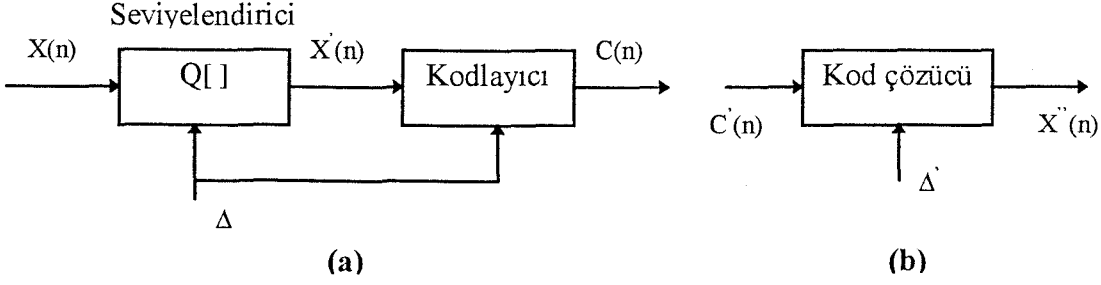
Bu temel fikir insanın haberleşme tarihinde çok önemlidir. Bu fikir ses işaretinin kaydedilmesi, gönderilmesi ve işlenmesi için gerekli olan düzenek ve sistemlerin kurulmasına neden olmuştur. Sayısal sistemler ses dalgasını, düzensiz olarak değişen genlikleri niteleyen elemanlar kümesi olarak gösterirler. Sayısal ses dalgasının genel olarak gösterimi şekil 12’de verilmiştir. Burada ses dalgası sürekli zaman değişimli sürekli bir fonksiyon olarak gösterilir. Genelde zamanda periyodik olarak örneklenir. $X_a(nT)$ örnek kaynağı üretilir. Böylece sayısal gösterimi elde edebilmek için gerekli olan sayısal değerli sonlu sayıdaki küme seviyelendirmek için gereklidir. Analog işaret band sınırlandırılır ve Nyquist frekansının en az iki katı bir örnekleme hızıyla örneklenir.



Şekil 12. Sayısal ses dalgasının genel blok diyagramı

1.4.1. Seviyelendirme

Ses dalgası alçak geçiren bir süzgeçten geçirilir ve uygun bir örnekleme hızıyla örneklendikten sonra $X(n)$ kaynağı elde edilir. Bu örnekler kümesini sayısal iletim kanalı üzerinden göndermek, hafızada depolamak, sayısal işaret işleme algoritmalarında onları giriş olarak kullanmak için örnek değerleri sonlu bir genlik kümesiyle seviyelendirilmesi gerekmektedir. Öyleki onlar sonlu bir semboller kümesiyle gösterilebilsin. Bu seviyelendirme ve kodlama işlemi şekil 13’te gösterilmiştir. Sonlu bir örnekler kümesi $\{X(n)\}$, sonlu bir semboller kümesi $\{C(n)\}$ olmak üzere iki kat söz konusudur. Seviyelendirme katı seviyelendirilmiş genlikler kümesini üretir. Bu $\{X'(n)\}=\{Q[X(n)]\}$ ile gösterilir. Her seviyelendirilmiş örnek bir kod kelimesiyle gösterilmek üzere kodlama katı oluşturulur. Bu kod kelimesi $C(n)$ ile gösterilir.



Şekil 13. Seviyelendirme ve kodlama işlemi, a) Kodlayıcı, b) Kod çözücü

Şekil 13 (a)'da Δ seviyelendiricinin seviyelendirme adım sayısını gösterir. Şekil 13 (b)'de eğer $C'(n)$, $C(n)$ kod kelimesiyle aynı ise hata yok demektir. İdeal kod çözücünün çıkışı seviyelendirilmiş örneklerle aynıdır, $X''(n)=X'(n)$.

Seviyelendirilmiş örnekleri ikili sayılarla göstermek en uygun olanıdır. B-bitlik ikili kod kelimelerini 2^B farklı seviyelendirme düzeyi ile göstermek mümkündür. İletim veya depolamak için gerekli olan bilgi kapasitesinin sayısal gösterimi için şu ifade kullanılır.

$$I = BF_s \quad (12)$$

F_s , örnekleme hızı (sembol/sn). B ise örnek başına düşen bit miktarıdır (bit/sembol). Verilen bir ses için min. örnekleme hızı örnekleme teoreminden dolayı sabittir. Böylece bit hızını azaltmanın tek yolu sembol başına düşen bit miktarını azaltmaktır. Bundan ötürü değişik seviyelendirme teknikleri geliştirilmiştir. Genelde $\{X(n)\}$ örnekleri sonlu bir genlik aralığına düşürülür.

$$|X(n)| \leq X_{\max} \quad (13)$$

Gamma ve Laplacian dağılımlar $X(n)$ 'in genliklerinin muhtemel bir fonksiyonu olarak düşünülür. Laplacian şiddeti ele alındığında, ses örneklerinin yalnız %0.35 aralığın dışına düşmektedir. Bu aralık şu şekilde verilmektedir.

$$-4\sigma_x \leq X(n) \leq 4\sigma_x \quad (14)$$

Böylece ses işaretinin tepeden-tepeye olan aralığı işaretin standart sapmasıyla orantılıdır. Örnek genlikleri giriş genlik aralığına bölünerek seviyelendirilir. Örneğin bütün $X(n)$ değerleri X_1 ile X_2 arasında olsun. Seviyelendiricinin çıkışı $X'(n)=Q[X(n)] = X'_2$ olduğu görülebilir. Seçilen 8 seviyelendirici durumunun her biri 3-bitlik ikili bir kod kelimesiyle etiketlenir. Bunlar genlik seviyelerini gösteren sembollerdir. Genliği X_1 ile X_2 arasında olan örnek için kodlanmış çıkış ikili olarak 101'dir. Burada etiketlendirme biçimi keyfi olarak yapılmıştır. Uniform seviyelendirici tanımı için;

$$X_i - X_{i-1} = \Delta \quad (15)$$

$$X'_i - X'_{i-1} = \Delta \quad (16)$$

denklemleri kullanılır. Δ , seviyelendirme adım sayısıdır. İki yaygın seviyelendirici karakteristiği vardır. Biri orjinin merdiven biçimli fonksiyonun yükselen kısmının ortası olduğu durumdur. Bu tip seviyelendiriciler “mid-riser” olarak adlandırılırlar. Diğer bir tip seviyelendirici de “mid-tread” olarak adlandırılır. Mid-riser seviyelendirici aynı sayıda pozitif ve negatif seviyelendirmeye sahip ve orjine göre simetrik konumdadırlar. Mid-tread mid-riser'in aksine pozitif seviyeden 1 fazla negatif seviyeye sahip. Ancak burada bir seviye sıfırdır. Mid-riser'da sıfır seviyesi yoktur. Kod kelimeleri ile seviyelendirilmiş örnekler arasındaki ilişki;

$$X'(n) = \frac{\Delta}{2} \text{isaret}(C(n)) + \Delta C(n) \quad (17)$$

ifadesi ile verilir. $C(n)$ 'in ilk biti 0 ise $C(n)$ 'in işareti +1'dir. $C(n)$ 'in ilk biti 1 ise $C(n)$ 'in işareti -1'dir. Benzer biçimde mid-tread için yukardaki eşitlikler şöyle yazılır.

$$X'(n) = \Delta C(n) \quad (18)$$

Δ ve B öyle seçilmelidir ki, $|X(n)| < X_{\max}$ olmak üzere şu ifade yazılır.

$$2X_{\max} = \Delta 2^B \quad (19)$$

$$X'(n) = X(n) + e(n) \quad (20)$$

$$-\frac{\Delta}{2} \leq e(n) \leq \frac{\Delta}{2} \quad (21)$$

$X(n)$; seviyelendirilmiş örnek, $e(n)$; seviyelendirme hatası veya gürültüdür.

1) Seviyelendirme gürültüsü durağan beyaz gürültü ise;

$$E[e(n)e(n+m)] = \begin{cases} \sigma_e^2, & m=0 \\ 0, & \text{d.d.} \end{cases} \quad (22)$$

2) Seviyelendirme gürültüsü giriş işaretiyle ilişkisiz ise;

$$E[X(n)e(n+m)] = 0, \quad \text{her } m \text{ için} \quad (23)$$

3) Seviyelendirme hatalarının dağılımı aynı uzunluklu aralıklar ve her seviyelendirme aralığı için uniform ise;

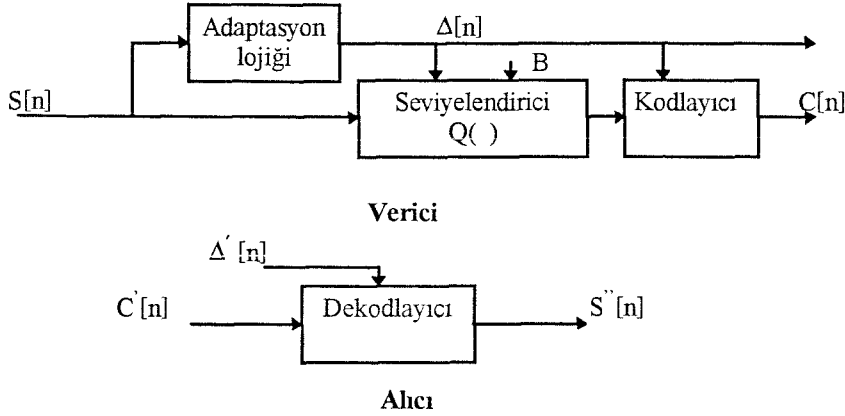
$$P_e(e) = \begin{cases} \frac{1}{\Delta}, & -\frac{\Delta}{2} \leq e \leq \frac{\Delta}{2} \\ 0, & \text{d.d.} \end{cases} \quad (24)$$

ifadeleri kullanılmaktadır.

Adaptif Seviyelendirme [6] : Ses işareti karakteristikleri zamanla değişen durağan olmıyan bir işarettir. Durağan olmamayı adreslemek için iyi bir yol adaptif seviyelendiriciler kullanmaktır. Öyleki seviyelendirici karakteristikleri ses işaretinin değişen özelliklerini karşılamak için dinamik olarak değiştirilir.

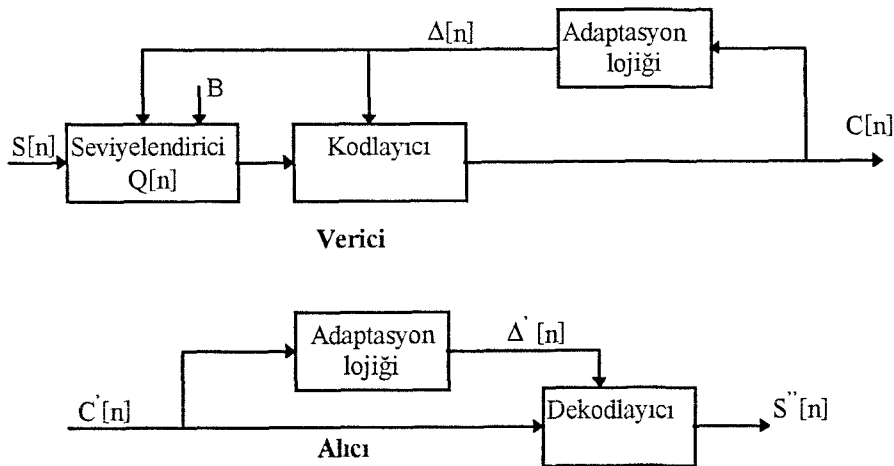
Adaptif seviyelendirmedeki ana fikir giriş işaretinin varyansını karşılamak için seviyelendiricinin adım sayısını değiştirmektir. Adaptasyon parametresi olan $\Delta(n)$ giriş işaretinden kestirilir ve seviyelendiriciyi kontrol için kullanılır. Çünkü alıcı bu değeri kestiremez ve $\Delta(n)$ dekodlama işlemini kontrol etmek için alıcıya gönderilir. Bu sistem feed-

forward adaptif PCM olarak adlandırılır. Çünkü seviyelendirici kontrol parametreleri kodlanmış örneklerle birlikte gönderilmek zorundadır. Blok diyagramı şekil 14'te verilmiş.



Şekil 14. Feed-forward adaptif PCM'nin blok diyagramı

Bu sisteme alternatif bir sistem feedback adaptif PCM olup seviyelendirilmiş çıkış ses esasına dayanmaktadır. Blok diyagramı şekil 15'te gösterilmiştir. Kodlanmış ses hem vericide hemde alıcıda elde edilebilmektedir. Sistem seviyelendirilmiş kontrol parametrelerini gönderme ihtiyacı duymaz. Feed-forward sistemlerde adaptasyon genellikle yavaş hızda çalışır, çünkü adaptasyon parametreleri her güncelleşmiş parametreden sonra gönderilmesi gerekmektedir. Aksi olarak feedback adaptasyon parametrelerin gönderilmesi gerekli değildir. Feedback adaptasyonlar genellikle eş zamanlıdır.



Şekil 15. Feedback adaptasyon PCM'nin blok diyagramı

Feed-forward adaptasyon; giriş işareti adaptasyon parametrelerini kestirmek için kullanılır. Yaygın olan bir yaklaşım adaptasyon işleminde sesin varyansını kullanmaktır. Varyans iki farklı yolla kullanılabilir. İlki, onun adım sayısında kullanılmasıdır. Diğeri ise işaretin varyansının kazanç adaptifinde kullanılmasıdır. İşaret zaman değişimli kazanç faktörüyle ölçeklendirilir. Bu sabit seviyelendiriciden geçirilmeden önce gerçekleştirilir. Bu işlem adaptif kazanç kontrolü olarak adlandırılır.

Feed-forward biçimindeki ses işaretinin varyansı, ses işaretinin kısa zaman enerjisi hesaplanarak kestirilebilir.

$$\sigma^2(n) = \sum_{m=-\infty}^{\infty} s^2(m)w(n-m) \quad (25)$$

Burada $w(n)$, veri miktarını veya kısa zaman enerjinin hesaplanmasında kullanılan sesin süresini sınırlandırmak için kullanılan bir pencere fonksiyonudur. Üstel pencere şöyle tanımlanır.

$$w(n) = \begin{cases} \alpha^{n-1}, & n \geq 1 \\ 0, & \text{d.d.} \end{cases} \quad (26)$$

Yaygın olarak $0 < \alpha < 1$ arasında bir değer seçilir. Fark denklemindeki bu sonuçlar varyansı hesaplamak içindir.

$$\sigma^2(n) = \alpha\sigma^2(n-1) + (1-\alpha)s^2(n) \quad (27)$$

Diğeri yaygın olan pencere dikdörtgen penceresidir. M pencere uzunluğunu göstermek üzere varyans ifadesi şöyledir.

$$\sigma^2(n) = \frac{1}{M} \sum_{m=n}^{m=n+M-1} s^2(m) \quad (28)$$

(26) eşitliği ses örneklerini içermemektedir. Böylece işlem süresi artmamaktadır. (28) eşitliği ise M adet işaret örneği içermekte olup M kadar örnek kodlaması gecikmesi oluşturmaktadır. (28) eşitliği hesaplanır ve n ile n+m-1 arasındaki süredeki ses örnekleri için uygulanır. Seçilen küçük değerli α ve M değerleri eşzamanlı varyans kestiriminde kullanılır. Dikdörtgen pencereler genellikle feed-forward sistemlerde kullanılırken üstel pencereler feedback sistemlerde tercih edilmektedir.

Adım sayısı ses işaretinin varyansı ile orantılıdır. Δ_0 , birim varyanslı işaret için seviyelendiricinin adım sayısını tanımlar. Ölçeklenmiş seviyelendirici adım sayısı Δ_0 , 0.01 ile 1 arasında olmak üzere adım sayısı şöyle hesaplanır.

$$\Delta(n) = \frac{\Delta_0 \sigma(n)}{2^{n-1}} \quad (29)$$

Seviyelendirici giriş işaretini elde etmek için kazanç işaretin varyansı ile ters orantılı olacak biçimde seçilir. G_0 sabit olup birim varyanslı giriş için kazanç değeridir. Zaman değişimli kazanç şöyle tanımlanır.

$$G(n) = \frac{G_0}{\sigma(n)} \quad (30)$$

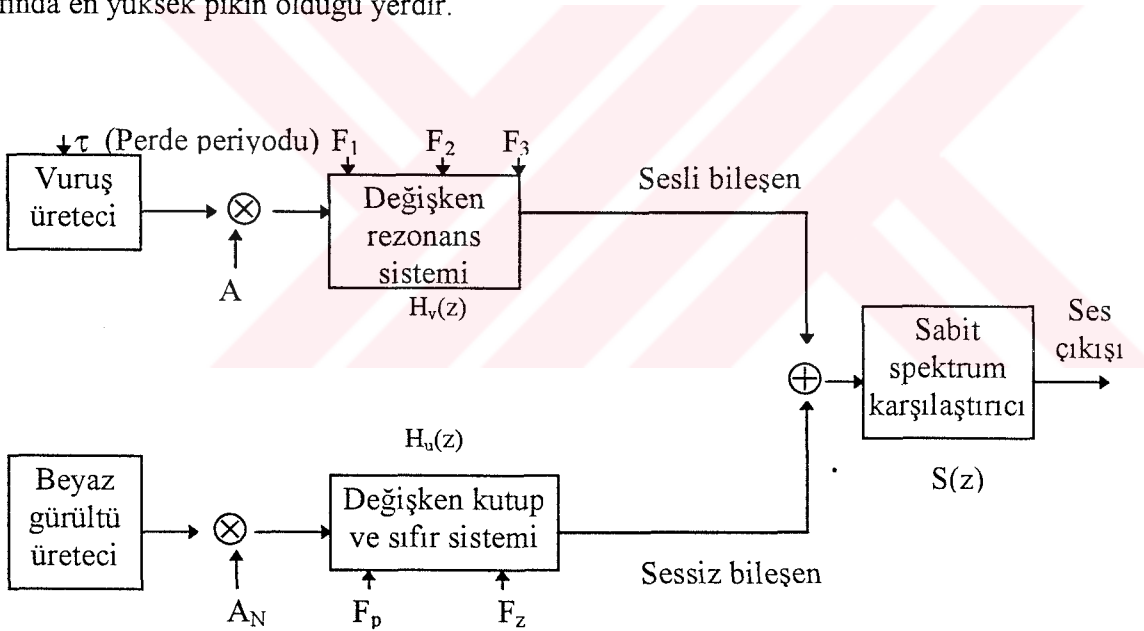
Düşük işaret enerjisi büyük bir kazanç ve yüksek işaret enerjisi düşük kazanç meydana getirmektedir.

