

1. GİRİŞ

Ses sentezi, diğer sentezleme biçimlerinde olduğu gibi daha basit parçalardan kompleks bir bütünün oluşturulmasıdır. Sentez önceleri elektronik devreler aracılığıyla yapılmaktayken günümüzde bu amaçla bilgisayarlar gibi dijital sistemler de kullanılmaktadır.

Ses sentezinin başlıca kullanım alanları yapay konuşma üretimi ve müziktir. Bu alanlar içerisinde ise kullanım olanakları sınırsızdır. Örneğin; konuşma yeteneğini kaybeden insanların dış dünyayla iletişiminin sağlanması ve görme yeteneğini kaybeden insanların yazılı metinleri algılayabilmesi gibi amaçlarla kullanımı son derece gereklidir. Ayrıca ses sentezi tekniklerinin ilerlemesi sayesinde bu tekniklere dayanan sayısız müzik enstrümanı üretilmiştir ve bu enstrümanların akustik enstrümanlara karşı birçok açıdan avantajlara sahip olduğu da görülebilir. Üstelik bu yöntemlerle akustik enstrümanlarda bulunmayan seslerin üretilmesi de mümkündür.

Sinüzoidal modelin de içinde bulunduğu Fourier dönüşümü tabanlı sentez metodlarının yanı sıra genlik, frekans ve faz modülasyonlu sentezleme, granüler sentez, fiziksel modellemeli sentez gibi çeşitli metodlar bulunmaktadır.

Sinüzoidal model, Fourier Teoremi'ne dayanır. Bu modelde ses; keyfi genlik, frekans ve faz değerlerine sahip sinüzoidal parçaların toplamıyla elde edilen dalga formlarıyla ifade edilir ve model, ortaya sinüs dalga tahminine dayalı bir analiz/sentez sistemi koyar [1]. Genlik, frekans, ve faz parametreleri; giriş işaretinin kısa-zamanlı Fourier dönüşümüne basit bir peak-picking algoritması (tepe noktalarının frekanslarının alınması) uygulanarak tahmin edilebilir. Sürekli değişmekte olan işaretlerin analizini mümkün hale getirmek için Kısa-Zamanlı Fourier Dönüşümü (KSFD) kullanılır. Tüm işaretin tek bir Fourier dönüşümü zamanla değişen frekans içeriğini ve harmonikleri karakterize edemediği için bu dönüşüme ihtiyaç vardır [1].

2. SES KAVRAMINA GENEL BAKIŞ

2.1 Ses Dalgaları

Ses; insan kulağının duyma aralığında bulunan, duyulabilmesi için yeterli genliğe sahip, katı,sıvı ve gaz ortamlarda yayılan basınç salınımları şeklinde ifade edilen mekanik dalgalardır. Yayılım sırasında dalgalar yansıyabilir, kırınımına uğrayabilir veya sönümlenebilir. Her organizma sesi ancak belli frekans aralıklarında algılayabilir.Sağlıklı bir insan kulağı için bu aralık 20-20000 Hz'dir. Diğer canlılarda bu aralık farklıdır. Örneğin; köpekler 20 kHz üstünü algılayabilmektedir fakat 40 Hz altını algılayamazlar.

Frekans; tekrar eden bir olayın birim zamanda gerçekleşme sayısıdır. Periyot ise tekrar eden bir olayın bir döngüsünün gerçekleşme süresidir. Frekans ile periyot çarpmaya göre birbirinin tersidir. Frekans birimi Hertz'dir ve f ile gösterilir. Frekans ve dalgaboyu arasındaki ilişki $f = c/\lambda$ şeklindedir. c ışık hızını ifade etmektedir. Dalgaboyu ise bir dalga örüntüsünün tekrarlanan birimleri arasındaki mesafedir.

Ses dalgaları uzamsal dalgalardır ve sıklıkla sinüzoidal düzlem dalgalar olarak da tanımlanırlar. Ses dalgaları frekans, dalgaboyu, dalga sayısı, genlik, faz, ses basıncı, ses hızı, ses derinliği ve yön gibi özellikleriyle karakterize edilirler. Cosinüs dalgaları da sinüzoidal dalga olarak adlandırılırlar. Çünkü cosinüs dalgaları da aslında $\pi/2$ faz farkına sahip sinüs dalgalarıdır. İnsan kulağı tekil sinüs dalgalarını net bir şekilde algılar. Bunun sebebi sinüs dalgalarının harmoniği olmayan tekil frekansların ifadeleri olmalarıdır. Işık sesi ve akort çatalının sesi saf sinüs dalgalarıdır. Birden fazla sinüs dalgadan oluşan işaretler insan kulağı için ya gürültüdür ya da ayırtedilebilir harmonikler içerir. Periyodik sesin ana frekansına temel harmonik denir. Ana frekansın tam katlarındaki bileşenlerin adlandırılması sırasıyla birinci harmonik, ikinci harmonik, üçüncü harmonik vs. şeklinde devam eder.

2.2 Fourier Serisi Açılımı ve Ayırık-Zamanlı ve Ayırık Fourier Dönüşümü

Sinüzoidal dalgalar sayısal işaret işlemede önemli yer tutmaktadır. Fransız matematikçi Joseph Fourier tarafından tanımlanan Fourier serileri ile periyodik bir fonksiyon ya da sinyalin sinüs ve cosinüs işaretlerinin ya da kompleks eksponansiyellerin toplamı olarak ifade edilebileceği görülmüştür. Bu toplam sonsuz bileşen içermektedir. Fakat kabul edilebilir bir hata payıyla bileşen sayısı sonlu hale getirilebilir. Elbette bileşen sayısı ne kadar fazla olursa Fourier serisine açılmış işaret asıl işarete o derece yakın olacaktır. Fourier serisi açılımları şu şekilde ifade edilir:

$$f(t) = \sum_{k=0}^{\infty} A_k \sin(k\omega t + \varphi_k) \quad (2.1)$$

Burada k. Bileşen için A_k genliği, φ_k da fazı ifade eder. Eşitlik aynı zamanda şu şekilde de yazılabilir:

$$f(t) = \sum_{k=0}^{\infty} (a_k \cos k\omega t + b_k \sin k\omega t) \quad (2.2)$$

Burada a_k ve b_k Fourier katsayılarıdır ve şu şekilde hesaplanır:

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

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

Fourier katsayıları kullanılarak genlik ve faz da hesaplanır:

$$A_k = \sqrt{a_k^2 + b_k^2}$$

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

Euler formülü kullanılarak periyodik sinyal kompleks eksponansiyeller şeklinde de ifade edilebilir:

$$e^{jnx} = \cos(nx) + j\sin(nx)$$

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{jnx} \quad (2.3)$$

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-jnx} dx$$

Fourier serileri, sonlu sınırlı periyotlarda tekrarlanan fonksiyonlar içindir. Bir fonksiyon periyodik değilse ve bütün uzayda tanımlıysa Fourier serisi anlamsızdır. Bu durumlarda, Fourier serilerinin genel biçimi olarak kabul edilecek olan Fourier dönüşümü ele alınmalıdır. Fourier dönüşümü Fourier serisinden türetilir. Fakat dönüşümün sonunda elde edilen, artık bir serinin terim katsayıları değil, bir fonksiyondur. Fourier dönüşümü, birbiriyle bir integral dönüşümü altında ilişkili olan iki uzay arasındaki dönüşümdür.

Ayrık-zamanlı Fourier dönüşümü (AZFD) çifti şu şekilde ifade edilir:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\omega n} d\omega \quad (2.4)$$

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} \quad (2.5)$$

Ayrık-zamanlı bir dizinin Fourier dönüşümü frekansın sürekli bir fonksiyonudur. Pratikte bilgisayar kullanımında sürekli frekansta çalışmak mümkün değildir. Bu sebeple dizinin yeniden elde edilebilir bir şekilde örneklenmesi gerekmektedir. N

sınırlı diziler için örnekleme ayrık Fourier dönüşümü (AFD) olarak adlandırılır. Ayrık Fourier dönüşüm çifti:

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi}{N}kn}, \quad 0 \leq n \leq N-1 \quad (2.6)$$

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, \quad 0 \leq k \leq N-1 \quad (2.7)$$

Ayrık-zamanlı Fourier dönüşümü düzgün dağılımlı $\omega = \frac{2\pi}{N}k$ örneklerle örneklendiği zaman ayrık Fourier dönüşümü değerleri elde edilir.

2.3 Perde Kavramı

Perde; sesin algılanan temel frekansını ifade eder. Perdeler peslik ve tizlikle karşılaştırılır. Bu bağlamda perde sübjektif bir kavramdır. Müzik; teknik olarak keyfi frekanslara sahip seslerle yaratılabilir. Müziğin tarih boyunca gelişiminin sonucu olarak Batı müziğinde sabit frekans değerlerine sahip 12 nota kullanılmaktadır. Notaların fiziksel sebebi mekanik sistemlerdeki titreşimler olduğuna göre 1 Hz saniyedeki bir tam titreşim olarak görülebilir. Notaların sabit frekans değerlerine gelinecek olursa örneğin; saniyedeki atım sayısı 440 ise La notası (A4) elde edilmiş olur. Diğer notalar da A4 notası referans alınarak şu formülle hesaplanabilir:

$$f = 440 \times 2^{n/12} \quad (2.8)$$

Burada n 'in pozitif olması A4'ten daha tiz sesleri, negatif olması da daha pes sesleri ifade eder. Örneğin C5 notası elde edilmek isteniyorsa $n = +3$ alınmalıdır. Çünkü A4 ve C5 arasında 3 yarım aralık ($A4 \rightarrow A\#4 \rightarrow B4 \rightarrow C5$) vardır. La notasının 440 Hz kabul edildiği referans perde konser perdesi olarak adlandırılır.

Telli bir enstrumanda tel uzunluğu ve frekans ilişkisi de şu şekildedir:

$$f = k \times \frac{\sqrt{\text{gerginlik}}}{\text{uzunluk} \times \sqrt{\text{özkütle}}}$$

2.4 Ses Gücü

Ses dalgalarının sahip oldukları güç watt cinsinden ölçülebilir. Ses sürekli yayılma halinde olduğundan dalgaının birim alan başına taşıdığı güç yoğunluğunun ölçülmesi gereklidir. Ses şiddeti yani sesin metrekare başına taşıdığı güç ise desibel cinsinden ölçülür. Desibel, belirli bir referans güce olan oranı belirten logaritmik ve boyutsuz bir birimdir. Ses şiddeti ölçümünde referans olarak insan kulağının algılayabildiği en düşük güç yoğunluğu olan 10^{-12} W/m^2 kullanılır ki bu da 0 dB'ye karşılık gelir. İnsan kulağı için üst sınır da (acı sınırı) 130 dB'dir.

3. SES SENTEZİ VE SENTEZLEME METODLARI

Sentez, daha basit parçalardan kompleks bir bütünün oluşturulmasıdır. Ses sentezi öncelikle analog ve dijital sentez olarak ayrılabilir. Analog sentez 60'larda Moog ve Buchla ile başlamıştır ve 80'lere kadar popüleritesini sürdürmüştür. Dijital sentez de ilk kez 1957'de Max V. Matthews tarafından başlatılmıştır fakat 80'lere kadar geniş kullanım bulamamıştır. Dijital sentez oldukça esnektir ve analog sentez de dahil olmak üzere birçok tekniğin uygulanmasına olanak tanır.

Dijital sentezleyiciler de donanımsal ve yazılımsal sentezleyiciler olarak ikiye ayrılabilir. Donanımsal sentezleyiciler aslında işlemci, bellek birimleri ve işletim sistemleriyle birer bilgisayardır. Sentez ve işleme için kullanılan sayısal işaret işleyiciler (DSP) barındırırlar. Ayrıca kontrol için müzikal klavye ve butonlar da sisteme dahildir.



Şekil 3.1 Robert Moog tarafından icat edilen Moog sentezleyici.

Yazılımsal sentezleyiciler genel amaçlı kullanılan bir bilgisayar ve bir ses kartı ile kullanılabilirler. Gelişen bilgisayar teknolojisi ile günümüzde gerçek-zamanlı ses üretimi mümkün hale gelmiştir. Yazılımsal sentezleyiciler esneklik ve ekonomik olma avantajlarına sahiptir. Buna karşın gürültü, ses gecikmesi, işlem gücü ve güvenilirlik potansiyel dezavantajlardır. Bunun yanı sıra yazılımsal sentezleyiciler kontrol için özel aygıtlara ihtiyaç duyabilirler. Bu durum yazılımsal ve donanımsal sentezleyicileri birbirine yaklaştırmaktadır.

Sayısal ses sentezinin temel problemleri Max V. Matthews tarafından şu şekilde ortaya konmuştur:

Ses sentezinin iki temel problemi; (1) bir ses fonksiyonunu tanımlamak için yüksek miktarda veri gereksinimi ve çok hızlı bir program ihtiyacı ve (2) karmaşık ses dizilerini tanımlamak için kullanılacak basit ve güçlü bir dil geliştirilmesidir [2].

Teknolojinin hızlı gelişimi ilk problemi kendiliğinden çözmüştür fakat bir müzikal parçanın örnek örnek yazılması çok fazla vakit alacağı için problem; mevcut sentez metodları arasından duruma göre en işlevsel olanların, kaynaklar da göz önünde bulundurularak seçilmesi gerekliliği boyutuna taşınmıştır.

