

79 107

**YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**SİNÜZOİDAL MODEL YARDIMIYLA MÜZİKAL SES
SENTEZİ**

Arş. Gör. Bülent BOLAT

**F.B.E. Elektronik ve Haberleşme Mühendisliği Anabilim Dalı Haberleşme Programında
Hazırlanan**

YÜKSEK LİSANS TEZİ

Yrd. Doç. Dr. Ünal Küçük

Tez Danışmanı: Yrd. Doç. Dr. Ünal KÜÇÜK

Doç. Dr. Ümit Aygölü

Çiğdem

Prof. Dr. Adilla Ataman

Hanım

İSTANBUL, 1998

TC. YÜKSEKÖĞRETİM KURULU
DOKÜMANİSTON BİRLİĞİ

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	iv
ÖZET.....	v
ABSTRACT.....	vi
1. GİRİŞ.....	1
2. SES ve MÜZİK.....	2
2.1 Akustik Dalgalar.....	2
2.2 Perde ve Periyodiklik.....	2
2.3 Sinüzoidal Dalgalar.....	4
2.4 Periyodik Seslerin Adlandırılması.....	5
2.5 Vuru Olayı ve Maskeleye.....	6
2.6 Ses Gücü.....	7
3. SES SENTEZİNE GENEL BAKIŞ.....	9
3.1 Ses Sentezi Yöntemlerinin Sınıflandırılması.....	9
3.2 İşlenmiş Kayıt (Doğrudan Sentez).....	11
3.3 Spektral Modeller.....	11
3.3.1 Toplamsal sentez.....	11
3.3.1.1 Kısa süreli Fourier dönüşümü.....	13
3.3.1.2 Faz kodlayıcı.....	15
3.3.1.2.1 Fourier dönüşümü ve süzgeç bankası.....	15
3.3.1.3 Sinüzoidal sentez.....	18
3.3.2 Çıkarımsal sentez.....	19
3.3.2.1 Doğrusal öngörü kodlaması.....	19
3.3.2.1.1 Doğrusal öngörü.....	20
3.3.2.1.2 Doğrusal öngörü kodlayıcı.....	21
3.4 Fiziksel Modeller.....	23
3.4.1 Fiziksel, işlevsel ve biçimsel yapılar.....	24
3.4.2 Uyarım.....	24
3.4.2.1 Doğrudan üretim.....	24
3.4.2.2 Belleksiz doğrusal olmayan model.....	25
3.4.2.3 Mekanik model.....	25
3.4.3 Rezonatör.....	26
3.4.3.1 Transfer fonksiyonu modeli.....	26
3.4.3.2 Mekanik model.....	26
3.4.3.3 Dalga kılavuzu modeli.....	26
3.4.4 Dalga kılavuzu modeli uygulamaları.....	27
3.4.4.1 Tek kamışlı nefesliler.....	27
3.4.4.2 Yaylı telliler.....	28

3.5	Fiziksel ve Spektral Modellerin Karşılaştırılması.....	29
4.	SİNÜZOİDAL SENTEZ.....	31
4.1	Sinüzoidal Ses Modeli.....	31
4.2	Model Parametrelerinin Kestirimi.....	32
4.3	Analiz ve Sentez İşlemlerinin Genel Yapısı.....	34
4.3.1	Genlik ve faz spektrumunun hesaplanması.....	35
4.3.2	Tepe bulma.....	37
4.3.3	Tepe uydurma.....	37
4.3.4	Sinüzoidal sentez.....	41
5.	UYGULAMA.....	46
5.1	Analiz Verisinin Hazırlanması.....	46
5.2	Pencereleme.....	46
5.3	Tepe Bulma.....	47
5.4	Tepe Uydurma.....	48
5.5	Analiz Verisinin Saklanması.....	48
5.6	Sinüzoidal Sentez.....	49
5.7	Uygulama Ortamı.....	49
6.	SONUÇ.....	50
	KAYNAKLAR.....	52
	EKLER.....	54
	Ek 1 Program Dökümü.....	55
	ÖZGEÇMİŞ.....	61

ÖNSÖZ

Başta Yrd. Doç. Dr. Ünal Küçük, Arş. Gör. Osman Nuri Ertürk, Arş. Gör. Cumhuriyet Erkut, Elk. ve Hab. Mh. Önder Çelik, Elk. ve Hab. Mh. Ahmet Akın olmak üzere, çalışmada emeği geçen herkese teşekkür ederim.



ÖZET

Sayısal ses sentezi yöntemlerinden, elektronik müzik ve ses sıkıştırma uygulamalarında yararlanılmaktadır. Elektronik müzikte kullanılacak sentez modellerinin, en az işlemle en yüksek kaliteyi sağlaması ve müzisyenin amacına uygun ses değişimlerini gerçekleştirmesi hedeflenmektedir. Ses sıkıştırma amaçlı kullanımlarda ise, modelin ses verisini bozmadan saklama alanını en aza indirmesi gerekmektedir. Bu çalışmada, hem elektronik müzik hem de ses sıkıştırma alanlarında kullanılan sinüzoidal sentez ele alınmıştır. Sinüzoidal sentez spektral modeller ailesinin bir üyesidir. Analiz temelli yapısı nedeniyle, sinüzoidal sentez modeli, her geçen gün daha çeşitli alanlarda kullanılmaktadır. Spektral modellerin kurulması yüksek bir işlem hacmi gerektirmekle birlikte, sinüzoidal modelde asıl işlem yükü analiz kısmındadır. Sentez kısmı daha az işlem yoluyla gerçekleştirilebildiği için, sinüzoidal model elektronik enstrümanlarda kullanılmaya elverişli bir yapıdadır. Yapılan uygulamalarda görülmüştür ki, harmonik yapıda sesler bu model ile iyi bir şekilde sentezlenebilmektedir. Ayrıca, bu model ile perde değiştirme ve geçiş etkisi gibi dönüşümler kolayca gerçekleştirilmiştir. Sıkıştırma açısından ise, uygulamalarda 164:13 gibi yüksek oranlar elde edilmiştir.



ABSTRACT

Digital sound synthesis models are used in computer music and sound compression applications. In computer music applications, synthesis models have to be designed to offer best quality with minimum computational cost and produce the expected sound effects. As for sound compression, the model must decrease the storage capacity but not the signal quality. In this work, the sinusoidal modelling technique which is used in both computer music and sound compression was realised. Sinusoidal synthesis is a member of spectral models family. Because it is based on analysing methods, the application fields of sinusoidal modelling is getting wider continuously. Although a considerable amount of computation is done to realise the analysing stage. Since the synthesis stage can be accomplished with less computation, sinusoidal modelling has a suitable structure for the simulation of musical instruments. As a result of the applications, it was seen that the harmonic sounds were synthesised well through sinusoidal modelling. In addition, the transformations such as pitch shifting or morphing effect were realised easily by using the designed model. When it comes to compression, the high ratios such as 164:13 were obtained.



1. GİRİŞ

Sunulan bu tez çalışmasında, sinüzoidal model incelenmiş ve bu modelle müzikal seslerin sentezlenmesine çalışılmıştır. Sinüzoidal sentez modeli, spektral modellerin bir alt koludur. Sinüzoidal model, bir ses işaretinin Fourier Dönüşümü uyarınca sinüslere ayrıştırılmasına ve bu sinüsler yardımıyla işaretin tekrar oluşturulmasına dayanır. Model, ses üzerinde müzikal değişimler yapılmasına izin veren bir yapıda olduğu için, elektronik müzik alanında geniş kullanım alanları bulmuştur. Ayrıca, analiz ile bulunan veriler asıl işarete göre daha az yer kapladığı için, sinüzoidal sentez bir ses sıkıştırma yöntemi olarak da kullanılabilir.

Çalışmanın ikinci bölümünde, akustiğin temel kavramları ve ses işareti ile ilgili tanımlar üzerinde durulmuştur. Toplamsal sentezin temel kabulleri ve sinüzoidal model için çok önemli olan sinüzoidal dalgalar bu bölümde anlatılmıştır.

Üçüncü bölüm, ses sentezi yöntemlerine ayrılmıştır. Sentez yöntemleri işlenmiş kayıt, spektral modeller ve fiziksel modeller olarak üç ana başlık altında ele alınmıştır. Bu ana başlıklar altında sentez yöntemleri anlatılırken, sinüzoidal model için önemli bir kavram olan Kısa Süreli Fourier Dönüşümü de incelenmiştir. Bu bölümün sonunda, spektral modeller ve fiziksel modeller arasında bir karşılaştırma yer almaktadır.

Dördüncü ve beşinci bölümler sinüzoidal senteze ayrılmıştır. Dördüncü bölümde modelin teorik kısmı anlatılırken, beşinci bölümde uygulamada dikkat edilmesi gereken hususlar ve modelin iyileştirilmesine yönelik çalışmalar işlenmiştir.

2. SES VE MÜZİK

2.1 Akustik Dalgalar

Akustik dalgalar, katı, sıvı ve gaz ortamlarda yayılan uzamsal dalgalardır. İnsan kulağı tarafından algılanabilen akustik dalgalar ses olarak nitelendirilir. Bir akustik dalga'nın insan kulağınca algılanabilmesi için frekansının 20-22050 Hz arasında olması gereklidir.

Akustik dalgalar, tıpkı elektromanyetik dalgalar gibi kırılma, yansıma, girişim, zayıflama özelliklerine sahiptir. Homojen ortamlarda yayılan akustik dalgalar, doğrusallık ve süperpozisyon teorilerini sağlarlar. Bu teoriler, ses analizi/sentezinde önemli yerlere sahiptir.

Ses analizi/sentezi yöntemlerinde sesin yapısının bilinmesi oldukça önemlidir. Kurulan modelin işlenecek olan sese uygun olması, uygulamada daha doğru sonuçlar verecektir. Bu çalışmada müzikal seslerin üzerinde durulacağına göre, öncelikle müzikal seslerin yapısı üzerinde durulması gereklidir.

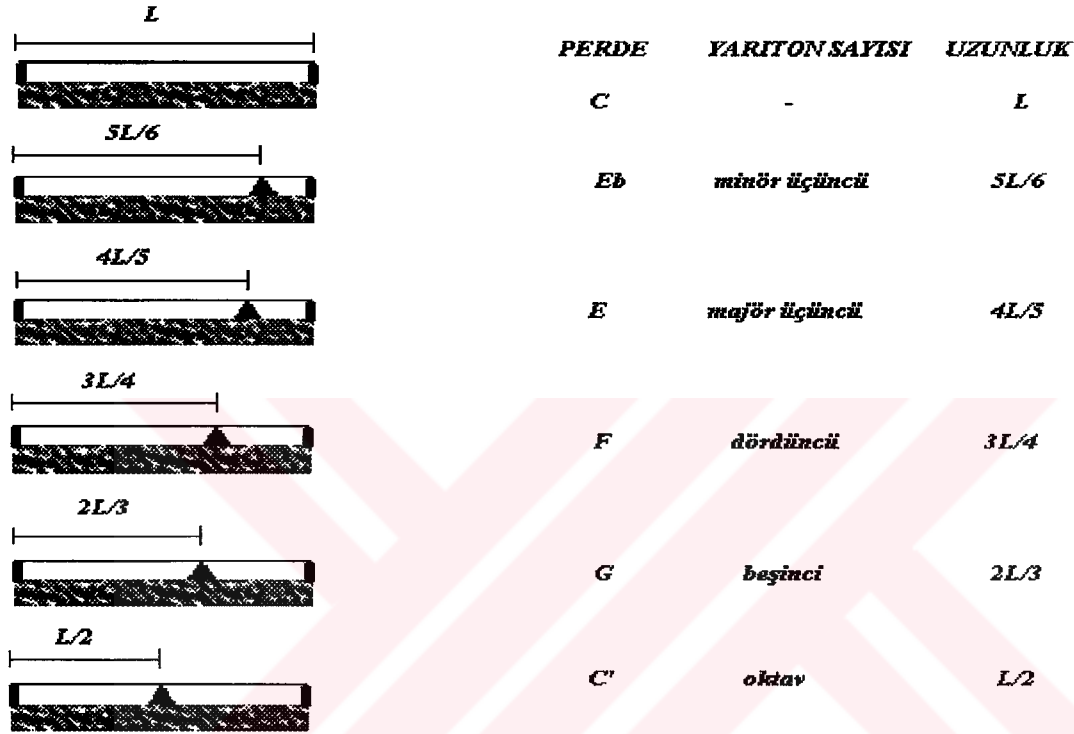
2.2 Perde ve Periyodiklik

Perde, sesin tizliğinin ölçüsüdür. Tanımlanabilir bir perdeye sahip olan müzikal sesler periyodiktir. Bu tanımdan görülebilir ki, perde ve periyot arasında bir ilişki söz konusudur. Bir siren, perde ve periyot arasındaki ilişkiyi anlamakta iyi bir örnektir.

Mekanik bir siren, hava üfleyen bir boru ve bu borunun önünde dönen delikli bir diskten oluşur. Deliklerden biri borunun ağzına denk geldiğinde, delikten geçen hızlı hava, atmosferde ilerleyen bir hava atımı oluşturur. Örneğin bu atımların sayısı saniyede 440 ise, siren konser perdesinden bir la notası ($f_0 = 440$ Hz.) üretir. Sirenin diski daha hızlı çevrildiğinde saniyedeki atım sayısı artar ve daha yüksek perdeden bir ses üretilir. Görüldüğü gibi, hava atımlarının periyodu küçüldükçe yani frekansı arttıkça, üretilen sesin perdesi yükselmektedir. Sirenin ürettiği periyodik hava atımları, havada ses dalgası adı

verilen periyodik yayılımlar oluşturur. Bu dalgalar kulağa ulaştığında, belli bir perdeden bir ses olarak algılanır.

Siren örneğinde, üretilen sesin perdesini belirleyen etkenler kolayca görülebilir. Perdeyi belirleyen sesin periyodu, başka bir deyişle frekansıdır. Sirenin ürettiği sesin frekansı ise



Şekil 2.1 Titreşen bir telin uzunluğu ve perdesi arasındaki ilişki

diskin saniyedeki dönme sayısı ile disk üzerindeki deliklerin sayısının çarpımına eşittir. Ancak bu formül, yalnızca siren örneği için geçerlidir. Değişik ses kaynakları için perde ve frekans arasındaki ilişki bu kadar basit değildir.

Titreşen bir telin frekansı ve perdesi arasındaki ilişki , Galileo tarafından bulunmuştur. Ancak Galileo'nun çalışmasında, yalnızca titreşen telin titreşim sayısı ile perdesi arasındaki ilişki üzerinde durulmuş, telin titreşim sayısını belirleyen etkenler araştırılmamıştır (Pierce, 1983).

1636'da Fransız rahip, filozof ve matematikçi Marin Mersenne titreşen bir telin frekansının aşağıdaki formülle ifade edilebileceğini göstermiştir:

$$frekans = k * \frac{\sqrt{gerginlik}}{uzunluk * \sqrt{\text{özkütle}}} \quad (2.1)$$

Şekil 2.1'deki gibi gergin bir tel olsun. Bu tele pena ya da parmakla vurulduğunda, tel do notası üretsinsin. Eğer bir katı tampon yardımıyla telin gerginliği değiştirilmeden uzunluğu azaltılırsa, üreteceği sesin perdesi yükselecektir. Örneğin telin uzunluğu üçte ikisine indirilirse, üreteceği sesin perdesi asıl notanın beşincisi (yedi yariton üstü) olacaktır.

Bir sesin perdesini belirleyen tek etken sesin periyodudur. Dengeleyici (equalizer) gibi cihazlar yardımıyla sesin bas ya da tiz bileşenlerinin kuvvetlendirilip zayıflatılmasıyla sesin perdesi değişmez. Bu gibi işlemler yalnızca sesin parlaklığını etkilemektedir.

2.3 Sinüzoidal Dalgalar

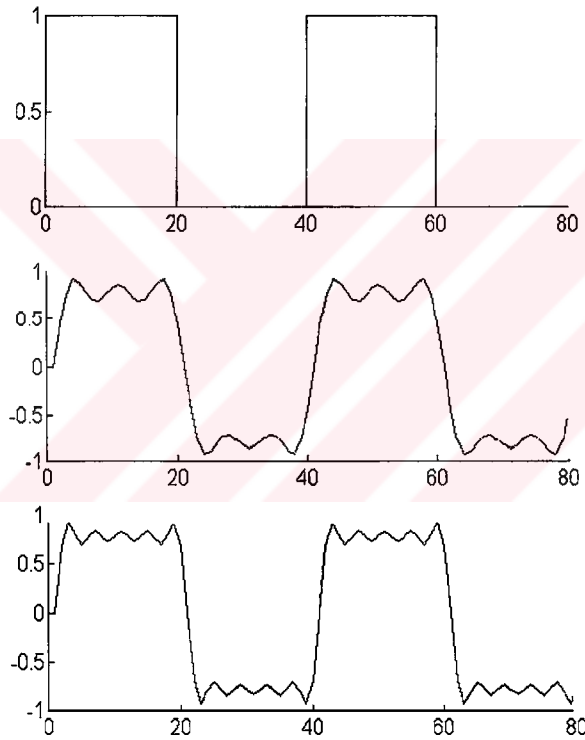
Önceki bölümde sesin, bir kaynağın hava basıncını değiştirmesiyle oluştuğundan bahsedilmişti. Müzikal enstrümanların oluşturduğu hava basıncı, genellikle periyodik değişimler gösterir. Bu değişimleri gözleminin en kolay yolu, ses kaynağına bir mikrofon tutmak ve mikrofonu da bir osiloskoba bağlamaktır.

Hava basıncındaki sinüzoidal değişimlerle oluşan ses dalgalarına sinüzoidal sesler ya da saf tonlar denir. Saf tonlar müzikal anlamda değer taşımasa da, analiz ve sentez yöntemleri açısından önemlidir. Bir sinüzoidal dalga, genlik, periyot ve fazıyla tanımlıdır. Genlik, sinüzoidal dalganın tepe değerini gösterir. Periyot, dalganın iki tepesi arasındaki süredir. Faz ise, dalganın başlangıç anında düşey eksenini kestiği yeri gösterir.