Feedback Adaptasyon; adaptasyon parametreleri kodlanmış ses işaretinden elde edilir. Bu işaret alıcıda elde edilebilmektedir. Adım sayısı veya kazanç gerekli değildir. Her iki adaptasyon sistemi (29) ve (30)'daki eşliklerdeki kuralları kullanır. Feedback sistemlerde, ses işaretinin varyansının kestirimi yalnız kodlanmış işaretin önceki değerlerine bağlıdır.

1.5. Formant Kestirimi

Düzgünleştirilmiş logaritmik tayf, kısmi giriş ses segmentinin rezonans yapısını gösterir [7]. Tayftaki pikler aslında formant frekanslarıyla ilişkilidirler. Bu formantlar logaritmik tayfta piklerin yerleştirilmesiyle kestirilebilir.

Ses üretimi için şekil 16'da verilen modeli göz önüne alalım. Bu model; perde periyodu, genlik ve en düşük üç formant frekanslı sesli ses ve genlik, tek bir sıfır ve kutuplu sessiz sesi göstermektedir. Bütün bu parametreler zaman değişimlidir. Bu parametrelerin kestirimi için her 10-20msn. için cepstrall [7] düzgünleştirilmiş logaritmik genlikli fonksiyonun hesaplanmasına bağlıdır. Eğer ses segmenti sesli ise, perde periyodu cepstrumdan kestirilir ve ilk üç formant frekansı spektrumdaki piklerden lojik esasına dayanan ses üretim modelini kullanarak kestirilir [7,8]. Sessiz ses durumunda, kutup tayfında en yüksek pikin olduğu yerdir.



Şekil 16. Ses üretimi için sayısal model

Sesli ses için sürekli durum sesli bölge transfer fonksiyonu şöyle ifade edilir.

$$V(z) = \prod_{k=1}^4 \frac{1 - 2e^{-\alpha_k T} \cos(2\pi F_k T) + e^{-2\alpha_k T}}{1 - 2e^{-\alpha_k T} \cos(2\pi F_k T)z^{-1} + e^{-2\alpha_k T}z^{-2}} \quad (31)$$

Bu denklem, sayısal bir kaskat rezanatorlerini tanımlar. Bunlar sıfır frekansta birim kazançlıdır. Ses genlikleri yalnız A_v genlik kontrolüne bağlıdır. İlk üç formant frekansı F_1 , F_2 ve F_3 zamanla değişmektedir. F_4 yaklaşık olarak 10kHz'lik bir örnekleme frekansında 4000Hz'lik bir sabit değerdir. Formant band genişliği α_k sabit olup sesin ortalama değeridir. Sessiz ses için sürekli durum transfer fonksiyonu şöyle ifade edilir.

$$V(z) = \frac{(1 - 2e^{-\beta T} \cos(2\pi F_p T) + e^{-2\beta T})(1 - 2e^{-\beta T} \cos(2\pi F_z T)z^{-1} + e^{-2\beta T}z^{-2})}{(1 - 2e^{-\beta T} \cos(2\pi F_p T)z^{-1} + e^{-2\beta T}z^{-2})(1 - 2e^{-\beta T} \cos(2\pi F_z T) + e^{-2\beta T})} \quad (32)$$

Burada F_p , düzgünleştirilmiş logaritmik tayfin en büyük piki olarak seçilir. Yaklaşık olarak 1000Hz'in üzerindedir. F_z empirik olarak şöyle formüle edilir.

$$F_z = (0.0065F_p + 4.5 - \Delta)(0.014F_p + 28) \quad (33)$$

$$\Delta = 20 \log_{10} \left| H \left[e^{j2\pi F_p T} \right] \right| - 20 \log_{10} \left| H \left[e^{j0} \right] \right| \quad (34)$$

Formant ve perde parametreleri lineer bir seviyelendirici ile seviyelendirilir. Bu doğrusal seviyelendirici her parametre aralığına ayarlanabilmektedir. Genlik parametreleri logaritmik bir seviyelendirici ile seviyelendirilir. Tablo 2'de bunun için bazı özet değerler verilmiştir.

Tablo 2. Formant kodlayıcılar için hesaplanan bazı değerler [9].

Parametre	Gereken bit/örnek
τ	6
F_1	3
F_2	4
F_3	3
$\text{Log}[A_v]$	2

Sesli /sessiz geçişleri uygun bir biçimde gösterebilmek için saniyede 100 adet ilave bir bit sesli/sessiz parametresi için göndermek gerekmektedir.

1.6. Sesin Doğrusal Öngörümlü Kodlanması

Sesin analiz tekniklerinden en güçlü olanlarından biri doğrusal öngörümlü analiz yöntemidir. Bu yöntem; perde, formant, tayf ve sesli bölge fonksiyonları gibi ana ses parametrelerinin kestirimi için üstün bir tekniğe sahiptir. Bu yöntemin önemi hem ses parametrelerinin oldukça doğru biçimde kestirilmesi ve hem de hesaplama hızının yüksek olmasındandır.

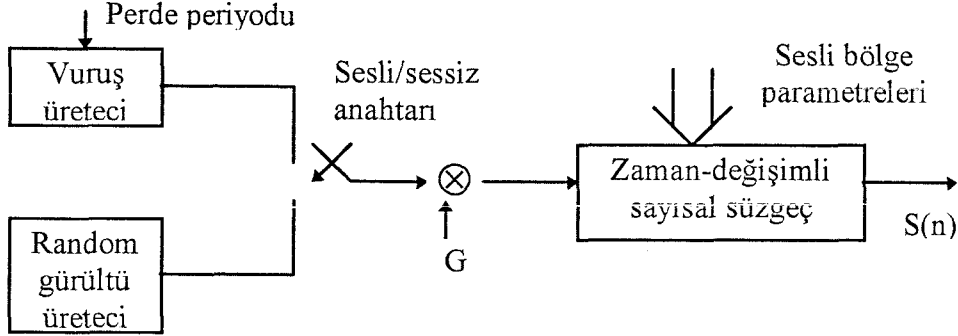
Doğrusal öngörümlü analizin arkasındaki ana fikir bir ses örneğinin, geçmiş ses örneklerinin doğrusal bir birleşimi olarak yaklaştırılmasıdır. Gerçek ses örnekleri ile doğrusal öngörülülerden birisi arasındaki karesel farkların toplamı min. yapılarak tek bir kestirici katsayılar kümesi belirlenebilir. Doğrusal öngörümlü yöntem, doğrusal zaman değişimli sistemi karakterize eden parametrelerin kestirimi için uygun bir yöntemdir. Doğrusal kestiriciler, ses dalgasının sayısal gösteriminde gerekli olan bit oranını azaltmak için kullanılabilir. Doğrusal öngörüm teknikleri uzun bir süredir mühendislik alanında kullanılmaktadır. Bunlar kontrol alanlarında, informasyon teorisinde, sistem kestirimi ve sistem tanımada kullanılmıştır. Ses uygulamalarında değişik doğrusal öngörümlü analiz formülasyonları şunlardır. Burada sadece ilk iki yöntemin benzerlikleri, farklılıkları ve detayları ele alınacaktır.

- 1) Kovaryans yöntemi [10].
- 2) Özilişki formülasyonu [11].
- 3) Latis yöntemi [12].
- 4) Ters süzgeç formülasyonu [13].
- 5) Spektrum kestirim formülasyonu [12].
- 6) Max. komşuluk formülasyonu [14].
- 7) İç çarpım formülasyonu [13].

1.6.1. Doğrusal Öngörüm Analizin Ana Prensipleri

Doğrusal öngörüm modelinin kısmi formu şekil 17’de gösterilmiştir. Modeldeki sayısal süzgecin sürekli-durum sistem fonksiyonu şöyledir.

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (35)$$



Şekil 17. Ses üretim modelinin blok diyagramı

Sistem, sesli ses için bir vuruş dizisiyle sessiz ses için ise bir random gürültü kaynağıyla uyarılır. Böylece bu sistemin parametreleri; sesli/sessiz sınıflandırması, sesli ses için perde periyodu, G kazanç parametresi ve sayısal süzgecin $\{a_k\}$ katsayılarıdır. Bu parametreler zamanla yavaş olarak değişmektedir. Eğer süzgeç derecesi p yeterince yüksek ise tüm-kutup modeli bütün sesin iyi bir gösterimini sağlar. Modeldeki $S(n)$ ses örnekleri $U(n)$ uyarma ile ilişkilidir.

$$S(n) = \sum_{k=1}^p a_k S(n-k) + GU(n) \quad (36)$$

Sistemin doğrusal katsayıları α_k olmak üzere o sistemin çıkışı;

$$S'(n) = \sum_{k=1}^p \alpha_k S(n-k) \quad (37)$$

ile verilir. P . dereceden doğrusal kestirici sistemin polinom ifadesi şöyledir.

$$p(z) = \sum_{k=1}^p \alpha_k z^{-k} \quad (38)$$

Kestirici hatası $e(n)$ şöyle ifade edilir.

$$e(n) = S(n) - S'(n) = S(n) - \sum_{k=1}^p \alpha_k S(n-k) \quad (39)$$

(31) eşitliğinden görüldüğü üzere kestirim hata kaynağı sistemin çıkışıdır. Sistemin çıkışının transfer fonksiyonu şöyledir;

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k} \quad (40)$$

Eğer $\alpha_k = a_k$ ise $e(n) = GU(n)$ olur. Böylece $A(z)$ kestirim hata süzgeci $H(z)$ sistemin ters süzgeci olmaktadır.

$$H(z) = \frac{G}{A(z)} \quad (41)$$

Doğrusal öngörüm analizin ana problemi kestirici katsayıların $\{\alpha_k\}$ kümesini doğrudan ses işaretinden belirlemektir. Ana yaklaşım, ses dalgasının kısa bölümleri üzerinde ortalama karesel hatayı min. yapan kestirici katsayılar kümesini bulmaktır. Kısa-zaman ortalama kestirim hatası şöyle ifade edilir.

$$E_n = \sum_m e_n^2(m) \quad (42)$$

$$E_n = \sum_m (S_n(m) - S_n'(m))^2 \quad (43)$$

$$E_n = \sum_m \left(S_n(m) - \sum_{k=1}^p \alpha_k S_n(m-k) \right)^2 \quad (44)$$

Burada $S_n(m)$ ses parçası (segmentidir). Öyleki bu segment, ses örnekleri civarında seçilir. E_n 'i sıfır yapan α_k 'yı $\partial E_n / \partial \alpha_k = 0$; $k=1, 2, 3, \dots, p$ yaparak bulabiliriz.

$$\sum_m S_n(m-i)S_n(m) = \sum_{k=1}^p \alpha_k' \sum_m S_n(m-1)S_n(m-k), \quad 1 \leq i \leq p \quad (45)$$

Buradaki α_k' , E_n değerini min. yapan α_k değerleridir. Bir tanım yapılırsa;

$$Q_n(i, k) = \sum_m S_n(m-i)S_n(m-k) \quad (46)$$

olur ve bunu kullanarak şu ifade bulunur.

$$\sum_{k=1}^p \alpha_k Q_n(i, k) = Q_n(i, 0), \quad i = 1, 2, 3, \dots, p \quad (47)$$

$$E_n = \sum_m S_n^2(m) - \sum_{k=1}^p \alpha_k \sum_m S_n(m)S_n(m-k) \quad (48)$$

$$E_n = Q_n(0, 0) - \sum_{k=1}^p \alpha_k Q_n(0, k) \quad (49)$$

Böylece toplam min. hata sabit bir bileşenden oluşmaktadır. Bu bileşen ise kestirici katsayılarına bağlıdır. Optimum kestirici katsayılarını çözmek için ilk olarak $Q_n(i, k)$ değerleri $1 \leq i \leq p$ ve $0 \leq k \leq p$ aralığında hesaplanmalıdır. Toplamın sınırları ve ses dalgası segmenti $S_n(m)$ 'in tanımından dolayı iki ana yaklaşım vardır.

1.6.2. Özilişki Yöntemi

Toplamın sınırlarını belirlemek için bir yaklaşım; dalga segmentinin $0 \leq m \leq N-1$ aralığı dışında sıfır olduğunu kabul etmektir.

$$S_n(m) = S(m+n)W(m) \quad (50)$$

Burada $W(m)$ sonlu uzunluklu penceredir. Genellikle Hamming penceresi kullanılır $0 \leq m \leq N-1$ aralığı dışında sıfırdır. P . dereceden kestirici için $e_n(m)$ $0 \leq m \leq N-1+p$ aralığında sıfırdan farklı olacaktır. Bu durum için;

$$E_n = \sum_{m=0}^{N+p-1} e_n^2(m) \quad (51)$$

ifadesi kullanılır. Basit olarak şu ifade yazılır.

$$Q_n(i, k) = \sum_{m=0}^{N-1-(i-k)} S_n(m) S_n(m+i-k), \quad 1 \leq i \leq p, 0 \leq k \leq p \quad (52)$$

(i-k) için özilişki fonksiyonu şöyle ifade edilir.

$$Q_n(i, k) = R_n(i-k) \quad (53)$$

$$R_n(k) = \sum_{m=0}^{N-1-k} S_n(m) S_n(m+k) \quad (54)$$

(23) ifadesi yeniden yazılabilir.

$$\sum_{k=1}^p \alpha_k R_n(|i-k|) = R_n(i), \quad 1 \leq i \leq p \quad (55)$$

Benzer biçimde min. ortalama karesel kestirim hatası şöyle ifade edilebilir.

$$E_n = R_n(0) - \sum_{k=1}^p \alpha_k R_n(k) \quad (56)$$

(55)'deki ifade matris biçiminde gösterilebilir.

$$\begin{bmatrix} R_n(0) & R_n(1) & R_n(2) & \dots & R_n(p-1) \\ R_n(1) & R_n(0) & R_n(1) & \dots & R_n(p-1) \\ R_n(2) & R_n(1) & R_n(0) & \dots & R_n(p-3) \\ \dots & \dots & \dots & \dots & \dots \\ R_n(p-1) & R_n(p-2) & R_n(p-3) & \dots & R_n(0) \end{bmatrix}_{P \times P} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \dots \\ \alpha_p \end{bmatrix}_{P \times 1} = \begin{bmatrix} R_n(1) \\ R_n(2) \\ R_n(3) \\ \dots \\ R_n(p) \end{bmatrix}_{P \times 1} \quad (57)$$

$P \times P$ özilişki matrisi Toeplitz matrisidir. Simetrik, ve diagonal eksen boyunca verilen bütün değerler birbirine eşittir.

1.6.3. Kovaryans Yöntemi

Ses segmentini tanımlamada ve toplamın sınırlarını belirlemede ikinci ana yaklaşım ortalama karesel hatayı sabitleştirmektir. Hata hesaplanır ve daha sonra $Q_n(i,k)$ hesaplanmasının etkisi düşünülür.

$$E_n = \sum_{m=0}^{N-1} e_n^2(m) \quad (58)$$

$$Q_n(i,k) = \sum_{m=0}^{N-1} S_n(m-i)S_n(m-k), \quad 1 \leq i \leq p, \quad 0 \leq k \leq p \quad (59)$$

Toplamın indeksini değiştirirsek;

$$Q_n(i,k) = \sum_{m=-k}^{N-k-1} S_n(m)S_n(m+k-i), \quad 1 \leq i \leq p, \quad 0 \leq k \leq p \quad (60)$$

$$\sum_{k=1}^p \alpha_k Q_n(i,k) = Q_n(i,0), \quad i = 1, 2, 3, \dots, p \quad (61)$$

ifadelerini elde etmiş oluruz. Bu en son ifade matris biçiminde yeniden yazılabilir.

$$\begin{bmatrix} Q_n(1,1) & Q_n(1,2) & Q_n(1,3) & \dots & Q_n(1,p) \\ Q_n(2,1) & Q_n(2,2) & Q_n(2,3) & \dots & Q_n(2,p) \\ Q_n(3,1) & Q_n(3,2) & Q_n(3,3) & \dots & Q_n(3,p) \\ \dots & \dots & \dots & \dots & \dots \\ Q_n(p,1) & Q_n(p,2) & Q_n(p,3) & \dots & Q_n(p,p) \end{bmatrix}_{P \times P} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \dots \\ \alpha_p \end{bmatrix}_{P \times 1} = \begin{bmatrix} Q_n(1,0) \\ Q_n(2,0) \\ Q_n(3,0) \\ \dots \\ Q_n(p,0) \end{bmatrix}_{P \times 1} \quad (62)$$

$P \times P$ matrisi simetrik ancak Toeplitz değildir. Diagonal eksenindeki elemanlar arasında şu ilişki vardır.

$$Q_n(i+1, k+1) = Q_n(i, k) + S_n(-i-1)S_n(-k-1) - S_n(N-1-i)S_n(M-1-k) \quad (63)$$

$\{Q_n(i,k)\}$ değerleri kovaryans matrisinin özelliklerine sahiptir. Özilişki yöntemi için ses işareti N noktalı bir pencere ile pencerelenir. $Q_n(i,k)$ değerleri kısa-zaman özilişki fonksiyonu kullanılarak elde edilir. Sonuç matrisi (ilişki matrisi) Toeplitz olup, kestirici katsayıların çözümü için bir yoldur.