3.1 Sentez Metodlarının Sınıflandırılması

Sayısal sentez metodları şu şekilde sınıflandırılır: mutlak algoritmalar, işlenmiş kayıtlar, spektral modeller ve fiziksel modeller [3].

3.1.1 Mutlak Algoritmalar, İşlenmiş Kayıtlar ve Örneklem

Bilgisayar müziğinin atası sayılabilecek ilk denemeler 1920'lerde Hindemith, Millhaud ve Toch gibi besteciler tarafından yapılmıştır. 1950'de Pierre Schaeffer Paris'te Studio de Musique Concrète'i kurmuştur. Musique concrète'de besteci kayıtlardan veya gerçek seslerden elde edilen elementlerle çalışır.

Bu metodlar ya mutlak algoritmalar ya da işlenmiş kayıtlara dayanır. Elde edilen verinin büyük yer kaplaması ve sınırlılık sebebiyle ve daha güçlü ve etkili metodların ortaya çıkmasıyla bu yöntemlerin kullanımı azalmıştır. Fakat basitlikleri sebebiyle diğer yöntemlere altyapı oluşturmak amacıyla kullanılabilirler.

Mutlak algoritmalar FM sentezi, dalga şekillendirme sentezi ve Karplus-Strong algoritması örnek gösterilebilir. İşlenmiş kayıtlarla senteze ise örneklem, çoklu dalga tablosu sentezi ve granüler sentez örnek gösterilebilir.

FM sentezinde biri taşıyıcı diğeri modüle edici iki sinusoidal osilatör, taşıyıcı frekansının module edici dalga formu tarafından module edileceği şekilde bağlanır.

Dalga şekillendirme sentezinde giriş işareti lineer olmayan bir şekillendirme fonksiyonuyla modifiye edilir.

Karplus-Strong algoritması işaretin periyodik olarak bilgisayardan okunduğu dalga tablosu sentezinden yola çıkılarak ortaya çıkmıştır. Aradaki fark her örnek okunduktan sonra dalga tablosunun değiştirilmesidir. Bu yöntemle başarılı davul ve çekilerek çalınan telli çalgı sesleri elde edilmiştir.

Örnekleme sentez kısa süreli kayıtların yeniden çalınmasına dayanır. Genellikle perde kaydırma, döngüye sokma ve diğer modifikasyonlar için kullanılır.

Granüler sentezde ses sinyalleri taneler veya ‘ses atomları’ olarak ele alınır. Sentetik ses işareti bu elementlerin zaman domeninde birleştirilmesiyle oluşur.

3.1.2 Spektral Modeller

Spektral modeller ses dalgalarının özelliklerinin dinleyici tarafından algılandığı şekle göre modellenmesine dayanır. Örnek olarak toplamsal sentez, faz kodlayıcı, kaynak-filtre sentezi, McAulay-Quatieri algoritması, Spektral Modellemeli Sentez (SMS), geçici modellemeli sentez (TMS), ters FFT sentezi gösterilebilir. İnsan sesi sentezi için ise CHANT ve VOSIM metodlarından bahsedilebilir. McAulay-Quatieri algoritması (Sinüsoidal Model) bölüm 5’te detaylı bir şekilde anlatılacaktır.

3.1.2.1 Toplamsal Sentez

Bu çalışmanın esas konusunu oluşturan Sinüsoidal Model toplamsal sentezin analiz tabanlı bir sunumudur ve toplamsal sentezi bütünüyle içermektedir. Metodun temelini basit dalga formlarının toplanmasıyla istenen sesin elde edilmesi oluşturur. Uzun zamandır kullanılan bir yöntemdir. Fourier teoremine dayandığı için Fourier sentezi olarak da adlandırılır. Detaylı bilgi 4. ve 5. bölümlerde verilecektir.

3.1.2.2 Faz Kodlayıcı

Faz kodlayıcı 1966'da Bell Laboratuvarlarında Flanagan ve Golden tarafından geliştirilmiştir. Her bir kodlayıcı çoklu paralel kanallardaki giriş işaretini farklı frekanslarda ifade eder. Kodlayıcılar karmaşık spektral bilgiyi basitleştirir ve kullanılan veri miktarını azaltır. Analizde KSFD veya filtre bankası kullanılabilir. Faz kodlayıcı harmonik ve statiklerle kullanıldığında ya da yavaşça değişen tonlarda en iyi sonuçları verir. Gürültü ve hızlı değişim durumlarında ise zorluklarla karşılaşılır.

3.1.2.3 Kaynak-Filtre Sentezi (Çıkarımsal Sentez)

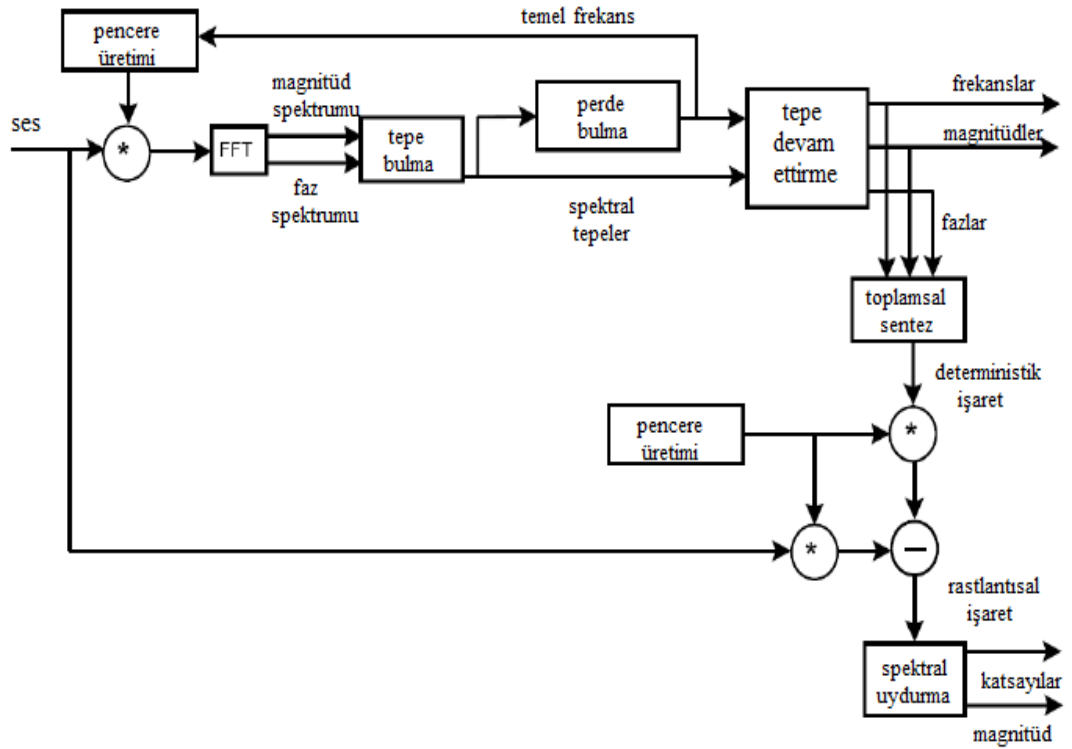
Kaynak-Filtre sentezinde dalgaformu bir uyarım sinyalinin zamanla değişen bir filtreden geçirilmesiyle elde edilir. Metod çıkarımsal sentez olarak da bilinir ve özellikle konuşma sentezi için kullanılmıştır. Düşünce, genişbandlı veya harmonik olarak zengin bir uyarım sinyalinin filtrelenerek istenen çıkışın elde edilmesidir. Bu bağlamda toplamsal sentezin karşıtıdır. Teoride keyfi, band limitli, periyodik dalga formu filtrelemeyle bir impuls dizisinden elde edilebilir. Konuşma sentezi yapılırken konuşmaya bağlı olarak kaynak ünlü harfler için impuls dizisi, ünsüz harfler için ise beyaz gürültü üretir [4].

Bu metod aynı zamanda fiziksel model olarak da yorumlanabilir. Çünkü insan ses üretimi mekanizması da bir uyarım sinyalinin konuşma organlarının hareketleriyle filtrelenmesi şeklindedir.

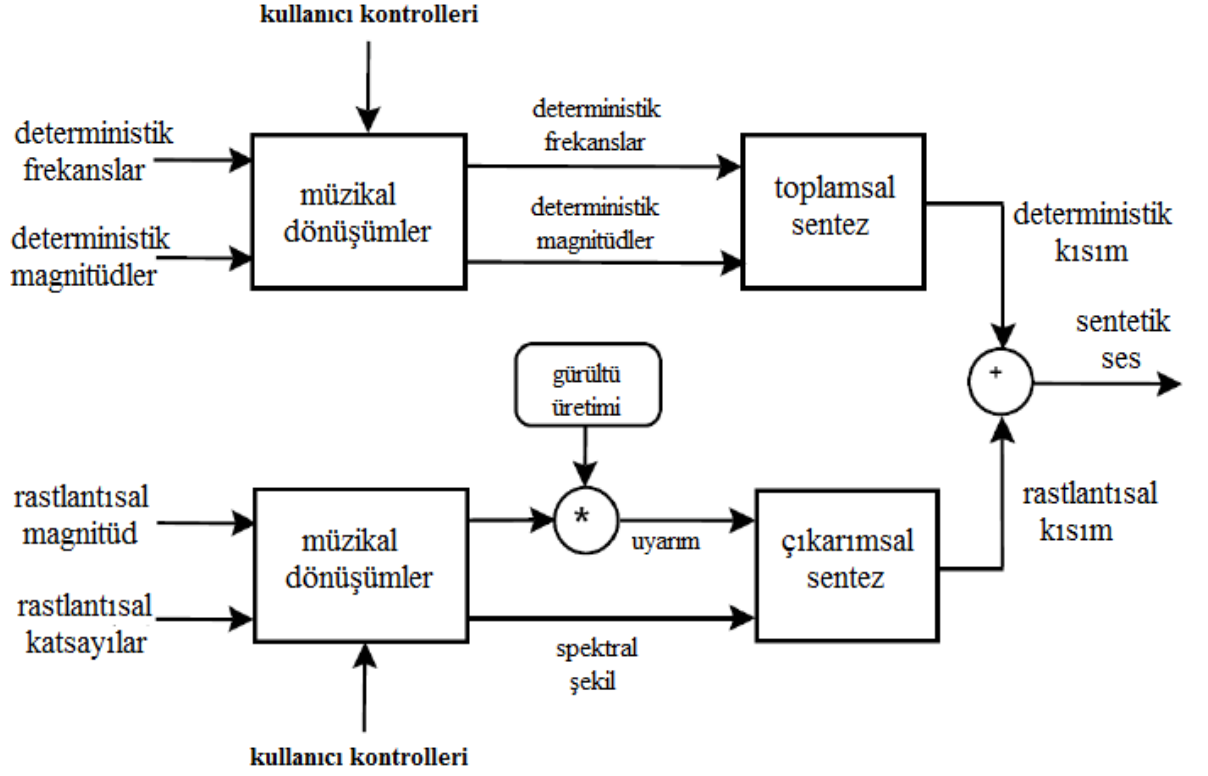
3.1.2.4 Spektral Modellemeli Sentez

1980'lerin sonunda CCRMA, Stanford Üniversitesi'nde Xavier Serra tarafından geliştirilmiştir. Serra'nın metodu ses sinyalinin belirleyici (deterministik, rastgele olmayan) ve rastlantısal kısımlarına ayrılmasına dayanır. Belirleyici kısım McAulay-Quatieri algoritmasıyla veya şiddet analiziyle elde edilir. Belirleyici kısmın orjinal sinyalden çıkarımıyla olasılıksal kısma karşılık gelen artakalan işarete ulaşılır. Model belirleyici kısmı lineer genlik ve frekans değerlerine sahip sinüzoidal parçalarla sınırlar. Rastlantısal kısım ise güç spektral derinliği ile tanımlanır. Bu sayede rastlantısal kısmın faz bilgisinin korunması gerekli değildir. Rastlantısal kısım her bir çerçeveden artakalan kısmın şiddet spektrum zarfıyla ifade edilebilir.

SMS metodu verinin azaltılması ve hesaplama ihtiyaçları konularında oldukça etkilidir. Genel bir metoddur ve pek çok sese uygulanabilir. Fakat zaman da kaymalar ve geçici işaretlerin yayılması gibi problemler mevcuttur.



Şekil 3.2 Spektral modelde analiz sistemi [5].



Şekil 3.3 Spektral modelde sentez sistemi [5].

3.1.2.5 Geçici Modellemeli Sentez

SMS metodunun devamı niteliğindedir. Bu methodda rastlantısal kısım geçici ve kalıcı gürültülü parçalar olarak ikiye ayrılır. Methodda impulsif sinyaller sinüzoidal karakterli frekans domenı sinyalleri olarak ele alınır. Bu sayede sinüzoidal model frekans domeninde uygulanmış olur ve impulsif sinyal parametrelerle ifade edilir [4].

3.1.2.6 Ters Hızlı Fourier Dönüşümü Metodu

Bu methodda toplamsal sentez frekans domeninde uygulanır. Sinyal parçaları KSFD çerçeveleri oluşturan spektral zarflar olarak toplanır. Dalgaformu her bir çerçevenin ters hızlı Fourier dönüşümü (ters FFT) hesaplanarak oluşturulur ve çakışma-ekleme methoduyla arka arkaya gelen çerçeveler birbirine eklenir.