Fransız matematikçi François Marie Charles Fourier kendi adıyla anılan ve periyodik bir dalganın belirlenebilir genlik, faz ve frekanslara sahip sinüzoidallerin toplamıyla ifade edilebileceğini gösteren bir teori geliştirmiştir. Buna göre herhangi bir $f(t)$ işareti,

$$f(t) = \sum_{k=0}^{\infty} A_k \sin(k\omega t + \phi_k) \quad (2.2)$$

şeklinde ifade edilir. Karmaşık bir dalganın Fourier gösterimi genellikle çok sayıda, hatta sonsuz sayıda bileşen gerektirir. Bununla birlikte, belli bir yaklaşıklıkla sonlu sayıda bileşen kullanılabilir. Örneğin birim genlikli ve f_0 frekanslı bir kare dalga, $f_0, 3f_0, 5f_0, 7f_0 \dots$ frekanslı, $1, 1/3, 1/5, 1/7 \dots$ genlikli ve eşit fazlı sinüzoidallerin toplamıyla gösterilebilir (Şekil 2.2). Burada da görülebilir ki, sonuç dalga şeklinin kalitesini kullanılan sinüzoidallerin sayısı belirlemektedir. Ne kadar fazla bileşen kullanılırsa, asıl işarete o kadar yakın bir sonuç elde edilir.



Şekil 2.2 Bir kare dalganın Fourier Serisi ile elde edilmesi: (a) Kare dalga; (b) Üç bileşenli yaklaşım; (c) Altı bileşenli yaklaşım

2.4 Periyodik Seslerin Adlandırılması

Periyodik seslerin sinüzoidal bileşenlerini isimlendirmede üç farklı sistem vardır. Bu çalışmada, Çizelge 2.1'deki sistemlerden ilki kullanılacaktır (Pierce, 1983).

Bazı sesler temel harmoniğin tam katı olmayan frekanslarda kısımlara sahiptir. Bu kısımlar, nonharmonik bileşenler olarak adlandırılacaktır.

Çizelge 2.1 Frekans bileşenlerinin isimlendirilme sistemleri

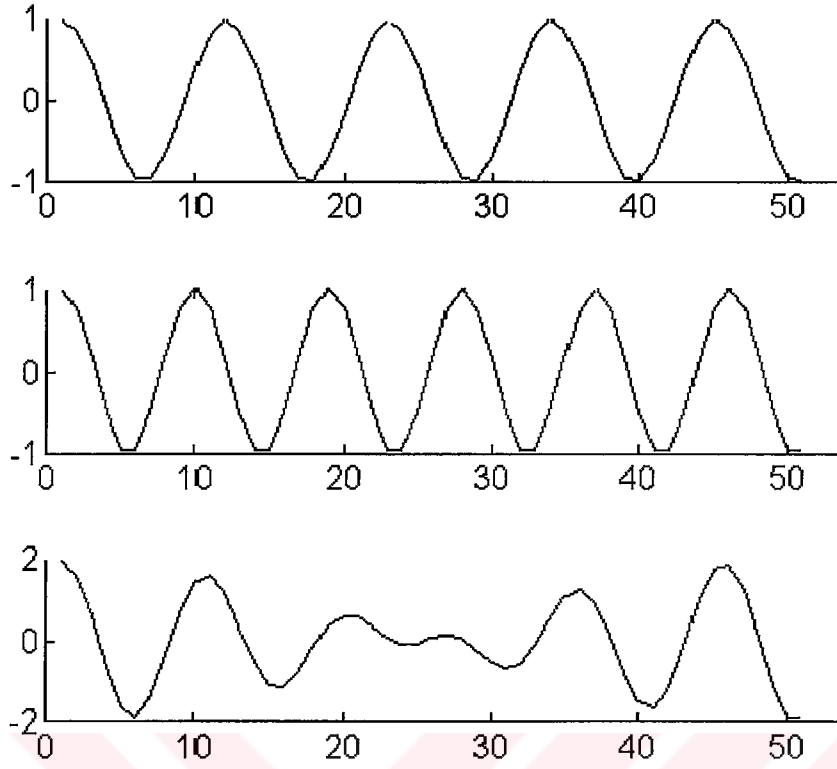
Frekans	Harmonikler	Üst tonlar	Bileşenler
f_0	Temel (ana) harmonik	Temel (ana) ton	Birinci bileşen
$2 f_0$	Birinci harmonik	Birinci üst ton	İkinci bileşen
$3 f_0$	İkinci harmonik	İkinci üst ton	Üçüncü bileşen
$4 f_0$	Üçüncü harmonik	Üçüncü üst ton	Dördüncü bileşen

2.5 Vuru Olayı ve Maskeleye

Vuru olayı, frekansları birbirine yakın olan iki farklı sesin aynı anda dinlenmesi sırasında oluşur. Eğer iki sesin frekansları birbirine yeterince yakınsa bu iki ses ayrı olarak algılanamayacak, sanki frekansı iki sesin frekanslarının ortasında olan ve genliği bu iki sesin frekansları arasındaki farka bağlı olarak salınım yapan bir tek ses varmış gibi algılanacaktır. Seslerden birinin frekansı diğerine doğru yavaşça yaklaştırıldığında, vuru olayı azalacak, frekanslar eşitlendiğindeyse tamamen kaybolacaktır (Şekil 2.4).

Eğer seslerin frekansları yakınlaştırılmak yerine yavaşça uzaklaştırılırsa, vuru olayı, önce yerini seste bir pürüzlenmeye bırakacak, frekansların arası arttırılmaya devam edilirse bir süre sonra iki ses ayrı olarak duyulacaktır. İki sesin ayrı olarak duyulamadığı bölgeye ‘kritik bant genişliği’ adı verilir (Pierce, 1983). Kritik bant genişliği, bir oktavin dörtte biri, diğer bir deyişle minör üçüncü kadardır.

Kritik bant genişliği, ses gücünün algılanmasında ve bir sesin diğeri tarafından maskelenmesinde önemlidir. Özetle, kritik bant genişliği kulağın frekanslara duyarlılığını gösterir. Düşük ve orta yükseklikteki seslerde kritik bant genişliğinden daha ayrı olan frekans bileşenleri farklı sınırlardan beyne iletilirken, kritik bant genişliği içinde yer alan frekans bileşenleri aynı sınırları kullanmaktadır.



Şekil 2.3 Yakın frekanslı iki sinüs ve toplamı.

Göze gelen parlak bir ışığın kısa bir süre için körlüğe neden olması gibi, yüksek genlikli bir ses de kulağın duymasını engeller. Ancak göze gelen parlak ışık neredeyse tüm renklerde algılamayı engellerken, yüksek ses yalnızca bazı frekansların duyulmasını engelleyebilir. Zayıf bir sesin algılanmasının güçlü bir ses tarafından engellenmesine 'maskeleme' adı verilir. Güçlü sese 'maskeleyici', zayıf sese ise 'maskelenen' denir. Düşük frekanslı seslerin maskeleme özellikleri yüksek frekanslı seslere göre daha fazladır. Kulağın bu özelliğinden yararlanılarak MPEG-3 ses kodlaması geliştirilmiştir. Bu kodlama tekniğinde temel olarak yapılan, sesin alt bantlara ayrılıp maskelenen alt bantların atılmasıdır. Böylece algısal bilgi kaybı olmadan veri azaltımına gidilmiş olur (Pan, 1993).

2.6 Ses Gücü

Bir ses kaynağı, watt cinsinden ölçülebilir toplam güce sahip ses dalgaları yayar. Ses dalgaları küresel olarak yayıldığına göre, kulağa gelen dalgaının metrekaresine başına taşıdığı güç, yani güç yoğunluğunu ölçmek daha verimli olacaktır. Ses şiddeti, sesin metrekaresine

başına taşıdığı güç olarak tanımlanır. Ancak güç yoğunluğundan farklı olarak desibel cinsinden ölçülür. Desibel birimini kullanabilmek için bir referansa ihtiyaç vardır. Ses şiddeti ölçülürken referans olarak insan kulağının algılayabildiği en düşük güç yoğunluğu olan 10^{-12} W/m^2 kullanılır

İnsan kulağı oldukça hassastır. Örneğin 1 wattlık 3500 Hz frekanslı bir saf tonun duyulabileceği mesafe, sesin her yönde eşit oranda yayıldığı varsayılırsa 564 kilometredir.

Ses şiddeti ile kaynaktan uzaklık arasındaki ilişki basittir. Yansımanın ve engellerin olmadığı durumda kaynağa olan uzaklığın iki katına çıkarılması ses şiddetinde 6 dB düşüğe sebep olur.

Eğer iki ses kritik bant genişliğinden daha ayrıksa, toplam şiddetleri tek tek şiddetlerinin toplamına eşittir. Ancak bu iki ses kritik bant genişliği içerisindeyse, toplam güç yoğunlukları, tek tek güç yoğunluklarının toplamına eşittir. Her iki sesin de eşit şiddette olduğu varsayılırsa, kritik bant genişliği dışındayken toplam ses şiddeti iki katına çıkarken, kritik bant genişliği içindeyken yalnızca 3 dB artacaktır (Pierce, 1983).

3. SES SENTEZİNE GENEL BAKIŞ

Müzikal seslerin sayısal sentezi, işaret işleme tekniklerinin gelişmesiyle birlikte hız kazanmıştır. Bilgisayar teknolojisinin ilerlemesi ve ucuzlamasıyla da ses işleme uygulamaları yaygınlaşmaya başlamıştır.

İlk sayısal ses sentezi denemeleri, 1957 yılında Bell Laboratuvarlarında Max V. Matthews tarafından başlatılmıştır. Matthews bu çalışmalarında ses sentezine ait iki temel problemi ortaya koymuştur:

- Bir sesi tanımlayan fonksiyonun büyük miktarda veri gerektirmesi;
- Karmaşık seslerin üretilebileceği güçlü ama basit bir dil geliştirilmesi.

İlk problem bilgisayar teknolojisinin hızlı gelişimiyle kendiliğinden çözülmüştür. Artan işlemci hızı ve ucuzlayan bellek birimleri büyük miktarda verinin işlenmesini kolaylaştırmıştır. İkinci problem ise henüz çözülmüş değildir. Bir sesi oluşturan verinin boyutları düşünülürse, müzikal bir parçanın örnek örnek yazılması oldukça uzun sürer. Bu yüzden ses örnekleri algoritmik ya da çözümsel olarak sentezlenmelidir. Böylelikle çok sayıdaki örnekler yerine bu örnek dizilerini üretebilecek az sayıda veriye gerek duyulacaktır (Moorer, 1977; Smith, 1991).

3.1 Ses Sentezi Yöntemlerinin Sınıflandırılması

İlk kullanılan yöntemler, bir bilgisayar programı yardımıyla istenen dalga şeklini oluşturmaya ya da önceden kaydedilmiş sesleri çalmaya dayalıdır. Böyle yöntemlere ‘doğrudan sentez’ ya da ‘işlenmiş kayıt sentezi’ adı verilir. İşlenmiş kayıt sentezi programları, çıkış dalga formunu verecek şekilde arka arkaya çalışan modüllerden oluşur. Bu yöntemlerde bir müzik parçası, enstrüman tanımı ve nota listesi ile oluşturulur. Enstrüman tanımı ile notaların başlangıç anları, süreleri ve güçleri belirlenir. Nota listesi ise üretilecek seslerin perdelerini tanımlar.

İkinci olarak, sesin spektral özelliklerinden yararlanmak ve istenen sesin dalga şekli yerine,

sesin frekans spektrumunu elde etmeye dayalı yöntemler önerilebilir. Bu yöntemlere genel olarak ‘spektral modeller’ denir. Bu yöntemlerde önemli olan sesi üreten kaynak değil, alıcıya ulaşan sestir.

Sesin kendisi yerine ses kaynağını taklit eden yöntemlere ‘fiziksel modeller’ adı verilir. Bu tür modellerde ses kaynağının mekanik ve akustik özelliklerinin matematiksel modelleri çıkarılmaya çalışılır (Smith, 1991).

Bu sınıflandırmayı farklı bir şekilde yapmak ta mümkündür. Karmaşık bir sesin basit bileşenlerle elde edildiği yöntemlere ‘toplamsal’ yöntemler, sesin daha karmaşık bir işaretin istenmeyen bölümlerinin atılmasıyla elde edildiği yöntemlere ise ‘çıkarımsal’ yöntemler adı verilir (Moorer, 1977). Ses sentezi yöntemlerinin sınıflandırılması Çizelge 3.1’de verilmiştir.

Ses sentezi yöntemlerinin sınıflandırılmasında çok katı kurallar yoktur. Örneğin Karplus-Strong modeli hem fiziksel modellere hem de çıkarımsal modellere dahil edilebilir. Ya da Doğrusal Öngörü Kodlaması çıkarımsal bir model, Sinüsler + Gürültü toplamsal bir model olduğu halde her ikisi de spektral modeller sınıfına dahildir.

Çizelge 3.1 Ses sentezi yöntemlerinin sınıflandırılması

İşlenmiş Kayıt	Spektral Model	Fiziksel Model
Musique Concrète	Frekans Domeni Dalga Tablosu	Ruiz Teller
Örnekleme	Toplamsal	Karplus-Strong
Vektör	Faz Kodlayıcı	Dalga Kılavuzu
Granüler	Çakışma/Toplama	Modal
Zaman Domeni Dalga	Sinüzoidal	Cordal-Anima
Tablosu	Sinüsler+Gürültü(Serra)	Mosaic
	Risset FM	
	Chowning FM Ses	
	Çıkarımsal	
	Doğrusal Öngörü Kodlaması	
	Ters FFT	

3.2 İşlenmiş Kayıt (Doğrudan Sentez)

Doğrudan sentez yöntemleri, kayıt aletlerinin geliştirildiği günlerden beri kullanılmaktadır. Bu tür yöntemler iki ana gruba ayrılır. İlk grupta Musique Concrète gibi karmaşık sesler üretebilen programlar yer alır. İkinci grup ise Vektör, Zaman Domeni Dalga Tablosu gibi önceden kaydedilmiş verinin çalınmasına dayanan yöntemlerden oluşur.

Bu yöntemlerde temel sorun, saklanması gereken verinin oldukça büyük yer kaplaması ve ses üzerinde fazla değişiklik yapılamamasıdır (Moorer, 1977; Smith, 1991).

3.3 Spektral Modeller

Spektral modellerde ana fikir doğal sesin spektral özelliklerini taklit etmektir. Yani bu modeller ses kaynağının değil, doğrudan sesin özelliklerini ele alır. Spektral modeller iki ana gruba ayrılır: Toplamsal sentez ve çıkarımsal sentez. Toplamsal sentezde spektral olarak basit bileşenlerin toplanmasıyla istenen sese ulaşılır. Çıkarımsal sentezde ise istenen ses spektral olarak daha zengin olan bir işaretten (örneğin beyaz gürültü) süzgeçler yardımıyla elde edilmeye çalışılır.

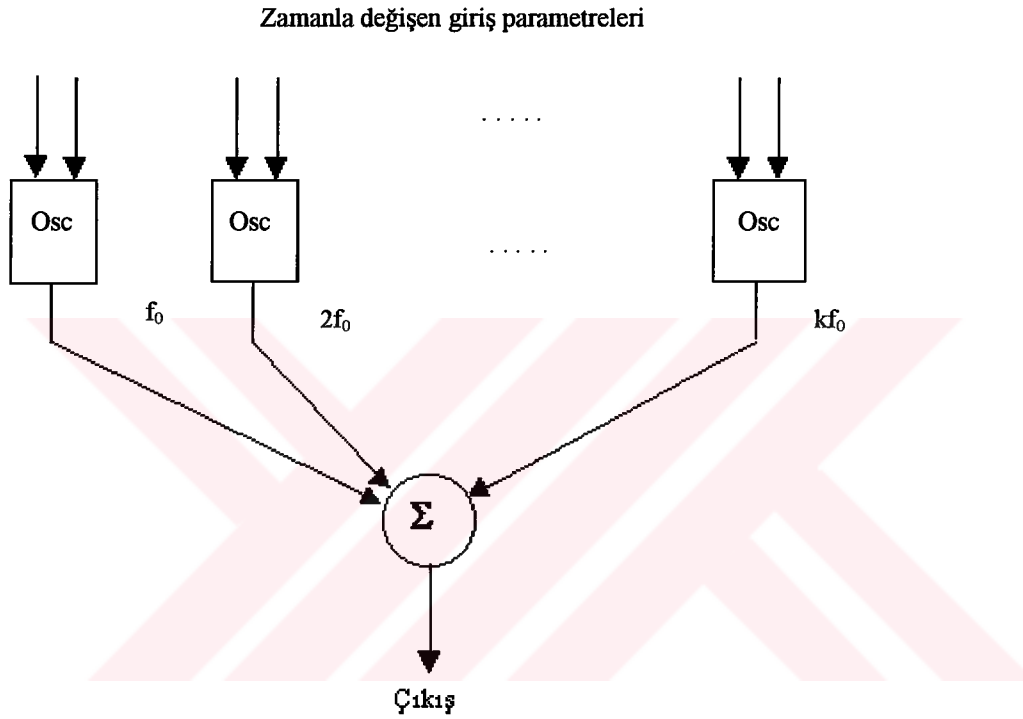
3.3.1 Toplamsal sentez

Toplamsal sentez, en eski yöntemlerden biridir. İlk toplamsal sentez uygulaması, 1906 yılında gerçekleştirilen Tellharmonium sentezcisidir. Tellharmonium sentezcisi, karmaşık sesler elde edebilmek için çok sayıda elektriksel ton üreteçleri kullanmaktaydı. Aynı fikirden yola çıkılarak yapılan bir başka toplamsal sentez enstrümanı da Hammond orgudur. Hammond orgu, elektromekanik ton çarkları ile karmaşık sesler üreten bir yapıya sahiptir (Roads, 1996).

Toplamsal sentezin temel yapısı, basit dalga formlarını üst üste ekleyerek karmaşık dalga formları elde etmektir. Alt bileşenlerin çoğu tek başına bir anlam ifade etmeyebilir ancak her biri toplam sesin kalitesini etkileyecektir. Toplamsal sentez yöntemlerinin en büyük avantajı sesin analizinde kullanılan Fourier dönüşümünün teorik alt yapısının

sağlamlığındadır. Bu nedenle toplamsal sentez kimi zaman Fourier Sentezi olarak adlandırılır.

Bilindiği üzere, Fourier teoremine göre Dirichlet koşullarını sağlayan her periyodik dalga formu sinüzoidal harmoniklerin toplamı ile ifade edilebilir. Bu sinüzoidallerin Fourier dönüşümü ile bulunan genlik ve faz bilgileri kullanılarak periyodik dalga formları oluşturabilecek bir düzenek yapılabilir (Şekil 3.1).