Kovaryans yöntemi için ise işaretin $-p \leq n \leq N-1$ aralığında olduğu kabul ediliyor ve sonuç ilişki matrisi simetrik olup ancak Toeplitz değildir.

1.7. Değişik Ses Parametreleri Arasındaki İlişkiler

Sesle ilgili diğer parametre kümeleri doğrudan doğruya doğrusal öngörümlü katsayılarından (DÖK) elde edilir. Kestirici polinomun kökleri şöyledir.

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k} = \prod_{k=1}^p (1 - z_k z^{-1}) \quad (64)$$

Bu ifadenin z_i kökleri $\{i=1, 2, 3, \dots, p\}$ $A(z)$ 'nin bir eşdeğer ifadesidir. Kökler z -uzayından s -uzayına şu şekilde dönüştürülebilir.

$$z_i = e^{s_i T} \quad (65)$$

Burada $s_i = \sigma_i + j\Omega_i$ 'dir. Eğer $Z_i = Z_{ir} + jZ_{ii}$ ise şu ifadeler bulunur.

$$\Omega_i = \frac{1}{T} \tan^{-1} \left(\frac{Z_{ii}}{Z_{ir}} \right) \quad (66)$$

$$\sigma_i = \frac{1}{2T} \log(Z_{ir}^2 + Z_{ii}^2) \quad (67)$$

(66) ve (67) denklemleri DÖK analiz sistemlerinin formant analizi için kullanılır.

Cepstrum : DÖK katsayılarına diğer bir alternatif ise LPC sisteminin birim vuruş cevabının cepstrumudur. Eğer $H(z)$ DÖK sistemin transfer fonksiyonu ise $h(n)$ sistemin birim vuruş cevabı ve $h'(n)$ kompleks cepstrumdur. Böylece $h'(n)$ şöyle ifade edilir.

$$h'(n) = \alpha_n + \sum_{k=1}^{n-1} \frac{k}{n} h'(k) \alpha_{n-k}, \quad 1 \leq n \quad (68)$$

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n} = \frac{G}{1 - \sum_{k=1}^p \alpha_k z^{-k}} \quad (69)$$

Transfer fonksiyonu (69) ile verilen tüm-kutup sistemin birim vuruş cevabı $h(n)$, DÖK katsayılarından çözülür.

$$h(n) = \sum_{k=1}^p \alpha_k h(n-k) + G\delta(n), \quad 0 \leq n \quad (70)$$

Burada $h(n)$ 'nin $n < 0$ için 0 olduğu farz edilir. G uyarma genlik değeridir. Süzgecin birim vuruş cevabının özilişki fonksiyonu şöyle tanımlanır.

$$R'(i) = \sum_{n=0}^{\infty} h(n)h(n-i) = R'(i) \quad (71)$$

$$R'(i) = \sum_{k=1}^p \alpha_k R'(|i-k|), \quad 1 \leq i \quad (72)$$

$$R'(0) = \sum_{k=1}^p \alpha_k R'(k) + G^2 \quad (73)$$

(71) ve (72) denklemleri kestirici katsayılarından $R'(i)$ 'yi belirlemek için kullanılabilir. Bunun terside yapılabilmektedir.

Ters süzgecin birim vuruş özilişki fonksiyonu kullanılarak $a(n)$ değerleri;

$$R_a(i) = \sum_{k=0}^{p-i} a(k)a(k+i), \quad 0 \leq i \leq p \quad (74)$$

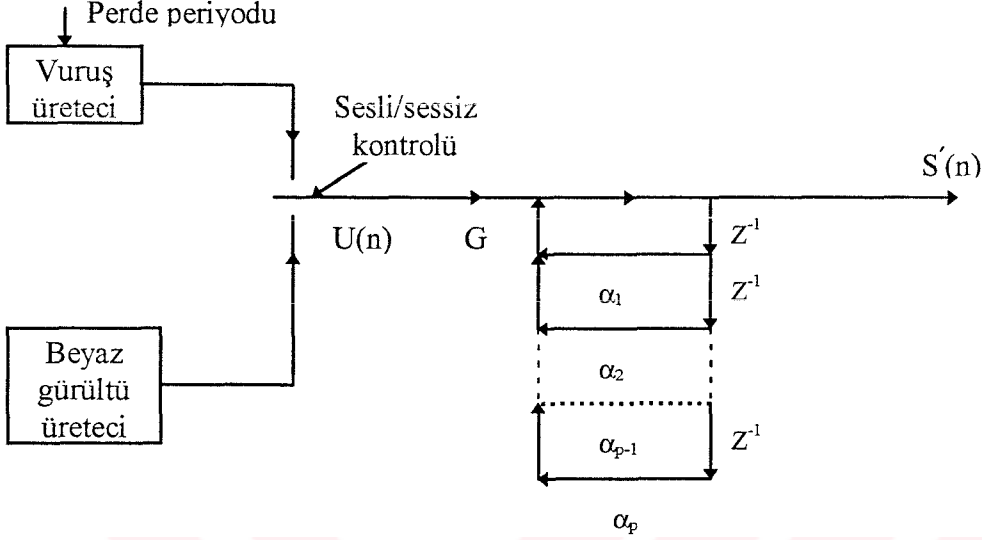
$$a(n) = \delta(n) - \sum_{k=1}^p \alpha_k \delta(n-k) \quad (75)$$

ile elde edilmektedir. Özilişki yöntemi için kestirici katsayıları parcor (yansıma) katsayılarından da elde edilebilir.

1.8. Doğrusal Öngörümlü Parametrelerden Sesin Sentezi

Ses, doğrusal öngörümlü analiz parametrelerinden birkaç farklı yolla sentezlenebilir. En basit yol bir sistem kurmaktır. Bu sistem analizde kullanılan parametrik gösterimle aynıdır. Şekil 18 böyle bir ses sentezleyicinin blok diyagramını göstermektedir. Sentezleyici tarafından ihtiyaç duyulan zaman değişimli kontrol parametreleri perde periyodu, sesli/sessiz anahtarı, kazanç veya rms ses değeri ve p kestirici katsayılarıdır. Vuruş üretici her perde periyodunun başında birim genlikli bir uyarma kaynağı olarak davranır. Bu sesli ses bölgeleri

için geçerlidir. Sessiz ses bölgeleri için ise beyaz gürültü genaratörü, 0 ortalama değerli birim standart sapmalı random örnekl bir uyarma kaynağı olarak davranır. İki kaynağın seçimi sesli/sessiz kontrolü ile yapılır.



Şekil 18. Doğrusal öngörümlü sentezleyicinin blok diyagramı

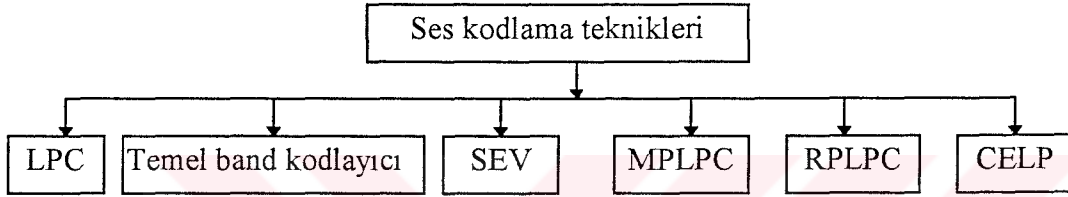
G kazanç kontrolü uyarmanın genliğini belirler. Sentetik ses örnekleri şu forma sahiptir.

$$S'(n) = \sum_{k=1}^p \alpha_k S'(n-k) + GU(n) \quad (76)$$

Her çıkış örneğini üretmek için toplam olarak p adet çarpma ve p adet toplama işlemi gerekmektedir. Sentez parametreleri zamanla değişebilmektedir. Parametreler genellikle sesli ses bölgesi boyunca belli bir aralıkta kestirilmesine rağmen, kontrol parametreleri her periyodun başında değiştirilir. Sessiz ses için çerçeve başına bu değerler bir kez değiştirilir. Sentezleyici DÖK sistemlerinin genişçe kullanılması, onun basit ve uygulama kolaylığından kaynaklanmaktadır.

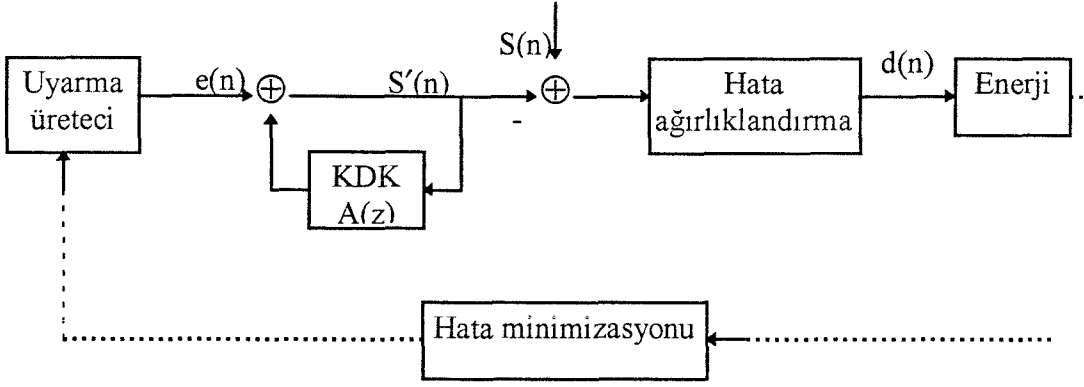
2. YAPILAN ÇALIŞMALAR

Bu çalışmada analiz-sentez ses kodlayıcılarına ilişkin kodlayıcı türleri incelenmiştir. Bu sınıfa giren kodlayıcılar kod-uyarmalı doğrusal öngörümlü kodlayıcılar (KUDÖK), çoklu darbe uyarmalı doğrusal öngörümlü kodlayıcılar ve düzenli darbe uyarmalı doğrusal öngörümlü kodlayıcılardır. Ses kodlama teknikleri şekil 19'da verilmiştir. Bütün bu kodlayıcılar, bit hızını azaltmak ve ses kalitesini artırmak için ses işaretinden elde edilebilecek bütün bilgileri kullanmaktadırlar.



Şekil 19. Ses kodlama çeşitleri [15].

Burada analiz-sentez işlemleri tartışıldı ve analiz-sentez tekniğine dayalı DÖK kodlayıcıları incelenmiştir. Düzenli darbe uyarmalı doğrusal öngörümlü kodlayıcılarda olduğu gibi analiz-sentez esaslı kodlayıcılarda uyarmalar sesli/sessiz sınıflandırılması ve perde belirlemesine gerek duymamaktadır [15]. Analiz-sentez kodlayıcılar ya yüksek bit hızlı uyarma modeli etkili perde uyarmalı doğrusal öngörümlü kodlayıcılar yada artık işaretin kodlanması için blok kodlama tekniğini etkili biçimde kullanan dalgabiçimli uyarmalı doğrusal öngörümlü kodlayıcılar olarak düşünülebilir. Bütün analiz-sentez kodlayıcılar uyarma işareti için genellikle 2 ile 6 adet gibi az sayıda parametre kullanırlar. Analiz-sentez işleminin iki önemli bileşeni hata ağırlıklı fonksiyon ve uzun dönem kestiricisidir. Hata ağırlıklı süzgeç kodlanmış gürültüyü maskeleyerek için kullanılır. Uzun dönem kestiricisi ise sesli sesin perde periyodunu belirlemek için kullanılır. Analiz-sentez kodlayıcı için analiz işleminin blok diyagramı şekil 20'de gösterilmiştir.



Şekil 20. Analiz-sentez doğrusal öngörümlü kodlayıcıda kullanılan analiz işleminin blok diyagramı

Analiz-sentez kodlayıcılarında uyarma üretici K farklı uyarma kaynağı, $e_k(n)$ üretebilmektedir. Genellikle bu K değeri (64-1024) aralığında seçilmektedir. Analiz işlemi K ayrı $S'_k(n)$ ses işaretini üretir. Bunu orijinal sestten çıkartır ve hata işaretindeki ağırlıklı enerjiyi hesaplar. Analizatör optimal uyarma kaynağını seçmek için K farklı sentez işlemi yapmaktadır. Uyarma kaynakları çoklu darbe uyarmalı doğrusal öngörümlü kodlayıcı için darbe uyarma, kendi kendine uyarmalı kodlayıcılar için perde uyarma ve KUDÖK kodlayıcılar için ise kod uyarma olarak adlandırılır.

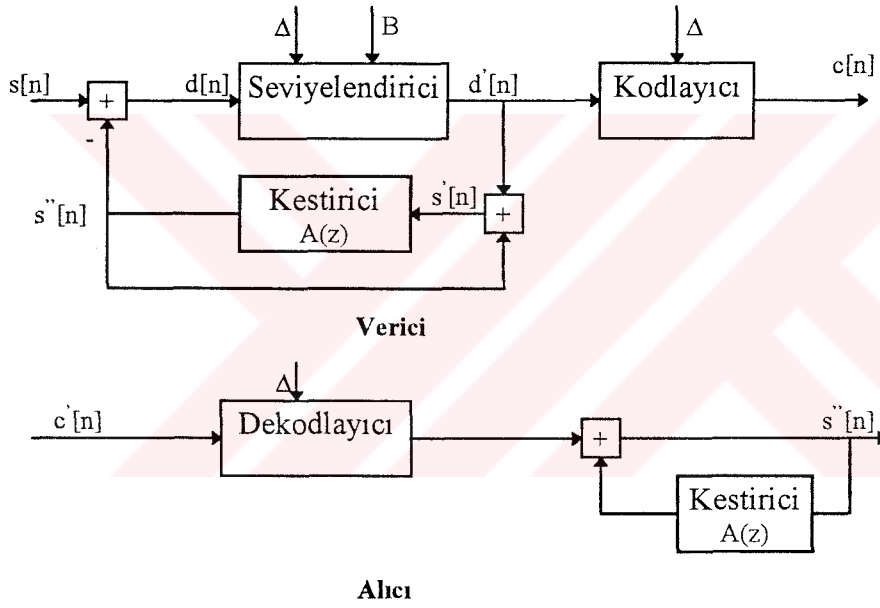
2.1. Sabit Kestirimli Dalgabiçimli Kodlama [15]

Vokal tırak etkilerini açıklamak için ses işaretini iki açıdan ele almak lazımdır. Bunlar uzun dönem ve kısa dönemdir. Zamandan bağımsız uzun dönem açısından bakıldığında sesin özellikleri ve kodlama için adaptif olmıyan stratejileri bize gösterir. Kısa dönem açısından bakıldığında zamanla sesin özelliklerinin yavaşça değişmesine vocal tract'ın mekanik özellikleri sebep olmaktadır.

Dalgabiçimli gösterim teorisi adaptif olmıyan diferansiyel kodlama esasına dayanmaktadır. Diferansiyel kodlama işaretin kendilerinden ziyade iki işaret arasındaki farkı kodlamayı göstermektedir. Adaptif olmıyan diferansiyel kodlayıcı ile açıklanabilecek olan sesin ilk uzun dönem özellikleri bitişik örnekler arasındaki oldukça yüksek olan ilişkidir. Özellikle sesli ses işaretinin komşu örnekleri arasındaki bu yüksek ilişkiden dolayı, her ses

örneği genellikle önceki ses örneklerinden veya kodlanmış ses örneklerinden kestirilmektedir. Vokal tırak süzgecinin yaklaşık doğrusal özelliğinden ötürü gerekli kestirim doğrusal bir süzgeç ile yapılabilir.

Diferansiyel kodlamada, ses dalgasının kısa dönem fazlalığı mümkün olduğu kadar atılır. Arta kalan işaret seviyelendirilir ve kodlanır. Mevcut ses örnekleri bir veya daha çok sayıda geçmiş örneklerin bir doğrusal kombinasyonu olarak kestirilir. Şekil 21 diferansiyel darbe kod modülasyonunun (DPCM) blok diyagramını göstermektedir. Fark işareti $d[n]$ genellikle orijinal işaretin genliğinden daha düşüktür. Diferansiyel kodlayıcının en basit formu sabit bir kestirici ve sabit bir uniform seviyelendirici kullanmasıdır.