3.1.2.7 Biçimlendirici Sentez

Birçok durumda Fourier dönüşümüyle elde edilen detaylar yerine spektrumun daha genel bir görüntüsünü ele almak yararlıdır. Spektral zarfların ana konseptlerinden biri şiddet spektrumu zarfındaki bir tepeye karşılık gelen biçimlendiricilerdir. Başka bir deyişle biçimlendirici spektrumdaki bir enerji yoğunluğudur. Merkez frekans, band genişliği, genlik ve zarfıyla ifade edilir. Biçimlendiriciler müzik enstrumanlarını tanımlamada kullanışlıdır ama konuşma ve şarkı söyleme sesi sentezinde de geniş kullanım alanı bulmuştur. CHANT ve VOSIM metodları biçimlendiricilerden yola çıkılarak geliştirilmiştir [4].

3.1.3 Fiziksel Modeller

Müzik enstrumanlarının fiziksel olarak modellenmesi ses sentezi, müzik akustikleri ve bilgisayar müziği araştırmalarında en aktif alanlardan biri haline gelmiştir. Fiziksel modellemeli uygulamalar kullanıcıya gerçek bir enstruman çalma hissi vermektedir.

Modelin hedefi; gerçek bir enstrumanın temel fiziksel davranışının taklit edilmesidir. Bu, gözlenen sistemin davranış ve etkileşimlerinin fizik yasaları çerçevesinde incelenmesi ve matematiksel formül ve denklemlerle ifade edilmesi yoluyla gerçekleştirilir.

Fiziksel metodlar gerçek-zamanlı ses sentezinde fazla kullanışlı değildir. Bunun birinci sebebi yüksek kaliteli sentez istendiğinde hesaplama maliyetinin önemli miktarda artmasıdır. İkinci sebep ise parametrelerin müzikal bir şekilde sezgisel olmaması, gerçek bir enstrumanın kontrol edildiği gibi kontrol edilmelerinin zor olmasıdır.

Fiziksel metodlar şu şekilde kategorize edilebilir:

1. Kısmi diferansiyel denklemlerin numerik çözümü
2. Kaynak-filtre modellemesi
3. Kütle-yay sistemlerinin titreştirilmesi
4. Şekilsel sentez
5. Dalga kılavuzu sentezi

3.1.3.1 Kısmi Diferansiyel Denklemlerin Numerik Çözümü

Bu metod özellikle yaylı enstrümanlar için kullanılır fakat titreşen herhangi bir objeye uygulanabilir. Temel prensip gözlenen objenin titreşimsel davranışının denklemlerle ifade edilmesidir. Bu dalga denklemleri objedeki sınırlı noktalar kümesi için çözülür ve bir fark denklemi elde edilir. Fark denklemlerinin kullanımı titreşen objedeki dalga propagasyonu simülasyonu olarak yorumlanabilecek bir yineleme denklemi verir. Bu metod en iyi sonuçları tek boyutlu titreştiricilerle verir çünkü boyut sayısı arttıkça hesaplama ihtiyaçları da hızla artacaktır [4].

3.1.3.2 Kaynak-Filtre Modellemesi

Bu metod özellikle insan sesinin üretim mekanizmasının modellenmesinde kullanılmıştır. Bu metod için etkili sayısal filtreleme teknikleri özellikle konuşma iletimi amaçlarıyla geliştirilmiştir. Teknik temel olarak bölüm 2.1.2.3 de bahsedilen metodun fiziksel model yorumudur.

3.1.3.3 Kütle-Yay Ağları

Akustik sistemin yay, kütle ve sürtünme elemanları gibi basit ideal mekanik elementler kullanılarak modellenmesine dayanır. CORDIS sistemi fiziksel modelle gerçek zamanlı olarak ses üretiminin ilk kez başarıldığı sistemdir. Temel elementler uygun miktarda kombine edilip birbirine bağlanırsa istenen örnekleme hızında akustik sinyal elde edilebilir. Her element tek boyutlu olarak alındığında önemli bir

basitleştirme gerçekleşmiş olur. Zamanla değişen etkileşimlerin modellenmesinde ayarlanabilir parametrelere sahip sürtünme ve yay elemanlarının paralel bağlanmasıyla elde edilen koşulsal hatlar kullanılır.

CORDIS sisteminin geliştiricileri çok boyutlu uygulamalar için de ANIMA sistemini geliştirmişlerdir.

3.1.3.4 Şekilsel Sentez

Bu metod önceleri Mosaic daha sonraları Modalys olarak adlandırılan bir uygulama yazılımıyla gerçekleştirilmiştir. Bu uygulamayla kullanıcı titreşen yapıları simüle edebilir. Kullanıcı programa yapıyı ifade eder ve program şekilsel veriyi hesaplar ve istenen noktaya çıkışı verir. Metod ses üreten herhangi bir objenin şekilsel veriyle tanımlanan titreşen yapılar şeklinde ifade edilebileceği düşüncesine dayanır. Alt yapılar eşleştirilebilir ve dış uyarımlara cevap verebilirler. Ayrıca bu eşleşmeler alt yapılar arası enerji akışına da olanak tanır. Tipik alt yapılar; gövdeler ve köprüler, yaylar, akustik tüpler, membran ve plakalar ve zillerdir. Metod istenen karmaşıklıkta yapıya uygulanabilir fakat karmaşıklık arttıkça hesaplama ihtiyacı da artacaktır [4].

3.1.3.5 Dalga Kılavuzu Sentezi

Sayısal dalga kılavuzları ve tekil gecikme döngüleri (SDL) gerçek zamanlı fiziksel modellemede en verimli sonuçları verirler. Sayısal dalga kılavuzu konsepti J.O. Smith tarafından geliştirilmiştir. Metod dalga denkleminin çözümüyle başlar. Sınırlı fark metoduyla dalga denklemi titreşen obje üzerindeki bir ayrık noktalar kümesinde çözülür. Her periyotta bir fiziksel değişken her noktada hesaplanır. Bu şekilde ayrıklaştırılmış titreşen objenin tüm davranışı gözlenebilir hale gelir [4].

4. KISA-SÜRELİ FOURIER DÖNÜŞÜMÜ

İşaret analizinde tüm işaretin tek bir Fourier dönüşümünün zamanla değişen frekans içeriğini ve harmonikler gibi değişiklikleri karakterize edememesi problemiyle karşılaşılır. Bu durumda ayrık-zamanlı kısa-zamanlı Fourier dönüşüm (ayrık-zamanlı KSFD) olarak bilinen yeni bir dönüşümden bahsetmek gerekir [6]. Ayrık-zamanlı KSFD bir pencere altında işaretin birbirinden bağımsız Fourier dönüşümlerini içerir. Ayrık-zamanlı KSFD şu şekilde elde edilir:

$$X(\omega, \tau) = \sum_{n=-\infty}^{\infty} x[n, \tau] \exp[-j\omega n] \quad (4.1)$$

$x[n, \tau]$ pencerelenmiş işareti ifade eder ve $x[n, \tau] = w[n, \tau]x[n]$ şeklindedir. Zamanla değişen spektral karakteristiklerin şiddeti şu şekilde verilir:

$$S(\omega, \tau) = |X(\omega, \tau)|^2$$

Pratik uygulamalarda KSFD'deki her bir Fourier dönüşümünün yerini ayrık Fourier dönüşümü alır. Buna da ayrık kısa-zamanlı Fourier dönüşüm (ayrık KSFD) denir.

4.1 Kısa-Zamanlı Analiz

KSFD ile ilgili iki görüş vardır. Bunlar: (1) Fourier dönüşümü görüşü ve (2) filtre bankası görüşüdür.

4.1.1 Fourier Dönüşümü Görüşü

Bir n anındaki ayrık-zamanlı KSFD şu şekildedir:

$$X(n, \omega) = \sum_{m=-\infty}^{\infty} x[m]w[n - m]e^{-j\omega n} \quad (4.2)$$

Burada $w[n]$ sadece $[0, N_w - 1]$ aralığında 0'dan farklıdır ve analiz penceresi olarak adlandırılır. $f_n[m] = x[m]w[n - m]$, $x[m]$ 'in n anındaki kısa-zamanlı kısmıdır. İşlem, pencere fonksiyonun ters çevrilip, örnekten örneğe kaydırılıp çarpılması ve pencerelenmiş kısmın Fourier dönüşümünün alınması şeklindedir. Pencere genişliği genellikle 20-30 ms ya da birkaç perde periyodu kadar seçilir [1].

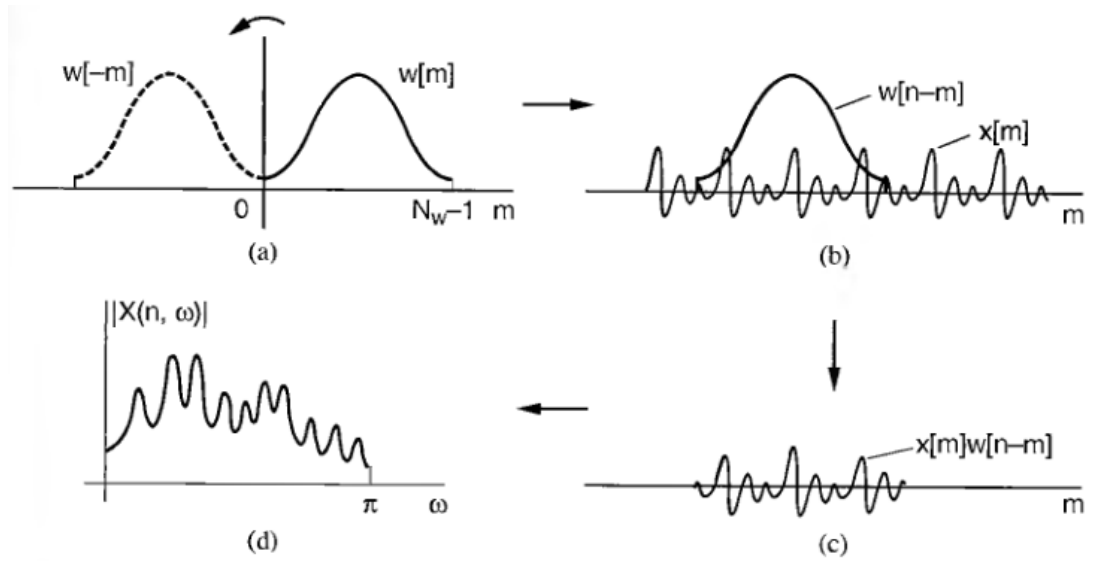
Ayrık-zamanlı KSFD – ayrık KSFD ilişkisi şu şekildedir:

$$X(n, k) = X(n, \omega) \Big|_{\omega = \frac{2\pi}{N}k} \quad (4.3)$$

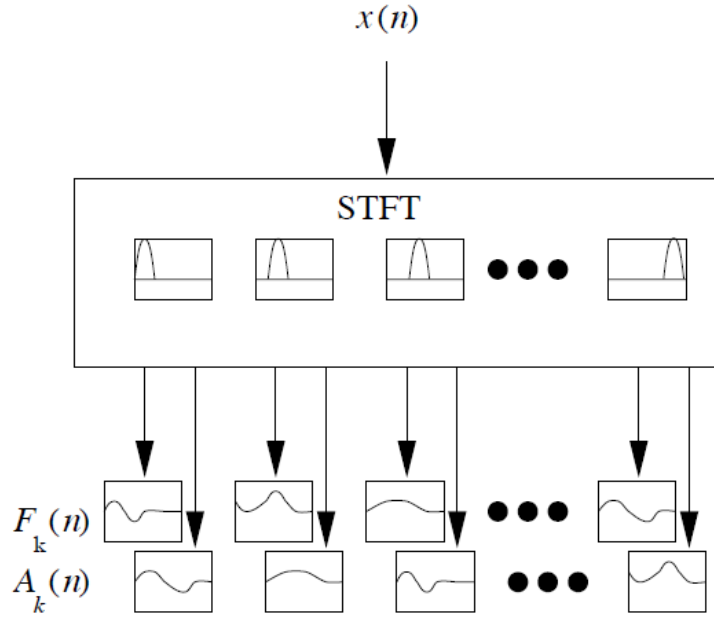
Burada ayrık-zamanlı KSFD, $\frac{2\pi}{N}$ 'lik frekans örnekleme aralığında örneklenmiştir. N ise frekans örnekleme faktörü olarak adlandırılır. Ayrık KSFD denklemi:

$$X(n, k) = \sum_{m=-\infty}^{\infty} x[m]w[n - m]e^{-j\frac{2\pi}{N}km} \quad (4.4)$$

Birçok uygulamada $X(n, k)$ 'nın örnekleri geçici örnek seyreltme faktörü L ile seyreltilir ve fonksiyon $X(nL, k)$ halini alır.



Şekil 4.1 Pencereleme ve pencerelenmiş parçaların Fourier dönüşümünün alınması.



Şekil 4.2 Toplamsal analiz tekniği.