Şekil 3.1 Fourier Dönüşümü ile periyodik dalga formlarının elde edilmesi

Matematiksel olarak, T periyotlu herhangi bir $f(t)$ dalga formu Fourier Teoremi uyarınca aşağıdaki gibi ifade edilebilir:

$$f(t) = \sum_{k=0}^{\infty} A_k \sin(k\omega t + \varphi_k) \quad (3.1)$$

Burada A_k k. sinüzoidalın genliği, ω dalga formunun temel harmoniği, φ_k da k. sinüzoidalın fazıdır. Bu eşitlik aşağıdaki gibi düzenlenebilir:

$$f(t) = \sum_{k=0}^{\infty} (a_k \cos k\omega t + b_k \sin k\omega t) \quad (3.2)$$

Burada a_k ve b_k Fourier katsayılarıdır ve aşağıdaki gibi hesaplanırlar:

$$a_k = \frac{2}{T} \int_0^T f(t) \cos \frac{2\pi t}{T} dt \quad (3.3)$$

$$b_k = \frac{2}{T} \int_0^T f(t) \sin \frac{2\pi t}{T} dt \quad (3.4)$$

(3.1) eşitliğindeki genlik ve faz bilgileri ise Fourier katsayıları yardımıyla şu şekilde bulunur:

$$A_k = \sqrt{a_k^2 + b_k^2} \quad (3.5)$$

$$\phi_k = \tan^{-1} \left(\frac{a_k}{b_k} \right) \quad (3.6)$$

Fourier Dönüşümü, gerçekten periyodik dalga formları ile çalışırken oldukça yararlıdır. Ancak çoğu müzikal ses tam periyodik değildir. Ayrıca zamanla frekansları değişime uğrar. Bu durumda, Fourier Dönüşümü yerine zaman bilgisinin de işin içine katılabileceği Kısa Süreli Fourier Dönüşümü(KSFD) kullanmak daha doğru olacaktır (Moorer, 1977; Roads, 1996; Serra, 1997; George ve Smith, 1997).

3.3.1.1 Kısa süreli Fourier dönüşümü

KSFD, temel Fourier Dönüşümünün bir uzantısıdır. Bu bağlamda, tıpkı Fourier Dönüşümündeki gibi ayrık zaman KSFD ve ayrık KSFD 'nden bahsetmek mümkündür.

Bir $x(n)$ dizisi için ayrık zaman KSFD farklı zaman aralıklarında gözlemlenmiş $x(n)$ 'in ayrık zaman Fourier Dönüşümlerinden oluşur. Zaman aralıkları, $x(n)$ dizisini kaydırılmış

bir pencereleme fonksiyonu $w(n_0-n)$ ile çarparak oluşturulur. $x(n)$ dizisinin n_0 anındaki ayrık zaman KSFD şu şekilde hesaplanır:

$$X(n_0, \omega) = \sum_{n=-\infty}^{\infty} x(n)w(n_0 - n)e^{-j\omega n} \quad (3.7)$$

Burada $w(n)$ dizisine ‘analiz penceresi’ ya da ‘analiz süzgeci’ adı verilir. $f_{n0}(n)=x(n)w(n_0-n)$ dizisi, $x(n)$ dizisinin n_0 anındaki kısa zaman dilimi olarak adlandırılır. n_0 anındaki kısa zaman dilimi bulunduktan sonra Fourier Dönüşümü alınırsa, $x(n)$ ‘in n_0 anındaki $X(n_0, \omega)$ KSFD bulunmuş olur. $X(n_0+1, \omega)$ ise analiz penceresinin zamanda bir kez kaydırılıp yukarıdaki işlemlerin tekrarlanmasıyla bulunur. (3.7) eşitliği genişletilecek olursa, $x(n)$ dizisinin her hangi bir anındaki ayrık zaman KSDF aşağıdaki gibi bulunur:

$$X(n, \omega) = \sum_{m=-\infty}^{\infty} x(m)w(n - m)e^{-j\omega m} \quad (3.8)$$

Sayısal işlemede, ayrık KSFD kullanılır. Ayrık zaman KSFD’nün ayrık zaman Fourier Dönüşümünden türetilmesi gibi ayrık KSFD de ayrık Fourier Dönüşümünden (AFD) türetilir. Ayrık zaman Fourier Dönüşümü ve AFD arasındaki ilişkiden yararlanılarak ayrık KSFD şöyle tanımlanır:

$$X(n, k) = X(n, \omega) \big|_{\omega=2\pi k/N} R_N(k) \quad (3.9)$$

ve,

$$R_N(k) = u(k) - u(k - N) \quad (3.10)$$

dir. (3.8) eşitliği, (3.9) eşitliğinde yerine konulduğunda,

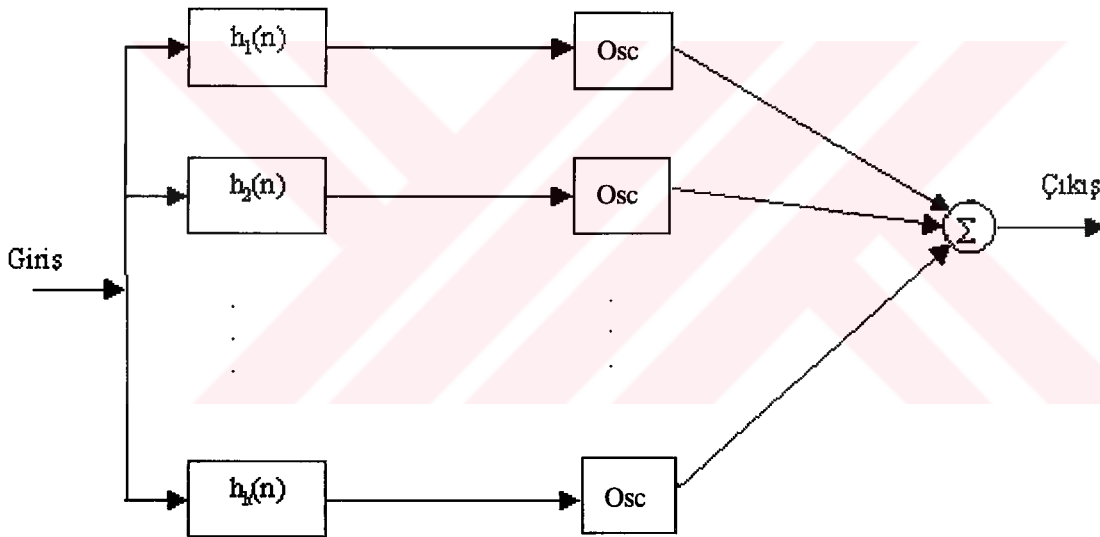
$$X(n, k) = \sum_{m=-\infty}^{\infty} x(m)w(n - m)e^{-j2\pi km/N} R_N(k) \quad (3.11)$$

elde edilir (Harris, 1978; Oppenheim ve Schafer, 1983).

3.3.1.2 Faz kodlayıcı

Faz kodlayıcı, J. L. Flanagan ve R. M. Golden tarafından AT&T Bell Laboratuvarlarında geliştirilmiştir. Temel fikri ise H. Dudley 'in kanal kodlayıcısından alınmıştır. Bu kodlayıcılar temelde geniş spektrumlu bir işareti dar spektrumlu kanallara bölerler. Daha dar spektrumlu işaretler, saklama ve iletim ortamlarında görece daha az yer kaplarlar.

Faz kodlayıcı basitçe bir dizi osilatörü süren süzgeç bankaları gibi düşünülebilir (Şekil 3.2). Her süzgeç asıl işaretin bir kısmını analiz eder; zamanla değişen genlik ve frekans bilgilerini hesaplar. Bu veriler bir dizi osilatörü sürmede kullanılır (Moorer, 1977).



Şekil 3.2 Faz kodlayıcının temel yapısı

Genlik ve frekans kontrol verileri uygun şekillerde değiştirilerek, asıl işareten farklı işaretler elde etmek mümkündür. Özellikle işaretin süresini değiştirmeden perdesini ya da perdesini değiştirmeden süresini değiştirmek olasıdır.

3.3.1.2.1 Fourier dönüşümü ve süzgeç bankası

Önceki bölümlerde bir $x(n)$ işaretinin KSFD aşağıdaki gibi tanımlanmıştır:

$$X(n, \omega) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-j\omega m} \quad (3.12)$$

Bu ifade iki şekilde yorumlanabilir:

- Fourier Dönüşümü Yaklaşımı: Eğer zaman indisi n sabit tutulursa ifade $w(n-m)x(n)$ pencerelemiş işaretin Fourier dönüşümü olacaktır.
- Süzgeç Bankası Yaklaşımı: Eğer açısal frekans ω , ω_k gibi sabit bir değerde tutulursa, ifade $x(n)$ işaretinin uygulandığı bir süzgeci tanımlar. Bu süzgecin dürtü yanıtı ise,

$$h(n) = w(n)e^{j\omega_k n} \quad (3.13)$$

dir.

Toplam indekisinde bir değişimle (3.10) eşitliği aşağıdaki gibi düzenlenebilir:

$$X(n, \omega) = \sum_{m=-\infty}^{\infty} x(n-m)w(m)e^{-j\omega(n-m)} \quad (3.14)$$

$$X(n, \omega) = e^{-j\omega n} \sum_{m=-\infty}^{\infty} x(n-m)w(m)e^{-j\omega m} \quad (3.15)$$

ω 'nın belirli bir değeri ω_k ile gösterilecek olursa, (3.15) eşitliği aşağıdaki gibi düzenlenebilir:

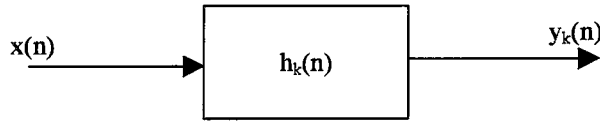
$$X(n, \omega) = e^{-j\omega n} \sum_{m=-\infty}^{\infty} x(n-m)h_k(m) \quad (3.16)$$

Burada $h_k(n)$ analiz süzgecidir ve aşağıdaki gibi tanımlanır:

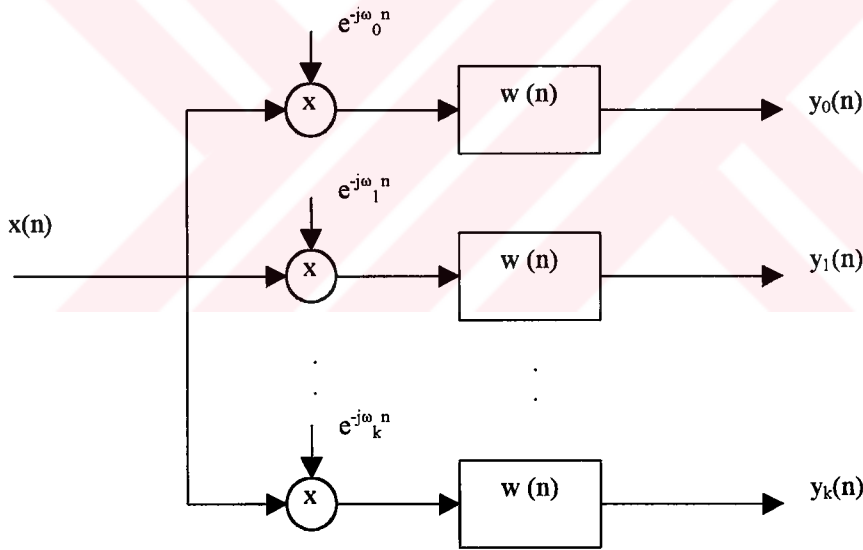
$$h_k(n) = w(n)e^{-j\omega n} \quad (3.17)$$

(3.16) eşitliği, $x(n)$ işaretinin dürtü yanıtı $h_k(n)$ olan bir süzgeçten geçirildiğini ve süzgeç çıkışının $e^{-j\omega_k n}$ ile modüle edildiğini gösterir. Dikdörtgen, Hamming gibi pencereler alçak geçiren süzgeç özellikleri taşıdığı ve $e^{-j\omega_k n}$ gibi karmaşık bir sinüzoidle çarpıldığı için, $h_k(n)$ süzgeci orta frekansı ω_k radyan frekansında olan bir bant geçiren süzgeçtir (Şekil 3.3).

(3.14) eşitliği farklı ω_k merkez frekanslarına sahip $h_k(n)$ süzgeçlerinin aynı $w(n)$ pencere fonksiyonu ile elde edildiğini gösterir. Bu durumda analiz süzgeç bankası Şekil 3.4'teki gibi kurulabilir.



Şekil 3.3 Faz kodlayıcının bir kanalı



Şekil 3.4 Faz kodlayıcının süzgeç bankaları

(3.16) eşitliği ve Şekil 3.4'te görüleceği gibi, faz kodlayıcı katsayıları bulunurken karmaşık çarpımlar yapılmaktadır. Karmaşık işlemler ise bilgisayar ortamında fazla işlemsel yük getirirler. İşlem hızını arttırmak için Euler 'in karmaşık sinüs gösterimi kullanılabilir:

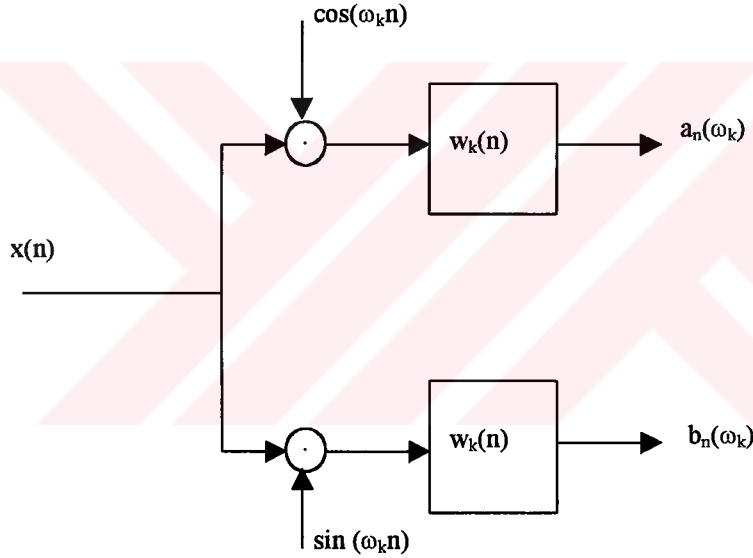
$$e^{-j\omega_k n} = \cos(\omega_k n) - j \sin(\omega_k n) \quad (3.18)$$

Böylece faz kodlayıcının bir kanalı Şekil 3.5'teki hale gelir. Kanal çıkışında bulunan a_n ve b_n katsayılarından sentez işleminin gereksinim duyduğu genlik, frekans ve faz değerleri bulunabilir:

$$|X(n, \omega_k)| = \sqrt{a_n^2(\omega_k) + b_n^2(\omega_k)} \quad (3.19)$$

$$\theta_n(\omega_k) = -\tan^{-1} \frac{b_n(\omega_k)}{a_n(\omega_k)} \quad (3.20)$$

Frekans bilgisi ise, faz bilgisinden dolaylı olarak elde edilir. Frekans, fazın zamana göre türevi alınarak hesaplanır.



Şekil 3.5 Gerçel işlemlerle gerçekleştirilebilen faz kodlayıcının bir kanalı

3.3.1.3 Sinüzoidal sentez

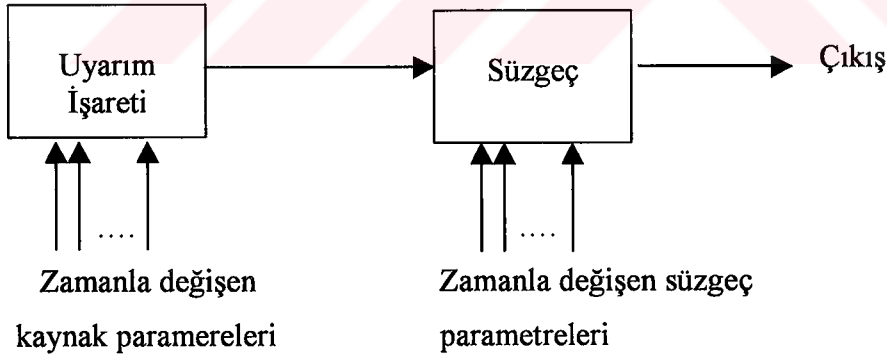
Bölüm 4 tamamen sinüzoidal senteze ayrıldığı için, burada sinüzoidal sentez ayrıntılarıyla anlatılmayacaktır. Bu bölüm, yalnızca sinüzoidal sentezin sentez yöntemleri arasında hangi gruba girdiğini gösterme amacını taşımaktadır.

Sinüzoidal sentez yöntemi analiz temellidir. Analiz işlemi, Fourier dönüşümüne dayalıdır. Fourier Teorisi'ne göre, periyodik bir işaret bir dizi sinüsün toplamıyla ifade edilebilir. O halde Fourier Dönüşümü ile bir işareti sinüzoidal bileşenlerine ayırmak mümkündür. Sinüzoidallerin genlik, frekans ve fazları bulunduğundan sonra bu değerler kullanılarak işaret yeniden elde edilebilir. Sentez işlemi, faz kodlayıcıda olduğu gibi, analiz ile bulunan veriler tarafından kontrol edilen bir dizi osilatörün çıkışlarının toplanmasından ibarettir.

3.3.2 Çıkarımsal sentez

Çıkarımsal sentez yöntemlerinde istenen işaret, spektral olarak zengin bir işaretten zamanla değişen süzgeçler yardımıyla elde edilir. Süzgeçten geçirilecek işaret ve süzgecin transfer fonksiyonu doğrudan belirlenebileceği gibi, istenen işaretin analiz edilmesiyle de bulunabilir.

Çıkarımsal sentezin temel yapısı, bir uyarım işareti ve bir süzgeçten oluşur (Şekil 3.6). Bu tanım her ne kadar fiziksel modelleri de içeriyorsa da, fiziksel modeller, sesin spektral özellikleri yerine ses kaynağının fiziksel özelliklerinden yararlanılarak kurulduğu için, bu sınıfa dahil edilmeyebilirler.



Şekil 3.6 Çıkarımsal sentezin temel yapısı

3.3.2.1 Doğrusal öngörü kodlaması

Doğrusal öngörü, güçlü ve genel bir spektral zarf kestirim yöntemidir. Doğrusal öngörü ile amaçlanan, frekans yanıtı elde edilmek istenen işaretin spektrumuna mümkün olduğunca

benzeyen bir süzgeç tasarlamaktır. Daha sonra bu süzgeç, beyaz gürültü gibi bir işaretle sürülerek istenen işaret elde edilmeye çalışılır. Daha iyi bir sonuç elde etmek için süzgeç genelde zamanla değişen yapıdadır. Kısa süreli spektrum yöntemlerinden farklı olarak, doğrusal öngörü giriş işaretinin spektrumunun zarfının tümünü ele alır (Haykin, 1986).

3.3.2.1.1 Doğrusal öngörü

Zaman dizileri analizinde önemli bir problem, ayrık zamanlı durağan bir rasgele sürecin bilinen bir kısmından yararlanarak gelecekteki değerlerinin bulunmasıdır. Tanım biraz özelleştirilirse, doğrusal öngörü, $u(n)$ zaman dizisinin n . elemanının kendinden önceki $u(n-1)$, $u(n-2)$, ... $u(n-M)$ adet değer ile bulunmasıdır. Bu durumda $u(n)$, kendinden önceki M elemanın bir fonksiyonudur:

$$u(n)=\theta(u(n-1), u(n-2), \dots u(n-M)) \quad (3.21)$$

Burada M , süzgecin derecesi, n , belirler. Öngörünün doğrusal olması için θ fonksiyonunun $u(n-1)$, $u(n-2)$, ... $u(n-M)$ örneklerinin doğrusal kombinasyonu olması gerektiği göz önünde tutulursa, denklem (3.22) şu şekilde yeniden oluşturulur:

$$u(n) = \sum_{k=1}^M w_k u(n-k) \quad (3.23)$$

Burada w_k değerleri ağırlık katsayıları olarak adlandırılırlar. Şekil 3.7'de bir doğrusal öngörü süzgeci görülmektedir. Ağırlık katsayıları Wiener süzgeç teorisi uyarınca, $r(k)$, $u(n)$ dizisinin $u(n-k)$ dizisi ile ilişkisini göstermek üzere (3.24) eşitliğindeki gibi bulunur:

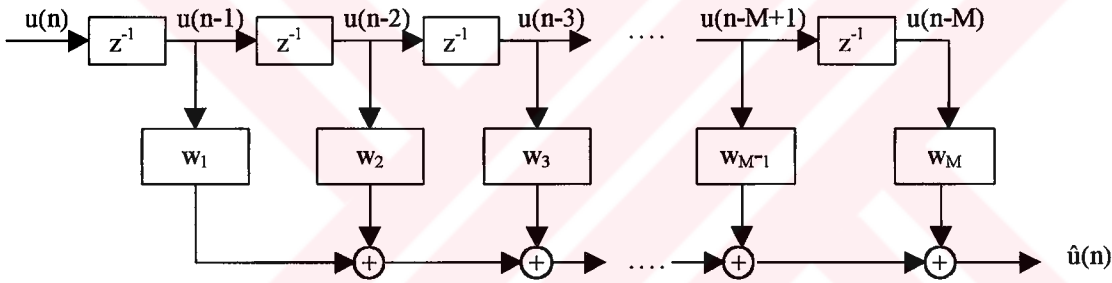
$$\begin{bmatrix} r(0) & r(1) & \dots & r(M-1) \\ r(-1) & r(0) & \dots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(-M+1) & r(-M+2) & \dots & r(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} = \begin{bmatrix} r(-1) \\ r(-2) \\ \vdots \\ r(-M) \end{bmatrix} \quad (3.24)$$

3.3.2.1.2 Doğrusal öngörü kodlayıcı

Doğrusal öngörü kodlayıcı doğrusal öngörü süzgecine oldukça benzer. Kodlayıcı süzgeci aşağıdaki gibi tanımlıdır:

$$y(n) = a_0 x(n) + \sum_{i=1}^M b_i y(n-i) \quad (3.25)$$

Burada $x(n)$ uyarım işareti, $y(n)$ istenen işaret, a_0 ve b_i değerleri de süzgeç katsayılarıdır. (3.25) eşitliği bir sayısal süzgeç tanımlamakla birlikte, $y(n)$ ' in kendinden önceki M değere bağlı olması nedeniyle bir öngörü olarak da nitelendirilebilir. Bu durumda $x(n)$ de n anında $y(n)$ ' in gerçek ve kestirilmiş değerleri arasındaki fark olarak adlandırılır. Öyleyse $x(n)$ ' i mümkün olduğunca küçültmek kestirilen işareti asıl işarete yaklaştıracaktır.



Şekil 3.6 Doğrusal öngörücü

Wiener-Kinchin teoremine göre bir işaretin özilişkisinin Fourier Dönüşümü, işaretin güç spektrumuna eşittir. Doğrusal öngörü süzgecinin güç spektrumu, transfer fonksiyonunun genliğinin karesidir:

$$|H(e^{j\omega})|^2 = \frac{a_0^2}{\left|1 + \sum_{i=1}^M b_i z^{-i}\right|^2} \quad (3.26)$$

N elemanlı bir işaretin d uzaklıklı özilişkisi, (3.27) eşitliğindeki gibi tanımlıdır. Ancak hatırlatılması gerekir ki, ses işareti yalnız gerçel değerler alabilir. Bunun için özilişki fonksiyonundaki eşlenik operatörüne gerek yoktur.

$$r(d) = r^*(-d) = \frac{1}{N-d} \sum_{k=0}^{N-d-1} y(k)y^*(k-d) \quad (3.27)$$

$y(n)$ ' in özilişkisinin Fourier Dönüşümünü almak yerine z dönüşümü de alınabilir. Sonuçta Fourier Dönüşümü, z dönüşümünün özel bir halidir. (3.26) ve (3.27) eşitliklerinin z dönüşümü alınıp, hata işaretinin istenen işaretle ilişkisiz olması koşulu altında Wiener-Kinchin teoremi uygulanırsa,

$$\frac{a_0}{\left|1 + \sum_{i=1}^M b_i z^{-i}\right|^2} \approx \sum_{d=-M}^M r(d) z^d \quad (3.28)$$

elde edilir. Eşitlik matris formunda yeniden düzenlenerek,

$$\begin{bmatrix} r(0) & r(1) & \dots & r(M) \\ r(1) & r(0) & \dots & r(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(M) & r(M-1) & \dots & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix} = \begin{bmatrix} a_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.29)$$

elde edilir. Bu matrisyel çarpım $M+1$ bilinmeyenli $M+1$ denklem tanımlar. Bu denklem takımının çözülmesiyle süzgeç katsayıları elde edilmiş olur.

(3.29) eşitliği ile elde edilen süzgeç, her zaman kararlı olmayabilir. Bu durumda yapılabilecek bir kaç basit müdahale ile süzgeç kararlı hale getirilebilir. İlk akla gelecek çözüm her N örnekte bir süzgeç katsayılarını yenilemektir. Böylece süzgeç kararsız da olsa kısa sürede kasayıları değişeceği için katsayılar yenilenene kadar çıkışı fazla büyümeyebilir.

İkinci yöntemse, süzgecin kararsızlığa yol açan kutuplarını bulup bu kutupları birim çemberin içine ya da üzerine yansıtmaktır (Haykin, 1986).

Doğrusal öngörü süzgecinin katsayılarının özilişki yöntemiyle bulunmasına maksimum entropi, tüm kutup ya da öz yinelemeli yöntem adları da verilmektedir.

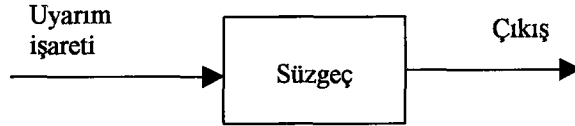
Doğrusal öngörü kodlaması, akustik işaretlerin büyük oranda fazlalık bilgi taşıdığı varsayımından yola çıkılarak geliştirilmiştir. Eğer işaretin n . örneği kendinden önceki M örnek tarafından belirlenebiliyorsa, enformasyon teorisine göre bu örnek değeri fazlalıktır, yani bu örneğin kaybedilmesiyle işaretin taşıdığı bilgi miktarı (enformasyon) değişmez. Süzgeç katsayılarıyla, mümkün olan en yüksek bilgi miktarı elde edildikten sonra, asıl işaret ve elde edilen işaret arasındaki fark olan $x(n)$ işaretinin maksimum entropiye sahip olmasına çalışılmalıdır. Spektral düzlemde söylemek gerekirse, fark işaretinin spektrumu olabildiğince düz olmalıdır.

Doğrusal öngörü kodlamasının avantajı, fark işareti $x(n)$ ' nin genellikle sentez süresinin büyük kısmında oldukça küçük olmasıdır. Yani $x(n)$ için gerekecek bellek miktarı asıl işarete oranla daha azdır. Örnek olarak insan sesi sentezinde ünlü harfler için $x(n)$ perde periyodu kadar periyotlu bir dürtü katarıdır. Ünsüz harflerde ise beyaz gürültü kullanılır. Fark işaretini ifade etmek için dürtü ya da gürültü kullanılacağını belirten bir bit ve dürtü için perde periyodu, gürültü içinse güç yoğunluğu bilgileri yeterli olacaktır.

3.4 Fiziksel Modeller

Fiziksel modeller, akustik müzikal enstrümanların ayrık zamanlı modelleridir. Sayısal dalga kılavuzu tekniklerinin keşfedilmesiyle bu modeller geniş uygulama alanları bulmuştur. Temel olarak fiziksel modeller bir uyarım işareti ve bu işaret tarafından sürülen zamanla değişebilen süzgeçten oluşur (Şekil 3. 7).

Fiziksel model ile ses sentezinde amaç, ses kaynağının fiziksel özelliklerini taklit edebilecek modeller geliştirebilmektir. Bu tür sentez tekniklerinin iki temel hipotezi vardır: İlk hipotez, iyi bir fiziksel modelin gerçek bir enstrüman gibi karmaşık sesler üretebilmesinin gerektiğidir. İkinci hipotez ise parametrik kontrolün kolay olmasıdır (Borin, De Poli ve Sarti 1992).



Şekil 3.7 Fiziksel modelin temel yapısı

3.4.1 Fiziksel, işlevsel ve biçimsel yapılar

Karmaşık bir sistem geliştirmenin kolay yolu, sistemi daha basit alt sistemlere bölmek ve bu alt sistemler arasındaki ilişkiyi kurmaktır. Alt sistemlere bölme izlenebilecek bir çok yol vardır ancak göz önünde tutulması gereken üç kriter mevcuttur: fiziksel benzerlik, işlevsel yapı ve biçimsel yapı.

Fiziksel benzerlik, alt sistemlerin enstrümandaki fiziksel parçalara denk düşmesini gerektirir. Örneğin bir kemanın alt bölümleri teller, yay, köprü ve gövdedir. İşlevsel yapıya göre bölme yapıldığında, müzikal enstrümanlar iki bölüme ayrılır: uyarım ve rezonatör. Uyarım, enstrümanın titreşmesini sağlayan kısım, rezonatör ise bu titreşime müzikal değeri katan kısımdır. İşlevsel ayrıma örnek vermek gerekirse, keman yayı uyarıma, gövde ve tellerin birleşimi de rezonatöre denk düşer. Biçimsel bölmeleme ise alt sistemleri basitliklerine göre ayırmaktır. Örneğin model doğrusal ve doğrusal olmayan alt sistemler halinde ayrılabilir (Borin, De Poli ve Sarti 1992; Smith, 1996b).

3.4.2 Uyarım

Uyarım, modelin doğrusal olmayan parçasıdır. Uyarım işaretinin fiziksel uyarıcıya benzemesinden çok, modele uygun olması gereklidir. Uyarım işaretinin elde edilmesinde üç temel yol vardır: doğrudan üretim, belleksiz doğrusal olmayan model ve mekanik model.

3.4.2.1 Doğrudan üretim

Sürekli uyarılan enstrümanlarda uyarım yalnızca rezonatörün başlangıç koşullarını belirleyen bir aygıt değil, rezonatör gibi sürekli etkin bir alt sistemdir. Pek çok durumda,

uyarım işaretini modellemeden istenen uyarım işaretini üretecek bir sistem tasarlanabilir. Bunun en kolay yolu, tablo bakışımı üreteçler kullanmaktır. Uyarım işaretinden alınan örnekler bir tabloda saklanır ve gereken örnek bu tablodan okunur. Bu yöntem daha çok geri beslemesi zayıf ve uyarım parametreleri fazla değişmeyen enstrümanlar için geçerlidir (Borin, De Poli ve Sarti, 1992; Smith, 1996b).

3.4.2.2 Belleksiz doğrusal olmayan model

Kimi zaman uyarım işaretini $y(t) = f[x(t), x_e(t)]$ gibi bir denklemle modellemek mümkündür. Burada $y(t)$ uyarım işareti, $x(t)$ rezonatörün yanıtı ve $x_e(t)$ icracının enstrümanla ilişkisine karşılık düşen dışsal kontrol parametreleridir. Fonksiyon belleksiz olduğu için model uyarım işareti üzerinde dinamik işlemlere izin vermez.

Genelde, $f(\cdot)$ fonksiyonu birçok fiziksel olaya bağlıdır. Örneğin klanet çalan kişinin dudaklarıyla kamışa uyguladığı baskı bile $f(\cdot)$ eğrisinde önemli değişikliklere sebep olabilmektedir. Bununla beraber, asıl problem kontrol parametrelerinin fazlalığı değil, işlem yoğunluğudur (Borin, De Poli ve Sarti, 1992).

3.4.2.3 Mekanik model

Uyarım işareti yay, kütle ve sürtünme elemanlarından oluşan mekanik bir ağ gibi de düşünülebilir. Yaylar genellikle doğrusal olmayan yapıdadırlar ve uyarım mekanizmasının elastik kısımlarına karşılık düşerler. Kütleler harekesiz kısımları, sürtünme elemanları da kayıpları temsil eder. Bu modelde, uyarım işareti tüm bu elemanların dinamiğinin oluşturduğu fark denklemlerinin çözümüyle elde edilir. Dikkat edilmesi gerekir ki, bu sistem belleksiz değildir.

Uyarım için en basit model, doğrusal olmayan bir yayın ucuna bağlanmış bir kütledir. Bu modele ayrıca bazı dışsal etkileri de eklemek mümkündür. Uyarımın mekanik modeli genel bir yaklaşımdır ve yay, kamış, dudak ve gırtlak gibi uyarım mekanizmalarına uygulanabilir (Borin, De Poli ve Sarti, 1992).

3.4.3 Rezonatör

Genellemenin dışına çıkmadan söylenebilir ki, rezonatörler nedensel, doğrusal ve zamanla değişen yapılardır. Doğrusal yapı, sentez işleminde kullanışlı kabuller yapmayı kolaylaştırır. Rezonatör modellenirken birden fazla yol izlenebilir. Bu modeller aşağıda incelenmiştir.

3.4.3.1 Transfer fonksiyonu modeli

Rezonatörü taklide yönelik yaklaşımlardan biri, rezonatörün fonksiyonel giriş çıkış ilişkisinden yararlanarak transfer fonksiyonunu kestirmektir. Bu tür modellerde rezonatörün fiziksel yapısı tamamen göz ardı edilir. Rezonatörün transfer fonksiyonu, analitik yollarla dürtü yanıtı ya da akustik empedansı ölçülerek bulunabilir.

Etkili alternatif yöntemlerden biri de rezonatör içinde yayılan, yansıyan ve iletilen dalgaları gözlemlemektir. Bu tür yöntemlerin en büyük eksikliği rezonatör üzerinde ölçümün yapıldığı yere göre farklı transfer fonksiyonları bulunabilmesidir (Borin, De Poli ve Sarti, 1992).

3.4.3.2 Mekanik model

Mekanik modellerin temel hipotezine göre, rezonatörün mekanik yapısı bir dizi farksal denklem tanımlar. Rezonatör uygulaması, bu denklem takımının çözülmesine dayalıdır. Bu denklemler uzaysal koordinatlar ve zamanda çözülür. Bazı uygulamalarda denklem takımı analitik olarak çözülebilse de genellikle sayısal çözümler bulunur. Bu model nefesli enstrümanların taklidinde pek başarılı değildir. Modelleme için kullanılması gereken denklem sayısı fazla ve yüksek derecedendir. Bu yüzden, modelin işlemsel yükü çok yüksektir (Borin, De Poli ve Sarti, 1992).

3.4.3.3 Dalga kılavuzu modeli

Dalga kılavuzu modeli, rezonatör için iyi bir modeldir. Dalga kılavuzu modeli, ortamda yayılan dalganın denkleminin analitik çözümüne dayanır. Sayısal dalga kılavuzları, bir

dalga kılavuzunda yayılan dalganın örneklenmesiyle elde edilir. Kayıplar ve saçılmalar belli noktalarda toplanarak, dalga kılavuzu kayıplı ve kayıpsız olarak iki kısma ayrılır. Kayıpsız kısım yalnızca yayılım gecikmesini içerir ve gecikme hatlarıyla temsil edilir. Kayıplı/saçılmalı kısım ise bir kaç düşük dereceli sayısal süzgeç ile oluşturulabilir. Dağınık dalga kılavuzu modelleri toplu modellerle birlikte kullanılabilir. Örneğin nefesli enstrümanların kamış kısmı toplu, boru kısmı da dağınık parametrelili modellerle modellenir (Borin, De Poli ve Sarti, 1992; Smith, 1996a; Valimaki, Huopaniemi, Karjalainen ve Janosy, 1996).

3.4.4 Dalga kılavuzu modeli uygulamaları

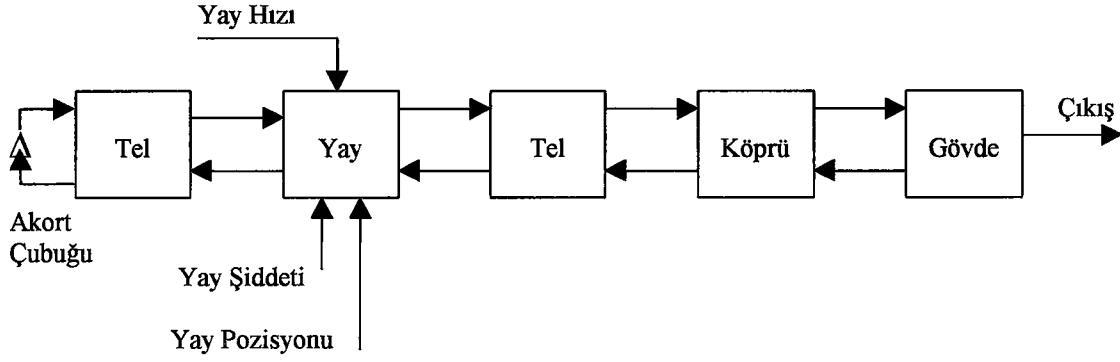
3.4.4.1 Tek kamışlı nefesliler

Şekil 3.8'de klarnetin temel dalga kılavuzu modeli görülmektedir. Gecikme hatlarında taşınan basıç örnekleri (p_b^+ ve p_b^-), boru içinde sağa ve sola doğru yayılan basınç dalgalarına karşılık düşmektedir. Eğer boru silindirikse, iki yönlü bir gecikme hattı yeterlidir. Ancak boru saksafonda olduğu gibi konikse, iki yönlü gecikme hattına erişim karmaşılaşacaktır.

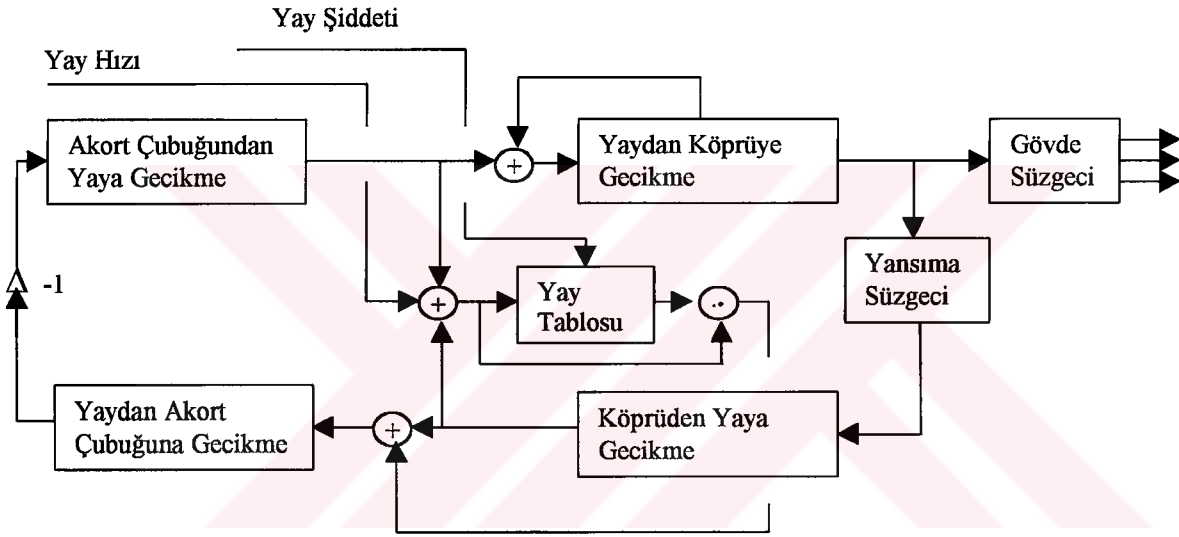
Yansıma süzgeci, çanak ve ton deliklerindeki kayıplar ile, dalganın boru içinde ileri geri hareketi süresince uğradığı zayıflamaya eşdeğerdir. Çıkış süzgeci yüksek geçiren yapıdadır. Birinci dereceden bir çıkış süzgeci, hoparlörlerdeki bas-tiz ayrımını yapan basit bir çapraz ağ gibi düşünülebilir. Örnek olarak klarnet için bu süzgecin kesim frekansı 1500 Hz dir.

Ton delikleri, çanak gibi bir yansıma geçirme süzgeci ile ya da saçılma matrisleriyle modellenir. En basit uygulamalarda, ton deliklerini modellemek yerine boru uzunluğunu, yani gecikme hattının uzunluğunu değiştirmek yeterli olacaktır.

Klarnetin sayısal dalga kılavuzu modelinde kamış, boruyu sonlandıran ve ağız basıncı ile ağızlık dengesi tarafından kontrol edilen nonlineer bir elemandır. Enstrümanı çalan kişinin ağızlığı kullanma şekli, kamışın sönümlenmesi, kamış aralığı gibi parametreleri etkilediği



Şekil 3.9 Temel yaylı telli modeli



Şekil 3.10 Yaylı telliler için ayrıntılı model

3.5 Fiziksel ve Spektral Modellerin Karşılaştırılması

Fiziksel modeller yapıları gereği karmaşık işlemler içermezler. Buna karşılık spektral modeller genellikle yoğun işlem hacimli algoritmalar içerirler. Yüksek işlem hacminden dolayı spektral modeller daha fazla bellek gereksinimi duyarlar. Bu da uygulamada fiziksel modellerin daha düşük maliyetli olması sonucunu doğurur.

Fiziksel modeller doğası gereği doğrusal olmayan kısımlar içerir. Spektral modellerde ise böyle bir şey söz konusu değildir. Bunun bir sonucu olarak, fiziksel modeller sesin atak kısmını spektral modellere göre daha iyi taklit ederler. Yine aynı sebepten dolayı, spektral

modellerle çınlama, gecikme gibi müzikal etkiler zor elde edilirken, fiziksel modellerle bu etkiler rahatlıkla elde edilmektedir. Buna karşılık spektral modellerle çok kolay yapılabilen geçiş etkisinin fiziksel modellerle yapılması imkansızdır.

Fiziksel ve spektral modellerin yapısal farklarının en önemlisi, fiziksel modellerin ses kaynağını, spektral modellerin ise ses alıcısını temsil etmesidir. Buradan yola çıkarak söylenebilir ki, spektral modeller geneldir; aynı model kullanılarak farklı özellikler gösteren sesler modellenebilir. Fiziksel modeller ise özeldir; bir modelle yalnız bir enstrüman ailesi modellenebilir.

Fiziksel ve spektral modeller arasındaki bir başka yapısal fark da spektral modellerin zaman–frekans düzleminde, fiziksel modellerin ise uzay–zaman düzleminde çalışmasıdır.

Bu bilgilerin ışığında söylenebilir ki, fiziksel modeller elektronik enstrümanlar gibi uygulamalarda daha kullanışlıdır. Spektral modeller ise hem elektronik enstrümanlarda, hem de ses sıkıştırma gibi uygulamalarda kullanılabilir.

4. SİNÜZOİDAL SENTEZ

4.1 Sinüzoidal Ses Modeli

Ses üreten her hangi bir mekanizma, zamanla değişen bir süzgeç ve bu süzgeci süren bir uyarım işareti ile modellenenbilir (McAulay ve Quatieri, 1986). Zamanla değişen süzgecin dürtü yanıtı $h(\tau; t)$, uyarım işareti de $e(t)$ ile gösterilirse, ses işareti şu şekilde elde edilir:

$$s(t) = \int_0^t h(t - \tau; t) e(\tau) d\tau \quad (4.1)$$

Uyarım işareti, bir dizi farklı genlik, frekans ve fazda sinüsün toplamıyla temsil edilebilir:

$$e(t) = \text{Re} \sum_{l=1}^{L(t)} a_l(t) e^{j \left[\int_0^t \omega_l(\sigma) d\sigma + \phi_l \right]} \quad (4.2)$$

Burada a_l l . sinüsün genliği, ω_l frekansı ve ϕ_l de başlangıç fazıdır. $L(t)$, t anında $e(t)$ işaretindeki sinüs sayısını gösterir. Zamanla değişen süzgecin transfer fonksiyonu ise

$$H(\omega; t) = M(\omega; t) e^{j\Phi(\omega; t)} \quad (4.3)$$

dir. Denklem (4.2) ve (4.3) denklem (4.1)'de yerine konursa, $s(t)$

$$s(t) = \sum_{l=1}^{L(t)} a_l(t) M(\omega_l(t); t) e^{j \left[\int_0^t \omega_l(\sigma) d\sigma + \Phi[\omega_l(t); t] + \phi_l \right]} \quad (4.4)$$

olarak elde edilir. Uyarım ve süzgecin genlik ve fazları birleştirilerek

$$s(t) = \sum_{l=1}^{L(t)} A_l(t) e^{j\psi_l(t)} \quad (4.5)$$

$$A_i(t) = a_i(t)M[\omega_i(t);t] \quad (4.6)$$

$$\psi_i(t) = \int_0^t \omega_i(\sigma) d\sigma + \Phi[\omega_i(t),t] + \phi_i \quad (4.7)$$

ifadeleri elde edilir.

4.2 Model Parametrelerinin Kestirimi

Ses dalgasının analizi ve sentezinde temel problem, dalga şeklinin geniş anlamda durağan bir parçasını temsil edecek parametreler kestirmek ve bu parametreler yardımıyla, asıl işarete mümkün olduğu kadar yakın olan yeni bir dalga oluşturmaktır. Ses işaretinin sinüzoidal bileşenlerine analitik olarak ayrılması zordur. Bu yüzden bazı kabuller ve idealize edilmiş şartların varlığı gereklidir. Eğer ideal şartlara göre kestirimler yapılabilirse, bu kestirimlerde yapılacak bazı değişikliklerle gerçek duruma yaklaşılabılır (McAulay ve Quatieri, 1986).

İlk adım olarak, ses işaretinin durağan kabul edilebileceği en uzun zaman aralığı T olsun ve işaret T süreli çerçevelere bölünsün. k. çerçevenin analiz penceresinin ortasını gösteren zaman indisi t_k ile gösterilsin. Analiz penceresi boyunca süzgecin ve uyarımın parametreleri sabit olsun. Bu durumda (4.7) eşitliği aşağıdaki gibi basileştirilebilir:

$$\psi_i(t) = \omega_i^k(t - t_k) + \theta_i^k \quad (4.8)$$

Burada k üst indisi k. çerçeveyi göstermektedir. Ses sentezi sayısal olarak yapılacağına göre, bundan sonraki işlemleri ayrık işaretler üzerinde yapmak daha doğru olacaktır. Bu bağlamda zaman indisi n olacak ve N+1 uzunluklu bir çerçeve boyunca $-N/2$ ile $N/2$ arasında değişecektir. Çerçevenin ortası $n=0$ 'a denk gelmektedir. Bu durumda, k. çerçevedeki yapay dalga şekli şu şekildedir:

$$s(n) = \sum_{l=1}^{L^k} \gamma_l^k e^{jn\omega_l^k} \quad (4.9)$$

$$\gamma_l^k = A_l^k e^{j\theta_l^k} \quad (4.10)$$

Burada γ_l^k k . çerçevedeki l . sinüzoidalın karmaşık genliğidir. Yapılmak istenen, yapay ses işareti $s(n)$ 'i gerçek işaret $y(n)$ 'e mümkün olduğunca benzetmektir. Bu benzetmede optimizasyon yöntemi olarak karesel ortalama hata kullanılabilir. Yukarıdaki cümleyi matematiksel olarak tekrarlamak gerekirse, yapılmak istenen aşağıda tanımlanan e^k işaretini sıfıra yaklaştırmaktır.

$$\begin{aligned} e^k &= \sum_n |y(n) - s(n)|^2 \\ &= \sum_n |y(n)|^2 - 2\operatorname{Re} \sum_n y(n)s^*(n) + \sum_n |s(n)|^2 \end{aligned} \quad (4.11)$$

Denklem (4.9), (4.11)'de yerine konursa hata işareti

$$e^k = \sum_n |y(n)|^2 - 2\operatorname{Re} \sum_{l=1}^{L^k} (\gamma_l^k)^* \sum_n y(n)e^{-jn\omega_l^k} + (N+1) \sum_{l=1}^{L^k} \sum_{i=1}^{L^k} \gamma_l^k (\gamma_i^k)^* \sin c(\omega_l^k - \omega_i^k) \quad (4.12)$$

olarak bulunur. Bu aşamada problem (4.12) eşitliğini minimum yapacak sinüzoidal dalgaları bulmaktır ancak, bu çözümü zor bir problemdir. Eşitliği minimum yapan sinüslerin arasında harmoniklik ilişkisinin bulunduğu varsayımı, problemi basitleştirecektir. Böylece yapay işaret,

$$s(n) = \sum_{l=1}^{L^k} \gamma_l^k e^{jnl\omega_0^k} \quad (4.13)$$

haline gelir. Burada $\omega_0^k = 2\pi/\tau_0^k$ dır ve τ_0^k da k . çerçevedeki perde periyodudur. Bu ideal koşullar altında (4.12) eşitliğinin son terimi şu şekilde basitleştirilebilir:

$$\operatorname{sinc}(\omega_l^k - \omega_i^k) = \operatorname{sinc}[(l-i)\omega_0^k] = \begin{cases} 1, l=i \\ 0, l \neq i \end{cases} \quad (4.14)$$

Bu durumda hata işareti e^k da

$$e^k = \sum_n |y(n)|^2 - 2(N+1) \operatorname{Re} \left[\sum_{l=1}^{L^k} (\gamma_l^k)^* Y(\omega_l^k) \right] + N \sum_{l=1}^{L^k} |\gamma_l^k|^2 \quad (4.15)$$

olur. Burada,

$$Y(\omega_l^k) = \frac{1}{N+1} \sum_n y(n) e^{-jn\omega} \quad (4.16)$$

ölçülen işaretin kısa süreli Fourier Dönüşümüdür. (4.16) eşitliği tekrar düzenlenirse,

$$e^k = \sum_n |y(n)|^2 + (N+1) \sum_{l=1}^{L^k} \left[|Y(\omega_l^k) - \gamma_l^k|^2 - |Y(\omega_l^k)|^2 \right] \quad (4.17)$$

elde edilir. Bu eşitliği minimum yapabilmek için,

$$\hat{\gamma}_l^k = Y(l\omega_o^k) \quad (4.18)$$

olmalıdır. Bu durumda minimum hata işareti aşağıdaki gibi olur:

$$e^k = \sum_n |y(n)|^2 - (N+1) \sum_{l=1}^{L^k} |Y(l\omega_o^k)|^2 \quad (4.19)$$

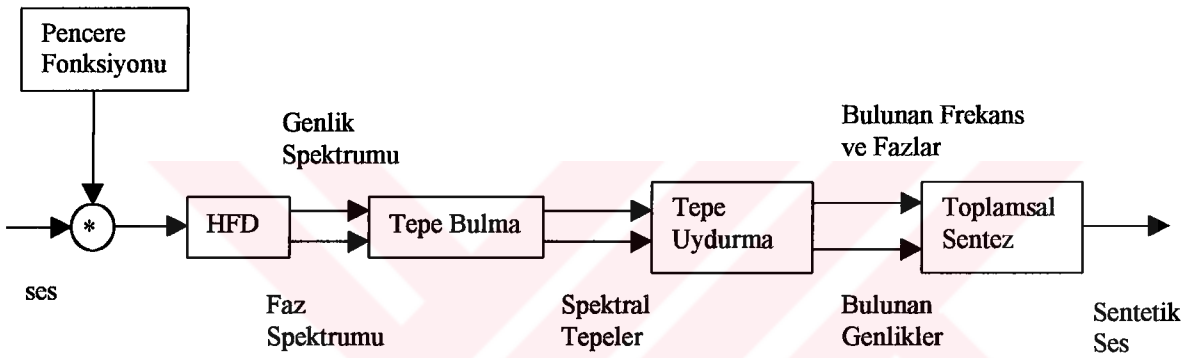
(4.16), (4.18) ve (4.19) eşitlikleri ideal kestiricinin yapısını ortaya koymaktadır. Bu sonuçlar göstermektedir ki, ideal kestirim için KSFD kullanılmalıdır.

4.3 Analiz ve Sentez İşlemlerinin Genel Yapısı

Analiz ve sentez işlemleri çerçeve temellidir. Yani sesin tamamı bir kerede değil, parça parça işlenir. Şekil 4.1'de genel analiz ve sentez yapısı gösterilmiştir. Öncelikle, ses işareti analiz penceresi ile çarpılarak çerçevelere bölünür. Her çerçevenin spektrumu, Hızlı

Fourier Dönüşümü (HFD) ile hesaplanır. Bulunan spektrumdan ses işaretinin sentezlenmesinde kullanılacak olan spektral tepeler çekilir. Bu spektral tepeler, çerçeveler arası geçişlerde çıttırtı oluşturmaması için bir eşleştirme algoritmasından geçirilir. Eğer analiz edilen sesin harmonik yapısı kuvvetli değilse, tepe bulma işleminden hemen sonra bir perde belirleme algoritmasıyla sesin perde periyodu bulunabilir ve bu bilgi ile de pencere boyu yeniden ayarlanabilir (perde eşzamanlı analiz, uyarlamalı pencere analiz).

Analizle bulunan genlik, frekans ve faz değerleri üzerinde müzikal değişimler yapılabilir. Bu işlemler, sentez aşamasından hemen önce, tepe bulma işlemiyle bulunan değerler üzerinde yapılır.



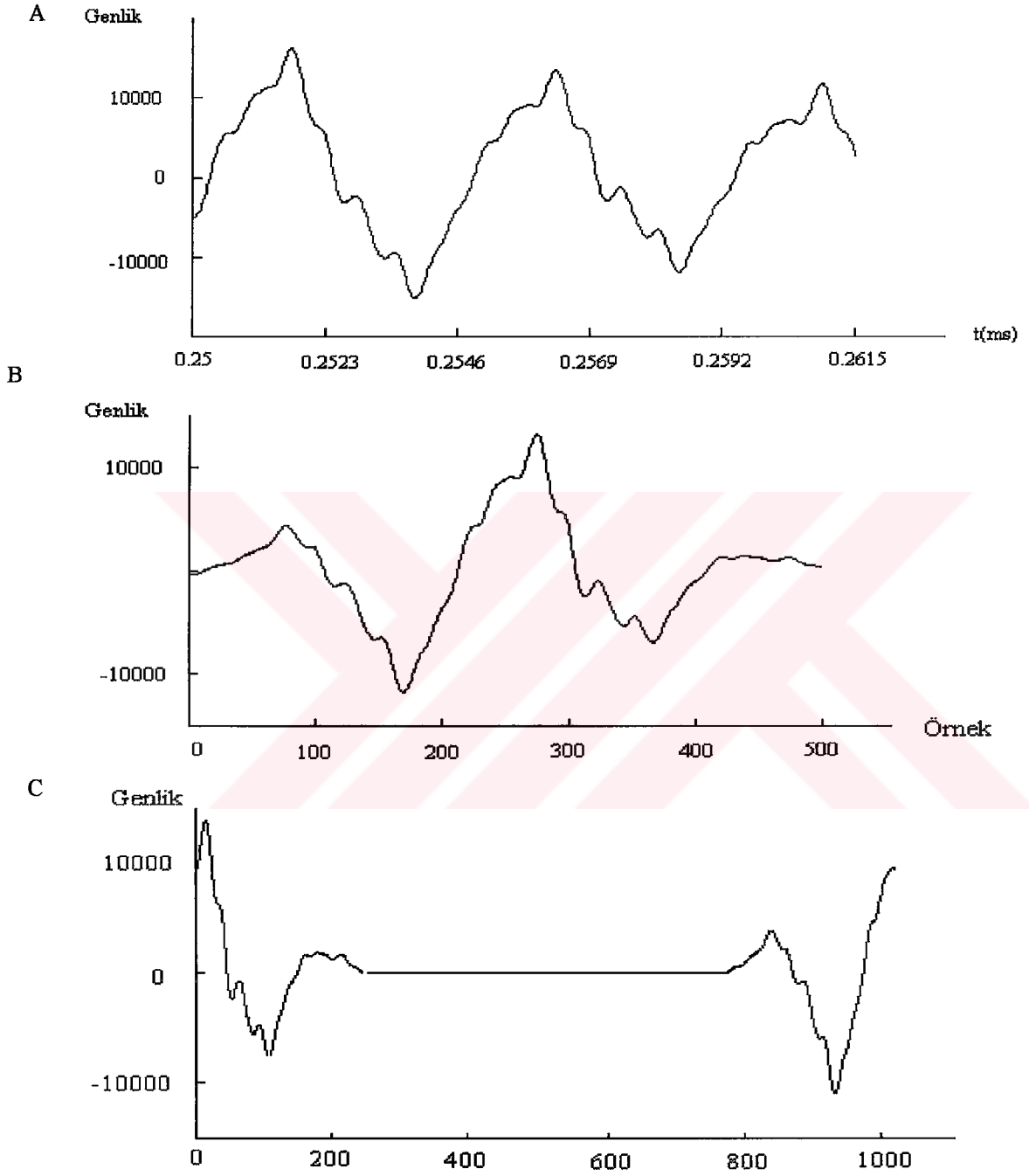
Şekil 4.1 Sinüzoidal analiz ve sentez işlemlerinin blok yapısı

Sentez aşaması, Bölüm 3'te anlatılan toplamsal sentezden ibarettir. Sinüzoidal sentez, analiz yoluyla bulunan genlik, frekans ve faz değerleriyle kontrol edilen bir dizi paralel osilatörden ibarettir (McAulay ve Quatieri, 1986; Macon, 1993; Serra, 1997).

4.3.1 Genlik ve faz spektrumunun hesaplanması

Bir pencerenin genlik ve faz spektrumunun hesaplanması analiz işleminin ilk adımıdır. KSFD'nün kontrol parametreleri (pencere boyu, pencere tipi, HFD boyu) işlenecek sese göre belirlenmelidir. Öncelikle, genlik spektrumundan sinüzoidal bileşenleri doğru bulabilmek için, pencerenin yeterince uzun olması gerekir. Faz bilgisi de önemli olduğu için, pencere uzunluğunun tek sayıda olması gereklidir. Ayrıca, spektral çözünürlüğü arttırmak için HFD alınırken işarete sıfır eklenebilir. Pencerelemeden kaynaklanan doğrusal faz ötelemesini karşılamak için, önce işaretin sonuna HFD tamponunu dolduracak

kadar sıfır eklenir. Daha sonra, asıl işaretin ilk yarısı koparılarak sıfırların sonuna eklenir (Şekil 4.2). Bu işlemin adı “sıfır fazlı pencereleme” dir (Macon, 1993; Serra, 1997).



Şekil 4.2 Sıfır fazlı pencereleme: (a) Asıl işaret; (b) Pencerelemiş İşaret; (c) HFD tamponundaki işaret

Analiz penceresinin türü belirlenirken, spektral bozulmayı en aza indirgeyebilmek için yan lobları yeterince düşük kazançlı pencereler kullanılmalıdır. Zaman domeninde çarpma frekans domeninde konvolüsyona denk düştüğüne göre, bozulmayı en aza indirmek için ana lobun da olabildiği kadar dar olması gerektiği açıktır. Analiz penceresi olarak Hamming, Hanning ve Blackman-Harris gibi pencereler kullanılabilir (Goodwin, 1997).

Harmonikliği kuvvetli seslerde, analiz boyunca pencere boyu perde periyodunun en az iki katı olacak şekilde seçilip sabit tutulabilir. Eğer sesin perdesi zamanla değişiyorsa perde periyodu her çerçeve için yeniden hesaplanmalıdır. Nonharmonik seslerde ise komşu bileşenlerin arasındaki en küçük frekans farkı perde periyodu olarak kullanılır. İnsan sesi uygulamalarında, perde periyodu genellikle 10 ms olarak alınır (McAulay ve Quatieri, 1986; Serra, 1997).

4.3.2 Tepe bulma

Bir çerçevenin spektrumu hesaplandıktan sonraki adım, genlik spektrumundaki tepelerin bulunmasıdır. Bir spektral tepe, genlik spektrumundaki yerel maksimumdur. Teorik olarak, genlik ve frekansı sabit olan sonsuz uzunluklu bir sinüzoidalın Fourier Dönüşümü $[0, \pi]$ aralığında bir dürtüden ibarettir. Bununla beraber, doğal sesler mükemmel periyodik değildir, bileşenlerinin frekansları zamanla yavaş ta olsa değişir ve frekans domeninde örtüşürler. Farklı bileşenler arasında etkileşimler oluşur ve spektral tepeler bir miktar hata yapmadan bulunamazlar. Ayrıca HFD'nün ayrık zamanlı Fourier Dönüşümünün örneklenmiş hali olduğu hatırlanırsa, bileşenin frekansı HFD'nün ilgili gözüne tam olarak denk gelmediği durumlarda da bir miktar hata olacağı görülür (Macon, 1993; Serra, 1997).

4.3.3 Tepe uydurma

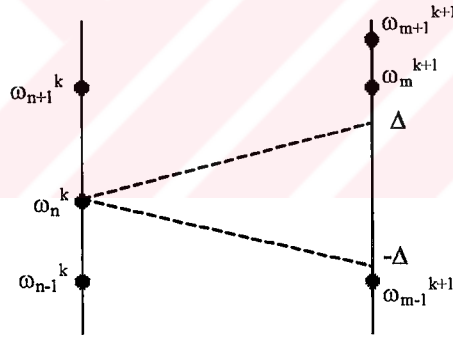
İşlenen ses tam periyodik değilse, bileşenlerin frekansları ve sayıları zamanla değişecektir. Bu durumda, bir çerçeveden diğerine geçerken tüm bileşenlerin bir sonraki çerçevedeki karşılıklarının bulunması gerekir. Böylelikle çerçeveler arası geçişte oluşabilecek kesiklikler ve çıtırtıların önüne geçilebilir. Tepe uydurma algoritmaları farklı şekillerde kurulabilir. Etkili bir algoritma, aşağıda bir örnek üzerinde anlatılacaktır.

k. çerçevede $\omega_0^k, \omega_1^k, \omega_2^k \dots \omega_{N-1}^k$ gibi N adet bileşen bulunmuş olsun. k+1. çerçevede de $\omega_0^{k+1}, \omega_1^{k+1}, \omega_2^{k+1} \dots \omega_{M-1}^{k+1}$ gibi M adet bileşen bulunsun. N ve M genellikle birbirlerinden farklıdır. Uydurma algoritması, k. çerçevedeki her ω_n^k frekanlı bileşene karşılık k+1. çerçevede ω_m^{k+1} gibi bir eş bulmaya çalışacaktır. Algoritma aşağıdaki üç adımdan ibarettir (McAulay ve Quatieri, 1986; Serra, 1997):

Adım 1: k. çerçevedeki $\omega_0^k, \omega_1^k, \dots \omega_{n-1}^k$ frekanslı bileşenler uydurulmuş olsun. Bu noktada yapılması gereken k+1. çerçeveden ω_n^k frekanslı bileşene uygun bir bileşen bulmaktır. Burada karar kuralı, karar eşiği Δ olmak üzere şöyledir:

$$|\omega_n^k - \omega_m^{k+1}| \leq \Delta ; \text{ tüm } m\text{'ler için} \quad (4.20)$$

Eğer bu eşitsizliği sağlayan her hangi bir ω_m^{k+1} bileşeni bulunmuyorsa, ω_n^k “ölü” olarak nitelendirilir ve k+1. çerçevede aynı frekanslı ve sıfır genlikli bir bileşene uydurulur (Şekil 4.3). Adım 1 ω_{n+1}^k frekanslı bileşenle tekrarlanır.



Şekil 4.3 ω_n^k uydurulamamış ve ölüdür.

Eğer bu eşitsizliği sağlayan bir ya da daha fazla bileşen bulunabiliyorsa,

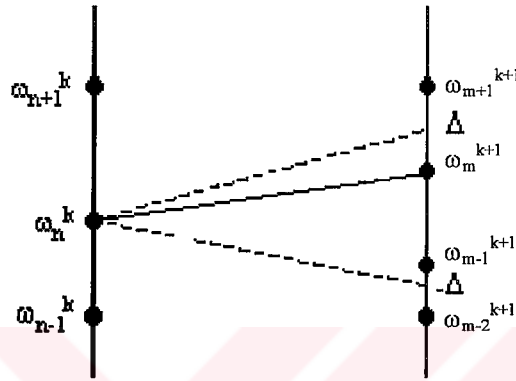
$$|\omega_n^k - \omega_m^{k+1}| < |\omega_n^k - \omega_i^{k+1}| ; \text{ tüm } i \neq m \text{ için} \quad (4.21)$$

eşitsizliğini sağlayan ω_m^{k+1} , ω_n^k için muhtemel eştir (Şekil 4.4).

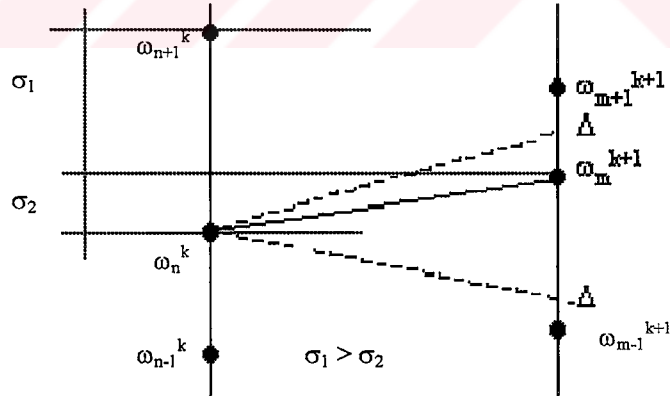
Adım 2. Bu adımda, ω_n^k frekanslı bileşenin ω_m^{k+1} gibi bir muhtemel eşi vardır. Eğer

$$|\omega_m^{k+1} - \omega_n^k| < |\omega_m^{k+1} - \omega_i^k| \quad ; i > n \text{ için} \quad (4.22)$$

sağlanıyorsa, ω_n^k ve ω_m^{k+1} kesin eştirler (Şekil 4.5). Bu iki bileşen elimine edilerek adım 1 ω_{n+1}^k ile devam eder.



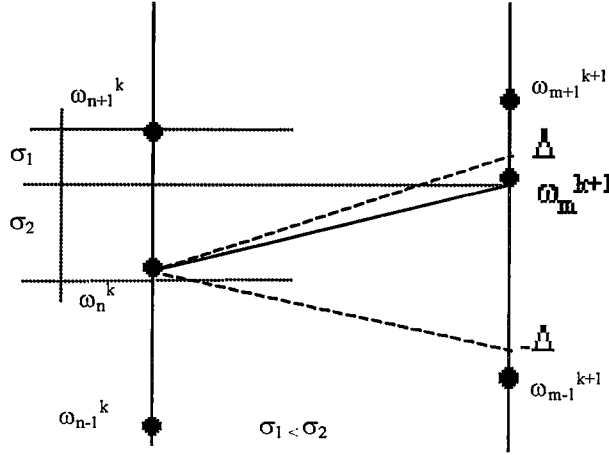
Şekil 4.4 ω_m^{k+1} , ω_n^k için muhtemel eştir.



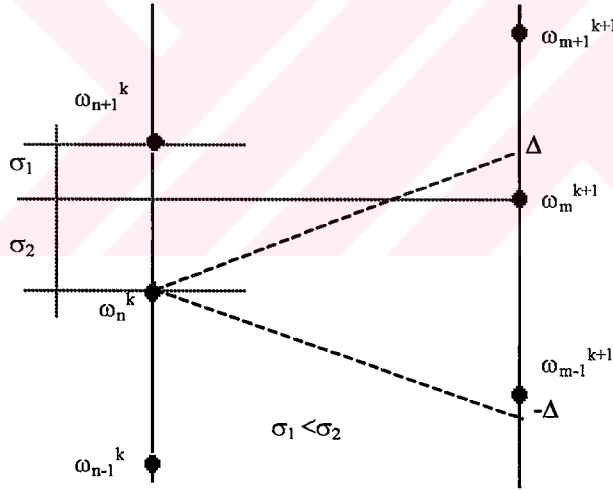
Şekil 4.5 ω_n^k ve ω_m^{k+1} eştir.

Eğer (4.22) eşitsizliği sağlanamıyorsa, ω_m^{k+1} ω_{n+1}^k 'ya daha yakındır. Bu durumda iki ek koşul vardır: Ya ω_{m-1}^{k+1} frekansı $\omega_n^k - \Delta$ 'dan daha küçüktür ve böylece ω_n^k ölüdür (Şekil

4.6); ya da, ω_{m-1}^{k+1} frekansı $\omega_n^k - \Delta$ 'dan büyüktür ve bu durumda ω_n^k ile ω_{m-1}^{k+1} kesin eştirler (Şekil 4.7).



Şekil 4.6 ω_m^{k+1} , ω_{n+1}^k ya daha yakındır ve ω_n^k için başka muhtemel eş yoktur; ω_n^k ölüdür.



Şekil 4.7 ω_m^{k+1} ω_{n+1}^k ya daha yakındır. ω_{m-1}^{k+1} ve ω_n^k eştir.

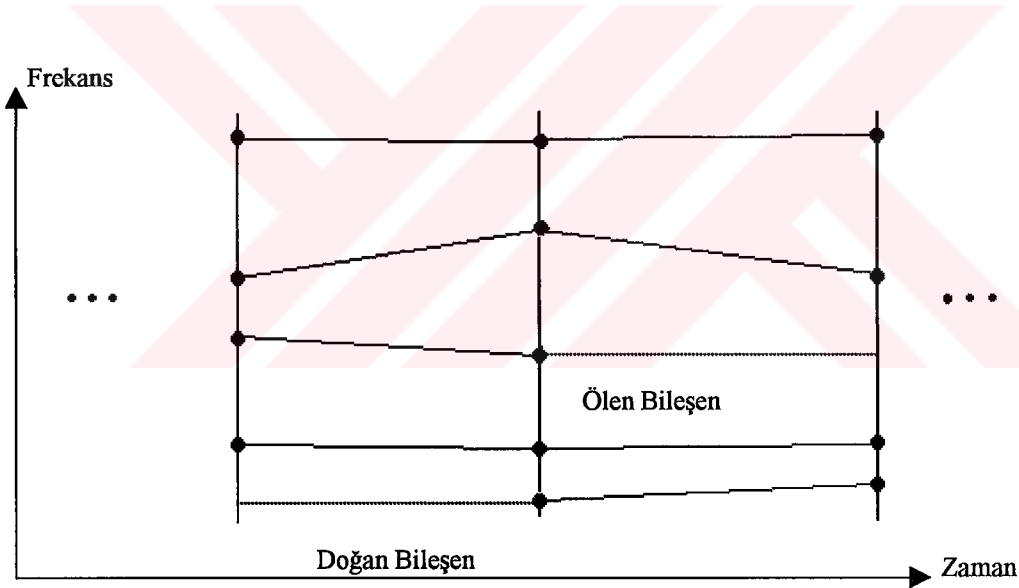
Adım 3. k. çerçevedeki tüm bileşenler denendikten sonra k+1. çerçevede k. çerçevedeki bir bileşene eşlenmemiş olan tüm bileşenlere “doğmuş” bileşenler denir ve bunlar, k. çerçeveye eklenecek aynı frekanslı fakat sıfır genlikli bileşenlere eşlenirler. Şekil 4.8’de bir uydurma çizelgesi örneği verilmiştir.

4.3.4 Sinüzoidal sentez.

Her çerçeve için genlik, frekans ve faz değerleri hesaplandıktan sonra, bu değerler yardımıyla bir çerçeve için yapay ses aşağıdaki gibi elde edilebilir:

$$\hat{s}(n) = \sum_{l=1}^{L(k)} \hat{A}_l^k \cos[n\omega_l^k + \hat{\theta}_l^k] \quad (4.23)$$

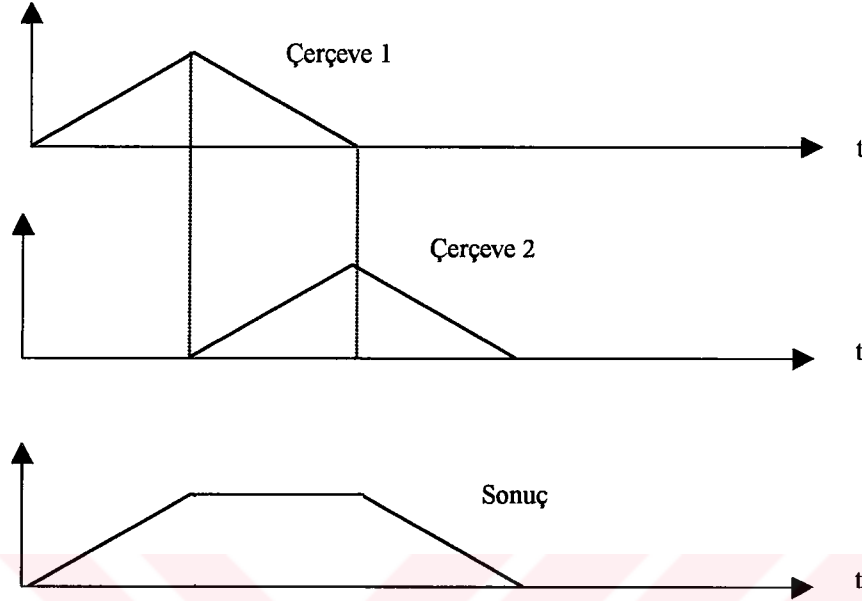
S, sentez çerçevesinin uzunluğu olmak üzere $n=0, 1, 2, \dots, S-1$ 'dir. Parametrelerin zamanla değişmesinden dolayı, bu yaklaşım çerçeveden çerçeveye geçişte süreksizliklere neden olur. Bu süreksizlikler kulak tarafından algılanır, yani yapay seste kalite kaybı oluşur. Ancak müzikal uygulamalarda kalite çok önemli olduğu için bu süreksizlikler önlenmelidir.



Şekil 4.8 Bir tepe uydurma çizelgesi.

Sık kullanılan yöntemlerden biri, çakışma-toplama yöntemidir (George ve Smith, 1997). Bu yöntemde, öncelikle bulunan genlik, frekans ve faz değerleri kullanılarak sentez çerçevesinin iki katı uzunluğunda üretilen sinüsler, üçgensel bir pencereyle çarpılır. Daha sonra, arka arkaya gelen ağırlıklandırılmış iki çerçeve, birinci çerçevenin ortası ikinci çerçevenin başlangıcına denk gelecek şekilde üst üste konur ve çakışan örnekler toplanır (Şekil 4.9).

Çakışma–toplama yöntemi, kısa çerçeve uzunluklu işaretlerde uygun bir çözümdür ancak çerçeve uzunluğu arttığında verimi azalır.



Şekil 4.9 Çakışma–toplama yöntemi

Başka bir çözüm de, tepe uydurma algoritmasıyla çerçeveler arasında kurulan ilişkiden yararlanmaktır (McAulay ve Quatieri, 1986; Serra, 1997; Goodwin, 1997). $(\hat{A}^k, \omega^k, \theta^k)$ ve $(\hat{A}^{k+1}, \omega^{k+1}, \theta^{k+1})$, k. ve k+1. çerçevelerdeki birbirine eşlenmiş bileşenleri gösterebilir. Zaman indisi n, $[0, S-1]$ aralığında olmak üzere, k. çerçeve için genlik interpolasyonu şu şekildedir:

$$\tilde{A}(n) = \hat{A}^k + \frac{(\hat{A}^{k+1} - \hat{A}^{(k)})}{S}n \quad (4.24)$$

Frekans ve faz değerleri, birbirlerine bağlıdır. Bulunan fazlar 2π 'ye göre modüle olduğu için, fazda doğrusal interpolasyon yapılamaz. Faz interpolasyonu, fazda sürekliliği sağlayacak en yumuşak eğriyle gerçekleştirilmelidir. Bu sorunla uğraşmadan önce, bir tanım yapmak gereklidir: Frekans, anlık fazın zamana göre türevidir. Buna göre, anlık faz şöyle tanımlanır:

$$\theta(t) = \int_0^t \omega(t) dt + \theta_0 \quad (4.25)$$

Anlık faz interpolasyon eğrisi aşağıdaki gibi kübik bir polinom olsun:

$$\tilde{\theta}(t) = \eta + \gamma t + \alpha t^2 + \beta t^3 \quad (4.26)$$

İşlem kolaylığından dolayı, şimdilik faz fonksiyonu sürekli zaman domenindedir. $t=0$ anı k . çerçevenin başlangıcına, $t=T$ anı da $k+1$. çerçevenin başlangıcına karşılık gelmektedir. Anlık fazın türevi frekansı verdiğine göre, fonksiyonun kendisi ve türevinin çerçeve sınırlarında çözülmesiyle polinomun katsayıları bulunabilir. $t=0$ anında fonksiyonun değeri, bileşenin ölçülen fazına eşittir. Yine aynı anda anlık fazın türevi de bileşenin ölçülen frekansına eşittir. Bu durumda,

$$\tilde{\theta}(0) = \eta = \hat{\theta}^k \quad (4.27)$$

ve

$$\tilde{\theta}'(t) = \gamma + 2\alpha t + 3\beta t^2 \quad (4.28)$$

dir. Buradan,

$$\tilde{\theta}'(0) = \gamma = \hat{\omega}^k \quad (4.29)$$

bulunur. Çerçevenin diğer sınırında, yani $t=T$ anında faz fonksiyonunun değeri $k+1$. çerçevenin ölçülen fazına, faz fonksiyonunun türevi de $k+1$. çerçevenin frekansına eşit olacaktır.

$$\tilde{\theta}(T) = \hat{\theta}^k + \hat{\omega}^k T + \alpha T^2 + \beta T^3 = \hat{\theta}^{k+1} + 2\pi M \quad (4.30)$$

$$\tilde{\theta}'(T) = \hat{\omega}^k + 2\alpha T + 3\beta T^2 = \hat{\omega}^{k+1} \quad (4.31)$$

Ölçülen faz $\hat{\theta}^{k+1}$ 2π 'ye göre modüle olduğu için, (4.30) eşitliğinde faz sürekliliğini sağlayabilmek için M bir tam sayı olmak üzere $2\pi M$ terimi eklenmiştir. Bu noktada M bilinmemektedir. Fakat (4.30) ve (4.31) denklemleri M'ye bağlı olarak çözülebilir. Bu iki denklem matris biçiminde yazılırsa,

$$\begin{bmatrix} \alpha(M) \\ \beta(M) \end{bmatrix} = \begin{bmatrix} \frac{3}{T^2} & \frac{-1}{T} \\ \frac{-2}{T^3} & \frac{1}{T^2} \end{bmatrix} \begin{bmatrix} \hat{\theta}^{k+1} - \hat{\theta}^k - \hat{\omega}^k T + 2\pi M \\ \hat{\omega}^{k+1} - \hat{\omega}^k \end{bmatrix} \quad (4.32)$$

elde edilir. Artık tek yapılması gereken M'nin bulunmasıdır. Çerçeve boyunca frekansların sabit ve bu aralıkta asıl işaretin durağan olduğu varsayılırsa, fazın doğrusal olması gerekir. Bu durumda M, $\tilde{\theta}''(t; M)$, $\tilde{\theta}(t; M)$ fonksiyonunun zamana göre ikinci türevi olmak üzere

$$f(M) = \int_0^T [\tilde{\theta}''(t; M)]^2 dt \quad (4.33)$$

fonksiyonunu minimum yapan tam sayıdır. Bu eşliği çözerken tam sayı olan M yerine sürekli x değişkenine bağlı $f(x)$ fonksiyonunu çözüp M'yi x'e en yakın tam sayı olarak almak daha kolay olacaktır. (4.33) eşitliğini minimum yapan x aşağıdaki gibi bulunur:

$$x = \frac{1}{2\pi} \left[(\hat{\theta}^k + \hat{\omega}^k T - \hat{\theta}^{k+1}) + (\hat{\omega}^{k+1} - \hat{\omega}^k) \frac{T}{2} \right] \quad (4.34)$$

(4.32) eşitliğindeki matrislerin çarpımıyla α ve β şöyle bulunur:

$$\alpha = \frac{3}{T^2} (\hat{\theta}^{k+1} - \hat{\theta}^k - \hat{\omega}^k T + 2\pi M) - \frac{1}{T} (\hat{\omega}^{k+1} - \hat{\omega}^k) \quad (4.35)$$

$$\beta = -\frac{2}{T^3} (\hat{\theta}^{k+1} - \hat{\theta}^k - \hat{\omega}^k T + 2\pi M) - \frac{1}{T^2} (\hat{\omega}^{k+1} - \hat{\omega}^k) \quad (4.36)$$

Faz fonksiyonu ayrık zamana çevrildiğinde aşağıdaki denklemler elde edilir:

$$x = \frac{1}{2\pi} \left[(\hat{\theta}^k + \hat{\omega}^k S - \hat{\theta}^{k+1}) + (\hat{\omega}^{k+1} - \hat{\omega}^k) \frac{S}{2} \right] \quad (4.37)$$

$$\alpha = \frac{3}{S^2} (\hat{\theta}^{k+1} - \hat{\theta}^k - \hat{\omega}^k S + 2\pi M) - \frac{1}{S} (\hat{\omega}^{k+1} - \hat{\omega}^k) \quad (4.38)$$

$$\beta = -\frac{2}{S^3} (\hat{\theta}^{k+1} - \hat{\theta}^k - \hat{\omega}^k S + 2\pi M) - \frac{1}{S^2} (\hat{\omega}^{k+1} - \hat{\omega}^k) \quad (4.39)$$

$$\tilde{\theta}(n) = \hat{\theta}^k + \hat{\omega}^k n + \alpha n^2 + \beta n^3 \quad (4.40)$$

Sonuç olarak k. çerçeve için sentez eşitliği şu şekildedir:

$$\tilde{s}(n) = \sum_{l=1}^{L(k)} \tilde{A}_l \cos[\tilde{\theta}_l(n)] \quad (4.41)$$

5. UYGULAMA

5.1 Analiz Verisinin Hazırlanması

Analiz edilecek ses, yeterli kaliteyi sağlayabilmek için en az 16 bit örneklenmelidir. İnsan kulağının duyma üst sınırının 22050 Hz. olduğu hatırlanırsa, örnekleme frekansı en az 44100 Hz. olmalıdır. Ses, dinamik açıklığı mümkün olan en yüksek seviyeyle kaydedilmelidir. Seste çınlama ve titreşim olmamalıdır. Mümkünse ses kayıtları yankısız odalarda yapılmalıdır. Eğer sentezlemenin amacı müzikal değişimlerse, ses tek bir kaynaktan alınmalıdır. Aynı anda birden fazla kaynağın bulunduğu seslerde yapılacak değişimler iyi sonuç vermeyecektir. Son olarak, kayıt sırasında arka plan gürültüsünün işaretin spektrumunu bozmaması için en az seviyede tutulması gereklidir.

5.2 Pencereleme

Analiz penceresi, bileşenlerin daha doğru olarak bulunabilmesi için uzun tutulmalıdır. Ancak, işaretin zarfı ve bileşenlerinin frekansları zamanla hızlı ya da keskin değişimler gösteriyorsa, uzun süreli çerçevelerle bunların takibi zorlaşır. Bu yüzden pencere boyu çok uzun olmamalıdır. Analiz penceresinin boyu, sesin perde periyodunun iki ila beş katı arasında olarak alınır. Sesin harmonikliği kuvvetli ve ses tek kaynaklıysa, pencere boyunun uyarlamalı olarak değiştirilmesi gerekli değildir. Ana harmoniğin frekansı perde frekansı olarak alınabilir ve bu değerle bulunan pencere boyu tüm veri için kullanılabilir. Çalışmada gitar gibi hızlı sönümlenen ve kuvvetli harmonik sesler için pencere boyu perde periyodunun iki buçuk katı olarak alınmıştır. Obua, klarnet gibi sesler, zamanla daha yavaş değişir ve bu sesler yapıları gereği bir miktar gürültü içerirler. Böyle sesler için, pencere boyu daha fazla tutulmalıdır. Bu tür seslerde, işlenen sese bağlı olarak pencere boyu perde periyodunun üç buçuk ile beş katı arasında alınmıştır.

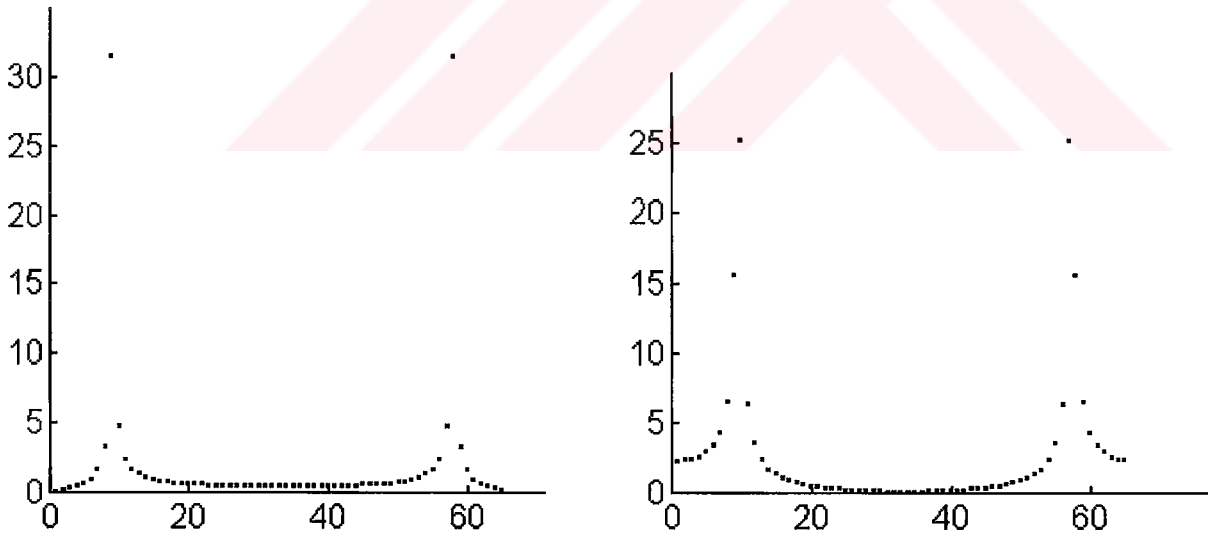
Analiz penceresi olarak yan lobları yeterince basık ve yumuşak eğime sahip pencere tipleri kullanılmalıdır. Hızlı değişen seslerde Hamming penceresi, daha kararlı seslerde ise Blackmann-Harris pencerelerinden biri kullanılabilir.

5.3 Tepe Bulma

İşaret pencerelendikten sonra, AFD ile genlik ve faz spektrumu hesaplanır. AFD ile hesaplanan spektrum hem pencerelemeden hem de AFD'nün örneklenmiş doğasından kaynaklanan bazı hatalar içerir. Pencerelemeden kaynaklanan bozulmalara müdahale edilemese de, AFD'nden kaynaklanan hatalar bir miktar giderilebilir. AFD, sürekli Fourier Dönüşümü'nden hareketle şöyle bulunur:

$$X(k) = X(\omega) \Big|_{\omega=2\pi k/N} \quad (5.1)$$

Burada görülebilir ki AFD, sürekli Fourier Dönüşümünün $k\omega_0$ gibi ayrık frekanslarda örneklenmesiyle elde edilir . Buna dayanarak AFD'nün bir harmonik analiz olduğu söylenebilir. Analiz edilen işaretin i. bileşeninin frekansı ω_i , $m\omega_0$ gibi AFD'nün temel frekansının tam katıysa, ayrık genlik spektrumundan işaretin genliği doğru olarak bulunabilir. Eğer bileşenin frekansı temel frekansın tam katı değilse, genlik spektrumundan bulunan değerler gerçek değerlerden farklı olacaktır (Şekil 5.1).



Şekil 5.1 İki sinüsün Ayrık Fourier Dönüşümleri. (a) $f = 8$ Hz, $A = 1$, (b) $f = 7,3$ Hz, $A = 1$

AFD'nden kaynaklanan bu bozulmayı interpolasyon ile azaltmak mümkündür. Bir sinüsün Fourier Dönüşümünün genlik spektrumu parbole benzediği için, spektrum üzerinde

parabolik interpolasyon yapılabilir. Spektrumdan bulunan maksimum ve bu maksimumla komşu iki noktadan geçen parabolün maksimum noktasının düşey koordinatı düzeltilmiş genliğe, yatay koordinatı ise düzeltilmiş frekansa karşı gelir. Bu düzeltmeyle kesirli frekanslar da elde edilmiş olur. Bu işlem sonuç kaliteyi olumlu etkilerken, analiz verisi saklanması sırasında kesirli sayıların tam sayılardan daha fazla yer kaplaması yüzünden saklama alanını arttırmaktadır. Sıkıştırma oranının kaliteden daha önemli olduğu durumlarda bu düzeltmeden vaz geçilebilir.

5.4 Tepe Uydurma

Tepe uydurma algoritması, gitar sesinin kuvvetli harmonik özelliği kullanılarak basitleştirilmiştir. Karar eşiği, peşpeşe iki harmoniğin aritmetik ortalaması olarak alınır. Böylece i . frekansa ait muhtemel eşlerin $i+1$. bileşene daha yakın olma olasılığı ortadan kaldırılmıştır. Artık i . frekansın muhtemel eşleri içinde kendine en yakın olanı o frekansın eşidir. Uygulamada görülmüştür ki, kayıt koşulları iyi olmak şartıyla gitar sesinin spektral bileşenleri yeterince ayrıktır ve frekansları zamanla yavaş değişir. Sözü edilen basitleştirme, sonuç kaliteyi etkilemeden işlem süresini kısaltmıştır.

5.5 Analiz Verisinin Saklanması

Analiz işleminden sonra, elde edilen değerlerin ileride kullanılmak amacıyla saklanması gereklidir. Saklanacak veriler yalnızca bileşenlere ait değerler değildir. Çerçeve uzunluğu, her çerçevedeki bileşen sayısı gibi değerler de sentez aşamasında gereklidir.

Saklanacak veri paketi iki alt bölüm halinde düşünülebilir. İlk bölümde çerçeve uzunluğu, çerçeve sayısı, çerçevelerdeki bileşen sayısı bulunur. İkinci bölüm ise analizle bulunan bileşenlere ait bilgileri içerir. Analiz dosyasının yapısı Şekil 5.2 ve Şekil 5.3'teki gibidir.

Çerçeve uzunluğu (2 bayt)	Çerçeve sayısı (iki bayt)	k. Çerçevedeki bileşen sayısı (Çerçeve sayısı * 1 bayt)
İşaretin DC seviyesi (2 bayt)	Bileşenlere ait veriler	

Şekil 5.2 Analiz verisinin saklanma şekli

1. çerçevenin 1. Bileşeninin genliği, frekansı, fazı (Her biri 4 bayt)
1. çerçevenin 2. Bileşeninin genliği, frekansı, fazı
...
2. çerçevenin 1. Bileşeninin genliği, frekansı, fazı (Her biri 4 bayt)
2. çerçevenin 2. Bileşeninin genliği, frekansı, fazı
...
...
Dosya sonu

Şekil 5.3 Bileşenlere ait verilerin saklanma şekli

5.6 Sinüzoidal Sentez

Sentez işlemi, tüm çerçevelerdeki bileşenlerin eşleri bulunduktan sonra 3.41 eşitliği uyarınca iç içe iki döngüden ibarettir. Sentez yoluyla elde edilen işaret, Microsoft Windows PCM ses dosyası formatına uygun olarak kaydedilir. Kaydedilen ses dosyası, her hangi bir ses dosyası ya da ses düzenleyicisi ile dinlenebilir.

Yapay sesin kalitesini ölçmek için kesin kurallar mevcut değildir. Uygulamada en iyi karar yöntemi algısaldır. Maskeleye etkisi yüzünden asıl seste duyulamayan ya da düşük genlikli bileşenler, yapay seste olmayabilir. Böyle durumlarda, karesel ortalama gibi hata kriterlerine göre hata oranı yüksek olduğu halde işaret hatasızmış gibi algılanabilir.

5.7 Uygulama Ortamı

Uygulama, hesaplanan değerlerin gözlemlenmesindeki kolaylıklar göz önünde bulundurularak Matlab ortamında yazılmıştır. Matlab'in satır yorumlamalı bir dil olmasından ve sinüzoidal modelin işlem yüksek işlem yükünden dolayı, uygulamanın çalışma süresi uzundur. Uygulamada kullanılan PCM dosya okuma ve yazma programcıkları (wwrite16.m ve wread16.m) The Mathworks Inc. firmasınca geliştirilmiştir.

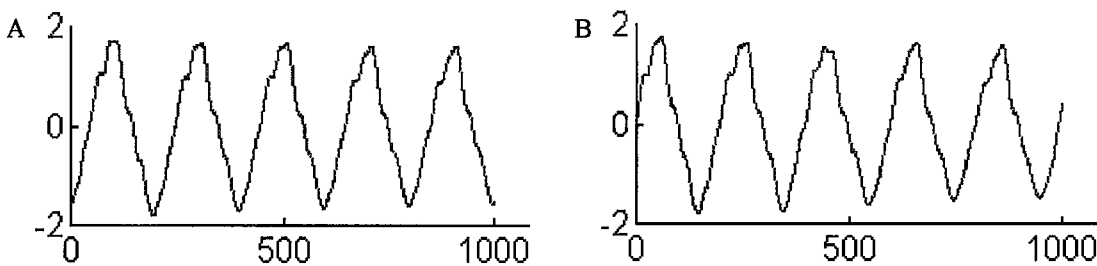
6. SONUÇ

Uygulama, farklı türden sesler üzerinde denenmiştir. Bu denemelerde, pencere uzunluğunun sonuç kaliteye etkisi, asıl işaretin boyutunun analiz verisinin boyutuna oranı ve modelin ne tür seslerde etkili olduğu gibi konular araştırılmıştır.

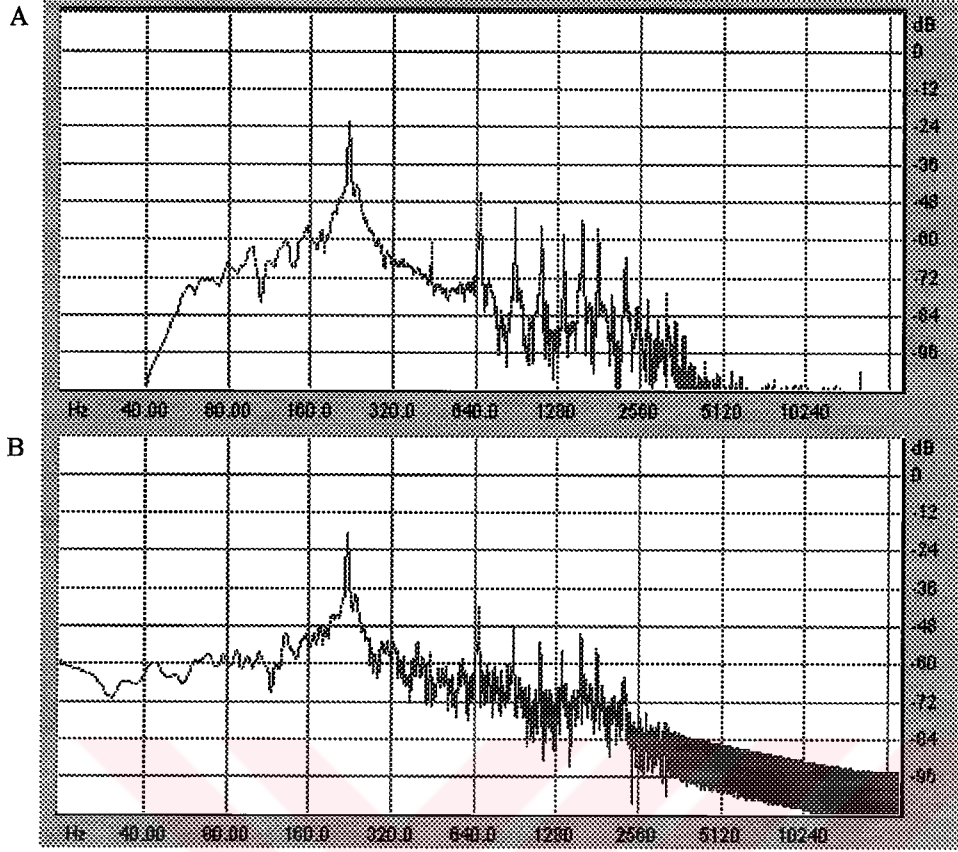
Uygulamada elde edilen sonuçlara göre, penalı telli enstrümanlar için pencere uzunluğu perde periyodunun iki buçuk katı olmalıdır. Ağaç nefeslilerde ise pencere boyunun perde periyodunun üç buçuk ila beş katına çıkarılması daha iyi sonuçlar vermektedir.

Penalı telli enstrümanlar gibi kuvvetli harmonik ve kararlı seslerde analiz verisi, asıl işarete oranla çok az yer kaplamıştır. Örnek vermek gerekirse, 86.1 Kb uzunluğundaki bir PCM ses dosyası analiz edilip saklandığında yalnızca 6.3 Kb yer kaplamıştır. Ancak analiz dosyası da sıkıştırılmaya açıktır. Daha yüksek sıkıştırma oranları için, Huffman türü kayıpsız kodlama tekniklerinden yararlanılabilir. Nefesli enstrümanlarda ise, gerek bileşenlerin fazla sayıda oluşu, gerekse bu tür seslerin gürültülü ve değişken yapıda oluşları nedeniyle sıkıştırma oranı daha düşüktür. Ayrıca pencere boyunun daha uzun olduğu düşük oktavlı sesler de yüksek oktavlı seslere göre daha iyi sıkıştırma oranları sağlamıştır.

Model, gürültüsüz ve kararlı seslerde çok az algısal kayba neden olmuştur. Gürültülü ve değişken seslerde ise algısal farklar ortaya çıkmıştır. Seste çınlama ve titreşim varsa, başarımda ses kaynağından bağımsız olarak hissedilir düşüşler gözlenmiştir. Şekil 6.1 ve Şekil 6.2’de sinüzoidal sentez ile elde edilmiş yapay bir ses, doğal sesle karşılaştırılmıştır.



Şekil 6.1 220 Hz. Gerçek ve yapay harp seslerinin zaman domeninde karşılaştırılması;
(a) Asıl ses; (b) Yapay ses



Şekil 6.2 220 Hz. Gerçek ve yapay harp seslerinin frekans domeninde karşılaştırılması;
(a)Asıl ses; (b) Yapay ses

Zaman domeninde karşılaştırmada , asıl ve yapay harp seslerinin ataktan hemen sonraki 1/44 sn.'lik kısımları; frekans domeninde karşılaştırmada ise, seslerin ortalarından alınan 32768 örneklik kısım kullanılmıştır.

KAYNAKLAR

- Borin, G., De Poli, G. ve Sarti, A., (1992), "Algorithms and Structures for Synthesis Using Physical Models", *Computer Music Journal*, 16(4): 30-42.
- George, E. B. ve Smith, M. J. T., (1997), "Speech Analysis/Synthesis and Modification Using an Analysis-by-Synthesis/Overlap-Add Sinusoidal Model", *IEEE Transactions on Speech and Audio Processing*, 5(5): 389-406.
- Goodwin, M. M., (1997), *Adaptive Signal Models: Theory, Algorithms and Audio Applications*, Doktora Tezi, University of California, Berkeley (yayımlanmamış).
- Harris, F. J., (1978), "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", *Proceedings of the IEEE*, 66(1): 51-83.
- Haykin, S., (1986), *Adaptive Filter Theory*, Englewood Cliffs, New Jersey.
- Macon, M. W., (1993), "Applications of Sinusoidal Modeling to Speech and Audio Signal Processing", Doktora Tezi, Georgia Institute of Technology (yayımlanmamış).
- McAulay, R. J. ve Quatieri, T. F., (1986), "Speech Analysis/Synthesis Based on a Sinusoidal Representation", *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-34(4): 744-754.
- Moorer, J. A., (1977), "Signal Processing Aspects of Computer Music: A Survey", *Proceedings of the IEEE*, 65(8): 1108-1137.
- Oppenheim, A. V. ve Schaffer, R. W., (1983), *Digital Signal Processing*, Englewood Cliffs, New Jersey.
- Pan, D. Y., (1993), "Digital Audio Compression", *Digital Technical Journal*, 5(2).
- Pierce, J. R., (1983), *The Science of Musical Sound*, Scientific American Books, New York.
- Roads, C., (1996), *Computer Music Tutorial*, MIT Press, MA.
- Serra, X., (1997), "Musical Sound Modeling with Sinusoids plus Noise", *Musical Signal Processing*, C. Roads (Derl.), Swets & Zeitlinger Publishers.
- Smith, J. O., (1991), "Viewpoints on the History of Digital Synthesis", *Proceedings of the ICMC*, Oct. 1991, Helsinki.
- Smith, J. O., (1996a), "Discrete Time Modeling of Acoustic Systems with Applications to Sound Synthesis of Musical Instruments", *Proceedings of the Nordic Acoustical Meeting*, 11-14 June 1996, Helsinki.

Smith, J. O., (1996b), "Physical Modeling Synthesis Update", Computer Music Journal, 20(2): 44-56.

Valimäki, V., Huopaniemi, J., Karjalainen, M. ve Janosy, Z., (1996), " Physical Modeling of Plucked String Instruments with Application to Real Time Sound Synthesis", Journal of Audio Engineering Society ,44(5): 331-353.



EKLER**Ek 1 Program Dökümü**

EK 1 Program Dökümü

```
%%%%%%%%%%
%SİNÜZOIDAL ANALİZ PROGRAMI%
%%%%%%%%%%
```

```
function analysis(file1,file2,S)
```

```
if findstr(file1,'.')==[]
    fname1=[file1,'.wav'];
end
```

```
data = wread16(file1);
samp_rate=44100;
trans=2*pi/samp_rate;
frate = floor(size(data,1)/S);
data=data*32767/max(data);
ctr=0;
```

```
%DC bileşen çıkarılıyor.
```

```
mn=mean(data);
data= data - mn;
mak=max(data);
```

```
%Pencereleme ve FFT
```

```
penc=hanning(S)./S;
for i=1:frate;
    parca=data((((i-1)*S)+1:i*S). *penc;
    stf(i,1:1024)=fft([parca((S+1)/2:S)' zeros(1,1024-S) parca(1:(S-1)/2)']);
end;
stfamp=abs(stf).*2;
stfang=angle(stf);
```

```
%Tepe bulma. Her pencere için FFT deki lokal maksimalar bulunuyor.
```

```
for i=1:frate;
    freq(i,1)=0;
    for j=3:511;
        if stfamp(i,j)>stfamp(i,j-1) & stfamp(i,j)>stfamp(i,j+1) & stfamp(i,j)>32
            ctr=ctr+1;
            angl(i,ctr) = stfang(i,j);
```

```
%Parabolik interpolasyon: Frekans ve genliklerde düzeltme.
```

```
x2 = (j-1) * samp_rate / 1024;
y2 = stfamp(i,j);
x1 = (j-2) * samp_rate / 1024;
y1 = stfamp(i,j - 1);
x3 = (j) * samp_rate / 1024;
```

```

        y3 = stfamp(i,j + 1);
        b = ((x1^2 - x3^2) * (y1 - y2) - (x1^2 - x2^2) * (y1 - y3)) / ((x1^2 -
        x3^2) * (x1 - x2) - (x1^2 - x2^2) * (x1 - x3));
        a = ((y1 - y2) - (x1 - x2) * b) / (x1^2 - x2^2);
        c = (y1 - a * x1^2) - (b * x1);
        freq(i,ctr) = -b / (2*a);
        amp(i,ctr) = (a * freq(i,ctr)^2) + (b * freq(i,ctr)) + c;
    end
end;

ca(i)=ctr;
ctr=0;
end;

%Analiz sonu.
yaz(S,frate,mn,ca,amp,freq,angl,file2);
fclose('all');

%%%%%%%%%%%%%%
%SİNÜZOİDAL SENTEZ PROGRAMI%
%%%%%%%%%%%%%%

function synth(fname1,fname2)

samp_rate=44100;
trans=2*pi/samp_rate;

if findstr(fname1,'.')==[]
    fname1=[fname1,'.dat'];
end
if findstr(fname2,'.')==[]
    fname2=[fname2,'.wav'];
end

% *.dat dosya açılıp değerler okunuyor

fid=fopen(fname1,'rb');
S=fread(fid,1,'uint');
frate2=fread(fid,1,'uint');
mn=fread(fid,1,'float');
for i=1:frate2
    ctr(i)=fread(fid,1,'uint');
end

for i=1:frate2

    for j=1:ctr(i)

        amp(i,j)=fread(fid,1,'float');
    end
end
end

```

```

        freq(i,j)=fread(fid,1,'float');
        angl(i,j)=fread(fid,1,'float');
    end
end
fclose(fid);

deterministik(frate2*S) = 0;
frsize=size(freq,2);
freq(frate2,frsize)=0;
amp(frate2,frsize)=0;
angl(frate2,frsize)=0;

for i= 1:frate2-1;

% Çerçeveden çerçeveye uydurma yapıyor.
    if ctr(i)>0
        for j=1 : frsize;
            if j<frsize
                uplev= (freq(i,j) + freq(i,j+1))/2;
            else uplev=freq(i,j) *1.5;
            end;

            if j>1
                dnlev= (freq(i,j) + freq(i,j-1))/2;
            else dnlev=freq(i,j)/2;
            end;

            uyumf(j)= freq(i,j);
            uyumg(j)=0;
            uyuma(j)=angl(i,j);
            index=0;

            for k=1 : frsize
                if freq(i+1,k)<uplev & freq(i+1,k)>=dnlev
                    index=index + 1;
                    olasifr(index)=freq(i+1,k);
                    olasiamp(index)=amp(i+1,k);
                    olasiang(index)=angl(i+1,k);
                end
            end
            if index==1
                uyumf(j)= olasifr(1);
                uyumg(j)= olasiamp(1);
                uyuma(j)= olasiang(1);
            else if index>1
                sonuc=bul(olasifr,freq(i,j)); % bul(A[],f) A vektörü içinde f e en
                % yakın olanla geri döner.
                uyumf(j)= olasifr(sonuc);
            end
        end
    end
end

```

```

        uyumg(j)= olasiamp(sonuc);
        uyuma(j)= olasiang(sonuc);
        end;
    end;
end;

% Genliklerde lineer, anlık fazda kübik interpolasyon yapılacak.
for r=1:frsize
    if (freq(i,r)>0 & freq(i,r)<22000)
        M= round((angl(i,r)+ trans*freq(i,r)*S-uyuma(r) + (S/2)*(trans*uyumf(r)-
        trans*freq(i,r)))/(2*pi));
        nu1= (3/S^2)*(uyuma(r)- angl(i,r)- trans*freq(i,r)*S +2*pi*M)-
        (trans*uyumf(r) - trans*freq(i,r))/S;
        nu2= -(2/S^3)*(uyuma(r)- angl(i,r)- trans*freq(i,r)*S + 2*pi*M)-
        (trans*uyumf(r) - trans*freq(i,r))/S^2;

        for m=1:S
            teta=angl(i,r) + trans*freq(i,r)* (m-1) + nu1* (m-1)^2 + nu2* (m-
            1)^3;
            ampli=amp(i,r) + (uyumg(r)- amp(i,r))* (m-1)/ S;

            deterministik(m + (i-1) * S) = deterministik(m + (i-1) * S) + ampli *
            cos(teta);
            end;
        end;
    end;
end;
end;

% yukarıdaki işlemle aynı. yalnızca son track için.
if ctr(frate2)>0
    i = frate2;
    uyumg=zeros(1,frsize);
    uyumf=freq(frate2,:);
    uyuma=angl(frate2,:);
    for r=1:frsize
        M= round((angl(i,r)+ trans*freq(i,r)*S-uyuma(r) + (S/2)*(trans*uyumf(r)-
        trans*freq(i,r)))/(2*pi));
        nu1= (3/S^2)*(uyuma(r)- angl(i,r)- trans*freq(i,r)*S +2*pi*M)-
        (trans*uyumf(r) - trans*freq(i,r))/S;
        nu2= -(2/S^3)*(uyuma(r)- angl(i,r)- trans*freq(i,r)*S + 2*pi*M)-
        (trans*uyumf(r) - trans*freq(i,r))/S^2; %faz uyumu için kübik interpolasyon

        for m=1:S
            teta=angl(i,r) + trans*freq(i,r)* (m-1) + nu1* (m-1)^2 + nu2* (m-
            1)^3;
            ampli=amp(i,r) + (uyumg(r)- amp(i,r))* (m-1)/ S;

```



```

        deterministik(m + (i-1) * S) = deterministik(m + (i-1) * S) + ampli *
        cos(teta);
    end;
end;

end

%sentetik data skala ediliyor..
maks=max(deterministik);
if max(deterministik)<-min(deterministik)
    maks=min(deterministik);
end

deterministik=deterministik*(32767-mn)/maks;

% Wav data hazır. Yazılıyor.
wwrite16(fname2, (deterministik'+mn) , [1 1 samp_rate 16]);
fclose('all');

function yaz(S,frate,mn,ca,amp,freq,angl,fname)

if findstr(fname, '.')==[]
    fname=[fname, '.dat'];
end
fid=fopen(fname, 'wb');
fwrite(fid, S, 'int');
fwrite(fid, frate, 'uint');
fwrite(fid, mn, 'float');

for i= 1:frate
    fwrite(fid, ca(i), 'uint');
end

for i=1:frate
    for j=1:ca(i)
        fwrite(fid, amp(i,j), 'float');
        fwrite(fid, freq(i,j), 'float');
        fwrite(fid, angl(i,j), 'float');
    end
end
fclose(fid);

%%%%%%%%%%%%%%
%BUL.M SENTEZ İÇİN YARDIMCI PROGRAM%
%%%%%%%%%%%%%%

function ret=bul(olasifr,freq)

```

```

n=size(olasifr,1);
if n==1
    n=size(olasifr,2);
end

temp(:)=abs(olasifr(:)-freq);

closest=min(temp);

for i=1:n
    if closest==temp(i)
        ret=i;
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%YAZ.M ANALİZ VERİSİNİ SAKLAMAK İÇİN YARDIMCI PROGRAM%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function yaz(S,frate,mn,ca,amp,freq,angl,fname)

if findstr(fname, '.')==[]
    fname=[fname, '.dat'];
end
fid=fopen(fname,'wb');
fwrite(fid,S,'int');
fwrite(fid,frate,'uint');
fwrite(fid,mn,'float');

for i= 1:frate
    fwrite(fid,ca(i),'uint');
end

for i=1:frate
    for j=1:ca(i)
        fwrite(fid,amp(i,j),'float');
        fwrite(fid,freq(i,j),'float');
        fwrite(fid,angl(i,j),'float');
    end
end
fclose(fid);
clear

```

ÖZGEÇMİŞ

Doğum tarihi	20.11.1973	
Doğum yeri	Kahramanmaraş	
Lise	1986-1989	Bahçelievler Lisesi
Lisans	1989-1996	Yıldız Teknik Üniversitesi Elektrik-Elektronik Fak. Elektronik ve Hab. Müh. Bölümü
Yüksek Lisans	1996-1998	Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektr. ve Hab. Müh. Anabilim Dalı, Haberleşme Programı
Çalıştığı kurum(lar)		
	1996-1997	Only Bilgisayar Genel Koordinatör
	1997-Devam ediyor	Yıldız Teknik Üniversitesi Elektr. ve Hab. Müh. Bölümü Telekomünikasyon Anabilim Dalı Araştırma Görevlisi

Yıldız Teknik Üniversitesi
Elektronik ve Haberleşme Bölümü
Araştırma Görevlisi