Şekil 21. Fark PCM'in blok diyagramı

Şekil 21'de $s''[n]$ kestirilen $s[n]$ değerini gösterir. Bu önceki kodlanmış değerlerden elde edilir.

$$s''[n] = \sum_{i=1}^p a_i s'[n-i] \quad (77)$$

Burada a_i değerleri $i=1, \dots, p$ doğrusal kestirici $A(z)$ 'nin katsayılarıdır. $S'[n]$ önceki kodlanmış ses örnekleridir. Kestirici $A(z)$ şöyle tanımlanır.

$$A(z) = \sum_{i=1}^p a_i z^{-i} \quad (78)$$

P kestirici katsayısıdır. Kestirici katsayıları genellikle sesin doğrusal öngörümlü analiziyle hesaplanır. Kestirim hatası minimize yapılarak a_i 'ler bulunur.

$$E = \sum_n (s[n] - s''[n])^2 \quad (79)$$

$$d[n] = s[n] - s''[n] \quad (80)$$

Kestirim hata işareti $d[n]$ olup, seviyelendirilmiş hata işareti şöyle gösterilir.

$$d'[n] = Q(d[n]) = d[n] + q_d[n] \quad (81)$$

Seviyelendirme hatası $q_d[n]$ olup seviyelendirilmiş işaret şöyledir.

$$s'[n] = s''[n] + d'[n] \quad (82)$$

$$s'[n] = s[n] + q_d[n] \quad (83)$$

Bu şunu gösterir ki, sistemin çıkış seviyelendirme hatası kestirim hata işaretinin seviyelendirme hatasıyla aynıdır.

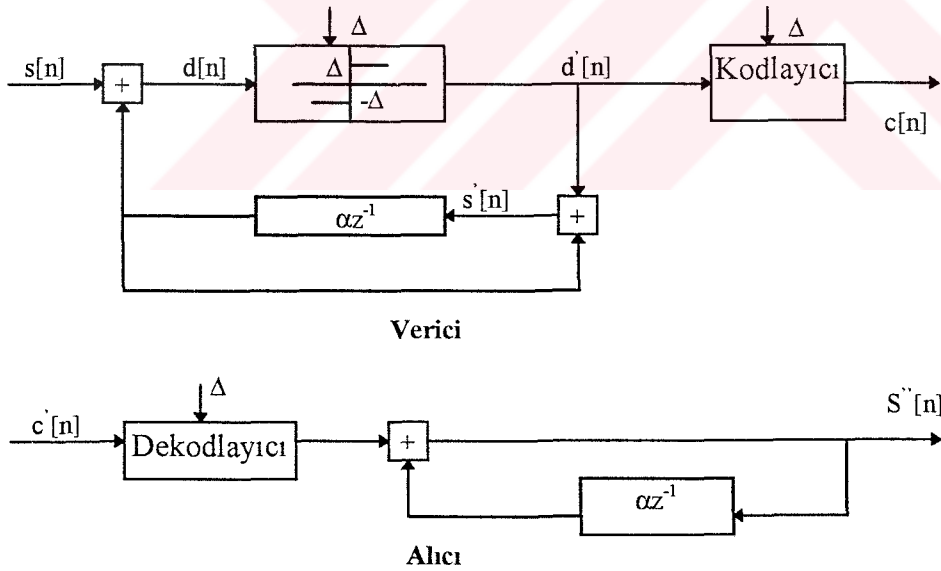
Adaptif Seviyelendirmeli DPCM : Sabit seviyelendirmeli DPCM aynı sayıda bit/örnek'li PCM ile karşılaştırıldığında işaret/gürültü oranında ortalama olarak 6dB'lik bir iyileştirme sağlamaktadır. APCM'de olduğu gibi adaptif seviyelendiriciler kodlayıcı performansında iyileştirme elde ettikleri için kullanılmaktadırlar. ADPCM'nin parametreleri

kestirici ve seviyelendirici parametrelerdir. M uzunluklu dikdörtgen pencere kullanılır, varyans her M örnek için güncelleştirilir.

Delta Modülasyonu [15] : Delta modülasyonu 1. dereceden kestiricili DPCM'nin özel bir halidir. Kaba seviyelendirmeyi kompanze etmek ve kodlanmış işaretin kalitesini korumak için delta modülasyonu giriş işaretinin örnekleme hızı olan f_{DM} Nyquist frekansının birkaç katından daha büyük seçilir.

$$f_{DM} = 2Rf_N \quad (84)$$

Nyquist hızı f_N olup R ise örnekleme oranı üzerinde bir değerdir. Böyle yüksek örnekleme hızı, istenilen hızda analog işaretin örneklenmesiyle elde edilebildiği gibi ses işaretinin aradeğerlenmesiyle de elde edilebilir. Böyle yüksek bir örnekleme hızı örneklenmiş ses işaretindeki örnekler arasındaki ilişkinin yüksek olması sonucunu doğurmaktadır. Şekil 22 delta modülasyonunun blok diyagramını gösterir.



Şekil 22. Delta modülatörünün blok diyagramı

Delta modülatöründeki α 1'den küçük seçilmek üzere, kestiricinin formu şöyledir.

$$P(z) = \alpha z^{-1} \quad (85)$$

Delta modülatörünün iyi tarafı basit yapıda olması, bit hızının da delta modülasyonu giriş işaretinin örnekleme hızına eşit olmasıdır.

Doğrusal Delta Modülasyonu [15]: Doğrusal delta modülasyonu (LDM) sabit iki seviyeli bir seviyelendiricinin kullanıldığı delta modülasyonu çeşitidir. LDM'nin eğimi T örnekleme periyodu olmak üzere Δ/T oranıyla verilir. Eğim Δ ve T değiştirilerek kontrol edilir. T'nin azalması örnekler arası ilişkinin artması ve kestirim kazancının artmasına sebep olur.

Delta modülasyonunda iki tip bozulma vardır. Granular gürültü ve eğim yükü gürültüsüdür. Granular gürültü, DM çıkışının hedef ses dalgası değeri etrafındaki osilasyonu sonucu oluşur. Granular gürültüyü azaltmak için adım sayısı değerini küçük seçmek gerekmektedir. Adım sayısı değerini büyük seçmek eğim yükü gürültüsünden korunmak için gereklidir. Bu yüzden adım sayısının seçimi LDM'nin performansı için karışık ve kritik bir öneme sahiptir.

Adaptif Delta Modülasyonu [15] : Adaptif delta modülasyonunda (ADM) adım sayısı her iki gürültüyü azaltacak biçimde dinamik olarak değiştirilir. Kodlayıcının basitliğini ve kenar bilgisini iletmeyi sağlayabilmek için geribeslemeli(feedback) adaptasyon tercih edilir. ADM Jayant tipi seviyelendirici kullanır. Bu seviyelendiricinin adım sayısı adaptasyon kuralı şöyledir.

$$\Delta[n] = M[n]\Delta[n-1] \quad (86)$$

$M[n]$ iki değerden birini alır. Burada $c[n]$ n. andaki çıkış kod değeri olmak üzere

$$M[n] = \begin{cases} M_1, & c[n] = c[n-1] = c[n-2] \\ M_2, & \text{d.d.} \end{cases} \quad (87)$$

ifadesi kullanılır. Buradaki $c[n]=c[n-1]=c[n-2]$ durumları işaretin ardışıl artmaları veya azalmalarını gösterir. Genelde $M_1M_2 \leq 1$ olacak şekilde $M_1 > 1$ ve $M_2 < 1$ olarak seçilir. ADM'nin parametre değerleri tablo 3'te verilmiştir.

Tablo 3. ADM'nin parametreleri

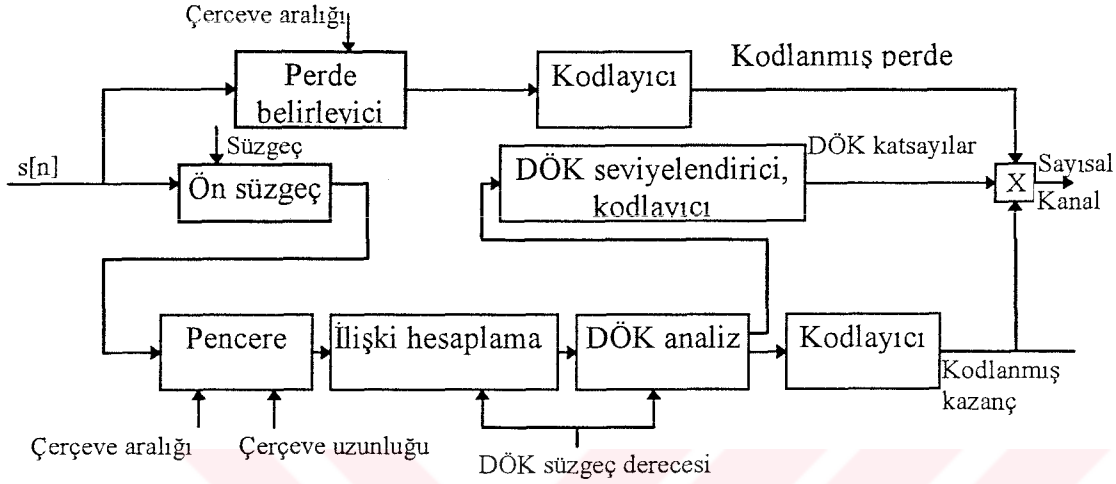
Parametreler	Aralık	Tipik değer
Aşırı örnekleme oranı (R)	1-8	4
Kestirici parametresi (α)	0.1-0.99	0.85
Max. adım sayısı çarpanı (M_1)	1-10	1.2
Min. adım sayısı çarpanı (M_2)	0-0.99	0.7
Min. adım sayısı (Δ_{min})	10-500	20

2.2. Perde Uyarmalı Doğrusal Öngörümlü Kodlayıcı [15]

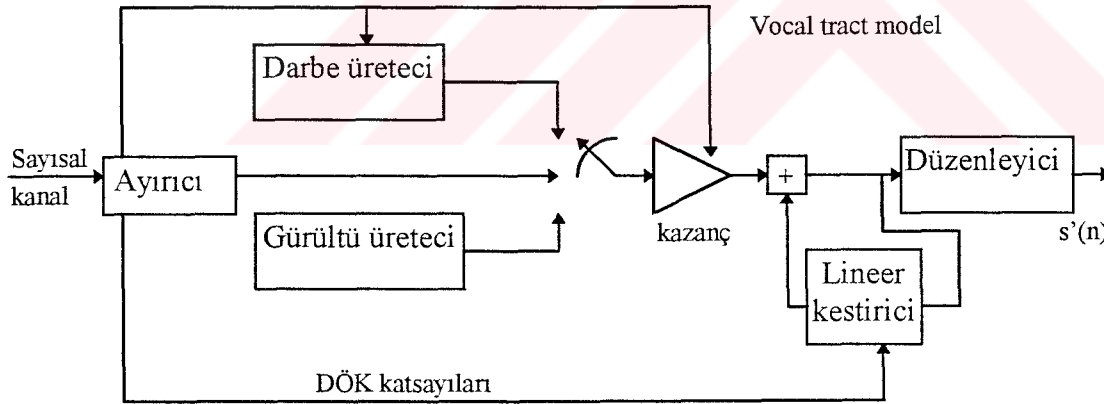
Doğrusal öngörümlü kodlama ses işaretleri için en popüler kodlama tekniklerinden birisidir. Bu ses kodlayıcıların değişik tiplerinde değişiklik göstererek kullanılan bir tekniktir. Bu teknik perde uyarmalı, sesli ses uyarmalı, dalgabiçimli, analiz sentez ve frekans domeni kodlayıcılarında kullanıldı. Perde uyarmalı DÖK ses kodlayıcıların gelişiminde önemli bir rol oynamıştır. Çünkü ses için kullanılan DÖK'ün ilk çeşiti idi. Perde uyarmalı kodlayıcıların avantajı düşük bit hızlarında çalışabilmeleridir. Bunların eksikleri ise kodlayıcının kalitesini sınırlandırmasıdır. Doğrusal öngörümlü analiz teknikleri DSP mikro işlemcileri ve diğer uygulama teknikleriyle uygunluk göstermektedir.

Perde Uyarmalı DÖK: Perde uyarmalı DÖK tamamen parametrik bir kodlayıcıdır. Kodlanmış ses, ses sentez modelinin zaman değişimli parametreleriyle karakterize edilir. Bu sentez modeli uyarma modeli ve vocal tract modeli diye iki bileşene sahiptir. DÖK teknikleri bu sentezleyicide vocal tract modeli parametrelerini elde etmek için kullanılır. Bütün DÖK'lerde vocal tract doğrusal bir zaman değişimli süzgeç ile modellenir. Doğrusal süzgeçnin parametreleri ses işaretinin doğrusal öngörümlü analiziyle elde edilir. Perde uyarmalı DÖK'lerde parametreler perde belirleyicisi kullanılarak bulunur. Şekil 23 ve 24'te

perde uyarmalı DÖK analiz ve sentezleyicisinin blok diyagramları gösterilmektedir. Vericide vokal tırak model parametreleri ve uyarma modeli parametreleri üretilir, seviyelendirilir, kodlanır ve gönderilir. Alıcıda ise kodlanmış parametreler yeniden elde edilir ve kodlanmış ses sentezi için kullanılır.



Şekil 23. Perde uyarmalı DÖK vericisi



Şekil 24. Perde uyarmalı DÖK alıcısı

Perde uyarmalı DÖK vericisi iki farklı analiz tipi uygular. Uyarma analizi (perde belirleyici) ve vokal tırak analizi (DÖK analizi). Perde analizi doğrudan giriş işaret $s[n]$ üzerinde çalışır. Perde belirleyicinin çıkışı her çerçeve için sesli/sessiz kararına ve sesli çerçeveler için de perde değerine karar verir. DÖK analizinde, ses öncelikle bir ön süzgeçten

geçirilir. Bu süzgecin amacı ses tayfının dinamik alanını azaltmaktır. Bu DÖK analiz algoritmalarının sayısal özelliklerini iyileştirir. Sonra çerçeveler pencerelenir ve ilişki analizine uygulanır. Her çerçeve için DÖK analizinin çıkışında bir kazanç parametresi ve DÖK süzgeç parametresi kümesi oluşur. Bunlar iletim veya depolamak için seviyelendirilir, kodlanır ve çoğullanır.

Basit ses sentezleyici uyarma işaretinden oluşur. Uyarma üretici; darbe ve gürültü üreticilerinden , sesli/sessiz ayırıcı anahtardan ve kazançtan ibarettir. Alıcıda sayısal kanalın girişinden alınan data; perde, kazanç ve DÖK katsayılarına ayrılır. Perde bilgisi darbe üreticinin darbe hızını kontrol etmek için kullanılır. Kazanç verisi uyarma işaretinin genliğini kontrol eder. DÖK katsayıları vokal tırak süzgecini kontrol etmek içindir. Vokal tırak mekanik bir sistemdir. Ses işaretinin yavaşça değişen karakteristikleri vokal tırak süzgecinin yavaşça değişen parametreleriyle modellenir.

Vokal Tırak Model: Vokal tırak modeli, vokal tırak süzgeci ve ters enson süzgecine sahiptir. Vokal tırak IIR süzgeci olarak modellenir. Bu süzgeç p parametresiyle karakterize edilir. Normalde p, saniyede 8000 örnekli ses için 10-12'dir.

Pencereleme; DÖK analizi parametrelerinin önemli bir kümesidir. Bunlar kullanılan pencere tipi ve uzunluğunu içerirler. Otokorelasyon yöntemi kullanıldığında pencere ses işaretine uygulanır. Pencere kenar etkilerini azaltmak için Hamming veya Hanning pencereler kullanılır. Keskin kenarlı pencereler iyi sonuç üretirler. Pencere uzunluğu, sesli ses için birkaç perde periyodunu içerecek şekilde (20-30msn.) seçilir. Buna uyarma işaretinin vokal tırak süzgeç parametrelerinin kestirimindeki etkilerini azaltmak için gerek duyulur.

Kazanç hesaplama; DÖK modelinde kazanç parametresi sentetik ses işaretini üretmek için kullanılır. Öyleki bu sentetik ses orijinal ses işaretiyle aynı enerjiye sahip olsun. Bu DÖK süzgeçsinin birim vuruş girişine olan cevabının enerjisinin orijinal sesin enerjisine uygunlaştırılarak elde edilebilir. $R(k)$ otokorelasyon katsayıları olmak üzere kazanç ifadesi şöyledir.

$$G = \left[R(0) - \sum_{k=1}^p a(k)R(k) \right]^{1/2} \quad (88)$$

DÖK Modeli Parametrelerinin Seviyelendirilmesi: Her analiz aralığında kestirici katsayıları, perde periyodu, kazanç ve voice parametresi gönderilir. Perde periyodu, kazanç ve voice parametreleri skaler seviyelendirici kullanılarak seviyelendirilir ve kodlanır. DÖK katsayılarının doğrudan seviyelendirilmesi her katsayı için yüksek sayıda bit (8-10bit) gerektirmektedir. Kullanılan form; yansıma katsayılarını, çizgi tayf çiftlerini, H(z)'nin birim vuruş cevabının ilk p örneklerini, otokorelasyon katsayılarını ve Cepstrum katsayılarını içermektedir.

DÖK kullanarak tayf kestirimi; ses işaretleri için H(z) sentez süzgecinin frekans cevabı ses tayfının zarfını izlememize olanak sağlamaktadır. Bu frekans domeninde ortalama karesel kestirim hatasıyla açıklanabilir.

$$E(z) = (1 - A(z))S(z) \quad (89)$$

E(z) kestirim artığının z dönüşümü ve S(z) ses işaretinin z dönüşümüdür. Parseval teoremi kullanılarak ortalama karesel hata şöyle ifade edilir.

$$E = \sum_n e^2[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 d\omega \quad (90)$$

$$E = \frac{G^2}{2\pi} \int_{-\pi}^{\pi} \frac{|S(e^{j\omega})|^2}{|H(e^{j\omega})|^2} d\omega \quad (91)$$

E'nin min. olduğu bölgeler güç tayfi oranının 1'den büyük olduğu yerler ve toplam hatanın 1'den küçük olduğu bölgelerdir.