Fourier dönüşümü yorumunda, $X(n, \omega)$, her n için Fourier dönüşümünün tüm genel özelliklerini taşır. Örneğin 2π 'yle periyodiktir ve gerçekte diziler için Hermitian simetrisi gösterir. Bir başka özellik de dizideki bir zamana kaymaya karşın frekans domeninde bir lineer faz faktörü oluşur. $x[n]$, n_0 örnekle kaydırılırsa:

$$\tilde{X}(n, \omega) = e^{-j\omega n_0} X(n - n_0, \omega).$$

4.1.2 Filtre Bankası Görüşü

KSFD; $w[n]$ 'in filtre impuls cevabı olduğu bir filtreleme operasyonu çıkışı olarak da düşünülebilir. Bu durumda $w[n]$ analiz filtresi olarak adlandırılır. Bu görüşte ω değeri ω_0 'da sabitlenir (Fourier dönüşümü görüşünde ise n sabitlenmişti). Denklem (4.2) yeniden yazılırsa:

$$X(n, \omega_0) = \sum_{m=-\infty}^{\infty} (x[m]e^{-j\omega_0 m})w[n - m] \quad (4.5)$$

Bu durumda KSFD, $x[n]e^{-j\omega_0 n}$ ile $w[n]$ dizisinin konvolusyonudur. Denklem şu şekilde yazılabilir:

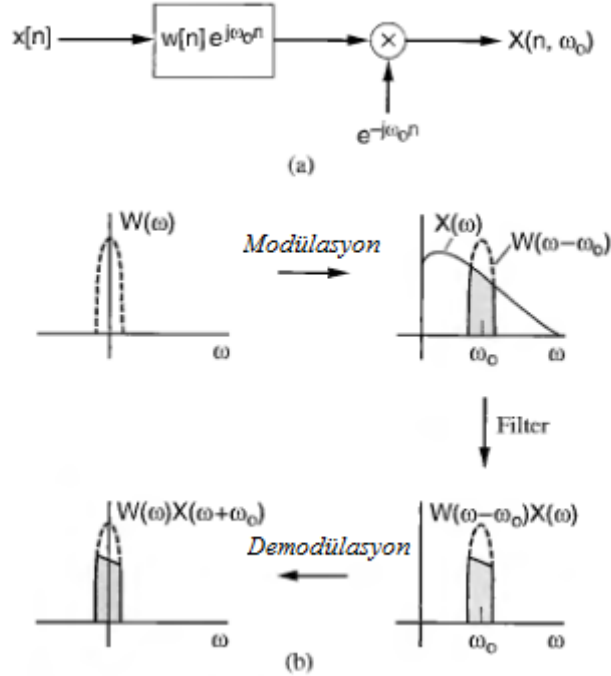
$$X(n, \omega_0) = (x[n]e^{-j\omega_0 n}) * w[n] \quad (4.6)$$

Burada $x[n]$, eksponansiyel modülasyondan sonra impuls cevabı $w[n]$ olan alçak geçiren filtreden geçmektedir. Denklem aşağıdaki gibi yazıldığında ise $x[n]$ önce filtreden geçip daha sonra modüle edilmektedir.

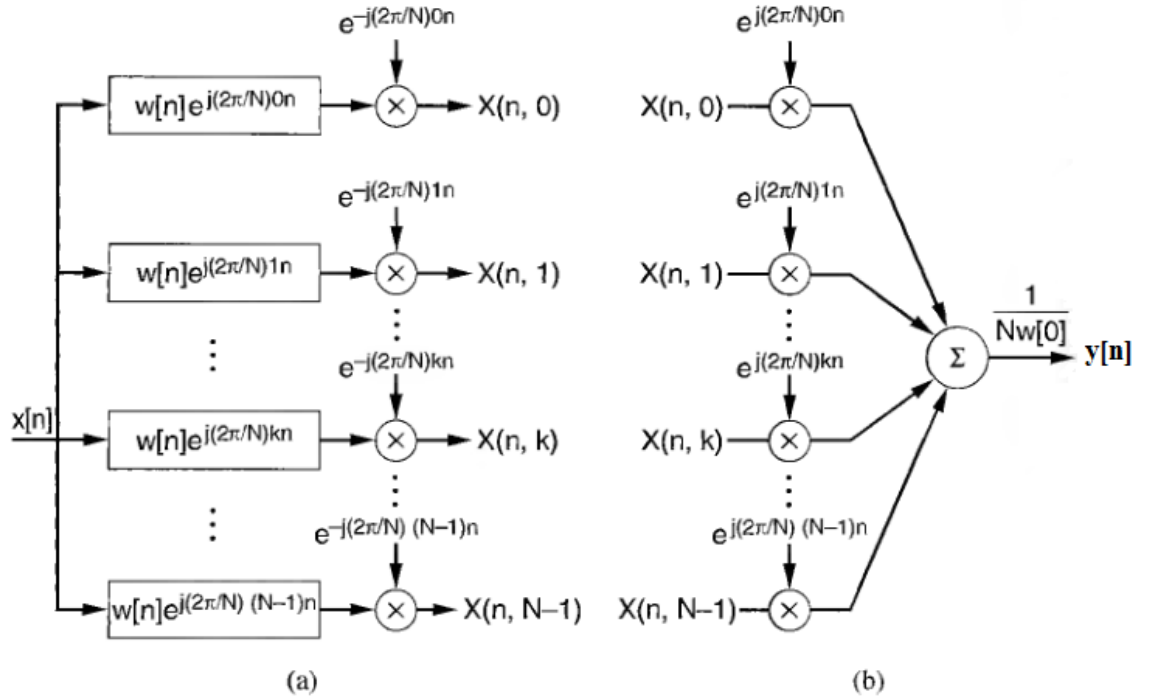
$$X(n, \omega_0) = e^{-j\omega_0 n} (x[n] * w[n] e^{j\omega_0 n}) \quad (4.7)$$

Ayrık KSFD denklemi de şu şekilde ifade edilir:

$$X(n, k) = e^{-j\frac{2\pi}{N}kn} (x[n] * w[n] e^{j\frac{2\pi}{N}kn}) \quad (4.8)$$



Şekil 4.3 ω_0 frekansında KSFD analizi.



Şekil 4.4 Ayırık KSFD'nin filtreleme görüşü.

4.1.3 Analiz Penceresi Seçimi

KSFD şu şekilde de ifade edilebilir:

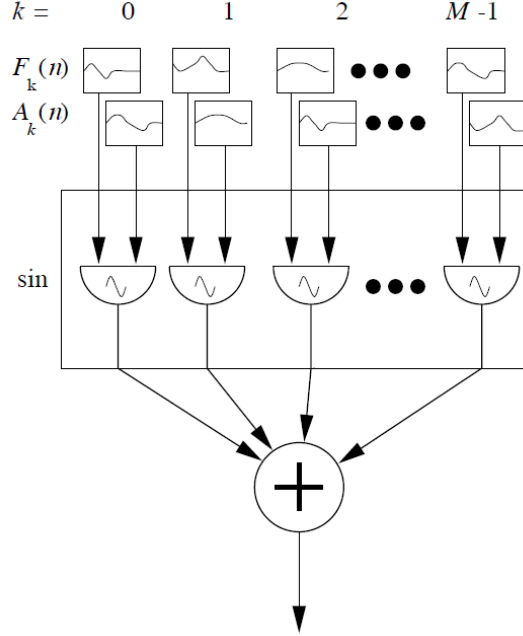
$$X(n, \omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\theta) e^{j\theta n} X(\omega + \theta) d\theta \quad (4.9)$$

Bu durumda $W(\omega)$ impulsa ne kadar yakın olursa $X(n, \omega)$ 'nın o kadar iyi frekans çözünürlüğüne sahip olması beklenir. Ancak frekans çözünürlüğü pencere uzunluğuyla ters orantılıdır. Bu durumda iyi frekans çözünürlüğü uzun analiz pencereleri, iyi zaman çözünürlüğü ise kısa analiz pencereleri gerektirir [6].

4.2 Kısa-Zamanlı Sentez

Ayırık KSFD sentez metodları geliştirmek için kullanılan genel yaklaşım mevcut birçok sentez denkleminde yola çıkmaktır. Daha sonra bir denklemin ayırık hali aday

sentez metoduna temel oluşturur. İki klasik metod Çakışma-Ekleme ve Filtre Bankası Toplamı metodlarıdır.



Şekil 4.5 Zamanla değişen toplamsal sentez.

Ayrık-zamanlı KSFD'nin tersine çevrilebilirliği KSFD'nin Fourier dönüşümü yaklaşımıyla belirlenebilir. Belirli bir n için $X(n, \omega)$ 'nin ters-Fourier dönüşümü alınır ve sonuç $w[0]$ 'a bölünürse ayrık-zamanlı KSFD için bir sentez denklemi olan şu denklem ortaya çıkar:

$$x[n] = \frac{1}{2\pi w[0]} \int_{-\pi}^{\pi} X(n, \omega) e^{j\omega n} d\omega \quad (4.10)$$

Ayrık-zamanlı KSFD'nin aksine ayrık KSFD her zaman tersine çevrilebilir değildir. Örneğin B band genişliğine sahip $w[n]$ için örnekleme aralığı $\frac{2\pi}{N}$, B'den büyük olursa bu durum ortaya çıkar. Ayrıca geçici örnek seyreltme faktörü L'nin analiz penceresi uzunluğu N_w 'den büyük olduğu durumda bazı $x[n]$ örneklerinin ayrık KSFD'nin hiçbir kısa zaman dilimine dahil olmadığı görülür. Bu durumlarda ayrık KSFD tersine

çevrilebilir değildir. Frekans-örnekleme ve zamanda örnek seyreltmede uygun değerler seçilirse ayrık KSFD tersine çevrilebilir hale getirilebilir.

4.2.1 Filtre Bankası Toplamı Metodu

Bu metod en iyi şekilde ayrık KSFD'nin filtre bankası görüşüyle açıklanır. Bu görüşte ayrık KSFD bir dizi filtrenin çıkışı olarak görülür. Filtre bankası toplamı metodunda her bir filtrenin çıkışı bir kompleks eksponansiyelle modüle edilir ve her bir anda toplanır. Bu şekilde orjinal dizinin zaman örneği elde edilir. Geçici örnek seyreltmeyle ilgilenmek için ayrık KSFD üzerinde geçici enterpolasyon filtrelemesi yapılır. Daha sonra bu çıkış üzerinde Filtre Bankası Toplamı metodu uygulanır [6].

Verilen bir ayrık KSFD, $X(n,k)$ için Filtre Bankası Toplamı metodu $y[n]$ dizisini sentezler.

$$y[n] = \frac{1}{Nw[0]} \sum_{k=0}^{N-1} X(n,k) e^{j\frac{2\pi}{N}kn} \quad (4.11)$$

$X(n,k)$ yerine yazılır ve filtre bankası görüşünde elde edilen denklemler göz önünde bulundurulursa denklem şu hale gelir:

$$y[n] = \frac{1}{Nw[0]} x[n] * \sum_{k=0}^{N-1} w[n] e^{j\frac{2\pi}{N}kn} \quad (4.12)$$

$$y[n] = \frac{1}{w[0]} x[n] * w[n] \sum_{r=-\infty}^{\infty} \delta(n - rN) \quad (4.13)$$

Eğer $y[n] = x[n]$ olması istenirse:

$$w[n] \sum_{r=-\infty}^{\infty} \delta(n - rN) = w[0] \delta[n] \quad (4.14)$$

Bu koşul $N_w \leq N$ için geçerlidir. $w[n]$ her N 'inci örnek 0 olacak şekilde seçilirse $N_w > N$ olarak da alınabilir.

Denklem (4.14) için her iki tarafın da Fourier dönüşümü alınır:

$$\sum_{k=0}^{N-1} W\left(\omega - \frac{2\pi}{N}k\right) = Nw[0] \quad (4.15)$$

Bu koşul analiz filtrelerinin frekans cevaplarını tüm band boyunca toplamının bir sabit vermesi gerektiğini gösterir. Bir dizi ve ayrık-zamanlı KSFD'sinden yola çıkılarak elde edilen genelleştirilmiş Filtre Bankası Toplamı metodu şu şekildedir:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\sum_{r=-\infty}^{\infty} f[n, -m]X(r, \omega) \right] e^{j\omega n} d\omega \quad (4.16)$$

$$\sum_{m=-\infty}^{\infty} f[n, -m]w[m] = 1 \quad (4.17)$$

$f[n, m]$ zamanla değişen sentez filtresi olarak adlandırılır. Denklem (4.17)'yi karşılayan her $f[n, m]$, denklem (4.16)'yı geçerli bir sentez denklemi kılar. $f[n, m] = \delta[m]$ gibi bir filtre için de bu denklem elde edilebilir.

Metodun genelleştirilmiş hali (4.16)'daki işlemlerin ayrıklaşmış halini gerçekleştirir:

$$y[n] = \frac{L}{N} \sum_{r=-\infty}^{\infty} \sum_{k=0}^{N-1} f[n, n - rL]X(rL, k)e^{-j\frac{2\pi}{N}kn} \quad (4.18)$$

Pratikteki önemi sebebiyle zamanla değişmeyen filtreler göz önüne alınır. Bu durumda $f[n, n - rL] = f[n - rL]$ alınır ve denklem şu hale dönüşür:

$$y[n] = \frac{L}{N} \sum_{r=-\infty}^{\infty} \sum_{k=0}^{N-1} f[n - rL]X(rL, k)e^{-j\frac{2\pi}{N}kn} \quad (4.19)$$

Bu denklem aşağıdaki koşul sağlandığı sürece geçerlidir. Her n için:

$$L \sum_{r=-\infty}^{\infty} f[n - rL]w[rL - n + pN] = \delta[p] \quad (4.20)$$

4.2.2 Çakışma-Ekleme Metodu

Çakışma-Ekleme metodunda KSFD'nin Fourier dönüşümü görüşü baz alınır. Bu methodda ayrık KSFD'deki her an için ters ayrık Fourier dönüşümü işlemi yapılır ve kısa-zaman dilimleri arasında bir çakışma ve ekleme işlemi yapılır. Metod analiz penceresinin, çakışma ve ekleme operasyonunun analiz penceresini sentezlenmiş diziden etkin bir şekilde elemesini sağlayacak şekilde oluşturulması durumunda işlevseldir [6].