2.3. Adaptif Öngörümlü Dalgabiçimli Kodlama [15]

Bu tip kodlayıcılar iki şekilde gösterilebilir. İlki adaptif DPCM sistemlerdir. Bunlar adaptif doğrusal kestirici kullanırlar ve yüksek kodlama kazancı elde ederler. Diğeri ise dalgabiçimli uyarımlı kodlayıcılardır. Bu sistemler perde uyarımlı kodlayıcılar olarak düşünülebilir. APC'ler dalgabiçimli kodlama ve parametrik kodlama arasında hayati bir bağa sahiptir. Adaptif öngörümlü kodlayıcılar bu sınıfın en önemli üyesidir. APC'de kestirici parametreleri ya giriş sesinden yada yeniden üretilen çıkış sesinden elde edilir.

Adaptif öngörümlü kodlayıcı sınıfının diğeri bir üyesi gürültü geri beslemeli kodlayıcılardır. Gürültü geri beslemeli kodlayıcıda sistemin maskeleyme özelliği yeniden üretilen ses işaretinin gürültü miktarını min. etmek için kullanılır. Gürültü tayfindaki enerji ses tayfindaki enerjiden daha azdır ve bu gürültü maskelenebilmektedir.

Adaptif Öngörümlü Kodlama: APC'de kestirici katsayıları ses işaretinden elde edilir ve kenar bilgisi olarak gönderilir. Seviyelendirici parametrelerinin sabit olduğu düşünülür. Adaptif geri beslemeli APC'de ise kestirici parametrelerini belirlemek için kodlanmış işaret kullanılır. Hem de kestirici, bit hızını etkilemeden herhangi bir hızda güncelleştirilebilir. Böylece kestirici katsayıları her örnek için yüksek hızda güncelleştirilir. Bu geri beslemeli sistemlerde hesaplama karmaşıklığını artırır.

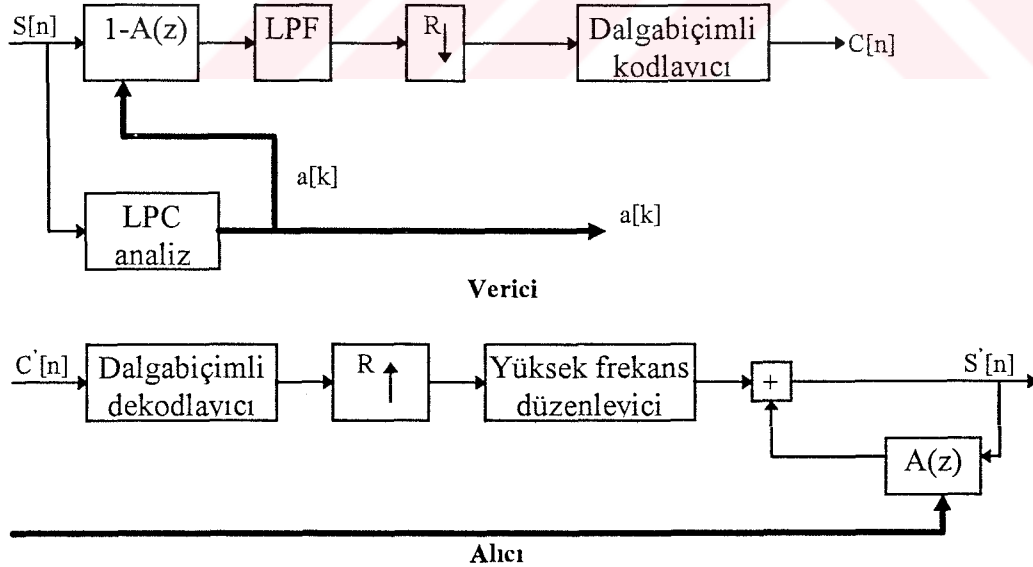
Perde Uyarımlı APC: Ses işaretini kestirmek ve kodlama kazancını artırmak için diğeri bir yol sesli sesler için perde fazlalığı kullanmaktır. Sesli ses süresi boyunca dalga yarı periyodiktir. Bu uzun dönem kestirici kullanılarak gerçekleştirilir. Kısa dönem kestirici gibi uzun dönem kestiriciside doğrusal bir kestiricidir. Kısa dönem kestiricisi komşu örnekleri kullanırken, uzun dönem kestiricisi geçmiş bir veya daha fazla perde periyodlu örnekleri kullanmaktadır. UDK'lar ses işaretinin perde fazlalıklarını uzaklaştırmak için kullanılır. UDK'lar ses spektrumunun harmonik yapısını gösterirler. Kısa zaman kestiricisi $A(z)$ ve uzun zaman kestiricisi $B(z)$ ile ifade edilmektedir.

$$A(z) = \sum_{k=1}^p a_k z^{-k} \quad (92)$$

$$B(z) = \sum_{k=-l}^l \beta_k z^{-\gamma-k} \quad (93)$$

β değerleri UDK katsayılarıdır ve kestiricinin derecesi $2l+1$ 'dir. Uzun dönem gecikmesi γ ile gösterilmiştir. Genelde $l=1$ olarak alınmaktadır.

Artık Uyarmalı LPC: Tüm APC'lerdeki ana problem uyarma işaretinin orijinal ses işaretinin hemen hemen bütün örneklerini içerecek şekilde kodlanabilmesidir. Uyarma işaretini kodlama için birkaç seviye kullanılsa bile uyarma için bit hızı doğrusal kestiricinin bit hızından daha büyüktür. LPC sentez süzgeşi, doğrudan LPC artık işaretinden elde edilen dalga ile uyarılır. Böyle bir yaklaşım sesli/sessiz ses ayırımına veya perde değerine ihtiyaç duymaktadır. Bu artık işaret perdesi enerji gibi büyük informasyona sahiptir. Şekil 25 RELP kodlayıcının blok diyagramını göstermektedir. Artık işaretin ana bandı alçak geçiren süzgeçle elde edilir. Ana band işaretin ana frekansını içermekte ve böylece Perde bilgisine sahip olunur. Diğer bir dalgabiçimli uyarmalı LPC ses uyarmalı kodlayıcıdır. Bu RELP'in basit halidir. Aralarındaki fark sesli ses uyarmalı kodlayıcı uyarma dalgabiçimli olarak artık işaret yerine ana band ses işaretini kullanmasıdır.



Şekil 25. Artık uyarmalı LPC'nin blok diyagramı

REL P sentezleyicide, ana band işareti dekodlanır ve uyarma işaretini üretmek için kullanılır. Sadece yüksek frekans bileşeninin üretilmesine ihtiyaç duyulmaktadır. Yüksek frekans bileşeni üreten yöntemler square-law gibi doğrusal olmayan özelliklere sahiptir.

2.4. Uyarma Modeli [15]

Analiz-sentez uyarma işareti blok üzerinden belirlenen bir kodlama işlemidir. Her blok için uyarma işareti farklı uyarma elemanlarının bileşimi olarak düşünülebilir.

$$e(n) = \sum_{k=1}^M \beta_k e_k(n) \quad (94)$$

Burada $e_k(n)$, k. uyarma bileşenidir. Uyarma bileşenleri üç tipten biri olabilir. Bunlar darbe, kod bloğu kaynağı ve uzun dönem kestiricinin çıkışı olabilir. Tipik KUDÖK bir kod bloğu kaynağı ve uzun dönem kestiricisi olmak üzere iki bileşene sahiptir. Her uyarma bileşeni bir γ_k indeksine ve β_k ilişkilendirme kazancı ile nitelendirilir. Verilen K uyarma kaynağı için $e_k(n)$, orijinal ses ile sentetik ses arasındaki farkı min. yapan $f_k(n)$ olarak seçilir. Optimum indeks γ_k ve ilişkilendirme kazanç değeri β_k , alıcıda uyarma kaynağını yeniden oluşturabilmek için alıcıya gönderilir.

2.5. Hata Ağırlıklandırma

Analiz-sentez doğrusal öngörümlü kodlayıcının önemli bir parametresi hata ağırlıklı süzgeçtir. Bu süzgeç kodlanmış hata işaretinin enerjisini dağıtmak için kullanılır. Hata ağırlıklı süzgeç analiz-sentez doğrusal öngörümlü kodlayıcıda Atal ve Schroeder [15] tarafından kullanıldı. Blok esaslı adaptif doğrusal öngörümlü kodlayıcının ağırlıklı süzgeci şöyle ifade edilmektedir.

$$W(z) = \frac{1 - A(z)}{1 - A(\alpha^{-1}z)} \quad (95)$$

Burada α , ağırlıklı sabit olup $0 \leq \alpha \leq 1$ aralığındadır. $A(z)$ kısa dönem kestiricisidir. $W(z)$ süzgeci ses tayfındaki çukurlardaki ve formant bölgelerindeki hataları vurgular. Tüm süzgecin birim vuruş tepkisi ise şöyledir.

$$H_w(z) = H(z)W(z) = \frac{1}{1 - A(\alpha^{-1}z)} \quad (96)$$

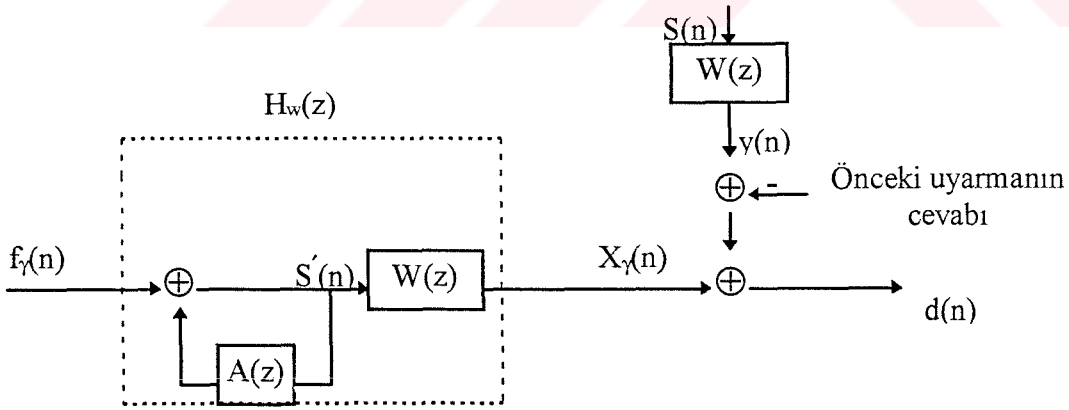
Ağırlıklı sentez süzgecinde iyi bir yaklaşıklık yapılarak şu ifade bulunur.

$$h_w(n) \approx 0, \quad n > 1 \quad (97)$$

Burada tipik olarak 1 değeri 20 alınmıştır. Yukardaki $h_w(n)$ ağırlıklı birim vuruş cevabını göstermektedir.

2.6. Analiz-Sentez İşlemi

Analiz-sentez her uyarma bileşeninin indeks ve kazancını elde etmek için kullanılır. Şekil 26 analiz-sentez kestiricinin analiz işleminin blok diyagramını göstermektedir.



Şekil 26. Kaynak analiz modelin blok diyagramı

Uyarma kaynağının her bileşeni, $d(n)$ enerjisi min. yapılarak elde edilir. Z dönüşümü kullanıldığında $D(z)$ elde edilir.

$$D(z) = Y(z) - \beta(\gamma) X_\gamma(z) \quad (98)$$

$Y(z)=S(z)W(z)$ ağırlıklı orijinal sesin z dönüşümüdür. $X_\gamma(z)$ ise verilen $f_\gamma(n)$ uyarma kaynağına ağırlıklı sistemin cevabının z dönüşümüdür. Verilen γ indeksi için $\beta(\gamma)$ ilişkilendirme kazancı, ağırlıklı ortalama karesel hatayı min. yapılarak elde edilir.

$$E_\gamma = \sum_{n=0}^{N-1} d^2(n) = \sum_{n=0}^{N-1} \left[y(n) - \beta(\gamma) \sum_{i=0}^{l-1} h_w(i) f_\gamma(n-i) \right]^2 \quad (99)$$

Burada $h_w(n)$ ağırlıklı birim vuruş cevabıdır. $\beta(\gamma)$ ise şöyle ifade edilir.

$$\beta(\gamma) = \frac{\sum_{n=0}^{N-1} y(n) X_\gamma(n)}{\sum_{n=0}^{N-1} X_\gamma^2(n)} \quad (100)$$

$$X_\gamma(n) = \sum_{i=0}^{l-1} h_w(i) f_\gamma(n-i) \quad (101)$$

Burada $X_\gamma(n)$ verilen uyarma fonksiyonuna sistemin ağırlıklı cevabıdır. Ortalama karesel hata $E(\gamma)$ ile verilir. Ortalama karesel hata min. yapılarak γ_k optimum indeks ve $e_k(n)$ uyarma bileşenleri elde edilir.

$$E(\gamma) = \left(\sum_{n=0}^{N-1} y(n)^2 \right) - \frac{\left[\sum_{n=0}^{N-1} X_\gamma(n) y(n) \right]^2}{\sum_{n=0}^{N-1} X_\gamma(n)^2} \quad (102)$$

Bu (102) nolu denklemdeki $y(n)$, verilen $X_\gamma(n)$ uyarma bileşenine sistemin cevabıdır. $X_\gamma(n)$ ile $y(n)$ belirlendikten sonra ortalama karesel hatayı minimum yapan γ_k indeksi ve $e_k(n)$ uyarma bileşeni için ortalama karesel hata olan $E(\gamma)$ değeri hesaplanabilmektedir.

2.7. Uzun Dönem Kestiricisi

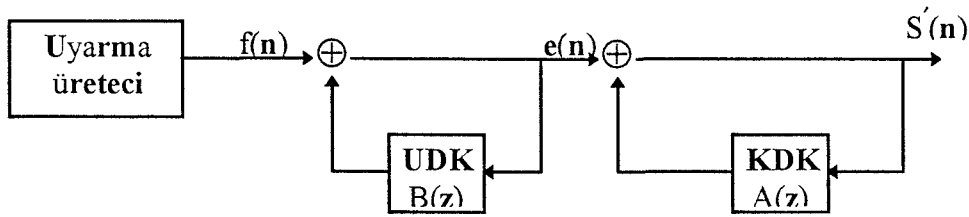
Şekil 27 iki kaskat sentez süzgeçli doğrusal öngörümlü kodlayıcı sentezleyiciyi göstermektedir. $A(z)$, kısa dönem kestirici (KDK), ses dalgasının formant yapısını modeller. $B(z)$, uzun dönem kestiricisi (UDK) ise sesin harmonik yapısını modeller. UDK daha çok perde periyodu olarak adlandırılır. UDK'nin genel formu şöyledir.

$$B(z) = \sum_{i=-N_1}^{N_2} \beta_i z^{-\gamma-i} \quad (103)$$

β_i , uzun dönem kestirici katsayıları veya kazançlarıdır. γ ise uzun dönem kestirici gecikmesi olarak adlandırılır. UDK bir uyarma kaynağı olarak düşünülebilir. Bu uyarma kaynağının çıkışı $e(n)$ ile ifade edilir.

$$e(n) = \beta e(n-\gamma) + f(n) \quad (104)$$

Buradaki $f(n)$ diğer uyarma bileşenlerinin toplamıdır. Uzun dönem kestirici parametreleri analiz-sentez işleminden elde edilir.



Şekil 27. Uzun dönem kestiricili bir DÖK sentezleyicinin blok diyagramı

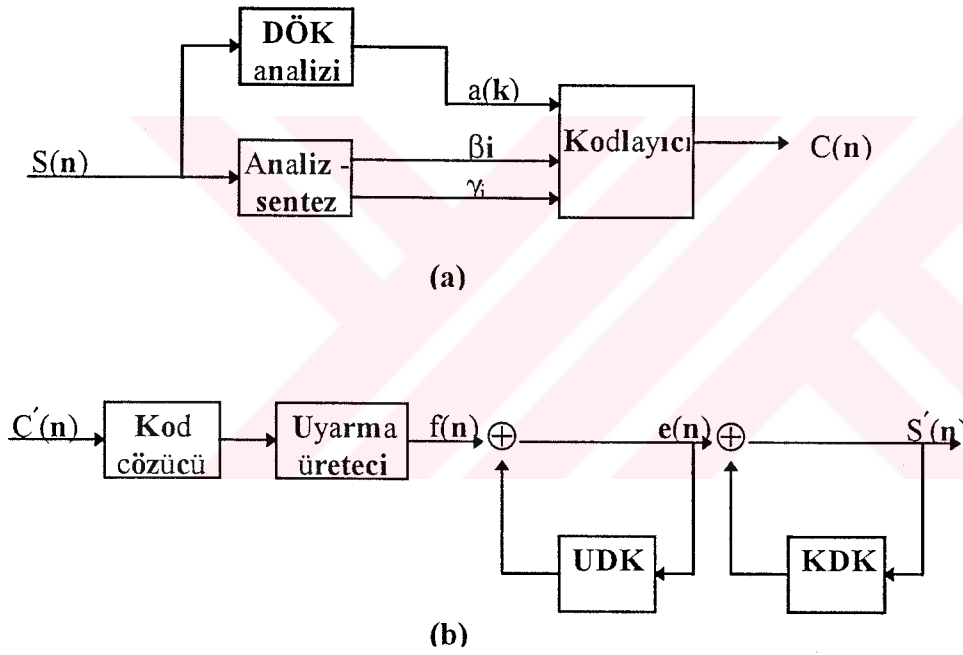
2.8. Çoklu Darbe Uyarmalı DÖK Uygulaması [16]

Çoklu darbe uyarmalı kodlayıcıda uyarma işareti ağırlıklı gecikmiş vuruşların toplamı olarak modellenir. Şekil 28'de çoklu darbe uyarmalı doğrusal öngörümlü kodlayıcının verici ve alıcıya ilişkin blok diyagramları verilmiştir. Verici tarafta ses işareti doğrusal öngörümlü

kodlayıcı katsayılarını, uzun dönem kestirici parametrelerini ve kaynak parametrelerini üretmek üzere analiz edilir. Kaynak parametreleri darbe konumları ve genliklerinden oluşmaktadır. Böylece çoklu darbe uyarma kaynağı şöyle ifade edilir.

$$f(n) = \sum_{k=1}^M \beta_k \delta(n - \gamma_k) \quad (105)$$

Bu tür kodlayıcıda; doğrusal öngörümlü kodlama için L, çerçeve boyu p, kestirici derecesine ilişkin parametreler ile α hata ağırlıklı faktör, N kaynak analiz çerçevesi boyu ve çerçeve başına M darbe sayısı parametrelerinden meydana gelmektedir.



Şekil 28. Doğrusal öngörümlü kodlayıcı, a) Verici, b) Alıcı

Parametrelere ilişkin tipik değerler tablo 4'te verilmiştir. Değişik ses dosyaları kullanılarak öncelikle doğrusal öngörümlü kodlayıcı parametreleri üretildi. Bunlar kullanılarak M, darbe sayısı 8,6,4,2, ve 1 alınarak çoklu darbe uyarmalı doğrusal öngörümlü kodlayıcıyla sentetik ses üretildi ve işaret gürültü oranları (İGO) hesaplandı. Uyarma çerçeve sayısı 60 ve 40 alınarak yukardaki işlemler tekrarlandı.

Tablo 4. Çoklu darbe uyarmalı kodlayıcı analiz-sentezleyicisinin parametreleri

Parametreler	Aralık	Tipik Değerler
Kestirici derecesi (P)	1-16	10
DÖK pencere uzunluğu (L)	160-360	240
DÖK çerçeve uzunluğu (I)	80-240	160
Hata ağırlık faktörü (α)	0-1	0.8
UDK uyarma çerçeve boyu (N)	20-60	40
Uyarma çerçevesi darbe sayısı (M)	2-20	8

2.9. Düzenli Darbe Uyarmalı Doğrusal Öngörümlü Kodlayıcı

Düzenli darbe uyarmalı kodlayıcı çoklu darbe uyarmalı doğrusal öngörümlü kodlayıcının özel bir halidir. Uyarma darbeleri zamanda eşit aralıklarla yerleştirilir. İlk darbe analiz çerçevesinde herhangi bir yere yerleştirilir. İlk darbenin γ_1 yeri belli olduktan sonra diğer darbelerin yerleri $\gamma_1 \pm kD$ ile belirlenmektedir. D, darbeler arasındaki mesafeyi belirtir. Kodlanmış parametreler ilk darbenin bulunduğu γ_1 yeri ile β_k darbe genlikleridir. Öyleki $k=1,2,3, \dots, N/D$ olup N uyarma analiz çerçevesinin uzunluğudur.

Analiz işlemi, ilk darbe konumunu ve ilişkilendirme kazanç değerlerini ağırlıklı hata enerjisini min. yapacak şekilde seçilmesini içermektedir. Bütün darbe genlikleri N/D adet doğrusal denklemlerin çözümünü elde edilir. Yüksek kalitede ses; D=4, çerçeve boyu=40 ve bit hızı=9.6kb/s değerlerinde elde edilmiştir [17]. Düzenli darbe uyarmalı doğrusal öngörümlü kodlayıcı çoklu darbe uyarmalı kodlayıcıya göre daha az hesaplama karmaşıklığına sahiptir.

“H-female.SIG” ve “L-male.SIG” gibi değişik ses dosyaları alınarak doğrusal öngörümlü kodlayıcı parametreleri üretilir. Darbe sayısı 8,5,4,2 ve 1 yapılarak sentetik ses üretilir. Her durum için işaret gürültü oranları hesaplanır. Çoklu darbe uyarmalı kodlayıcı ile düzenli darbe uyarmalı kodlayıcı arasında karşılaştırma da yapılabilir. Düzenli darbe uyarmalı doğrusal öngörümlü kodlayıcı için alınan tipik parametre değerleri tablo 5’te verilmiştir.

Tablo 5. Düzenli darbe uyarmalı kodlayıcı analiz sentezleyicisinin parametreleri

arametreler	Aralık	Tipik Değerler
Kestirici derecesi (P)	1-16	10
DÖK pencere uzunluğu (L)	160-360	240
DÖK çerçeve boyu (I)	80-240	120
Hata ağırlık faktörü (α)	0-1	0.8
UDK uyarma çerçeve boyu (N)	20-60	40
Darbe aralığı (D)	2-10	4

2.10. Kod Uyarmalı Doğrusal Öngörümlü Kodlayıcı Uygulaması

Çoklu darbe uyarmalı doğrusal öngörümlü kodlayıcının 9.6kb/s ve daha büyük bit hızlarında etkili bir ses kodlayıcı olduğu bilinmektedir. Çoklu darbe uyarmalı kodlayıcıdaki bitlerin büyük bir kısmı uyarma parametrelerinin kodlanması için kullanılmaktadır. Bit hızını 9.6kb/s'nin altına çekmek için uyarma parametrelerinin sayısını azaltmak gerekmektedir. Kod uyarma modeli düşük sayıdaki parametrelili uyarma modeli için çok etkilidir. Kod uyarmalı doğrusal öngörümlü kodlayıcı (KUDÖK) [18,19], 9.6kb/s bit hızının altındaki hızlarda iyi kalitede bir diğer analiz-sentez esaslı ses kodlayıcısıdır. Kod uyarma modelinde, kod tablosu olarak geniş bir random veya deterministik kod kelimeleri kullanılır.

Kod uyarmalı kodlama, vektör kuantalama teknikleri gibi örneklenmiş giriş işaretinin örnek bloklarını kullanan bir çerçeve esasına dayalı tekniktir. KUDÖK kodlama analiz-sentez arama işlemi, ağırlıklı vektör seviyelendirme (VQ) ve doğrusal öngörümlü işlemlerine dayanmaktadır. Ses işaretinin kısa dönem tayfının veya formant yapısını modellemek için 10. dereceden bir doğrusal öngörüm süzgeci kullanılır. Uzun dönem işaret periyodu veya perde adaptif kod tablosu ile modellenir. Kısa dönem doğrusal öngörüm VQ'dan artı kalan sabit bir stokastik kod tablosu kullanılarak seviyelendirilir. Kod uyarmalı kodlayıcının hesaplama yükünü iki kod tablosu arama işlemi oluşturmaktadır. Hesaplama karmaşıklığı kod tablolarının aranmasına bağlıdır. Federal Standart 1016 [20], 8kHz'lik bir örnekleme hızı ve 7.5ms'lik 4 adet alt çerçeveden oluşan bir 30ms'lik çerçeveyi kullanır. Kod uyarmalı kodlayıcı analizi 3 ana fonksiyondan oluşmaktadır.

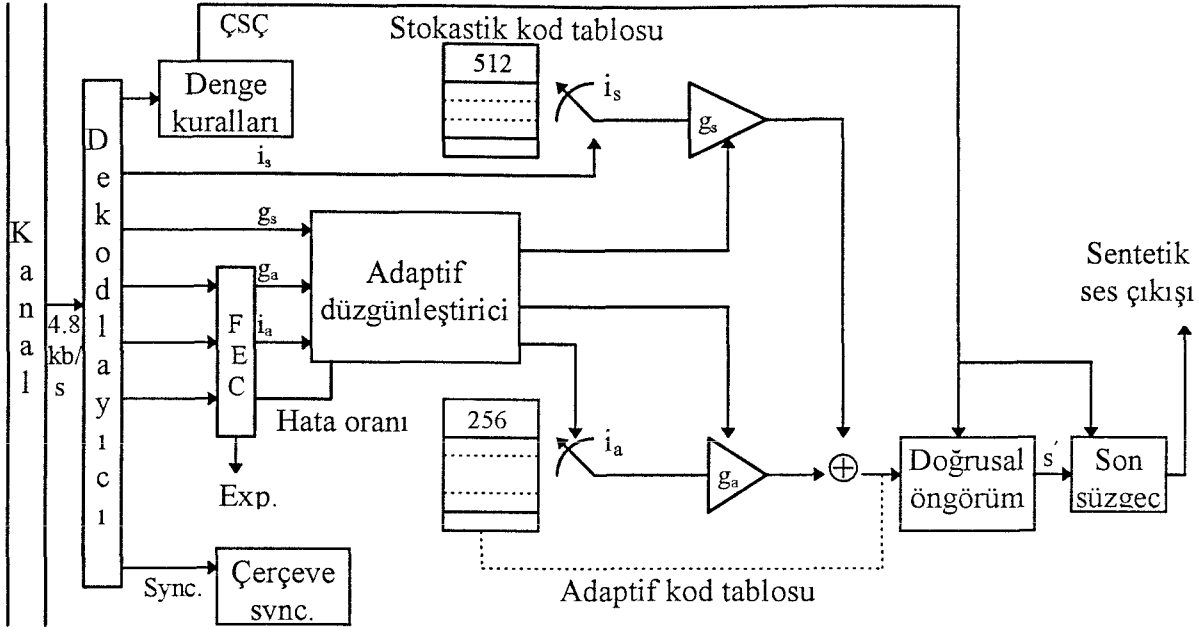
- 1) Kısa dönem doğrusal öngörümleme.
- 2) Uzun dönem adaptif kod tablosu arama.
- 3) Stokastik kod tablosu arama.

Kod uyarmalı kodlayıcı sentezleyicisi 3 sentez fonksiyonuna ilaveden birde ses kalitesini iyileştirmek için alıcı tarafta enson bir süzgeçten ibarettir. Kodlayıcı parametreleri stokastik kod tablosu indeksi ve kazancı, adaptif kod tablosu indeksi ve kazancı ile 10 adet çizgi tayf parametrelerinden oluşmaktadır.

2.10.1. KUDÖK Alıcısı

Kod uyarmalı doğrusal öngörümlü kodlayıcısının alıcısı şekil 29'da gösterilmiştir. Çerçeve senkronizasyonu yapıldıktan sonra alıcı KUDÖK parametrelerini dekodlamaya tabi tutar. Alıcı paralel bir kazanç biçimli kod uyarmalı doğrusal öngörümlü süzgeç yardımıyla sesi sentezler. Uyarma sabit bir stokastik kod tablosu ve adaptif kod tablosu kullanılarak biçimlendirilir. Her iki kod tablosu 60 örnekli kod kelimelerinden oluşur. Örneklerin gecikmesi zamanda perde gecikmesi olarak adlandırılır. Bu gecikme adaptif kod tablosu indeksi ile ilişkilidir. En sonda sentetik çıkış sesi iyileştirmek içinde adaptif bir son süzgeç ilave edilir.

Vericiden gönderilen kodlanmış parametreler alıcıda dekodlanarak, vericide yapılan işlemlerin ters işlemleri alıcıda yapılmaktadır. Alt çerçeve sentezi için çizgi tayf frekansları ara değerlendirme yapılır. Başlangıç uyarması için stokastik kod kelimeleri ve adaptif kod kelimeleri yapılandırılarak uyarma güncelleştirilir. Doğrusal öngörümlü çıkış sesi bir en son süzgeçinden geçirilir. Son süzgeçlemeden önce çıkış sesi yüksek geçiren süzgeçten geçirilir. Doğrusal öngörümlü sentezlemede çizgi tayf frekansları kestirici katsayılarına giriş sesin özilişkisi kullanılarak dönüştürülür. Son süzgeç kutup ve sıfır fonksiyonlarından oluşmaktadır.



Şekil 29. Federal standart 1016 kod uyarmalı kodlayıcı alıcısı [20].

Sentezleyicide giriş ses işaretinin kırılmış olup olmadığı da test edilir. Kırılma yapılmışsa bunun yapıldığı çerçeveler tespit edilir. Kırılmışların sayısı bulunur ve onlara ilişkin genlik değerleri toplanır. Bu işlemlerle bulunacak olan değerler enerji hesabında ve delta enerjisi gibi yerlerde kullanılır.

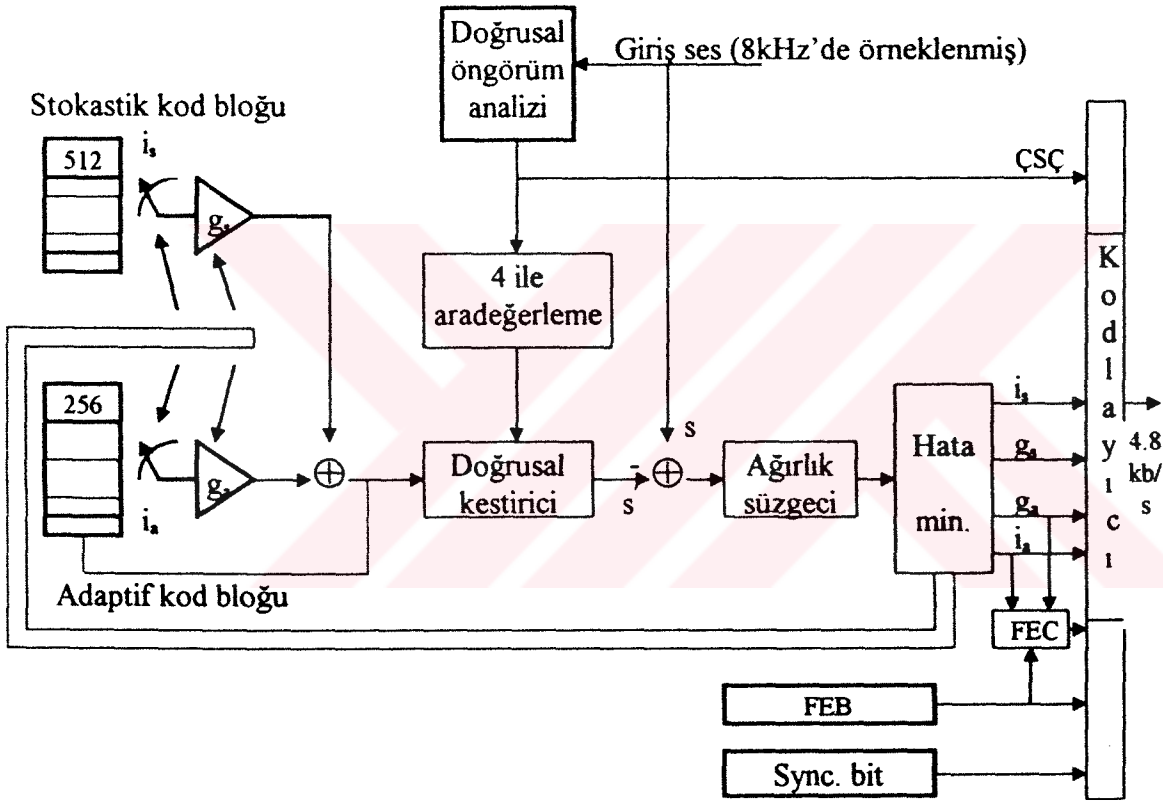
Stokastik kod bloğunun uzunluğu 1082 olarak alınmıştır. Vericiden gönderilen stokastik kod bloğuna ilişkin indeks ve kazancı alınarak stokastik kod bloğunun dağılımı hesaplanır.

Adaptif kod bloğunun yapılandırılmasında başlangıç değerleri olarak adaptif kod bloğunun uzunluğu, başlangıç kod kelimesi, adaptif gecikme ve artık işaretin boyu esas alınır. Stokastik kod tablosunda olduğu gibi kod kelimeleri vericiden gelen indeks ve kazanç değerleri kullanılarak belirlenir. Adaptif kod bloğu stokastik kod bloğundan farklı olarak bir geri besleme yoluyla kod kelimeleri yeniden güncelleştirilir.

Bütün bu işlemlerden sonra sentezleyicinin çıkışından analizin girişindeki orijinal sese ilişkin sentetik ses üretilmiş olunur.

2.10.2. KUDÖK Vericisi [20]

KUDÖK vericisi şekil 30'da gösterilmiştir. Giriş sesin s çıkartılır ve fark algılanarak bir düzeyde ağırlıklandırılır. Ağırlıklı hata, analiz-sentez hata minimizasyonunda kullanılır. Arama işlemi min. hatayı veren adaptif ve stokastik kod tablosu indeksleri ve kazançlarını bulur. Doğrusal kestirim süzgeci geleneksel açık çevrim doğrusal öngörüm analiz teknikleri giriş ses işareti üzerinden belirlenmektedir. Kodlayıcı parametrelerinin iletimi için Federal Standart 1016'nın belirlediği esaslar altında kodlanır.



Şekil 30. Federal standart 1016 KUDÖK vericisi [20].

2.10.2.1. Doğrusal Öngörüm Analizi

Kısa dönem doğrusal öngörüm analizi 30ms'lik bir Hamming penceresi ve 10. dereceden bir otokorelasyon analizi kullanılarak çerçeve başına bir kere koşulur. Doğrusal öngörüm analiz kestiricisi katsayılarında a_i yerine $a_i \gamma^i$ koyulur. Böylece kutuplar z düzleminde γ ağırlık faktörüyle orjine doğru kaydırılır. Katsayıları $1/A(z)$ olan doğrusal

öngörüm süzgecini tanımlar ve $0 < \gamma < 1$ aralığında $\gamma = 0.994$ değerini alır. Ağırlık süzgeci $A(z)/A(z/\gamma)$ olup $\gamma' = 0.8$ alınmıştır. Kodlayıcı gecikmesi doğrusal öngörüm analizi ile belirlenir. Kod uyarmalı kodlayıcıda toplam 105ms'lik bir gecikme mevcuttur. Doğrusal kestirici 34 bitle kodlanır. Çizgi tayf çiftleri çerçeve başına yalnız bir kere gönderilirken 4 alt çerçevenin herbiri için doğrusal aradeğerlemeye tabi tutulurlar.