Metod bir dizi ve ayrık-zamanlı KSFD'si arasındaki ilişkiden yola çıkılarak oluşturulur:

$$x[n] = \frac{1}{2\pi w[0]} \int_{-\pi}^{\pi} \sum_{p=-\infty}^{\infty} X(p, \omega) e^{j\omega p} d\omega \quad (4.21)$$

$$W(0) = \sum_{n=-\infty}^{\infty} w[n]$$

$$y[n] = \frac{1}{W(0)} \sum_{p=-\infty}^{\infty} \left[\frac{1}{N} \sum_{k=0}^{N-1} X(p, k) e^{j\frac{2\pi}{N}kn} \right] \quad (4.22)$$

Köşeli parantez içindeki terim bir ters AFD'dir ve her p için $f_p[n] = x[n]w[p - n]$, $N > N_w$ ifadesini verir. Denklem yeniden yazılır ve $y[n] = x[n]$ olduğu düşünülürse:

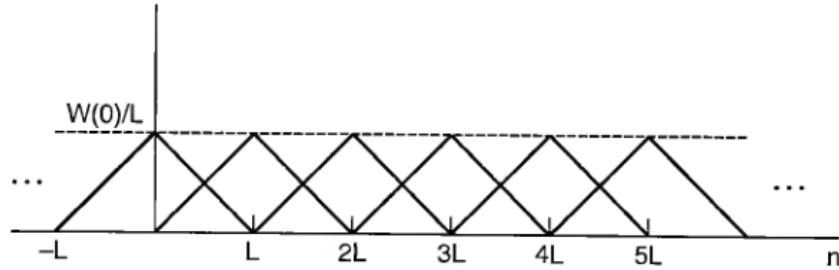
$$y[n] = x[n] \left(\frac{1}{W(0)} \right) \sum_{p=-\infty}^{\infty} w[p - n]$$

$$\sum_{p=-\infty}^{\infty} w[p - n] = W(0)$$

Denklemin her zaman doğru olduğu görülür. Ayırık KSFD L faktörüyle örnek seyreltilmesine maruz kalırsa aşağıdaki denklemlere ulaşılır:

$$\sum_{p=-\infty}^{\infty} w[pL - n] = \frac{W(0)}{L} \quad (4.23)$$

$$x[n] = \frac{L}{W(0)} \sum_{p=-\infty}^{\infty} \left[\frac{1}{N} \sum_{k=0}^{N-1} X(pL, k) e^{j\frac{2\pi}{N}kn} \right] \quad (4.24)$$



Şekil 4.6 Zaman domeninde çakışma-ekleme koşulu.

4.3 Kısa-Süreli Fourier Dönüşümü Magnitudü

Uygulamalarda önemli yer tutan spektogram, KSFD magnitudünün (KSFDM) karesi alınarak elde edilir. İnsan kulağının sesi spektograma benzer bir biçimde algıladığı öne sürülmektedir. Spektogramda ne kadar bilginin kaybolduğu ve sinyalin ne kadarının geri elde edilebileceği tartışma konusudur.

Faz fonksiyonunun KSFD'den çıkarılması ve sadece büyüklük cinsinden bir gösterim, zaman ölçekli modifikasyon ve hız arttırma gibi uygulamalarda kullanışlıdır. KSFD ve KSFDM, analiz açısından birbirine benzer olmasına karşın iki dönüşüm sentez açısından birbirinden çok farklıdır. KSFDM ve kısa-zamanlı öz ilinti fonksiyonu, $r[n, m] = f_n[m] * f_n[-m]$ arasındaki ilişki şu şekildedir:

$$r[n, m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(n, \omega)|^2 e^{j\omega m} d\omega$$

$$|X(n, \omega)|^2 = \sum_{m=-\infty}^{\infty} r[n, m] e^{-j\omega m}$$

Burada m , öz ilinti gecikmesi ve $f_n[m]$, kısa-zaman dilimidir. KSFDM her zaman sinyalin eşsiz bir gösterimi değildir. Örneğin $x[n]$ ve $-x[n]$ aynı KSFDM'ye sahiptir. Sinyalde ve analiz penceresinde bazı koşullar kullanılarak bu sorun aşılabilir. Bu koşullar aşağıdaki gibidir:

1. Analiz penceresi $w[n]$ sıfır değeri içermeyen, N_w uzunluklu bir dizi.
2. $x[n]$ dizisi en fazla $N_w - 2$ arka arkaya gelen sıfır örnekle tek taraflı ve sıfır olmayan ilk değerini taşıyor.

Eğer bir parçanın ilk $N_w - 1$ değerinden son örnek bulunabilirse, bu işlem tekrarlanarak tüm $x[n]$ işareti elde edilebilir. Bunu gerçekleştirmek için $x[n]$ 'in $n - 1$ anına kadar elde edildiği varsayılır. Böylece n anındaki ilk $N_w - 1$ örnek bilinir. Aşağıdaki denklemlerle $x[n]$ elde edilir:

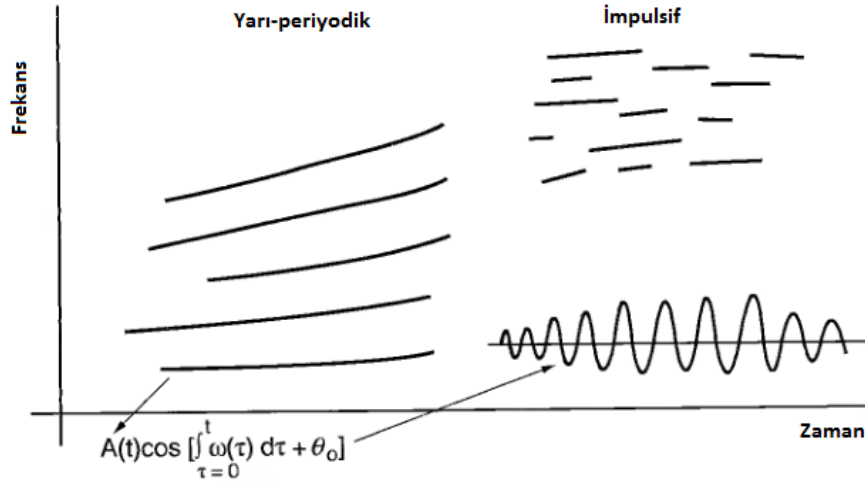
$$r[n, N_w - 1] = (w[0]x[n])(w[N_w - 1]x[n - (N_w - 1)])$$

$$x[n] = \frac{r[n, N_w - 1]}{w[0]w[N_w - 1]x[n - (N_w - 1)]} \quad (4.25)$$

5. SİNÜZOİDAL ANALİZ/SENTEZ MODELİ

Sinüzoidal modelde sentetik işaret; keyfi genlik, frekans ve faz değerlerine sahip sinüzoidal parçalardan oluşur ve model belli sinüs dalgalarının tahminine dayalı bir analiz/sentez sistemi oluşturur. Her bir sinüs dalgası zamanla değişen bir zarfla ve zamanla değişen bir frekans izi integralinin faz eşiyle ifade edilir. Konuşmada sesli harflerde sinüzoidal frekans çizgileri kabaca harmoniktir ve uzun süre devam ederler. Sessiz harfler gibi gürültülü, süreksiz sesler harmonik değildirler ama sinüs dalga toplamı olarak ifade edilebilirler. Fakat tutarlı bir faz yapısı içermezler, frekanslar keyfi değerlidir ve izler kısa sürelerle rastgele ilerler [1].

Göreceğimiz model frekans izlerini frekans eşlemeyle belirlenmesine, faz uyumluluğunu bir kaynak ve vokal filtre faz modeliyle sağlanmasına ve gürültü benzeri parçaların belirleyici ve rastlantısal kısımları toplamsal tahminine dayanır.



Şekil 5.1 Yarı-periyodik ve impulsif sesler için frekans çizgileri [1].

5.1 Sinüzoidal Konuşma Modeli

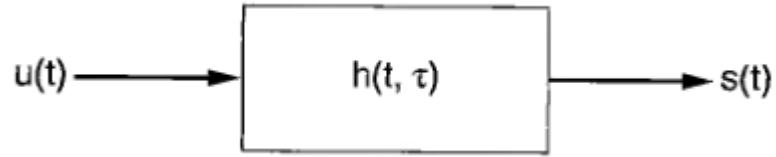
Lineer konuşma üretim modelinde, sürekli-zamanlı konuşma dalgaformu $s(t)$ 'nin lineer zamanla-değişen bir filtreden geçirilen kaynak uyarım işaretinin çıkışı olduğu

düşünülür. $h(t, \tau)$ vokal bölgesini modeller. Konuşma işareti $s(t)$, uyarım işareti $u(t)$ ve faz fonksiyonu $\phi_k(t)$ matematiksel olarak şu şekilde gösterilir:

$$s(t) = \int_0^t h(t, t - \tau) u(\tau) d\tau$$

$$u(t) = \text{Re} \sum_{k=0}^{K(t)} a_k(t) \exp[j\phi_k(t)] \quad (5.1)$$

$$\phi_k(t) = \int_0^t \Omega_k(\sigma) d\sigma + \phi_k$$



Şekil 5.2 $s(t)$ işareti $u(t)$ uyarım işaretinin bir filtreyle sürülmesiyle elde edilir.

Burada $K(t)$, t anındaki sinüzoidal parça sayısıdır. Sinüs dalga modeli keyfidir ve periyodik, impulsif ya da beyaz gürültü biçiminde olmak zorunda değildir. K . Terim için $a_k(t)$ ve $\Omega_k(t)$ zamanla değişen genlik ve frekansı ifade eder. Analizi basitleştirdiği için dalgaın reel kısmı kullanılır. Vokal alanı transfer fonksiyonu faz ve şiddet cinsinden şu şekilde yazılır:

$$H(t, \Omega) = M(t, \Omega) \exp[j\Phi(t, \Omega)] \quad (5.2)$$

Uyarım parametreleri vokal bölgesinin impuls cevabı boyunca sabitise $s(t)$ şu şekle getirilebilir:

$$s(t) = \text{Re} \sum_{k=0}^{K(t)} a_k(t) M(t, \Omega_k(t)) \exp \left[j \left(\int_0^t \Omega_k(\sigma) d\sigma + \Phi(t, \Omega_k(t)) + \phi_k \right) \right] \quad (5.3)$$

Bu denklem kısaca yazılmak istenirse:

$$s(t) = \sum_{k=1}^{K(t)} A_k(t) \exp[j\theta_k(t)] \quad (5.4)$$

$$A_k(t) = a_k(t)M[t, \Omega_k(t)]$$

$$\theta_k(t) = \phi_k(t) + \Phi(t, \Omega_k(t)) = \int_0^t \Omega_k(\sigma) d\sigma + \Phi(t, \Omega_k(t)) + \phi_k \quad (5.5)$$

Notasyonu basitleştirmek için her bir frekans eğrisi boyunca sistem genlik ve frekansları şu şekilde yazılabilir:

$$M_k(t) = M[t, \Omega_k(t)]$$

$$\Phi_k(t) = \Phi(t, \Omega_k(t)) \quad (5.6)$$

Denklem (5.4) konuşmadan bağımsız genel sinüs dalga modelidir ve konuşmadan bağımsız olarak her türlü işarete uygulanabilir. Denklem (5.5) ise konuşmaya bağlıdır [1].

5.2 Sinüs Dalga Parametrelerinin Tahmini

$x[n] = A \cos(\omega_0 n + \phi)$ şeklinde tekil bir ayırık-zamanlı sinüs dalga düşünelim. l 'inci çerçeve için \hat{A}^l ve $\hat{\omega}_0^l$ tahmini genlik ve frekans değerlerini ifade eder. l 'inci çerçevedeki sinüs dalga şu şekilde sentezlenebilir:

$$\hat{x}^l[n] = \hat{A}^l \cos(\hat{\omega}_0^l n), \quad lL \leq n < (l+1)L$$

Bu analiz ve sentez gerçekleştirildiğinde çerçeve sınırlarında bir dalgaformu süreksizliği elde edilir. Bu süreksizlik 100 Hz'lik bir cızırtı yaratır. Faz parametresi için:

$$\hat{\theta}^l[n] = \sum_{m=LL}^n \hat{\omega}_0^l + \hat{\theta}^{l-1} = (n - lL)\hat{\omega}_0^l + \hat{\theta}^{l-1}, \quad lL \leq n < (l+1)L \quad (5.7)$$

Bu durumda l 'inci çerçeve için tahmin:

$$\hat{x}^l = \hat{A}^l \cos(\hat{\theta}^l[n])$$

Bu işlem 0 fazla yapıldığında \emptyset kadar faz sapması olur.

Parametrelerin tahmininin analitik olarak gerçekleştirilmesi zordur. Bu sebeple ideal varsayımlara dayanan bir 'ideal tahmin edici' geliştirilir.