2.10.2.2. Adaptif Kod Tablosu Arama

Adaptif kod tablosunu arama; ağırlıklı hata işaretinin min. karesel kestirim hata kriteri kullanılarak kapalı çevrim analizi ile güncelleştirilir. 256 kod kelimeyi bir adaptif kod tablosunun kodlanması için 8 bit ayrılmıştır. Hesaplama karmaşıklığını azaltmak için kod tablosu alt kümelerle ayrılarak arama işlemi yapılmalıdır. Her tek numaralı alt çerçeve için kodlama 128 tam sayı ve 128 tam sayı olmayan gecikmelerden ibarettir. Bu gecikmeler 20 ile 147 örnek arasında olmaktadır. Her çift alt çerçeve için gecikmeler araştırılır ve 6 bitle kodlanır. Bu büyük ölçüde hesaplama karmaşıklığını azaltır ve ses kalitesinde fazla bir kayıp olmamaktadır. Adaptif kod tablosundaki indeks ve kazanç her çerçeve için 4 kez gönderilir.

Tamsayı ve tamsayı olmayan gecikmeli adaptif kod kelimeleri şöyle yapılandırılır. Üst üste gelmiş kod kelimelerinden oluşan adaptif kod tablosu r ile gösterilsin.

$$r = [r(-147), r(-146), \dots, r(-1)] \quad (106)$$

Adaptif kod tablosu üzerinde bulunulan kod kelimesini r' ile gösterelim.

$$r' = [r'(0), r'(1), \dots, r'(59)] \quad (107)$$

Ayrıca r'' ; r ve r' kullanılarak oluşturulur.

$$\begin{aligned} r'' &= [r(-147), r(-146), \dots, r(-1), r'(0), r'(1), \dots, r'(59)] \\ r'' &= [r''(-147), r''(-146), \dots, r''(-1), r''(0), r''(1), \dots, r''(59)] \end{aligned} \quad (108)$$

Tam değerli bir gecikme için r' kod kelimesi M örnekli gecikmiş önceki uyarma işaretinin tekrarlanmasıyla yapılandırılır. Tam değerli gecikmeli r' kod kelimesi rekursif olarak üretilir.

$$r_M'(i) = r_M''(i) = r_M''(i-M), \quad i=0, 1, 2, \dots, 59 \quad M=20, 21, \dots, 147 \quad (109)$$

Tamsayı olmayan gecikmeler için kod kelimeleri aradeğerleme ile biçimlendirilir. Verici ve alıcıda sentezleme için kullanılan aradeğerleme, Hamming pencerele sinc fonksiyonunun N ağırlıklı 40 noktalı bir aradeğerlemedir.

$$h(k) = 0.54 + 0.46 \cos(k \Pi / 6N),$$

$$W_f(j) = h(12(j+f)) \frac{\text{Sinc}((j+f)\Pi)}{(j+f)\Pi}, \quad j = \frac{-N}{2}, \frac{-N}{2}+1, \dots, \frac{N}{2}-1 \quad f = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4} \quad (110)$$

Rekursif bir aradeğerleme formülü $M+f$ gecikmeli r' kod kelimesini hesaplamak için kullanılabilir.

$$r'_{M+f}(i) = r''_{M+f}(i) = \sum_{j=\frac{-N}{2}}^{\frac{N}{2}-1} W_f(j) r''_{M+f}(i-M+j),$$

$$i = 0, 1, 2, \dots, 59 \quad M = 20, 21, \dots, 147 \quad f = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4} \quad (111)$$

Sonuçta kod tablosu arama işlemi tamamlandıktan sonra r adaptif kod tablosu seçilen e uyarma vektörüyle güncelleştirilir. Uyarma vektörü ölçeklendirilmiş stokastik ve adaptif kod kelimelerinin toplamıdır.

$$r(i) = r(i+60), \quad i = -147, -146, \dots, -61 \quad (112)$$

$$r(i) = e(i+60), \quad i = -60, -59, \dots, -1 \quad (113)$$

Doğrusal öngörüm sentez süzgecini uyarma için kullanılan $e(0)$ zaman domeninde ilk örnektir. Tamsayı olmayan gecikmeler kullanıldığı zaman kodlayıcı gürültüsü azaltılır. Çünkü onlar perde kestirimini iyileştirir ve gürültülü stokastik uyarma bileşenini azaltır.

2.10.2.3. Stokastik Kod Tablosu Arama

Stokastik kod tablosu arama geleneksel MSPE kriterini kullanan kapalı çevrim analiziyle gerçekleştirilir. 512 kod kelimeli kod tablosunu kodlamak için 9 bit kullanılır. Kod tablosu indeks ve kazancı çerçeve başına 4 kez gönderilir. Bu kod tablosu 0 ortalamalı birim varyanslı beyaz Gauss gürültüsü örneklerinden oluşmaktadır. Bu şekildeki kod tablosu belirsiz değildir ve hızlı arama işlemlerine sahip olup arama hesap yükünü başarılı olarak azaltmaktadır.

2.10.3. Hata Düzeltme

Hata kodlama hata düzeltme katının esasını oluşturmaktadır. Toplam adaptif hata düzeltme sistemi ileri hata düzeltme (FEC), düzgünleştirme ve parametreyi kodlamayı kapsar. İleri hata düzeltme tek hatayı belirleyen ve düzelten bir Hamming(15,11) ile yapılır. Kod uyarımalı kodlayıcının karakteristikleri tablo 6 ile verilmiştir.

Tablo 6. Kod uyarımalı kodlayıcının bazı karakteristikleri

	Doğrusal kestirici	Adaptif kod tablosu	Stokastik kod tablosu
Zaman	30ms	$30/4=7.5$ ms	$30/4=7.5$ ms
Parametereler	10 ÇSÇ	1 kazanç, 1 gecikme, 256 kod kelimesi	1 kazanç, 1 indeks, 512 kod kelimesi
Hız	1.133.33bit/s	1.600bit/s	1.866.67bit/s

2.10.4. Kod Tablosu Arama Yöntemleri

Hesaplamayı azaltmak için ardışıl iki kod tablosunu arama katları mevcuttur. Adaptif kod tablosunu aramanın ilk katı ağırlıklı doğrusal kestirimle önceki çerçevelerin kodlanmış

hatalarının toplamıdır. Stokastik kod tablosu ikinci katı ilk kat ile süzgeçlenmiş adaptif kod tablosu VQ uyarma arasındaki farktır. $L=60$ uzunluklu s vektörleri orijinal ses işaretini, s' vektörleri sentetik ses işaretini ve e vektörleri ağırlıklı hata işaretlerini göstermektedir. Bulunulan kattaki uyarma vektörü v ve önceki katın uyarma vektörü u ile gösterilsin. Uzunluğu N olan bir kod tablosu için uyarma kaynağı $1 \leq i \leq N$ aralığında olan bir i kod tablosu indeksi ve g_i kazanç parametresiyle karakterize edilir. Uyarma vektörü $v^{(i)}$ şöyle ifade edilir.

$$v^{(i)} = g_i x^{(i)} \quad (114)$$

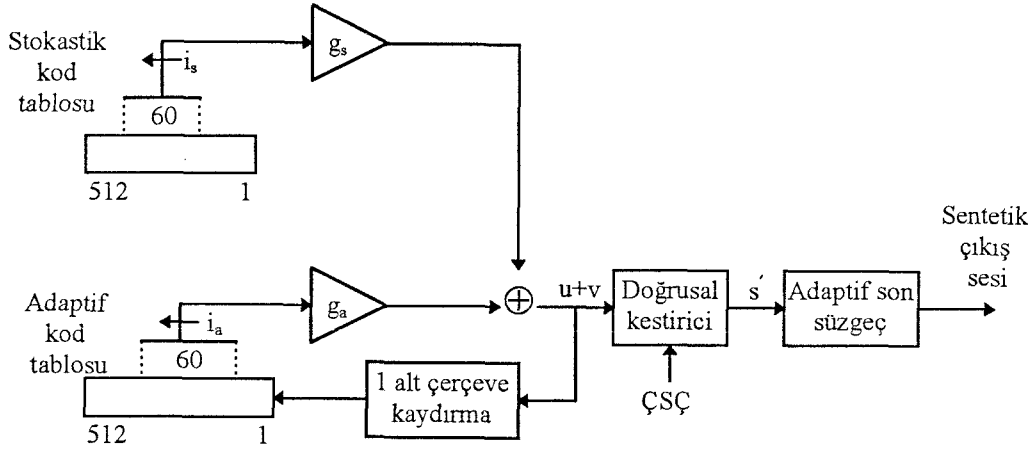
Üst indis i , $x^{(i)}$ kod tablosu vektörünün kod tablosu indeksini gösterir. H ve W , $L \times L$ boyutlu alt üçgensel matrislerdir. H 'ın satırları doğrusal öngörüm süzgecinin birim vuruş cevaplarıdır, W 'nın satırları ise hata ağırlıklı süzgecinin birim vuruş cevabıdır.

$$H = \begin{bmatrix} h_0 & 0 & \dots & 0 \\ h_1 & h_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ h_{L-1} & h_{L-2} & \dots & h_0 \end{bmatrix}_{L \times L} \quad W = \begin{bmatrix} w_0 & 0 & \dots & 0 \\ w_1 & w_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ w_{L-1} & w_{L-2} & \dots & w_0 \end{bmatrix}_{L \times L} \quad (115)$$

Şekil 31'de gösterildiği üzere sentetik ses, doğrusal öngörüm süzgecinin birim vuruş cevabı ile uyarma işaretinin katlanması, sıfır giriş cevabının toplamıdır.

$$s'^{(i)} = H[u + v^{(i)}] + s'^{(0)}, \quad 1 \leq i \leq N \quad (116)$$

Burada u ilk arama katındaki sıfır vektörü veya ikinci arama katındaki ölçeklenmiş uyarma vektörüdür.



Şekil 31. KUDÖK sentezleyicisi

Şekil 32’de ağırlıklı hata işareti gösterilmiştir.

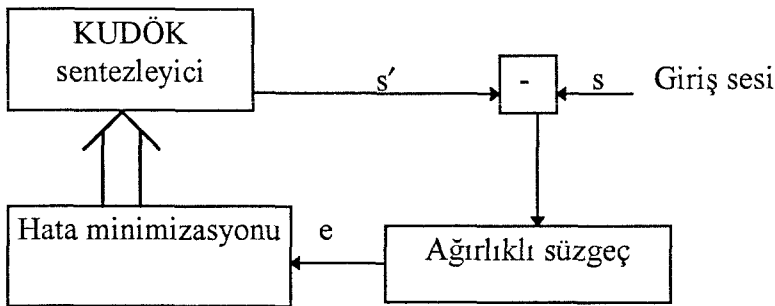
$$e^{(i)} = w[s - s'^{(i)}] = e^{(0)} - WHv^{(i)} \quad (117)$$

$$e^{(0)} = W[s - s'^{(0)}] - WHu \quad (118)$$

Böylece i kod kelimesi için ağırlıklı hata $e^{(i)}$ ve süzgeçlenmiş kod kelimesi $y^{(i)}$ ile verilebilir.

$$e^{(i)} = e^{(0)} - g_i y^{(i)} \quad (119)$$

$$y^{(i)} = WHx^{(i)} \quad (120)$$



Şekil 32. KUDÖK analizi

E_i kod kelimesi için toplam karesel hatayı vermektedir.

$$\begin{aligned} E_i &= \|e^{(i)}\|^2 = \langle e^{(i)}, e^{(i)} \rangle = e^{(i)T} e^{(i)} \\ &= e^{(0)T} e^{(0)} - 2g_i y^{(i)T} e^{(0)} + g_i^2 y^{(i)T} y^{(i)} \end{aligned} \quad (121)$$

T transpozudur. E_i hem g_i kazanç faktörü ve hemde i indeksinin bir fonksiyonudur. Verilen i için optimal kazanç E_i 'nin g_i 'ye göre türevini sıfır yaparak belirlenebilir.

$$\frac{\partial E_i}{\partial g_i} = -2y^{(i)T} e^{(0)} + 2g_i y^{(i)T} y^{(i)} = 0 \quad (122)$$

Böylece min. ortalama karesel hata kazancı, hedeflenen ile süzgeçlenmiş kod kelimesinin çapraz-ilişkisinin süzgeçlenmiş kod kelimesinin enerjisine oranıdır.

$$g_i = \frac{y^{(i)T} e^{(0)}}{y^{(i)T} y^{(i)}} \quad (123)$$

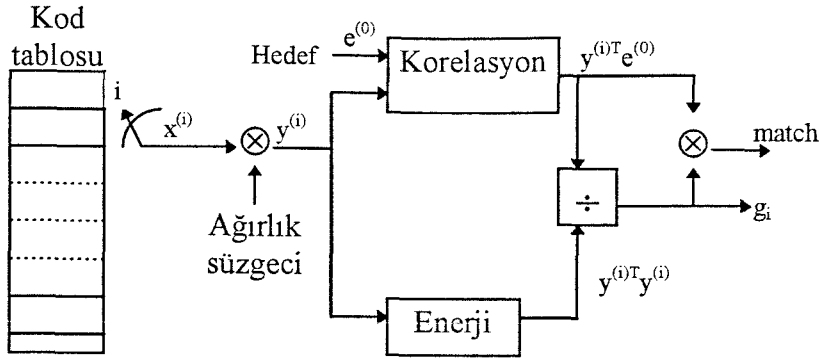
E_i ifadesinde ilk terim i kod indeksinden bağımsızdır.

$$m_i = g_i \left(2y^{(i)T} e^{(0)} - g_i y^{(i)T} y^{(i)} \right) \quad (124)$$

(123) ifadesi (124) ifadesinde yerine konulursa m_i şu forma dönüşür.

$$m_i = \frac{\left(y^{(i)T} e^{(0)} \right)^2}{y^{(i)T} y^{(i)}} \quad (125)$$

Böylece (125) eşitliğinde ve şekil 32'de gösterildiği gibi kod bloğunu arama işlemi m_i 'yi max. yapan i kod kelimesini bulur. Kazanç faktörü kod kelimesinin genliğiyle ilişkilidir.



Şekil 33. Kod bloğu arama

Bu uyarma yöntemi uzun dönem kestiricisi etkin olduğunda stokastik kod tablosu kazancını adaptif olarak azaltır. Bu stokastik kod tablosu uyarma bileşenlerinin azaltılmasına bağlı olarak adaptif kod tablosu uyarma bileşenlerini artırır. Kod uyarmalı kodlayıcıda ses kalitesi sesli ses bölgelerindeki seviyelendirilmiş gürültünün azaltılmasından dolayı artırılır. Uzun dönem kestiricisi etkin olmadığında stokastik kod tablosu kazancı artırılır. Bu sentezlenmiş ses ile giriş sessiz ses segmentleri arasındaki uygunluğu sağlar. Uzun dönem kestiricisinin etkisiyle perde kestiriminden önce ve sonra hedef vektörleri ölçülebilir. R normalizeli çapraz ilişkiyi ve g' stokastik kod tablosu kazancını göstermektedir.

$$R = \frac{\langle W[s - s^{(0)}], W[s - s^{(0)}] - WHu \rangle}{\|W[s - s^{(0)}]\|} \quad (126)$$

$$g' = \begin{cases} 0.2g_i, & |R| < 0.04 \\ 1.4g_i\sqrt{|R|}, & |R| > 0.81 \\ g_i\sqrt{|R|}, & \text{d.d.} \end{cases} \quad (127)$$

Böylece stokastik uyarma i indeksi ve g_i kazancı ile karakterize edilir. Kod tablosunu arama işlemi mümkün olduğunca azaltılmış olunur. Analiz çerçevesinin boyunun artırılması kodlayıcı performansını azaltır. Tipik bir kod tablosu uzunluğu 1024 olup $\log_2 1024 = 10$ bit gerektirmektedir.

Kod uyarmalı kodlayıcı modelinin kötü tarafı, arama işleminin yüksek hesap karmaşıklığına sahip olmasıdır. 40 uzunluklu 1024 girişli bir kod bloğu için yaklaşık olarak saniyede 500 milyon çarpma-toplama işlemine gerek duyulmaktadır. Etkili kod bloğu arama için halen değişik ve hızlı kod bloğu arama teknikleri üzerinde araştırmalar yapılmaktadır. Tablo 7’de KUDÖK kodlayıcısının tipik değerleri gösterilmektedir.

Tablo 7. KUDÖK analiz ve sentezleyicinin tipik parametreleri [15].