Öncelikle sürekli-zaman uzunluklu pencerelerle analiz edilen bir çerçeveler dizisi olarak düşünülür. l . çerçeve için pencere merkezi t_l anıdır. Parametrelerin pencere süresince yani $I = [t_l - \frac{T}{2}, t_l + \frac{T}{2}]$ aralığında sabit olduğu varsayılır. Bu durumda k. genlik ve frekans fonksiyonları zamandan bağımsız olarak yazılabilir. Faz fonksiyonu ve denklem (5.4)'ten yola çıkarak faza bağlı dalga formu I aralığında şu şekilde yazılır:

$$\theta_k^l = \Omega_k^l(t - t_l) + \theta_k^l \quad (5.8)$$

$$\theta_k^l = \phi_k^l + \Phi(t_l, \Omega_k^l)$$

$$s(t) = \sum_{k=1}^{K^l} A_k^l \exp(j\theta_k^l) \exp[j\Omega_k^l(t - t_l)], \quad t_l - \frac{T}{2} \leq t \leq t_l + \frac{T}{2}$$

Bu durumda önce işaret $t' = t - t_l$ ile orjine kaydırılır, sonra da ayrık zamana dönüştürülür:

$$s'(t') = \sum_{k=1}^{K^l} A_k^l \exp(j\theta_k^l) \exp(j\Omega_k^l t'), \quad -\frac{T}{2} \leq t' \leq \frac{T}{2}$$

$$s[n] = \sum_{k=1}^{K^l} \gamma_k^l \exp(j\omega_k^l n), \quad -\frac{N_w-1}{2} \leq n \leq \frac{N_w-1}{2} \quad (5.9)$$

Denklem (5.9) parametrelerin bulunması için kullanılacak ayırık-zamanlı durağan analiz modelidir. $\gamma_k^l = A_k^l \exp(j\theta_k^l)$ k. parçanın k. kompleks genliğidir.

Bu noktada problem sentetik dalga formu ve $y[n]$ ile gösterilen orjinal dalga formunun uyuşmasıdır. Bu durumda da kaynak işaretin türü önemlidir.

5.2.1 Yarı-Periyodik İşaretler

Kaynak işaret ve sentetik işaret arasındaki uyum için faydalı bir kriter karesel ortalama hatadır.

$$\begin{aligned} \epsilon^l &= \sum_{n=-(N_w-1)/2}^{(N_w-1)/2} |y[n] - s[n]|^2 \\ \epsilon^l &= \sum_{n=-(N_w-1)/2}^{(N_w-1)/2} |y[n]|^2 - 2\operatorname{Re} \sum_{k=1}^{K^l} (\gamma_k^l)^* \sum_{n=-(N_w-1)/2}^{(N_w-1)/2} y[n] \exp(-jn\omega_k^l) \\ &\quad + N_w \sum_{k=1}^{K^l} \sum_{m=1}^{K^l} \gamma_k^l (\gamma_m^l)^* \operatorname{sinc}(\omega_k^l - \omega_m^l) \end{aligned} \quad (5.10)$$

$$\operatorname{sinc}(x) = \frac{\sin(N_w \frac{x}{2})}{N_w \sin(\frac{x}{2})}$$

Bu durumda eşitliği minimize eden sinüzoidal dalgalar bulunmalıdır. Ancak ϵ^l 'deki parametreler lineer olmadığı için bu zordur. Giriş işaretinin ideal olduğu, yani rastgele olmadığı ve harmonik olduğu varsayımıyla tahmin edici oluşturulur. Bu

tahmin edici daha sonra ideal olmayan durumlarda kullanılır. İdeal durumda denklem şu şekilde yazılabilir:

$$s[n] = \sum_{k=1}^{K^l} \gamma_k^l \exp(jnk\omega_0^l)$$

Burada $\omega_0^l = \frac{2\pi}{Pl}$ ve P^l analiz penceresi süresince sabit olduğu varsayılan perde periyodudur. İdeal koşullar altında $\omega_k^l = k\omega_0^l$ iken sinc fonksiyonu şu hali alır:

$$\text{sinc}(\omega_k^l - \omega_m^l) = \text{sinc}[(k - m)\omega_0^l] = \delta[k - m]$$

$l = m$ durumunda $\delta[k - m] = 1$, $k \neq m$ durumunda $\delta[k - m] = 0$ 'dır. Karesel ortalama hata şu duruma gelir:

$$\epsilon^l = \sum_{n=-(N_w-1)/2}^{(N_w-1)/2} |y[n]|^2 - 2N_w \text{Re} \sum_{k=1}^{K^l} (\gamma_k^l)^* Y(\omega_k^l) + N_w \sum_{k=1}^{K^l} |\gamma_k^l|^2 \quad (5.11)$$

$$Y(\omega) = \frac{1}{N_w} \sum_{n=-(N_w-1)/2}^{(N_w-1)/2} y[n] \exp(-jn\omega) \quad (5.12)$$

Denklem (5.11) tamamlanır ve $\gamma_k^l = Y(k\omega_0^l)$ alınırsa karesel ortalama hata:

$$\epsilon^l = \sum_{n=-(N_w-1)/2}^{(N_w-1)/2} |y[n]|^2 - N_w \sum_{k=1}^{K^l} |Y(k\omega_0^l)|^2 \quad (5.13)$$

Hata, konuşma bandındaki tüm harmonik frekansların seçilmesiyle minimize edilir.

KSFD'nin $|Y(\omega)|^2$ olduğu düşünülürse, altta yatan sinüs dalgaların frekansları $|Y(\omega)|^2$ 'nin tepe noktalarına karşılık düşecektir. Sinüs dalgaları sabit frekanslı olursa ve harmonik olmazsa bu durumda bir zaman dilimi için KSFD:

$$Y(\omega) = \sum_{k=1}^{K^l} \gamma_k^l \text{sinc}(\omega_k^l - \omega)$$

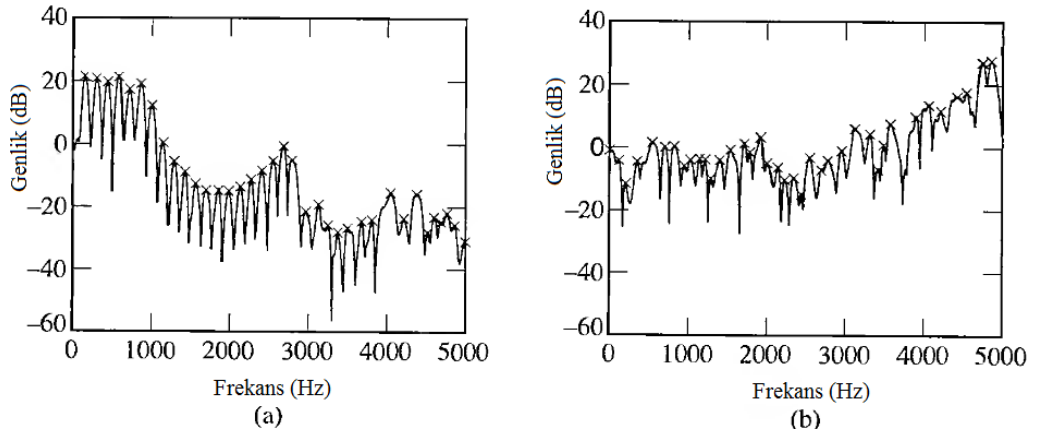
Analiz penceresi yeterince geniş olursa:

$$|\omega_m^l - \omega_k^l| \geq \frac{4\pi}{N_w}, \quad m \neq k$$

Bu durumda Fourier dönüşümü magnitudü:

$$|Y(\omega)|^2 \approx \sum_{k=1}^{K^l} |\gamma_k^l|^2 \text{sinc}^2(\omega_k^l - \omega).$$

Spektral magnitud tepe noktaları sinüs dalga frekanslarına, bu frekanslardaki KSFD örnekleri kompleks genliklere karşılık gelir.



Şekil 5.3 (a) Yarı-periyodik; (b) gürültülü ve impulsif konuşma işaretinin tipik KSFD magnitudü.

5.2.2 Gürültülü ve İmpulsif İşaretler

Eğer güç spektrumunun ardışık frekanslarla yavaşça değişmesini sağlayacak yeterince yakın frekanslar seçilirse sessiz harfler gibi gürültülü ve impulsif kaynaklı işaretler

için de yarı-periyodik işaretlerde kullanılan analiz yöntemi kullanılabilir. Eğer pencere genişliği en azından 20 ms olursa, örneklerin birbirinden 100 Hz'den uzak kalmaması sağlanır ve dolayısıyla bu koşul da sağlanmış olur.

5.2.3 Analiz Sistemi

Analizde KSFD'nin dörtgen bir pencere kullanılarak hesaplandığı varsayılır. Bu durumda $y[n]$ 'in KSFD'si:

$$Y(lL, \omega) = \sum_{m=-\infty}^{\infty} w[lL - m] y[m] \exp(-jm\omega)$$

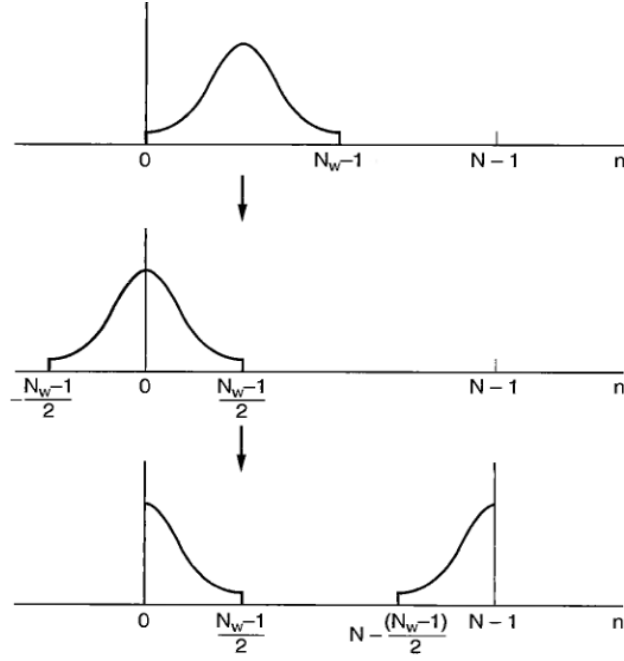
Dörtgen pencerenin zayıf yan kulak performansı sebebiyle burada Hamming penceresi kullanılır. Pencere şu şekilde hesaplanır ve normalize edilir:

$$w[n] = 0.54 + 0.46 \cos\left(2\pi \frac{n}{N_w}\right), \quad 0 \leq n < N_w$$

$$\sum_{n=-\infty}^{\infty} w[n] = 1$$

Pencerelemeden sonra işaretin ayrık Fourier dönüşümü alınır. Yüksek kaliteli analiz ve sentez için 1024 noktalı AFD yeterlidir. Pencere genellikle $0 \leq n < N_w$ aralığındadır. Lineer faz hatasını azaltmak için pencerelenmiş işaret dairesel şekilde kaydırılır. Bu durumda pencere $-(N_w - 1)/2 \leq n \leq (N_w - 1)/2$ aralığında değerler alır. Kaydırmadan sonra ideal tahmin edici sinüs dalga frekanslarını ($\hat{\omega}_k^l$) elde eder. Bu değerler aracılığıyla genlik ve faz parametreleri de bulunur.

$$\hat{A}_k^l = |Y(lL, \hat{\omega}_k^l)|, \quad \hat{\theta}_k^l = \angle Y(lL, \hat{\omega}_k^l).$$

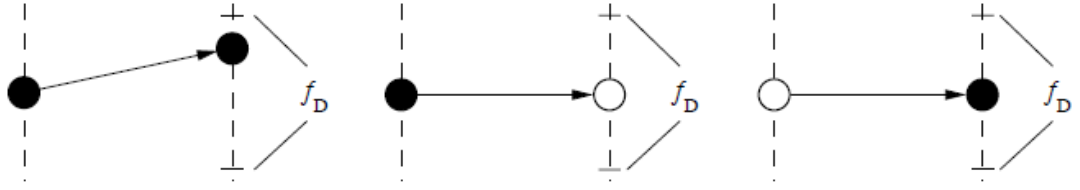


Şekil 5.4 Lineer faz hatasını azaltmak için pencerelenmiş parçanın dairesel olarak kaydırılması.

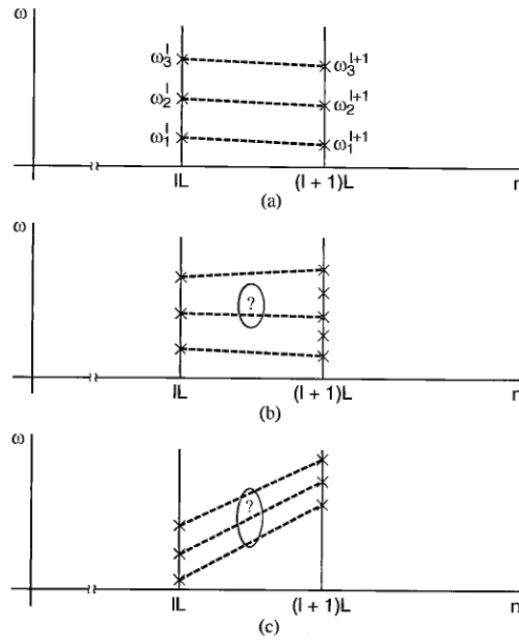
5.2.4 Çerçveden Çerçeveye Tepe Uydurma

Genlik, frekans ve faz parametreleri bulunduktan sonraki problem; bu parametrelerin uygun çerçevelere denk düşmesini sağlamaktır. Eğer çerçveden çerçeveye tepe sayısı değişmiyorsa bir çerçeve ve ardından gelen üzerindeki parametrelerin tahmini, tepelerin frekansa göre düzenlenmesini gerektirir. Ama pratikte perde değişikçe tepelerin yeri de değişir. Hızlı değişen işaretlerde tepelerin sayısı da değişir [7].

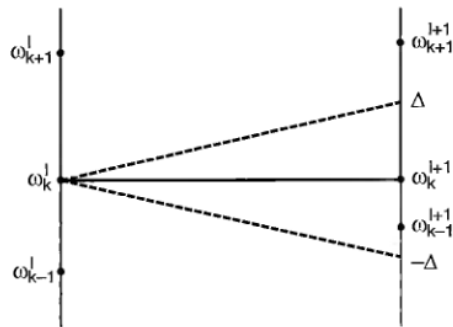
Spektral tepelerin çerçveden çerçeveye yerlerinin ve sayılarının hızlı değişimi sebebiyle sinüzoidal parçaların doğumu ve ölümü konseptlerinden söz edilir. Bu doğum ve ölüm işlemiyle beraber spektral tepelerin eşleştirilmesi genelde zor bir problemdir. Uygulanabilecek bir yöntem; birbirine en yakın komşular olan frekanslar için sinüs dalga izlerini belirlemektir. Bu izler mevcut çerçevede belli bir eşleşme frekans aralığına, $[-\Delta, \Delta]$, denk düşmelidir. Bu eşleştirme işlemi; takip eden frekanslar uygun aralığa düşmediğinde ortaya çıkan doğum ve ölüm gibi olaylara izin vererek dinamik hale getirilir.



Şekil 5.5 sinüzoidal parçaların doğım ve ölümü.



Şekli 5.6 Frekans eşleştirmedeki problemler (a) yavaşça değişen perde; (b) hızlı değişen perde; (c) hızlı yarı-periyodik/impulsif karakter değişimleri



Şekil 5.7 Frekans eşleşme aralığı koşulu.

5.3 Sentez

Her bir çerçeve için genlik-, frekans ve faz değerleri bulunduğundan sonra l 'inci çerçevedeki sentetik işaret şu denklemle elde edilir:

$$\hat{s}[n] = \sum_{k=1}^{K^l} A_k^l \cos(n\omega_k^l + \theta_k^l), \quad n = 0, 1, 2, \dots, L - 1$$

Burada L sentez çerçevesi uzunluğudur.

Parametreler zamanla değiştiği için bu doğrudan yaklaşım çerçeve sınırlarında süreksizliklere sebep olur. Bu da sentetik işaretin kalitesini ciddi oranda düşürür. Bu sebeple bir çerçevedeki parametreler ve bir sonraki çerçevede elde edilen parametreler arasında interpolasyon işlemi gerçekleştirilir [7].

5.3.1 Kübik Faz İnterpolasyonu

Keyfi bir çerçevede (l) belirlenen parametreler bir sonraki çerçevedeki ($l + 1$) karşılık gelen parametrelerle ilişkilendirilmiştir. k . frekans izi için söz konusu parametreler $(A_k^l, \omega_k^l, \theta_k^l)$ ve $(A_k^{l+1}, \omega_k^{l+1}, \theta_k^{l+1})$ ile ifade edilir. Bu durumda genlik interpolasyonu:

$$A_k^l[n] = A_k^l + (A_k^{l+1} - A_k^l) \left(\frac{n}{L}\right), \quad n = 0, 1, \dots, L - 1 \quad (5.14)$$

Faz interpolasyonu için ise bu basit yaklaşım söz konusu değildir. Çünkü faz 2π 'nin katları olarak ilerler ve faz ölçümlerinde 2π 'lik süreksizlikler olur. Bu durumda öncelikle faz, zaman değişkeni t 'nin sürekli bir fonksiyonu olarak yazılır. Burada $t = 0$, l 'inci çerçeve merkezi; $t = T$, $l + 1$ 'inci çerçeve merkezidir.

$$\theta(t) = \zeta + \gamma t + \alpha t^2 + \beta t^3 \quad (5.15)$$

Denklemdaki bilinmeyenlerin elde edilmesi için dört koşula ihtiyaç vardır. Vokal bölgesinin yavaşça değiştiği varsayımı altında faz fonksiyonu şu şekilde yazılabilir:

$$\theta(t) = \int_0^t \Omega(\sigma) d\sigma + \phi + \Phi[t, \Omega(t)]$$

Yavaşça değişen $\Phi[t, \Omega(t)]$ varsayımı altında:

$$\dot{\theta}(t) = \Omega(t) + 0 + \dot{\Phi}[t, \Omega(t)] \approx \Omega(t)$$

Bu durumda l ve $l + 1$ 'inci çerçeveler için:

$$\dot{\theta}^l \approx \Omega^l$$

$$\dot{\theta}^{l+1} \approx \Omega^{l+1}$$

Bu durumda dört koşul elde edilirken $t = 0$ 'da:

$$\theta(0) = \zeta = \theta^l$$

$$\dot{\theta}(0) = \gamma = \Omega^l$$

$t = T$ 'de:

$$\theta(T) = \theta^l + \Omega^l T + \alpha T^2 + \beta T^3 = \theta^{l+1} + 2\pi M$$

$$\dot{\theta}(T) = \Omega^l + 2\alpha T + 3\beta T^2 = \Omega^{k+1} \quad (5.16)$$

M burada bilinmeyen bir değerdir. α ve β , M 'e bağlı ifade edilir. Bu durumda M bulunmalıdır. $\ddot{\theta}(t) = 0$ olduğu düşünülürse:

$$f(M) = \int_0^T [\ddot{\theta}(t; M)]^2 dt$$

Bu durumda da $f(x)$ sürekli değişen x göz önünde bulundurularak minimize edilir ve M ; x 'e en yakın tam sayı değerini alır.

$$x^* = \frac{1}{2\pi} \left[(\theta^l + \Omega^l T - \theta^{l+1}) + (\Omega^{l+1} - \Omega^l) \frac{T}{2} \right]$$

Bulunan M^* değeri yerine yazılırsa:

$$\theta(t) = \theta^l + \Omega^l t + \alpha(M^*)t^2 + \beta(M^*)t^3 \quad (5.17)$$

Gözlenen frekans izinin bir noktada doğduğu varsayılır. Bu durumda faz interpolasyon koşullarının sağlandığından emin olmak için başlangıç fazı $\theta^l = \theta^{l+1} - \Omega^{l+1}T$ olarak alınır. T ; ayrık zamanda $l + 1$ 'den l 'inci çerçeveye geçen örnek sayısıdır. Bu durumda $M^* = 0$ 'dır. Bir izin doğumundaki koşullar:

$$A^l = 0$$

$$\Omega^l = \Omega^{l+1}$$

$$\theta^l = \theta^{l+1} - \Omega^{l+1}T$$

Bir frekans çizgisi ortadan kalkarken de benzer bir prosedür uygulanır.

Bu işlemler sonucunda her bir frekans izi ani, açılmış bir fazla eşleşir. Bu faz hem hızlı hem de yavaş faz değişimleri için geçerli bir sonuç verir. Ayrık zamanda nihai sentetik dalga formu l 'inci çerçevede şu şekildedir:

$$\hat{s}[n] = \sum_{k=1}^{K_l} A_k^l [n] \cos(\theta_k^l [n])$$

5.3.2 Çakışma-Ekleme İnterpolasyonu

Kübik faz interpolasyonu hesaplamalarının gerçekleştirilmesi zor olduğu için alternatif olarak bu metodla tatmin edici sonuçlar alınmıştır. l 'inci çerçevede tahmin edilen parametrelerle dalga formu $0 \leq n < L$ aralığında şu şekilde üretilebilir:

$$\hat{s}^l[n] = \sum_{k=1}^{K^l} A_k^l \cos(n\omega_k^l + \theta_k^l) \quad (5.18)$$

Aynı aralıkta $l + 1$ 'inci çerçevede elde edilen parametreler ile işaret şu şekildedir:

$$\hat{s}^{l+1}[n] = \sum_{k=1}^{K^{l+1}} A_k^{l+1} \cos[(n - L)\omega_k^{l+1} + \theta_k^{l+1}] \quad (5.19)$$

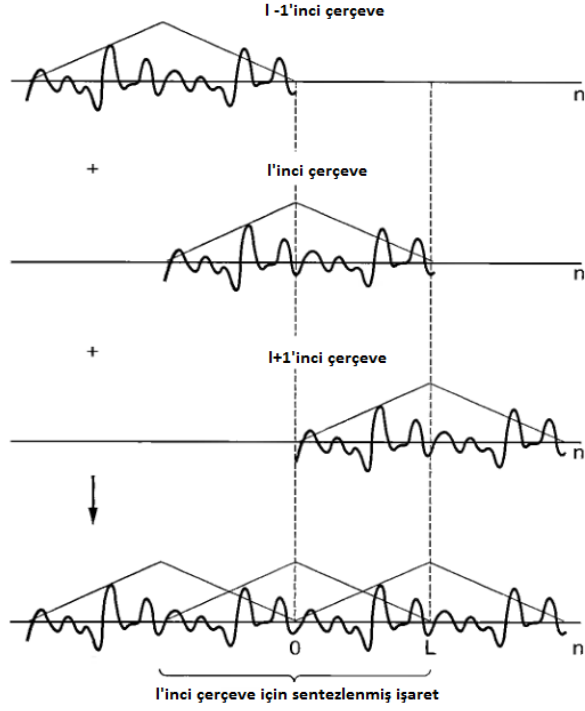
Eğer çerçeve aralığı L yeterince kısa olursa, $s^l[n]$ ve $s^{l+1}[n]$ benzer olur ve interpolate edilmiş sentetik işaret, yukarıdaki dalga formları üçgen bir pencereyle ağırlıklandırılarak üretilir. l 'inci çerçeve çakıştığında üçgenin aşağı eğimli kısmında kalan sinüs dalga üretilir ve $l + 1$ 'inci çerçevedeki üçgenin yukarı eğimli kısmında kalan işarete eklenir. Bu işlem şu şekilde ifade edilir:

$$\hat{s}[n] = \left(1 - \frac{n}{L}\right) \hat{s}^l[n] + \left(\frac{n}{L}\right) \hat{s}^{l+1}[n], \quad n = 0, 1, \dots, L - 1.$$

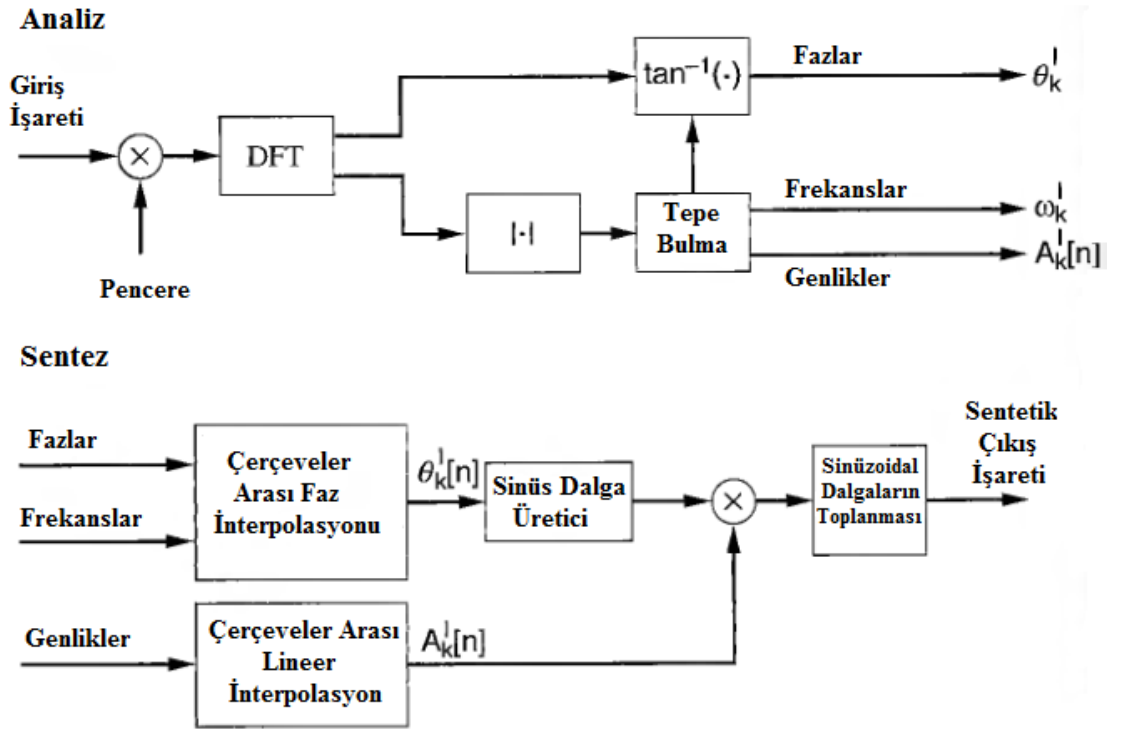
Hızlı Fourier dönüşümü kuantizasyonu yeterince küçük (~ 30 Hz) ve çerçeve oranı yeterince yüksek (~ 100 Hz) olursa bu işlem yüksek kaliteli çıkış verir.

Eşleştirme algoritmasında frekans eşleştirme penceresi sıfır alınırsa bu interpolasyon işlemi gerçekleşmiş olur. Çünkü bu durumda her iki çerçevede sinüs dalgaları doğar ve ölür ve bu ikili çerçeve gruplarında frekans sabit kalır.

Ekleme-çakışma metodu modifiye edilmiş bir KSFD'den sinyal sentezi olarak da yorumlanabilir. Bu durumda en küçük kareler yaklaşımı kullanılabilir.



Şekil 5.8 Çakışma ve ekleme.



Şekil 5.9 Sinüzoidal analiz/sentez sistemi blok diyagramı.

6. UYGULAMA

Metodun uygulaması iki kısımdan oluşmaktadır. İlk kısımda temel olarak kaydedilmiş bir ses dosyası analiz edilip gerekli genlik, frekans ve faz parametreleri elde edilmekte, ikinci kısımda ise bu parametrelerin kullanılmasıyla sentetik ses dosyası oluşturulmaktadır.

Uygulama analiz kısmındaki işlem yükü, programın kullanım kolaylığı, içerdiği hazır algoritmalar ve sağladığı interaktif ortam sebebiyle MATLAB ortamında gerçekleştirilmektedir.

Analizde kullanılacak kaydedilmiş ses dosyasının 44100 Hz hızında, 16 bit örnekleme kaliteli bir sentez için yeterlidir. Ses kaydının gürültüsüz ve yankısız bir ortamda yapılması da önemlidir. Analiz penceresi seçimi yapılırken pencere uzunluğu için optimum değer bulunmalıdır. İyi frekans çözünürlüğü uzun pencereleri, iyi zaman çözünürlüğü de kısa pencereleri gerektirir. Bu sebeple eğer hızlı değişen bileşenler içeren bir ses analiz edilecekse pencere boyunun çok uzun tutulmaması gerekir. Zamanla daha yavaş değişen seslerde ise pencere boyu uzun tutulmalıdır. Ses hızlı değişiyor ve birden fazla kaynak içeriyorsa pencere uzunluğunun da işaretteki değişime göre adaptif alınması gereksinimi ortaya çıkabilir. Bu çalışmada piyano sesleri analiz edilmiştir. Hızlı sönümlenen ve kuvvetli harmonik sesler için pencere boyu perde periyodunun iki buçuk katında alınmış [8] ve tüm analiz boyunca sabit tutulmuştur. Bu durumda 261.626 Hz perde frekansına sahip C4 notası için pencere uzunluğu hesabı şu şekildedir:

$$44100 \times 2.5 \div 261.626 \cong 421$$

Kullanılacak pencere tipi Hann penceresi olabilir. Daha iyi yan kulak performansı gerekiyorsa Hamming ve Blackman-Harris pencerelerinden biri kullanılabilir.

Pencerelenen işaret daha sonra AFD alınır. Kaliteli sentez için 1024 noktalı Fourier dönüşümü yeterlidir. Genlik ve faz spektrumu elde edildikten sonra bir tepe bulma

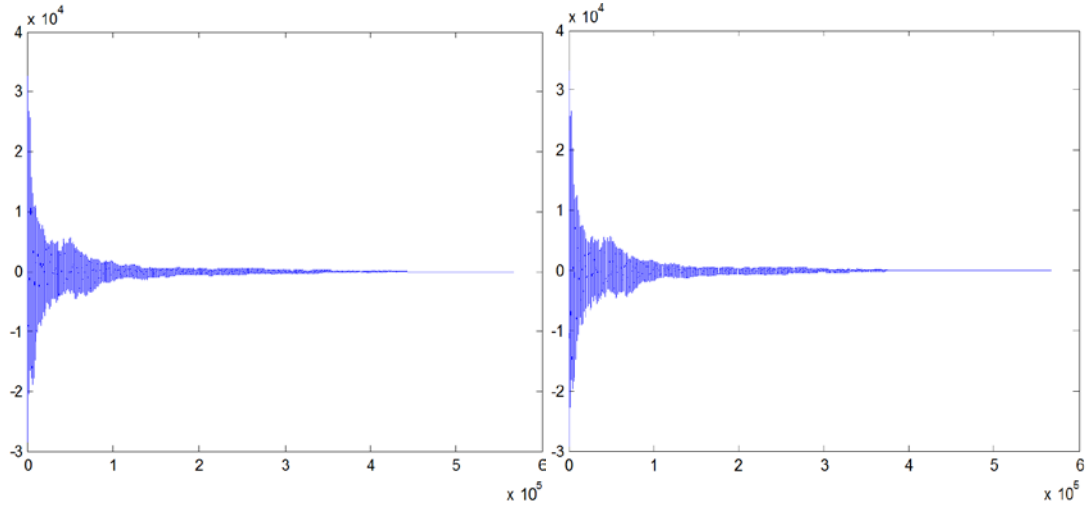
algoritmasıyla yerel tepe noktaları bulunur. Daha sonra bu veriler oluşturulan bir veri dosyasında her bir çerçeve için saklanır.

Uygulamanın ikinci kısmında sentez denklemi kullanılarak analizde elde edilen veriler uygun çerçevelere yerleştirilir. Çerçeveler arası devamlılığın sağlanabilmesi için Bölüm 5’te bahsedilen kübik faz interpolasyonu ve çakışma-ekleme interpolasyonu metodlarından biri kullanılır. Elde edilen sentetik işaret bir Microsoft Windows WAV (Waveform Audio File Format) dosyasına yazılır. Sentezde amaç analiz edilen sese mümkün olduğunca yakın bir ses elde etmek olduğu için oluşabilecek hataların mümkün olduğunca giderilmesi gerekmektedir. Kaydedilen ses dinlenerek orjinal sese yakınlığına karar verilebilir.

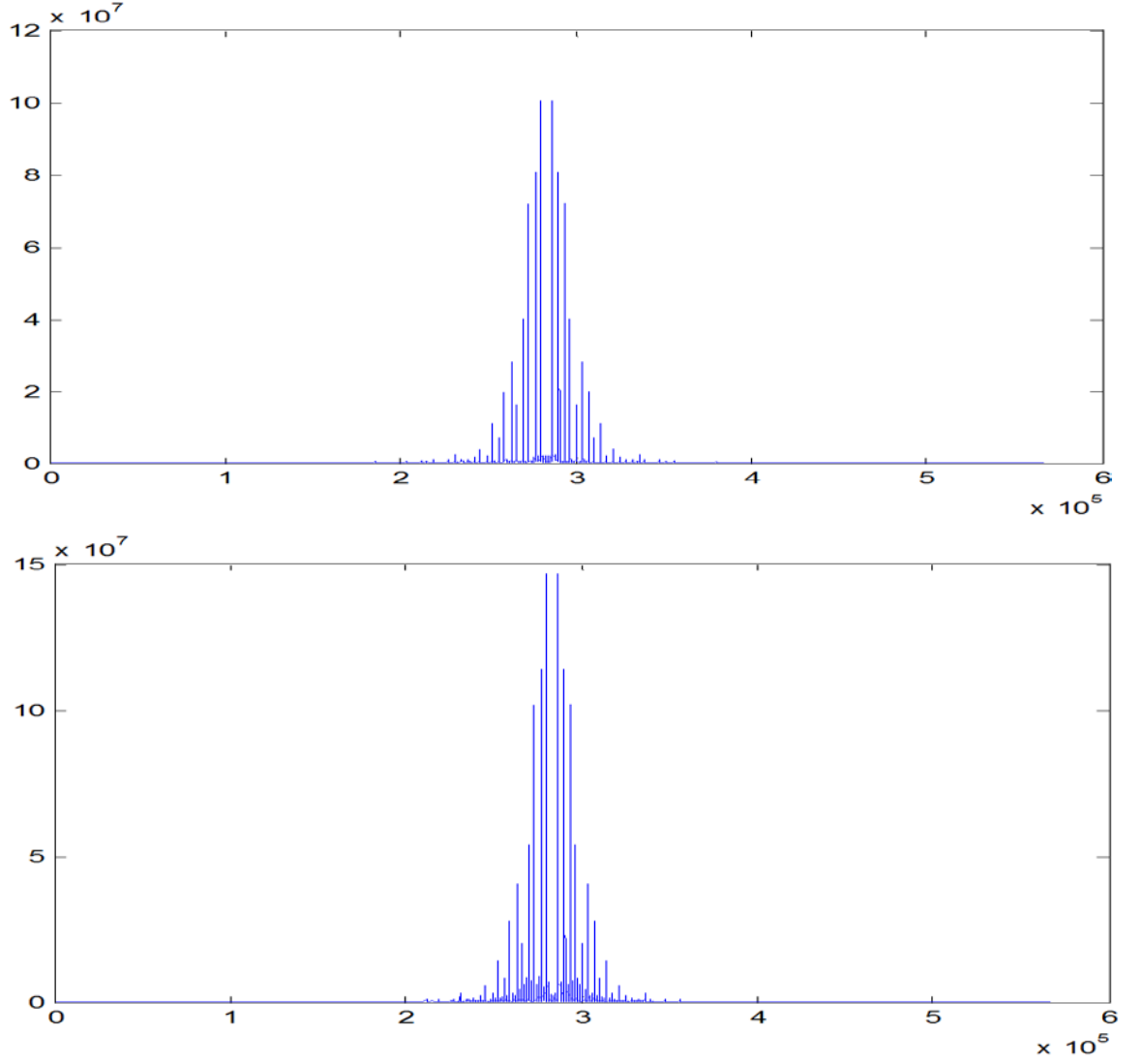
7. SONUÇ

Bu çalışmada ses sentezi konusu, çeşitli ses sentezleme metodları ve sinüzoidal modelin bu metodlar arasındaki konumu incelenmiş, sinüzoidal modelin teorik altyapısı ortaya konmuş ve uygulamada nasıl sonuçlar verdiği araştırılmıştır.

Uygulamada piyano sesi üzerinde çalışılmıştır. Farklı oktavlardaki notalar denenerek sesteki karakter değişimlerinin yarattığı sonuçlar, pencere tipinin ve uzunluğunun sonuç üzerindeki etkileri araştırılmıştır. Pencere uzunluğu kısa tutulduğunda işlem süresinin uzadığı ve sentezde istenen sonucun elde edilemediği görülmüştür. Pencere uzunluğunun gereğinden uzun tutulduğu denemelerde bir noktaya kadar başarılı sentez gerçekleşmiştir. Aynı şekilde daha pes ve daha tiz notalarla yapılan denemelerde de optimum pencere uzunluğu kullanımının gerekliliği görülmüştür. Öte yandan sentez kalitesini artırmak için gürültü etkisini azaltmanın gerekliliği de anlaşılmıştır. Keman ve çello gibi yaylı enstrumanlarla yapılan denemelerde gürültü etkisinin arttığı gözlenmiştir. Şekil 7.1 ve 7.2 piyano sesi içinelde edilen sonuçlara ait grafikler mevcuttur.



Şekil 7.1 Gerçek ve sentetik piyano seslerinin zaman domeninde karşılaştırılması.



Şekil 7.2 Gerçek ve sentetik piyano seslerinin frekans domeninde karşılaştırılması.

Çalışmanın devamında gidilebilecek noktalar olarak birden fazla kaynak içeren daha karmaşık sesler üzerinde çalışmak için önemli bir başlangıç noktası oluşturulmuştur. Buna ek olarak sinüzoidal modelden yola çıkarak oluşturulmuş, modelin deterministik ve rastlantısal uzantısı olan spektral model gibi daha gelişmiş metodların yanısıra fiziksel modellerle karşılaştırma yaparak farklı enstruman sesleri için en uygun sentez metodlarına ulaşmak mümkündür. Sinüs dalga yaklaşımıyla sentezlemeye ek olarak tek bir kayıttan farklı sinyallerin ayrımı, çoklu perde tahmini, perde modifikasyonu, konuşma kodlaması ve buna benzer farklı uygulamalar da söz konusudur.

KAYNAKÇA

- [1] Quatieri, T.F., (2002), Discrete-Time Speech Signal Processing Principles And Practice, Prentice-Hall, New Jersey.
- [2] M. V. Mathews,(1969), "The Technology of Computer Music", Cambridge, MA: MIT Press, 1969.
- [3] Smith, J.O., (1991), "Viewpoints on the History of Digital Synthesis", Proceedings of the ICMC, Oct. 1991, Helsinki.
- [4] Tolonen, T., Valimaki, V. ve Karjalainen, M., (1998), "Evaluation of Modern Sound Synthesis Methods", Helsinki University of Technology.
- [5] Serra, X., (1997), "Musical Sound Modelling with Sinusoids plus Noise", Musical Signal Processing, C. Roads (Derl.) Swets & Zeitlinger Publishers.
- [6] Nawab, S.H. ve Quatieri, T.F., (1988), "Short-Time Fourier Transform", Advanced Topics in Signal Processing (J. S. Lim and A. V. Oppenheim, eds.), ch. 6, Prentice-Hall.
- [7]McAulay, R.J. ve Quatieri, T.F., (1988), "Speech Processing Based on a Sinusoidal Model", The Lincoln Laboratory Journal, Volume 1, Number 2.
- [8] Bolat, B., (1998), "Sinüzoidal Model Yardımıyla Müzikal Ses Sentezi", Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi.

ÖZGEÇMİŞ

Ad Soyad : Cemal Kozanoğlu

Doğum Tarihi : 30.07.1988

Doğum yeri : İstanbul

Lise : 2002 - 2006 Silivri Lisesi

Staj Yaptığı Yerler : DHMİ, İstanbul (6 hafta),

KOÇSİSTEM, İstanbul (6 hafta)