Parametreler	Aralık	Tipik Değerler
Kestirici derecesi (p)	1-16	10
DÖK pencere uzunluğu (L)	160-360	240
DÖK çerçeve uzunluğu (I)	80-240	120
Hata ağırlık faktörü (α)	0-0.99	0.8
Kod kelimesi uzunluğu (N)	20-60	40

“H-female.SIG” ve “L-male.SIG” ses dosyaları kullanılarak doğrusal öngörüm kodlayıcı parametreleri üretildi. Aynı ses hem erkek ve hem de bayandan alınmıştır. H-female.SIG ve L-male.SIG ses dosyaları “ Go and sit on the bed ” cümlesinden oluşmaktadır. Standart parametre değerleri ile değişik değerleri alınarak KUDÖK uygulaması yapıldı. Yeniden elde edilen ses dinlendi ve işaret gürültü oranları hesaplandı. Uyarma çerçeve boyu 20, 40 ve 60 alınarak yukardaki işlemler tekrarlandı. Uyarma çerçeve boyunun değişimiyle KUDÖK kodlayıcının kalitesi araştırıldı. Kod tablosunun boyu 64, 128, 512 ve 1024 alınarak çıkış sesleri yeniden dinlenerek işaret gürültü oranlarına bakıldı.

Tablo 8. Ses kodlama tekniklerinin performanslarının karşılaştırılması [21].

Yöntem	Bit hızı (kb/sn.)	İşlem sayısı/çerçeve	İGO(dB)	İşitsel kalite
LPC	4.8	20200	14.87	orta
MPELPC	6.4	11200	12.44	orta
RPELPC	6.4	9200	9.95	kötü
CELP	2.4	7000	16.21	iyi

3. BULGULAR

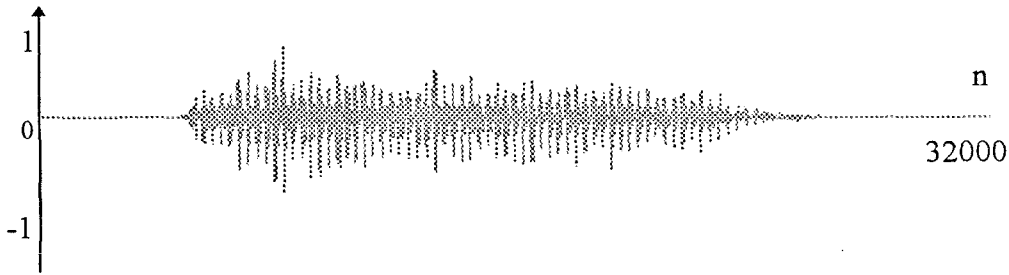
Federal Standart 1016'ya ilişkin kod uyarmalı doğrusal öngörümlü kodlayıcısının analizinde kullanılan giriş ses dosyaları "ia.WAV" ve "icellofma.WAV" dosyaları ele alınmıştır. Ancak bu ses dosyaları dışında pek çok ses dosyalarıda denenmiştir. Burada örnek olarak bu iki giriş ses dosyaları gösterilmiştir.

Kod uyarmalı doğrusal öngörümlü ses kodlayıcısının sentezleyicisinde de yukardaki iki giriş ses dosyasına karşılık elde edilen çıkış ses dosyaları da gösterilmiştir. "ia.WAV" giriş ses dosyası için çıkış ses dosyası "oa.WAV" ve "icellofma.WAV" giriş ses dosyası için çıkış ses dosyası "ocellofma.WAV" dosyaları ile verilmiştir.

Giriş ve çıkış ses dosyaları şekil 34'te verilmiştir.



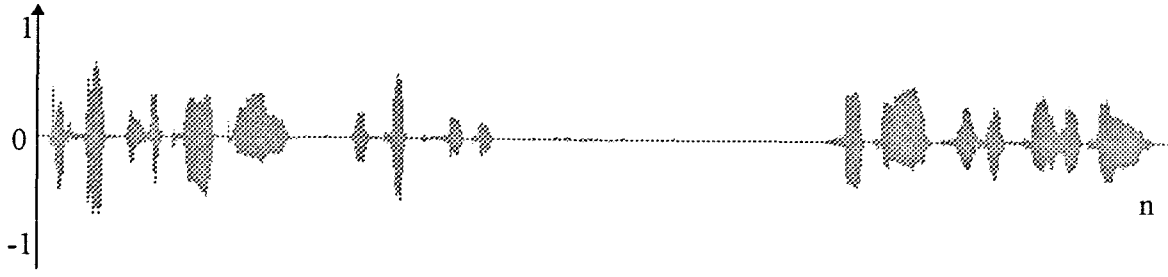
(a)



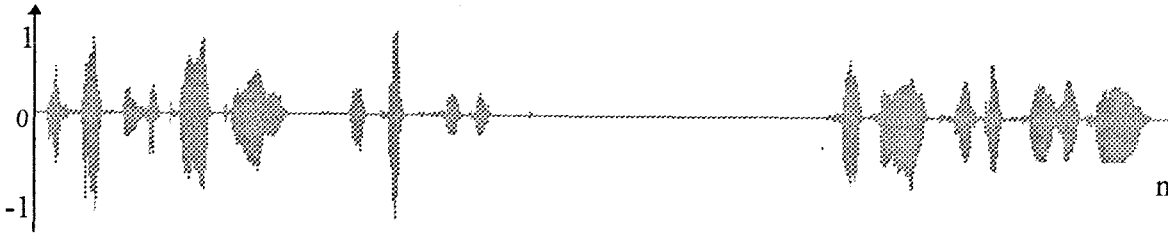
(b)

Şekil 34. KUDÖK giriş çıkış ses dosyaları,
a) "ia.wav" giriş ses dosyası,
b) "oa.wav" çıkış ses dosyası

Şekil 34'ün devamı;



(c)



(d)

Şekil 34. c) “iceloffma.wav” giriş ses dosyası, d) “oceloffma.wav” çıkış ses dosyası

Şekil 34'teki (a) ve (b) şekilleri söylenen “ a “ harfine ait giriş ve sentezlenmiş seslerdir. Sentezlenmiş ses genlikleri ile orijinal ses genlikleri arasında bir fark oluşmaktadır. Ancak her iki ses dinlendiğinde bu fark sadece hışırtı olarak kulağa gelmekte ve ses anlaşılabilir düzeyde olmaktadır. Diğer (c) ve (d) şekilleri söylenen “ This chapter is a programming secret “ cümlesinden oluşmaktadır. Bunda da dikkati çeken heceler arasındaki geçişlerde bazı bozulmaların olmasıdır. Bu durum, sentezlenen sesin dinlenmesinde bazı harflerin tam olarak anlaşılmasını engellemektedir.

Tablo 9'da “h-female.SIG” ses dosyasına ilişkin çoklu darbe uyarımlı kodlayıcıda elde edilen değerler verilmiştir. Tablolarda öncelikle aynı cümleyi söyleyen erkek sesine ilişkin sonuçlar verildi, daha sonrada bayan sesine ilişkin elde edilen değerler verilmiştir. Kestirici derecesi $P=10$, dök pencere uzunluğu $L=240$, dök çerçeve uzunluğu $I=160$, hata ağırlık faktörü $\alpha=0.8$, uyarma darbe sayısı $M=8,6,4,2$ alınarak İGO ve segment İGO hesaplandı.

Tablo 9. Çoklu darbe uyarmalı kodlayıcı için İGO ve SİGO değerleri

M	8	6	4	2
İGO (dB)	18.32	17.31	15.39	12.59
SİGO (dB)	14.48	13.01	11.55	9.40

Perde uyarmalı doğrusal öngörümlü kodlayıcılar 2.4kb/s'de çalışırken KUDÖK'ler 4.8-16kb/s aralığında çalışmaktadırlar [15]. Çoklu uyarma darbe sayısı olan M azaltıldığında işaret gürültü oranı ve segment işaret gürültü oranı azalmaktadır.

Düzenli darbe uyarmalı doğrusal öngörümlü kodlayıcı için dök çerçeve uzunluğu $I=120$ olarak alınmıştır. Ancak düzenli darbe uyarmalı kodlayıcı uyarma darbe sayısı $M=6$ için çalışmamaktadır. Çerçeve uzunluğunun uyarma darbe sayısına oranı bir tamsayı değerinde olmalıdır. Tablo 10'da düzenli darbe uyarmalı ses kodlayıcının "L-male.SIG" ses dosyası için elde edilen İGO ve SİGO verilmiştir.

Tablo 10. Düzenli darbe uyarmalı kodlayıcı için İGO ve SİGO değerleri

M	8	5	4	2
İGO (dB)	11.85	10.69	9.94	8.45
SİGO (dB)	10.60	9.38	8.45	7.28

Tablo 11'de "H-female.SIG" ses dosyasının kod uyarmalı doğrusal öngörümlü kodlayıcıya uygulanarak elde edilen sonuçlar verilmiştir.

Tablo 11. Kod uyarmalı kodlayıcı için İGO ve SİGO değerleri

N	20	40	60
İGO (dB)	11.23	12.01	10.16
SİGO (dB)	8.65	6.07	5.22

Aynı biçimde “L-male.SIG” giriş ses dosyası için N=40 durumunda İGO=11.23dB ve SİGO=8.65dB ile N=60 durumunda İGO=9.48dB ile SİGO=7.01dB değerleri elde edilmiştir. Uyarma çerçeve uzunluğu artıkça işaret gürültü oranları azalmaktadır. Erkeğe ait ses ile kadına ait ses dosyaları için bu değerler arasında fark vardır. Bu tip kodlayıcılarda kadın ses dosyaları için işaret gürültü oranları daha yüksek buna karşılık segment işaret gürültü oranları ise erkek ses dosyalarına göre daha düşüktür.

“L-male.SIG” için giriş ses dosyası 44KB’lik yer kaplarken çıkış ses dosyası 3KB’lik yer kaplamaktadır. Hem çoklu darbe uyarımlı ve hem de düzenli darbe uyarımlı kodlayıcılar için sıkıştırma oranı $44/3=14.6$ olmaktadır.

G(728) olarak adlandırılan KUDÖK’ler 16kb/s’de etkin olarak çalışmaktadırlar. “a.WAV” giriş ses dosyası için 119KB’lik yer ayrılırken çıkış ses dosyası için 8KB’lik yer gerekmektedir. Kod uyarmalı doğrusal öngörümlü kodlayıcı için sıkıştırma oranı $119/8=14.9$ olmaktadır.

KUDÖK kodlayıcılar diğer kodlama türlerine göre biraz daha fazla sıkıştırma yaptıkları görülmektedir.

4. İRDELEME

Ses işlemede, sesin analizi ve tanınması öncelikle ele alınması gereken konular olmalıdır. Yapılan bu çalışmada sesin analizi için Kod Uyarmalı Doğrusal Öngörümlü Kodlayıcı ve diğer bazı kodlayıcılar incelenmiştir. Bu kodlayıcıların birbirlerine göre olan üstünlükleri değişkendir. Ancak kod uyarmalı kodlayıcı düşük bitlerde diğerlerine göre daha iyi kalitede ses sentezleyebilmektedir.

KUDÖK kodlayıcıda stokastik kod tablosu ve adaptif kod tablosunu arama işlemi en fazla zaman alan kısımdır. Ayrıca hata eşiğini de belirlemek önemli bir parametre olarak görülmektedir. Genelde kestirici katsayılarının 8 ile 16 arasında seçilmesi bu tip kodlayıcılar için oldukça iyi sonuç vermiştir. Bu aralığın dışındaki kestirici değerlerinde kazanç değeri oldukça düşüktür.

Sessiz ses bölgeleri düşük genlikli ve yüksek frekanslıdır. Bu durum gürültüden etkilenmeden kodlama işlemine tabi tutmak oldukça zor olmaktadır. Sesli ses bölgeleri için böyle bir sorun olmamaktadır.

5. SONUÇLAR

Ses kodlama tekniklerinin genel yapıları ile bu sistemlerde yapılan matematiksel işlemler incelenmiştir. Ayrıca çoklu darbe uyarmalı, düzenli darbe uyarmalı ve kod uyarmalı doğrusal öngörümlü kodlayıcıların analiz ve sentez kısımları araştırılmıştır.

Ses işaretlerinin kodlanmasında kestirici katsayıları kadar Kod Uyarmalı Doğrusal Öngörümlü Kodlayıcıda stokastik ve adaptif kod tablosunu arama tekniğide önem kazanmaktadır. Sayısal ses kodlama geleneksel analog ses kodlayıcılarına göre çok daha fazla üstünlüğe sahiptirler. Düşük veri hızlarında kod uyarmalı doğrusal öngörümlü kodlayıcı etkin hata düzeltme özelliğine sahiptir. Düşük hızlı ses kodlayıcılarında ses performansını ölçmek oldukça zordur.

Ses kalitesini artırmada alıcı tarafta yer alan adaptif süzgecin etkisi büyük olmaktadır. Ayrıca arka plan gürültüsü ve kanal hataları ses performansını olumsuz yönde etkilemektedir. Doğrusal öngörümlü kodlama analiz yöntemi, ses işareti parametrelerinin doğru olarak temsil edilmesinde yeterlidir.

Analiz sentez esaslı kod uyarmalı doğrusal öngörümlü kodlama düşük bit hızlarında diğer kodlayıcılara göre daha yüksek kalitede ses sentezleyebilmektedir.

6. ÖNERİLER

Federal Standart 1016 tabanlı Kod Uyarımalı Doğrusal Öngörümlü Kodlayıcısının gerçek zamanda kuvvetli bir kodlayıcı olduğu görülmüştür. Federal Standart 1016 arama işlemleri için esneklik gösterebilmektedir. Daha etkili stokastik ve adaptif kod tablosu arama algoritmaları Federal Standart 1015 ve 1016 kullanılarak gerçekleştirilebilir.

Yüksek kalitede ses üretebilmek için gürültü ve kanal hatalarının belirlenip düzeltilmesi hususunda daha etkili incelemeler yapılmalıdır.

Federal Standart 1016 4.800bit/s'lik ses kodlamada pratik olarak uygulanabilmektedir.



7. KAYNAKLAR

1. Yong, M., A new LPC Interpolation Technique for CELP Coders, IEEE Transactions on Communications, 42, 1 (1994) 34-37.
2. Dudley, H., Remaking Speech, J. Acoust. Soc. Am., 11 (1939) 169-177.
3. Flanagan, J.L., Speech Analysis, Synthesis and Perception, Second Edition, Springer Verlag, New York, 1972.
4. Gold, B., Computer Program for Pitch Extraction, J. Acoust. Soc. Am., 34, 7 (1962) 916-921.
5. Barnwell, T.P., Brown, J.E., Bush, A.M., Patisaul, C.R., Pitch and Voicing in Speech Digitization, Res. Rept. Georgia Inst. of Tech., August 1974.
6. Robert, V. Bruce, Bell, Little Brown and Co., Boston, 1973, 144.
7. Noll, A.M., Cepstrum Pitch Determination, J. Acoust. Soc. Am., 41 (1967) 293-309.
8. Oppenheim, A.V., Schafer, R.W., Homomorphic Analysis of Speech, IEEE Trans. On Audio and Electroacoustics, 16 (1968) 221-226.
9. Rosenberg, A.E., Schafer, R.W., Rabiner L.R., Effects of Smoothing and Quantizing the Parameters of Formant-Coded Voiced Speech, J. Acoust. Soc. Am., 50, 6 (1971) 1532-1538.
10. Atal, B.S., Hanauer, S.L., Speech Analysis and Synthesis by Linear Prediction of the Speech Wave, J. Acoust. Soc. Am., 50, 2 (1971) 637-655.
11. Markel J.D., Gray, A.H., On Autocorrelation Equations as Applied to Speech Analysis, IEEE Trans. On Audio and Electroacoustics, AU-21 (1973) 67-79.
12. Burg, J., A new Analysis Technique for Time Series Data, Proc. NATO Advanced Study Institute on Signal Proc., Enschede Netherlands, 1968.
13. Markel, J.D., Gray, A.H., Linear Prediction of Speech, Springer-Verlag, New York, 1976.

14. Itakura, F.I., Saito, S., Analysis - Synthesis Telephony Based Upon the Maximum Likelihood Method, Proc. 6th Int. Congress on Acoustic, Tokyo, 1968 17-20.
15. Thomas, P.Barnwell III, Nayebi, K., Craig H.R., Speech Coding A Computer Laboratory Textbook, John Wiley & Sons. Inc., New York, 1994.
16. Rabiner, L.R., Schafer, R.W., Digital Processing of Speech Signals, Prentice-Hall Inc., New Jersey, 1978.
17. Atal, B.S., Schroeder, M.R., Predictive Coding of Speech and Subjective Error Criteria, IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-27 (1979) 247-254.
18. O'Shaughnessy, D., Speech Communication: Human and Machine, Addison-Wesley, 1987.
19. Deller, J.R., Proakis, J.G., Hansen, J.H.L., Discrete - Time Processing of Speech Signals, MacMillan, 1993.
20. National Communications System TIB 92 - 1, Details to Assist In Implementation of Federal Standard 1016 CELP, Virginia, January 1992.
21. Kroon, P., A Class of Analysis-by-Synthesis Predictive Coders for High Quality Speech Coding at Rates Between 4.8 and 16 kbits/s, IEEE Journal on Selected Areas in Communications, 6, 2 (1988) 353-363.

8. ÖZGEÇMİŞ

Salim KAHVECİ; 1975 yılında Trabzon'un Akçaabat ilçesinde doğdu. İlk öğrenimini Erikli İlkokulu'nda tamamladı. Ortaokulu Trabzon İmam Hatip Lisesi'nde ve liseyi Trabzon Yomra Fen Lisesi'nde okudu. 1992 yılında Karadeniz Teknik Üniversitesi Elektrik-Elektronik Mühendisliği Bölümünü kazandı. 1996 yılında aynı bölümden Elektronik Mühendisi ünvanı ile mezun oldu. Aynı yıl Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektronik Anabilim Dalında yüksek lisans öğrenimine başladı. Yabancı dil olarak İngilizce bilmektedir.



**T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ**