





**UYARLANMIŞ ITAKURA-SAITO UZAKLIĞI İLE  
SES SİNYALLERİNİN BENZERLİK ANALİZİ**

**Atacan ERDİŞ**

**YÜKSEK LİSANS TEZİ  
İSTATİSTİK ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**KASIM 2019**



Atacan ERDİŐ tarafından hazırlanan "UYARLANMIŐ ITAKURA-SAITO UZAKLIĐI İLE SES SİNYALLERİNİN BENZERLİK ANALİZİ" adlı tez çalışması aŐađıdaki jüri tarafından OY BİRLİĐİ ile Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Ana Bilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

**Danışman:** Prof. Dr. M. Akif BAKIR

İstatistik Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum. ....

**Başkan:** Doç. Dr. Ceylan TALU YOZGATLIGİL

İstatistik Ana Bilim Dalı, Orta Dođu Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum. ....

**Üye :** Doç. Dr. Hülya OLMUŐ

İstatistik Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum. ....

Tez Savunma Tarihi : 25.11.2019

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiđini onaylıyorum.

.....  
Prof. Dr. Sena YAŐYERLİ

Fen Bilimleri Enstitüsü Müdürü



## ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Atacan ERDİŞ

25.11.2019





UYARLANMIŞ İTAKURA-SAITO UZAKLIĞI İLE  
SES SİNYALLERİNİN BENZERLİK ANALİZİ  
(Yüksek Lisans Tezi)

Atacan ERDİŞ

GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
Kasım 2019

ÖZET

İntihal (Plagiarism), kısaca bir eseri izin almadan ve/veya kaynak göstermeden alıntılama durumu, etik olmamasının yanı sıra hukuken suç teşkil etmektedir. Günümüzde bilimsel ve sanatsal eserlerde intihal suçunun işlenmesine ilişkin telif yasaları ve bu yasalar gereği uygulanan hükümler mevcuttur. Bilimsel intihallerin aksine sanat alanında yapılan intihallerin belirlenmesi, sezgisel olması sebebiyle tartışmaya açık ve bir hayli zordur. Özellikle müzik eserlerindeki intihallerin tespiti, diğer sanat dallarına nazaran daha çok değişkenin dikkate alınması gerektiğinden kesinlikle kolay olmamaktadır. Müzik eserlerindeki intihalin tespiti, alanında uzman müzisyenlerden oluşturulan kurullar tarafından titizlikle gerçekleştirilmektedir. Her ne kadar temeli müziğin bilimine dayansa da tespit argümanı nicellikten uzaktır. Bu çalışmada amaç ses sinyallerinin benzerliğine ilişkin nicel bir karar argümanı geliştirmektir. Nitekim ses sinyallerinin fiziksel olarak dalga formunda olması, varoluşundaki rastgelelik ve zaman-bağımlı kronolojik sıralı yapısı, bu sinyallere fizik, matematik ve istatistik bilimlerinin perspektiflerinden bakılabileceğini işaret etmektedir. Bu çalışmada durağan zaman serileri arasındaki benzerliğin tespitinde kullanılan Fourier dönüşümü tabanlı Itakura-Saito uzaklığı, ses sinyalleri gibi durağan olmayan sinyallerin benzerliğinin tespitinde kullanılabilmesi adına uyarlanarak, ses sinyallerinin benzerliğine ilişkin nicel bir argüman geliştirilmiştir. Önerilen yöntemin başarısı, hipotetik ve gerçek ses sinyallerinde sınanmış olup, anlamlı sonuçlar elde edilmiştir. Ayrıca, yöntemin başarısının testini derinleştirmek adına, durağan olmayan sinyallerin alt karakteristik yapılarının belirlenmesinde kullanılan Ampirik Mod Ayırımı (EMD) yöntemi ile elde edilen İçsel Mod Fonksiyonları (IMF) arasındaki benzerlik yapıları irdelenmiş ve elde edilen sonuçların önceki bulguları destekler nitelikte olduğu görülmüştür.

Bilim Kodu : 20506  
Anahtar Kelimeler : Ses Sinyalleri, Ses Sinyallerinin Benzerliği, Kısa-Zamanlı Fourier Dönüşümü, Itakura-Saito Uzaklığı, Uyarlanmış Itakura-Saito Uzaklığı, Zaman Serileri, Pencere Fonksiyonları, Ampirik Mod Ayırımı, Grupsal Ampirik Mod Ayırımı  
Sayfa Adedi : 61  
Danışman : Prof. Dr. M. Akif BAKIR



SIMILARITY ANALYSIS OF SOUND SIGNALS WITH MODIFIED  
ITAKURA-SAITO DISTANCE

(M.Sc Thesis)

Atacan ERDİŞ

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

November 2019

ABSTRACT

Plagiarism, in short, is the situation where a production is dictated without permission and/or giving any reference. This is a type of crime not only from legal aspect but also from ethical side. There exists Copyright laws and provisions applied according as these laws for committing plagiarism in both scientific and art products. In contrast to scientific plagiarism, the detection of those in art products is extremely difficult and is done intuitively which makes it to be open to discussion. Especially, regarding other branches of art, detection of plagiarism in musical products can not easily be done due to the requirement of considering multiple variables. Plagiarism detection in music products are performed meticulously by a committee constituted with expert musicians. Although its basis depends on science of music, the detection argument is far from being quantitative. In this study, the basic purpose is to develop a quantitative argument regarding to the similarity between sound signals. In reality, sound signals are physically in wave form. Randomness in their existence and its time-dependent chronological sequential structure allow these signals to be observed from the perspectives of Physical, Mathematical and Statistical science. In this study, Fourier transformation based Itakura-Saito distance used for the detection of similarity between stationary time series is modified for the usage in detection of similarity between non-stationary signals like sound signals. Success of the suggested method is tested on real and hypothetical sound signals, and then meaningful results are obtained. Furthermore, in order to deepen the test of the success of method, the similarity structures between Intrinsic Mode Functions (IMF) obtained by the Empirical Mode Decomposition (EMD) method used for determining the sub-characteristic structure of the non-stationary signals are examined, and it has been seen that the results are matched with the previous results.

Science Code : 20506  
Keywords : Sound Signals, Similarity of Audio Signals, Short-Time Fourier Transform, Itakura-Saito Distance, Modified Itakura-Saito Distance, Time Series, Window Function, Empirical Mode Decomposition, Ensemble Empirical Mode Decomposition  
Number of pages : 61  
Supervisor : Prof. Dr. M. Akif BAKIR



## TEŐEKKÜR

Yüksek lisans öğrenimim sürecinde sahip olduđu engin bilgileri bana aktaran, önerileriyle beni yönlendiren, İstatistik bilimine duyduğum hayranlığın artmasını sağlayan, her zaman ilgisini ve desteğini gördüğüm, öğrencisi olmaktan gurur duyduğum danışman hocam sayın Prof. Dr. M. Akif BAKIR' a teşekkürlerimi sunarım.

Aynı kurumda yer almaktan onur duyduğum, her zaman ilgisini ve desteğini gördüğüm hocam sayın Doç. Dr. Filiz KARDİYEN' e teşekkürlerimi sunarım

Her anımda yanımda olan, birçok fedakârlıklar göstererek beni destekleyen ve bugünümün mimarları olan annem Leyla ERDİŐ'e, babam Eyyüp ERDİŐ' e ve Kardeşim Aycan ERDİŐ' e bana karşı duydukları sarsılmaz inançlarından dolayı teşekkürlerimi sunarım.



## İÇİNDEKİLER

	Sayfa
ÖZET .....	iv
ABSTRACT .....	v
TEŞEKKÜR .....	vi
İÇİNDEKİLER .....	vii
ÇİZELGELERİN LİSTESİ .....	ix
ŞEKİLLERİN LİSTESİ .....	x
SİMGELER VE KISALTMALAR .....	xi
<b>1. GİRİŞ</b> .....	<b>1</b>
<b>2. SES SİNYALLERİ VE ZAMAN SERİLERİ</b> .....	<b>7</b>
2.1. Ses, Ses Dalgası ve Ses Sinyali .....	7
2.2. Genlik, Frekans ve Dalga Boyu .....	8
2.2.1. Ses ile ilgili diğer kavramlar .....	12
2.2.2. Müzikal benzerlik.....	13
2.3. Zaman Serileri Analizinin Temel Kavramları ve Genel Çerçevesi.....	14
2.3.1. Zaman serilerinde amaç.....	14
2.3.2. Zaman serileri analiz yöntemleri .....	15
2.3.3. Zaman serilerinin durağanlığı.....	17
2.3.4. Durağanlık ölçüsü .....	19
<b>3. SES SİNYALLERİ ARASINDAKİ BENZERLİĞİN TESPİTİNE İSTATİSTİKSEL YAKLAŞIM</b> .....	<b>23</b>
3.1. Itakura-Saito Uzaklığı (ISU) .....	24
3.2. Fourier Dönüşümü .....	25
3.2.1. Sürekli zamanlı Fourier dönüşümü.....	26
3.2.2. Ayrık zamanlı Fourier dönüşümü .....	26
3.3. Kısa Zamanlı Fourier Dönüşümü (STFT) .....	28





	<b>Sayfa</b>
3.4. Pencere fonksiyonları.....	30
3.4.1. Sabit pencere fonksiyonları .....	30
3.4.2. Ayarlanabilir pencere fonksiyonları.....	34
<b>4. MÜZİKAL BENZERLİĞİN TESPİTİ İÇİN UYARLANMIŞ ITAKURA-SAITO UZAKLIĞI (UISU) .....</b>	<b>37</b>
4.1. Simülasyon Çalışması .....	38
4.2. Gerçek Müzik Sinyallerinin Benzerliği .....	40
<b>5. AMPİRİK MOD AYRIŞIMININ UYARLANMIŞ ITAKURA-SAITO UZAKLIĞINA ETKİSİ .....</b>	<b>45</b>
5.1. Ampirik Mod Ayrışımı (EMD).....	45
5.2. Grupsal Ampirik Mod Ayrışımı (EEMD) .....	50
5.3. Ses Sinyallerinin Benzerliğinin IMF'ler Arasındaki Benzerlik Üzerinden Tespiti .....	51
<b>6. TARTIŞMA .....</b>	<b>53</b>
<b>7. SONUÇ VE ÖNERİLER .....</b>	<b>55</b>
<b>KAYNAKLAR .....</b>	<b>57</b>
<b>ÖZGEÇMİŞ .....</b>	<b>60</b>



## ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. Notaların oktavlarına göre frekansları (Suits, n.d.).....	10
Çizelge 4.1. Hipotetik sinyallerin ISU ve UISU karşılaştırması .....	39
Çizelge 4.2. Gerçek ses sinyallerinin ISU ve UISU karşılaştırılması .....	42
Çizelge 5.1. IMF'ler arasındaki uyarlanmış Itakura-Saito uzaklıkları .....	51





## ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Düşük ve yüksek frekanslı dalga örnekleri .....	9
Şekil 2.2. Genlik, frekans ve dalga boyu niceliklerinin dalga fonksiyonu üzerinde gösterimi.....	9
Şekil 2.3. Sesin dijitalleştirilmesi ve insan kulağı tarafından algılanması .....	12
Şekil 2.4. Orijinal Fidayda (a) ve nota süreleri değiştirilmiş Fidayda eserlerinin bir ölçülük nota dizgesi.....	13
Şekil 3.1. Spektral sızıntı problemi .....	27
Şekil 3.2. STFT ile sinyalin analizi (Jaber and Bicker, 2018) .....	29
Şekil 3.3. Dikdörtgen pencere fonksiyonu ( $L = 64$ ) .....	30
Şekil 3.4. Hamming pencere fonksiyonu ( $L = 64$ ) .....	31
Şekil 3.5. Hann pencere fonksiyonu ( $L = 64$ ) .....	32
Şekil 3.6. Blackman pencere fonksiyonu ( $L = 64$ ) .....	33
Şekil 3.7. Gaussian pencere fonksiyonu ( $L = 64$ ) .....	34
Şekil 3.8. Kaiser pencere fonksiyonu ( $L = 64, \alpha_k = (0.5, 2.5, 12)^l$ ).....	36
Şekil 4.1. Hipotetik ses sinyalleri .....	39
Şekil 4.2. Karşılaştırma için seçilen ezgilerin zaman-genlik grafikleri.....	42
Şekil 5.1. IMF lerin elde edilşinin basit diyagramatik gösterimi .....	47
Şekil 5.2. EMD algoritmasının akış diyagramı .....	48
Şekil 5.3. Ampirik Mod Ayrışımı (EMD) .....	49



## SİMGELER VE KISALTMALAR

Bu çalışmada kullanılan simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

### Simgeler

### Açıklamalar

$T$	İndis kümesi
$\lambda$	Dalga boyu
$\gamma_k$	Otokovaryans fonksiyonu
$\rho_k$	Otokorelasyon fonksiyonu
$\phi_k$	Kısmi otokorelasyon fonksiyonu
$R_X$	$X_t$ sürecinin otokorelasyon matrisi
$R_Y$	$Y_t$ sürecinin otokorelasyon fonksiyonu
$\nabla$	Fark operatörü
$d_I S^*$	Itakura-Saito uzaklığı
$\omega$	Dalganın frekansı
$A$	Dalganın genliği
$\sharp$	Diyez
$b$	Bemol
$sFs$	Örnekleme frekansı

### Kısaltmalar

### Açıklamalar

<b>ACF</b>	Otokorelasyon Fonksiyonu
<b>A</b>	La
<b>B</b>	Si
<b>C</b>	Do
<b>D</b>	Re
<b>E</b>	Mi
<b>EMD</b>	Ampirik Mod Ayrışımı
<b>EEMD</b>	Grupsal Ampirik Mod Ayrışımı





<b>Kısaltmalar</b>	<b>Açıklamalar</b>
<b>F</b>	Fa
<b>FT</b>	Fourier Dönüşümü
<b>FFT</b>	Hızlı Fourier Dönüşümü
<b>G</b>	Sol
<b>ISU</b>	Itakura-Saito Uzaklığı
<b>WN</b>	Beyaz Gürültü (White Noise)
<b>PACF</b>	Kısmi Otokorelasyon Fonksiyonu
<b>STFT</b>	Kısa-Zamanlı Fourier Dönüşümü
<b>UISU</b>	Uyarlanmış Itakura-Saito Uzaklığı



## 1. GİRİŞ

İnsanlar eski çağlardan beri doğadaki çeşitli sesleri taklit etmiş ve bu taklitleri yeni sesler üretmek amacıyla kullanmışlardır. Nesnelere ile etkileşim sonucu çıkan sesler ve insan doğasında bulunan merak ve yaratıcılık ilk enstrümanların doğuşunu motive etmiştir. Enstrümanların doğuşu ile birlikte sesler belirli şekillerde bütünleşerek ezgi halini almış ve bu olaya müzik denmiştir. Ezgiler birbirinden bağımsız ve bağımlı olmak üzere kültürden kültüre, insandan insana büyük ölçüde değişmekte ve farklılaşmaktadır. Günümüzde belki de sayılamayacak kadar farklı ezgi mevcuttur. Bu ezgilerin kültürden kültüre, hatta insandan insana değişkenlik gösteriyor oluşu, ayrıca sürekli farklılaşıyor oluşu, bu olayın merkezinde rasgelelik olgusunun varlığına işaret etmektedir.

Kültürel farklılık her ne kadar farklı ezgiler ortaya çıkarsa da, kültürlerin birbirinden etkilenmesi durumu benzer ezgi yapıları ortaya çıkarabilmektedir. Ezgilerdeki bu benzerlik bazı durumlarda esinlenmenin bir adım ötesine geçerek referans vermeksizin alıntı yapma boyutuna geçmektedir. Alıntının illegal biçimde yapılması durumu aşırma (plagiarizm) olarak adlandırılmaktadır. Günümüzde eser sahiplerinin haklarını koruyan telif yasaları bulunmakta ve plagiarizm bu yasalar gereğince suç olarak kabul edilmektedir. Plagiarizm olduğu şüphesi ile açılan bir davada bir karara varabilmek için ise alanında uzman müzisyenlerden oluşan bir kurula başvurulmaktadır. Bir insanın işitsel becerisi benzer sesleri algılayabilmektedir. Usta müzisyenler ise müzik teorisi ile eğitilmiş işitsel beceriyi birleştirerek ezgilerin benzerliği konusunda yüksek kesinlikte bir karar verebilmektedir. Ancak bu karar doğru olsa dahi, sezgisel argümanlara dayalı olması sebebiyle tartışmaya açıktır ve davalı taraf haklı olarak sonuca itiraz edebilmektedir. Bu nedenle, eserlerin benzerliğine ilişkin daha kesin sonuçlara ve dolayısıyla nicel kanıtlara gereksinim duyulmaktadır.

Müzikal benzerlik konusu, uluslararası literatürde çeşitli boyut ve yöntemlerle ele alınmış olmakla birlikte, hem yöntemsel hem de Türk müziği bağlamında (bilgimize göre) ulusal literatürde herhangi bir çalışma mevcut değildir. Doğrudan müzikal melodi benzerliklerini (plagiarizm) belirlemeye yönelik olmasa da, veri tabanlarındaki müzik eserlerini tarz benzerliklerine göre sınıflandırma çerçevesinde dikkate değer sayıda çalışma yapılmış ve

sınıflandırmaya yönelik çeşitli yöntemler önerilmiştir. Müzikal eserlerin karakteristik yapısının son derece karmaşık oluşu, yapısal anlamdaki benzerliklerinin araştırmasını bir hayli zorlaşmaktadır. Bu nedenle, literatürde yapısal anlamdaki benzerliğin tespitine ilişkin sınırlı çalışma mevcuttur. Bu çalışmada eserlerin tarz benzerliğini belirlemeye yönelik olmamakla birlikte, literatürdeki bazı önemli çalışmalar ve yöntemsel çeşitlilik dikkate alınarak, melodilerdeki yapısal benzerliğin tespitine yönelik yeni bir yöntemin elde edilmesi hedeflenmektedir.

Weihls ve ark. (2007) denetimsiz sınıflandırma yöntemini (unsupervised classification method) kullanarak sinyallerden elde edilen ve bir müzik parçasını karakterize eden kısa, orta ve uzun dönem diye tanımladıkları bir takım özelliklere (değişkenler) dayalı olarak müzik parçalarını sınıflandıran bir yöntem geliştirmiştir.

Park ve Lee (2013), melodi karşılaştırması için MIDI notasyonlarına dayanan tek boyutlu sayısal karaktere (string) daha fazla odaklanan, bütüncül bir benzerlik yaklaşımı önermişlerdir. Çalışmalarında koşullu öklit uzaklığının (Conditional Euclidean Distance) melodi benzerliği için daha iyi bir uzaklık ölçüsü olduğunu öne sürmüşlerdir.

Vempala (2012)'deki çalışmasında bir notası değiştirilmiş iki melodinin dinletildiği dinleyicilerin melodilerin benzerliğine ilişkin verdikleri yanıtlara dayalı olarak benzerliğin psikolojik olarak algı boyutunda ölçümlenmesi üzerine yoğunlaşmıştır. Çalışmada melodik kontur, ses perdeleri arası uzaklık (pitch distance) ve ses perdesinin yönü (pitch direction) psikolojik olarak melodi benzerliğinde karar vermeyi büyük ölçüde etkileyen faktörler olarak bulunmuştur.

Toiviainen ve Eerola (2002), melodi benzerliği için farklı bir yaklaşım olarak çoklu betimleme (multiple representation) ve özörgütlü haritalamaya (self organizing mapping) dayalı hesaplamalı bir model önerisinde bulunmuştur.

Bir diğer çalışmada ise Urbano ve ark. (2012)'de melodinin perde zamanları düzleminde ve dizi hizalaması (sequence alignment) gösterimlerinin olduğu geometrik melodik gösterim yaklaşımı ile melodiler arası benzerliği belirlemeye yönelik farklı bir öneride bulunmuştur.

Melodi karşılaştırmasına yönelik bir başka yöntemsel öneri ise Xiao ve ark. (2014) tarafından yapılmıştır. Önerilen yöntemde melodilerin doğru eşleştirilmesi Grafiksel İşleme Birimi hızlandırıcısı (Graphical Processing Unit acceleration) yaklaşımı ile gerçekleştirilmiştir.

Cambouropoulos (2001)'deki çalışmasında özel olarak monofonik melodiler için bir benzerlik algoritması önermiştir. Cambouropoulos ve ark. (2005) tarafından yapılan bir diğer çalışmada ise kesikli sembollerin karakter dizisi olarak temsil edilen melodik bir diziye uygulanabilen etkin bir örüntü algoritması geliştirmiştir.

Ahonen ve ark. (2011), sembolik polifonik müzikteki MIDI dosyalardan ayrıştırılan ikili (binary) chromagramların değerlerinin eşleştirilmesine (matching) dayanan bir yöntem önermişlerdir.

Yine uluslararası literatürde doğrudan aşırma belirlemeye yönelik yöntemsel ve software nitelikli bazı çalışmalar da mevcuttur.

Shin ve ark. (2014)'de müzikal benzerliğin düzeyini ölçen bir sistem amaçlanmış olup Sistem IT teknolojisini kullanarak aşırma karar verme yöntemlerini içermektedir. Bu çalışmada aşırma durumunu analiz etmek için JFugue müzik stringleri kullanılarak müzikal içerik için aşırma analizi yapan ve notaları gruplandırarak karakter dizisine dönüştüren abstract syntax tree (AST) leri veren C++ tabanlı bir program sistem tasarımı amaçlanmıştır. Sistem iki melodi arasındaki yapısal benzerliği IT teknolojisi ile analiz etmektedir.

Velankar M.R. ve ark. (2015), melodinin insan beynindeki algısını esas alarak bilgisayar tabanlı bir melodi benzerlik algoritması geliştirmiştir. Çalışmada, sentetik olarak üretilen bir melodinin nota, süre ve ölçek gibi parametreleri değiştirilerek benzer melodiler üretilmiş ve referans melodiye göre benzer melodilerin dinleyici tarafından nasıl algılandığı anlaşılmasına çalışılmıştır. İki ezgi arasındaki uzaklığı ölçmek için nota dizilişlerinde, ölçek ve sürelerde değişiklikler yapılmıştır. Uzaklık ölçüsü, dikkate alınan parametrelerin ağırlıklı uzaklığı olarak hesaplanmıştır. Önerilen model ritim, tını vb. parametreleri

dikkate almamış ve bunların melodiden bağımsız olduğunu varsaymıştır. Ancak çalışmada önerilen bu yeni yaklaşımın diğer yaklaşımlara göre performansı test edilmemiştir.

Müzikal aşırmaı belirleme konusundaki diđer bir alıřma ise Suneja ve Bansal (2015) tarafından yapılmıřtır. MATLAB’da kodlanan benzerlik ölçüsü algoritmalarıyla aşırma problemi üç boyutta ele alınmıřtır: Müzikal melodinin aynen aşırılması, aynı melodinin uzunlukları farklılařtıđında aşırmanın belirlenmesi ve rastgele seilmiř iki müzik arasındaki aşırmanın belirlenmesi. alıřmada müzikler arası benzerliđi belirlemek için Dynamic time Warping (DTW), Longest Common Subsequence (LCSS) vb. 5 uzaklık ölçüsü denenmiřtir. Belli bir melodi seti üzerinde yapılan deneysel alıřma sonucunda DTW uzaklık ölçüsünün diđer ölçülere göre daha başarılı biçimde müzikal benzerliđi belirlediđi sonucuna ulařılmıřtır.

Mongeau M. ve Sankoff D. (1990), müzikal serilerin karşılaştırılmasını amaçlayan alıřmalarında benzerlik ölçüsünü hesaplayan bir dinamik programlama algoritması geliřtirerek, bu algoritmayı Mozart’ın bir ana temasının bir dizi varyasyonunda uygulamıřtır.

Slaney M. ve ark. (2008), benzerlik ölçüsü olarak kullanılan Öklidiyen (Euclidean) uzaklıđın, her bir özelliđin eřit öneme sahip olduđunu varsayması özelliđinden yola ıkarak, boyutlara farklı ađırlıklar vermenin eřitli yöntemlerini önermiřlerdir. Önerdikleri sistemin performansını k-en yakın komřuluk (kNN) sınıflayıcısının sonuçlarını kullanarak ölçmüřler ve anlamlı sonuçlar bulmuřlardır.

Logan B. ve Salomon A. (2001), řarkıları ses içeriklerine dayanarak karşılařtıran bir yöntem önermiřtir. Yöntemde, her bir řarkı için spektral özelliklerine K-Ortalama kümelemesi uygulanarak tanıtım müziđi elde edilmiřtir. Elde edilen tanıtım müzikleri Yer Deđiřtirme Uzaklıđı (Earth Mover’s Distance) ve Kullback Leibler (KL) uzaklıđının simetrik formu kullanılarak geliřtirilen ölçü ile karşılařtırılmıřtır. Kullanılan ölçünün ses sinyallerindeki bozulmaya karşı sađlam olduđu belirtilmiřtir.

Haas W. ve ark. (2013), alıřmalarında iki müzik parası arasındaki harmonik uzaklıđı ölçen

yeni bir geometrik uzaklık fonksiyonu önermiş ve bu fonksiyon başka harmonik uzaklıklarla karşılaştırarak test etmiştir. Önerilen fonksiyon en iyi performansa sahip olmasa da ölçünün etkin bir uzaklık ölçüsü olduğu gösterilmiştir.

Fiziksel olarak ses, dalga formundadır ve zaman olgusu dikkate alındığında, ses sinyallerinin zamanın belirli bir aralığında, kronolojik olarak sıralanmış olduğu görülmektedir. Bu nedenle, ses sinyalleri istatistiksel olarak zaman serileri formuna sahiptir ve zaman serisi analiz yöntemleri ile de analiz edilebilmektedir. Zaman serisi yaklaşımı ile, ses sinyallerinin karakteristik yapısı ortaya çıkarılarak sinyaller arası benzerlikler de araştırılabilmektedir. Zaman serileri bakış açısıyla ses sinyallerini incelemek, sinyaller arasındaki ilişki yapılarının da incelenmesini beraberinde getirecektir. İki ses sinyali arasındaki uzaklık, bu sinyallerin benzerliği hakkında sonuç çıkarımı yapabilme olanağını sağlayacaktır. Bu benzerlik yapılarının incelenmesi için çeşitli uzaklıkların revize edilerek kullanılabilmesi mümkündür.

Bu çalışmada, ses sinyalleri dijital ortama aktarıldıktan sonra zaman serileri perspektifi ile irdelenerek, sinyaller arasındaki benzerlikler istatistiksel uzaklık ölçülerinin uyarlanması aracılığı ile tespit edilmektedir. Bu çalışmada, durağan zaman serileri arasındaki benzerliğin bir ölçüsü olan Itakura-Saito uzaklığı, durağan olmayan zaman serilerinde kullanılabilir forma uyarlanarak, ses sinyalleri arasındaki benzerliğin tespiti için kullanışlı hale getirilmiştir. Bu sebeple çalışma, müzik eserlerindeki aşırımların belirlenmesine yönelik alternatif ve nicel bir yöntem ortaya koymaktadır. Önerilen uzaklık, simülasyon ve gerçek verilerle test edilmiş olup, anlamlı sonuçlar elde edilmiştir. Ayrıca, durağan olmayan zaman serilerinin sahip olduğu alt karakteristik yapıların incelenmesine olanak sağlayan Ampirik Mod Ayırıştırması (EMD) yöntemi ile de sınanarak, sonuçların güvenilir olup olmadığı incelenmiş ve güvenilir olduğu görülmüştür.





## 2. SES SİNYALLERİ VE ZAMAN SERİLERİ

Çalışmanın bu bölümünde ses sinyalleri ve zaman serileri ile ilgili bazı temel kavramlar üzerinde durulacaktır.

### 2.1. Ses, Ses Dalgası ve Ses Sinyali

Ses, atmosferde canlıların işitme organları tarafından algılanabilen periyodik basınç değişimleridir. Fiziksel boyutta ses, katı, sıvı veya gaz ortamlarda oluşan basit bir mekanik düzensizliktir. Bir maddedeki moleküllerin titreşmesi sonucunda oluşur ve aynı zamanda bir enerji türüdür. Sesin var olabilmesi için, bir ses kaynağının (beyni ve kulağı uyurabilecek nitelikteki etkenler), bir alıcı sistemin (bu durumda kulak ve beyin) ve bir iletici ortam (etkenlerin ses kaynağından alıcı sisteme kadar kulağı uyarmaya yetecek şiddette iletilmesi gerekir.) Bu öğelerden herhangi biri yoksa ses de yoktur. Başka bir deyişle; ses mekanik bir dalga olduğundan yayılması için bir ortama ihtiyaç duymaktadır.

Ses dalgası en önemli boyuna dalga örneklerindedir. Bu dalgalar herhangi bir ortamda (yani gazlar, katılar ve sıvılar) ortamın özelliklerine bağlı olan bir hızla, ortamlarda sıkışma ve genleşme şeklinde boyuna ilerleyerek yayılırlar. Ses dalgası bir ortamda yayılırken ortamın parçacıkları dalganın hareket doğrultusu boyunca “yoğunluk” ve “hacim” değişiklikleri üreterek titreşir. Frekanslarına göre boyuna mekanik dalgalar üç gruba ayrılır;

1. İşitilebilir Dalgalar: İnsan kulağının duyarlılık sınırları içinde olan ses dalgalarıdır. Bu dalgalar 20 Hz ile 20.000 Hz frekansları arasındadır. Bu duyarlılık sınırları içerisinde ses dalgaları değişik yollarla yaratılabilir. Örneğin, müzik aletleriyle, boğazdaki ses telleriyle ve hoparlör ile.
2. Ses Altı (İnfrasonik) Dalgalar: İşitilebilir mertebenin altındaki frekansta olan boyuna dalgalarıdır. Deprem dalgaları bu dalgalara örnektir.
3. Ses Üstü (Ultrasonik) Dalgalar: İşitilebilir mertebenin üstünde frekansları olan boyuna dalgalarıdır. Örneğin; bu dalgalar bir kuartz kristaline alternatif elektrik alanının uygulanmasıyla elde edilebilir. ("Ses", 2018)

Ses Sinyali (Audio), en yalın tanımlamayla, sesin “elektiriksel formu” dur. İnsan kulağı

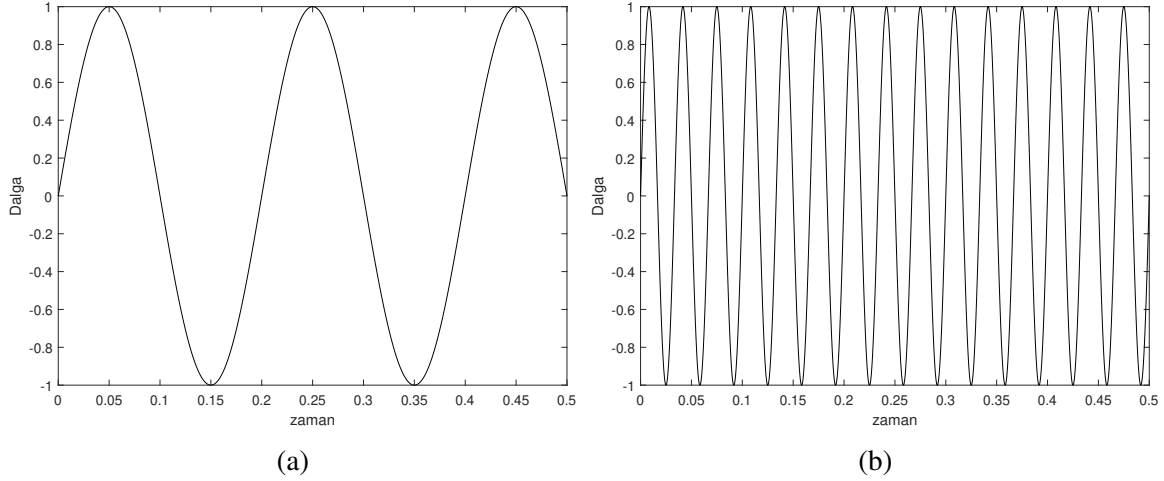
tarafından algılanamaz. Ses sinyali, sesin, kaydedilebilmesi, uzak mesafelere iletebilmesi ve saklanıp yeniden üretebilmesi için bir “çevirimden geçmiş” halidir. Analog ve Dijital olmak üzere iki çeşittir. Analog ses sinyali, sesin birebir kopyalanmış (tespit edilmiş) biçimidir. Dijital ses sinyali ise ‘analog ses sinyalinin’ birim zamanda belli örnekler alınması ve örneklerin sadece 1 ve 0 değerlerinden oluşan bir koda dönüştürülmesiyle elde edilmiş sayısal biçimidir.

İnsan kulağı ne analog ne de dijital ‘ses sinyalini’ duyamaz, sesi duyar. Ancak sesi kaydetmek, iletmek ve saklayıp yeniden üretmek istenildiğinde “ses sinyaline” çevrilmeli ve duymak istenildiğinde de yeniden “sese” dönüştürülmelidir. Sesler kulaklarımıza kaynaklarından çıktıkları güçle (ses seviyesi ile) ulaşamaz. Sesin ilerlediği yollar üzerindeki engeller, ses şiddetinin zayıflamasına sebep olur. Ayrıca kulak kepçesinin ve başın şekli, saçlar ve hatta giysiler, ses dalgalarını zayıflatma etkisi gösterebilir.

Doğada seslerin yönünü, çıkış kaynaklarının konumunu ve kaynağın bulunduğu ortamı bu parametreler ve akustiğin prensipleri sayesinde algılarız. Bir başka deyişle, sesin kaynaktan çıktıktan sonra nereden ne kadar yansıdığı ve soğurulduğu (emildiği), sesin yönünün ve şiddetinin anlaşılmasına katkıda bulunur. Örneğin bir konser salonunda icra edilen müziğin oluşturduğu ses dalgası, dinleyiciye, doğrudan sahneden değil konser salonun duvarlarından ve tavanından da yansımış hali ile gelmektedir. Müziğin farklı ortamlarda farklı etki yaratması da akustik ortam içindeki ses dağılımından ileri gelmektedir.

## **2.2. Genlik, Frekans ve Dalga Boyu**

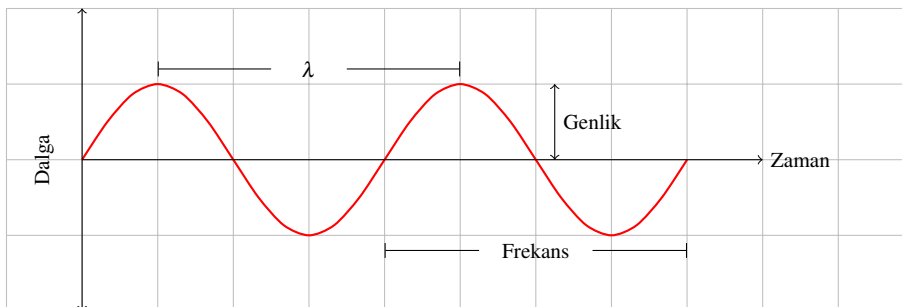
Ses sinyalleri frekansı, boyu, periyodu ve hızı ile karakterize olur. Bir saniye içerisindeki titreşim sayısına frekans denir. Bir dalganın frekansı, dalganın hava veya başka bir ortam içinden geçerken ortamdaki partiküllerin ne sıklıkta titreştiğine bağlıdır. Frekans ileri geri titreşimlerin zamana bağlı olarak ölçülmesi ile hesaplanır. Birimi ise Hertz (Hz) dir. (1 Hertz = 1 döngü/saniye). Şekil 2.1 de aynı şiddete (genliğe) sahip, farklı frekanslı sinyallerin karakteristik yapısını gösterilmektedir.



Şekil 2.1. Düşük ve yüksek frekanslı dalga örnekleri

Genlik (amplitude), dalgaların dikey büyüklüğünün bir ölçüsüdür. Ses dalgalarını oluşturan sıkışma ve genleşmeler arasındaki fark, dalgaların genliğini belirler. Örneğin bir müzik enstrümanının telindeki titreşimin genliği ne kadar fazla ise ortam tanecikleri (örneğin hava molekülleri) tarafından taşınan enerji de o kadar fazladır. Enerji ne kadar fazla ise sesin şiddeti de o kadar büyük olacaktır. Örneğin titreştirilen bir enstrüman teli, yaptığı periyodik salınım hareketi ile hava moleküllerinin belli bir frekansta sıkışmasını ve genleşmesini sağlar. Bu şekilde teldeki enerji havaya iletilmiş olur. Enerjinin miktarı, teldeki titreşim genliğine bağlıdır. Eğer tele fazla enerji yüklenirse, tel daha büyük bir genlikle titreşir. Bu durum, titreşen tüm cisimler için geçerlidir.

Dalga boyu, bir ses dalgasının oluşması için sesin aldığı yoldur. Başka bir deyişle dalga boyu, bir dalganın ardışık iki tepe veya iki çukur noktası arasındaki mesafedir. Dalga boyu genellikle  $\lambda$  (lambda) ile gösterilir. Şekil 2.2 de bir ses dalgasını karakterize eden parametreler gösterilmektedir.



Şekil 2.2. Genlik, frekans ve dalga boyu niceliklerinin dalga fonksiyonu üzerinde gösterimi

Ses hızı, sesin var olduğu ortama ve ortamın sıcaklığına bağlı olarak değişkenlik gösterir. Dolayısıyla, sesin yayılabilmek için ortama gereksinim duyduğu aşıkârdır. Böylece, ses boşlukta yayılamaz. Bunun nedeni, boşlukta titreşen bir cismin sıkışıp genişlemesine yol açabilecek atom ya da molekül gibi taneciklerin olmayışıdır.

Ses sinyalleri bir spektral tondan meydana gelebildiği gibi birden fazla spektral tonun karışımından da meydana gelebilmektedir. Burada spektral ton, nota kavramının matematiksel karşılığı olarak kullanılmıştır. Notalar matematiksel olarak belirli frekansa sahip sinüzoidal dalgalar ile modellenenmektedir. Tablo 2.1 de notaların oktavına göre frekansları verilmektedir.

Çizelge 2.1. Notaların oktavlarına göre frekansları (Suits, n.d.)

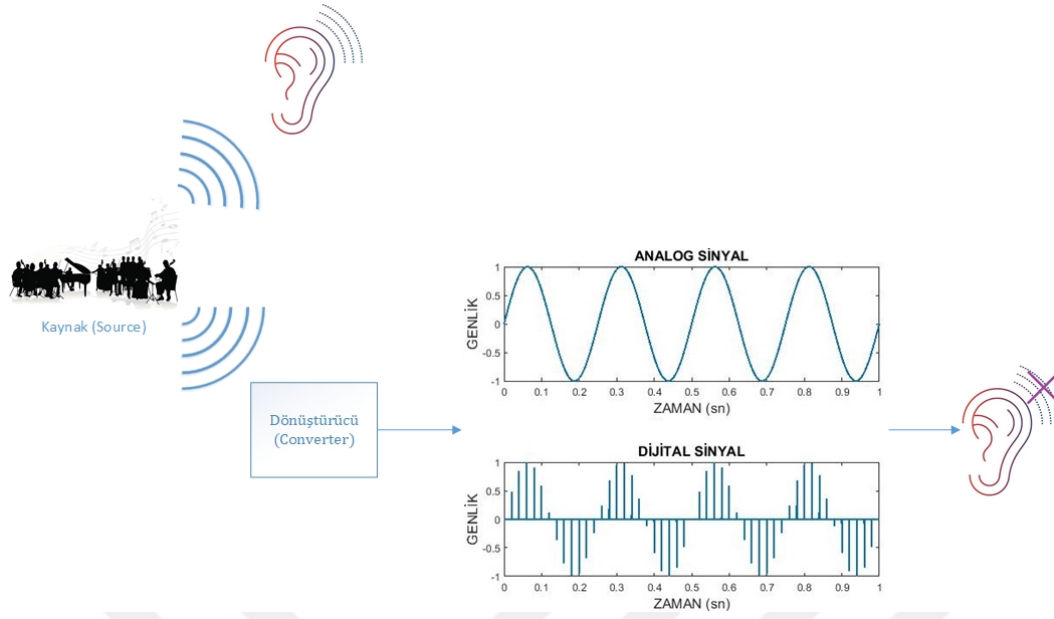
Notalar	OKTAV								
	0	1	2	3	4	5	6	7	8
C	16,35	32,70	65,41	130,80	261,60	523,30	1047,00	2093,00	4186,00
C#	17,32	34,65	69,30	138,60	277,20	554,40	1109,00	2217,00	4435,00
D	18,35	36,71	73,42	146,80	293,70	587,30	1175,00	2349,00	4699,00
E <sub>b</sub>	19,45	38,89	77,78	155,60	311,10	622,30	1245,00	2489,00	4978,00
E	20,60	41,20	82,41	164,80	329,60	659,30	1319,00	2637,00	5274,00
F	21,83	43,65	87,31	174,60	349,20	698,50	1397,00	2794,00	5588,00
F#	23,12	46,25	92,50	185,00	370,00	740,00	1480,00	2960,00	5920,00
G	24,50	49,00	98,00	196,00	392,00	784,00	1568,00	3136,00	6272,00
G#	25,96	51,91	103,80	207,70	415,30	830,60	1661,00	3322,00	6645,00
A	27,50	55,00	110,00	220,00	440,00	880,00	1760,00	3520,00	7040,00
B <sub>b</sub>	29,14	58,27	116,50	233,10	466,20	932,30	1865,00	3729,00	7459,00
B	30,87	61,74	123,50	246,90	493,90	987,80	1976,00	3951,00	7902,00

Ses sinyalleri, sürekli zamanlı analog sinyal sınıfına girmektedir ve bu sınıftaki sinyaller doğrudan bilgisayar ve elektronik sistemlerde değerlendirilemez. Her ne kadar Tablo 2.1 de verilen nota frekansları ve sinüzoidal fonksiyonlar aracılığı ile dijital ortamda ses sinyalleri oluşturulabilse de, herhangi bir kaynaktan meydana gelen ses sinyallerinin analiz edilebilmesi için sayısallaştırılması (dijital ortama aktarılması) gerekmektedir. Bu nedenle ses sinyallerine, analog sinyalin dijital sinyale dönüştürülme işlemi olan A/D (Analog/Dijital) dönüştürme süreci uygulanmalıdır. Bu dönüştürülme süreci örnekleme ve nicelleştirme şeklinde iki aşamalı olarak gerçekleştirilir. Örnekleme, sürekli bir sinyalden

belli zaman aralıklarıyla örnek toplanması işlemine denir. Örneğin, bir  $f(x) = \sin(x)$  fonksiyonu,  $(-\pi, \pi)$  aralığında sonsuz değer almakla birlikte fonksiyonun grafiğini çizilebilmek için sonsuz değere ihtiyaç yoktur. Bunun için fonksiyonu sadece belli değerlerde hesaplayarak, yani başka bir deyişle örnekleyerek orjinal fonksiyona benzeyen grafik çizilebilmektedir. Ses sinyalleri de aynı mantık ve amaçla örneklenmektedir.

Örnekleme aşamasında zaman dijitalleştirilirken, nicelleştirme aşamasında genlik dijitalleştirilmektedir. Örnekleme aşamasında örnekler rastgele zaman aralıklarıyla toplanmaz, belli bir periyot ve dolayısıyla frekans ile alınır. Örneğin, CD kalitesinde ses elde etmek için bir ses sinyalinden saniyede 44100 adet örnek toplanması gerekir. Bu, sesin 44100 hertzlik frekans ile örneklendiğini ifade eder. Örnekleme frekansı, sinyalin gerçek frekansından bağımsızdır.

Örnekleme periyodu, bant genişliğinin belirlenmesi (başka bir deyişle frekans spektrumunun (tayf) olabilecek en büyük frekans aralığında sınırlanması) ve örnekleme frekansının ( $f_s$ ) maksimum frekansın ( $f_{max}$ ) iki katından büyük veya eşit ( $f_s \geq 2f_{max}$ ) olacak şekilde belirlenmesi şeklinde iki adımda seçilir. Burada seçilen örnekleme frekansına “Nyquist Değeri” denmektedir. Ses sinyalleri için örnekleme frekansı uygulamada yaygın olarak 44 kHz (44100 Hz) olacak şekilde alınır. Örnekleme frekansı seçildikten sonra zaman dijitalleştirilmiş olmaktadır. Zaman dijitalleştirildikten sonra zamanın her bir anında sesin şiddeti ölçülerek nicelleştirme işlemi gerçekleştirilir. Şekil 2.3’de sesin dijital ortama aktarılması ve insan kulağı tarafından algılanabilme durumu gösterilmektedir.



Şekil 2.3. Sesin dijitalleştirilmesi ve insan kulağı tarafından algılanması

Ses sinyalinin dijital ortama aktarmak için yapılan örnekleme işleminde, sinyalden alınan her bir örnek için bilgisayarda bir alan ayrılmaktadır. Ayrılan bu alan boyut olarak 8bit, 16bit, 32bit, ... biçimindedir. Bu değerlerin 8 in katları ya da 2 nin kuvvetleri olduğu kolayca görülmektedir. Bunun nedeni, örneği en az 1 bayt (byte) ve katları olarak ifade etme isteğinden ileri gelmektedir. Örnekleri temsil eden verinin boyutu arttıkça, genlik aralığı genişleyecek ve örneklenebilecek benzersiz genlik değeri sayısını arttıracaktır.

### 2.2.1. Ses ile ilgili diğer kavramlar

Ses, fiziksel olarak genlik, frekans vb. niceliklere sahip olmasının yanı sıra müzikal olarak da bazı niceliklere (diklik, yoğunluk, süre ve tını) sahiptir. Diklik, bir sesin ne kadar tiz ya da pes olduğunu ifade eder ve fiziksel olarak frekans kavramının müzikal karşılığıdır. Her nota (Do, re, mi, ...) farklı bir dikliğe sahiptir ve bulunduğu oktava göre değişmektedir. Yoğunluk, bir sesin gürlüğünü ifade etmektedir ve müzikte nüans olarak da kullanılır. Süre, bir sesin ne kadar sürdüğünü ifade eder ve müzikte ikinin katları (birlik, ikilik, dördlük, sekizlik,...) biçiminde tanımlanır. Tını ise, bir sesin rengini ifade eder. Örneğin aynı oktavda, aynı notayı, aynı yoğunlukta ve aynı uzunlukta çalan bir bağlamadan çıkan ses ile bir kemandan çıkan ses arasındaki fark "tını" farkıdır. Dört özellik içinde en karmaşık olanı budur. Akustik olarak tını, sesin harmonik yapısına bağlı olarak değişir.

### 2.2.2. Müzikal benzerlik

İki müzik eserinin benzer olabilmesi için, eserin önemli bir bölümünde ve/veya tamamında ki notaların, yakın frekanslarda ve sürelerde olması gerekmektedir. Bir başka deyişle, eserlerin benzer olabilmesi için frekans ve süre bazında ilişkili olmaları gerekmektedir. Aksi halde benzerlikten söz etmek doğru olmayacaktır. Müzikal benzerlik kavramını açıklayabilmek adına, Şekil 2.4 (a)'da Bulguru Kaynadırlar (Fidayda, Hüdayda) eserinin giriş kısmının bir parçasının nota dizgesi, Şekil 2.4 (b)'de ise bu parçadaki notaların sürelerinde değişiklik yapılarak oluşturulan bir diğer nota dizgesi verilmiştir.



Şekil 2.4. Orijinal Fidayda (a) ve nota süreleri değiştirilmiş Fidayda eserlerinin bir ölçülük nota dizgesi

Şekil 2.4'de nota dizgeleri verilen eserler müzik teorisi bazında birbirinden tamamen farklıdır. Ayrıca, bir enstrüman tarafından icra edildiğinde, bu iki eserin işitsel olarak da tamamen farklı olduğu görülecektir. Çalışmada, müzikal benzerlik ve benzemezlik kavramları bu durum dikkate alınarak irdelenmiştir.

Ses sinyallerinin dijitalleştirilmesi sonucu elde edilen yapı, genliklerinin zamana bağlı değişimi olacaktır. Zaman bağımlı yapısı ve oluşum sırasına göre (kronolojik olarak) sıralanmış olması ses sinyallerinin zaman serisi formunda olduğu gerçeğini ortaya koyar. Bu nedenle ses sinyalleri, zaman serileri bakış açısıyla istatistiksel olarak değerlendirilebilmekte ve analiz edilebilmektedir.

### 2.3. Zaman Serileri Analizinin Temel Kavramları ve Genel Çerçevesi

Stokastik süreç ve zaman serisi kavramları birbiriyle yakın ilişkili kavramlardır.  $\{\dots, Y_1, Y_2, \dots, Y_t, Y_{t+1}, \dots\} = \{Y_t\}_{t=-\infty}^{\infty}$  biçiminde zaman indeksli rastgele değişkenler serisine stokastik süreç denir. Genellikle sürekli-zaman stokastik süreçleri  $\{Y_t : 0 \leq t < \infty\}$ , kesikli-zaman stokastik süreçleri ise  $\{Y_t : t = 1, 2, \dots\}$  biçiminde tanımlanır. Stokastik sürecin  $n$  gözlemlilik bir gerçekleşmesi ise  $\{Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n\} = \{Y_t\}_{t=1}^n$  biçimindeki gözlemlenmiş veri dizisidir. Bu veri dizisine "örneklem fonksiyonu", "gerçekleşme" ya da "yörünge" denmektedir.

$(\Omega, \mathcal{U}, P)$  bir olasılık uzayı,  $T$  de indis kümesi olsun.  $\Omega \times T$  çarpım uzayından reel sayılara değer eşleyen,

$$\begin{aligned} X(\dots) : \Omega \times T &\longrightarrow \mathbb{R} \\ (\omega, t) &\longrightarrow X(\omega, t) \end{aligned} \tag{2.1}$$

biçiminde tanımlı fonksiyona zaman serisi denir ve genellikle  $X_t, X_t(\omega)$  veya  $X(\omega, t)$  gösterimleri ile temsil edilir.

Bu tanıma göre bir zaman serisi, her sabit  $t$  için bir rasgele değişkendir.  $\omega$  sabit kabul edilirse zaman serisi  $t$  nin reel değerli bir fonksiyonudur. Bu reel değerli fonksiyona zaman serisinin bir yörüngesi adı verilir (Akdi 2010). Bir başka deyişle zaman serisi belli bir stokastik sürecin yörüngesi ya da zaman boyunca elde edildikleri sıraya göre indekslenen rasgele değişkenlerin topluluğu olarak da tanımlanabilmektedir. Tüm mümkün yörüngelerden oluşan yığın, stokastik süreçlerde ve zaman serilerinde topluluk (ensemble) olarak adlandırılır. (Bakır, 2019)

#### 2.3.1. Zaman serilerinde amaç

Zaman serilerinin modellenmesinin amacı, gözlemlenmiş verinin türediği stokastik sürecin olasılıksal davranışını tanımlamak ve aynı zamanda gözlemlenmiş veriyi kullanarak zaman serisi modelinin zaman bağımlılık gibi ölçümlerini tahmin etmek olduğundan, stokastik



süreçteki rastgele değişkenlerin ortak davranışına ilişkin varsayımlar yapılması gerekir. Zaman serisi analizlerinin amacı iki grupta toplanabilir:

1. *Serideki gözlemlerin zamana bağımlı (dinamik) özelliklerini anlamak:*

Stokastik sürecin bir gerçekleşmesi olan zaman serisi verisinin içerdiği bilgiyi kullanarak veriyi üreten stokastik sürecin olasılık yapısını anlamak ve bu yapıya uygun model kurmak. Bunun için de verilerden bir model kurarak, serinin olasılık yapısı (stokastik süreç) açıklanmaya çalışılır.

2. *Geçmişteki gözlemlerin davranış biçiminden hareketle, gelecekteki gözlemleri öngörmek.*

Örneğin istatistik, ekonometri, finans, meteoroloji, geofizik, ... gibi alanlarda temel amaç genellikle öngörü yapmakken, sinyal işleme, kontrol mühendisliği, iletişim mühendisliği vb. alanlarda sinyal yakalama ve tahmindir.

### **2.3.2. Zaman serileri analiz yöntemleri**

Zaman serisi modelleri değişik biçimlere sahip ve farklı stokastik süreçleri temsil ettiğinden, çeşitli biçimlerde sınıflandırılabilirler. Zaman serileri analizine ilişkin en genel sınıflandırma, analizin zaman bazlı (time domain) ya da frekans bazlı (frequency domain) yapılmasına göre zaman bölgesine dayalı yöntemler ve frekans bölgesine dayalı yöntemler olarak yapılan sınıflandırmadır.

#### Zaman bölgesine dayalı yöntemler (Time domain):

Zaman bölgesi yaklaşımı zaman boyunca komşu noktalar arasındaki korelasyonu cari değerlerin geçmiş değerlere bağımlı olduğu önvarsayımı ile açıklar. Zaman bölgesi yaklaşımında, rastgele değişkenin gelecekteki değerleri cari ve geçmiş değerlerin parametrik bir fonksiyonu olarak modellenir. Zaman bölgesi yöntemleri aşağıdaki gibi sıralanabilir:

1. Düzgünleştirme (smoothing)
2. Bileşenleri ayrıştırma yaklaşımı (Decomposition)
3. Otoressive (AR) modelleri
4. Hareketli ortalama modelleri (MA)

5. Entegre modeller (I)
6. Box-Jenkins yaklaşımı (ARMA, ARIMA)

#### Frekans bölgesi yöntemler (Frequency domain):

Frekans bölgesi yöntemlerinde modeller zamana göre değil frekanslara göre oluşturulur. Literatürde spektral analiz yöntemleri olarak da bilinmektedir. Belli bir fonksiyon zaman bölgesinden frekans bölgesine ya da tersine frekans bölgesinden zaman bölgesine Fourier (ya da Laplace) transformasyon işlemleri ile dönüştürülebilir. Bir zaman bölgesi grafiği sinyal değişmelerini zaman boyunca gösterirken, frekans bölgesi grafiği her bir frekans bandına ne sıklıkla sinyal düştüğünü gösterir.

#### Parametrik-parametrik olmayan yöntemler:

Zaman dizisi analizi dizinin belli bir yapıya sahip olup olmasına göre parametrik ya da parametrik olmayan yöntemlerle yapılabilir.

- Parametrik yöntemler:

Durağan stokastik sürecin yapısının AR ve MA gibi yöntemleri kullanarak parametrelerle tanımlanabildiğini varsayar. Hedef, stokastik süreci tanımlayan modelin parametrelerini tahmin etmektir.

- Parametrik olmayan yöntemler:

Bu yöntemler ise stokastik sürecin belli bir yapıya sahip olduğunu varsaymaksızın sürecin spektrumunu ya da kovaryansını tahmin eder.

#### Model formuna göre yöntemler:

Zaman serileri, sahip oldukları model yapısına göre iki sınıfa ayrılmaktadır.

- Doğrusal zaman dizisi yöntemleri
- Doğrusal olmayan zaman dizisi yöntemleri

### 2.3.3. Zaman serilerinin durağanlığı

Bir stokastik süreçten elde edilen gözlemlere dayalı olarak sürecin yapısı hakkında istatistiksel çıkarsama yapmak için, sürecin yapısı hakkında bazı varsayımların yapılması gerekmektedir. Bu varsayımlardan en önemlisi “durağanlık” varsayımdır. Durağanlık temel olarak zaman boyunca sürecin davranışını yöneten olasılık kurallarının değişmemesi olarak tanımlanabilir. Başka bir deyişle durağanlık, sürecin “istatistiksel dengede” olması demektir.

Durağan zaman serileri, istatistiksel özellikleri (ortalaması, varyansı, otokorelasyon yapısı, ...) zamana bağlı olmayan serilerdir. Dolayısıyla, durağan bir süreçte, sürecin yapısı (olasılık yapısı) zaman boyunca değişmemektedir. Ayrıca trend, mevsimsellik ya da sistematik davranış göstermeyen, sadece rastgele bileşeni olan seriler durağan serilerdir.

#### Kesin durağanlık:

Olasılık dağılımı zaman boyunca aynı kalan zaman serisine kesin durağan seri denir. Daha açık ifade edecek olursak, herhangi bir zaman periyodu kümesi için rastgele değişkenler setinin geldiği olasılık dağılımı aynıdır. Örneğin 2000-2005 arasında serinin dağılımı, 2006-2016 arasındaki alt serinin olasılık dağılımıyla aynıdır. Kesin durağanlığı karakterize eden unsur olasılık dağılımının zaman boyunca aynı kalmasıdır. Burada dikkat edilmesi gereken husus sadece dağılımın aynı olması ile birlikte dağılımın biçimi hakkında (normal vb.) herhangi bir varsayım olmamasıdır.

$T$  bir indis kümesi olmak üzere,  $\{Y_t, t \in T\}$  bir zaman serisi olsun.  $\forall n, k, t_1, t_2, \dots, t_n, t_1 + k, t_2 + k, \dots, t_n + k \in T$  ve  $\forall y_1, y_2, \dots, y_n \in \mathbb{R}$  için

$$F_{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}}(y_1, y_2, \dots, y_n) = F_{Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k}}(y_1, y_2, \dots, y_n)$$

koşulu sağlanıyorsa  $\{Y_t, t \in T\}$  zaman serisi kesin (güçlü) durağandır denir. Başka bir deyişle olasılık dağılımı zaman boyunca aynı kalan zaman serisine kesin durağan seri denir.

Kesin durağanlığın günlük yaşam problemlerine ilişkin zaman serileri için sağlanmasının

oldukça zor olduğu aşikârdır. Bu nedenle genellikle daha zayıf bir durağanlık tanımı olan “zayıf durağanlık” veya “ikinci derece durağanlık” kullanılır.

### Zayıf durağanlık

Bir  $\{Y_t, t \in T\}$  zaman serisi için zayıf durağanlık koşulları aşağıdaki gibidir.

- i. Serinin birinci momenti sabittir ( $E[Y_t] = \mu$ )
- ii. Serinin varyansı sonlu ve sabittir.  $Var(Y_t) = \sigma^2 < \infty, \forall t \in \mathbb{Z}$
- iii.  $Kov(Y_t, Y_s)$  otokovaryansı  $t$  ile  $s$  arasındaki uzaklığın ( $|t - s|$ ) bir fonksiyonu olmalıdır. Başka bir deyişle  $Y_t$  ile  $Y_{t+k}$  arasındaki kovaryans ( $Kov(Y_t, Y_{t+k})$ ) zaman boyunca değişmemelidir.

Zayıf durağanlık durumunda sürecin dağılımına ilişkin varsayım yapılmamakta ve süreç ortalaması ve varyansı ile tanımlanmaktadır. Trend ve/veya mevsimselliğin olduğu zaman serileri durağan olmayan serilerdir. Trend yoksa stokastik süreç "ortalamada durağan", mevsimsellik ve sistematik bir değişme yoksa "varyansta durağan" dır. Zayıf durağanlık, kovaryans durağanlık olarak ta isimlendirilmektedir (Bakır, 2019).

Müzik eserlerinin icrasının önemli bir kısmını oluşturan tempo, tempo değişiklikleri ve dinamikler ses sinyallerine, zaman serisi perspektifinden bakabilmek adına önemli ipuçları vermektedir. Kısaca, tempo eserin icra hızını (largo: çok ağır (dakikada 40-60 vuruş), adagio: çok yavaş yakın (dakikada 60-76 vuruş), andante: ağıra yakın (dakikada 76-108 vuruş), moderato: orta hızda (dakikada 108-120 vuruş), allegro: hızlı (dakikada 120-168 vuruş), presto: çok hızlı (dakikada 168-192 vuruş)), tempo değişiklikleri eser hızını (accelerando: aşamalı hızlandırma, rallentando: aşamalı yavaşlama, ritardando: altına yazılan notanın değerinden uzun çalınması, ritenuto: anında yavaşlama, a tempo: aşırı hızlanma/yavaşlama sonucu orijinal hıza dönülmesi, rubato: tempoda isteğe bağlı değişim, piu mosso: biraz hızlanarak, poco: yavaş, poco a poco: yavaş yavaş) ve dinamikler ise sesin şiddetini (crescendo: sesin şiddetinin kademe kademe artırılması, decrescendo: sesin şiddetinin kademe kademe azaltılması, piano: yumuşak, forte: sert) karakterize etmektedir (Birgül ve Çoğulu, 2014). Müzik icrasındaki bu kısıtlar sinyalin varyansında değişkenlik

meydana getirmektedir. Bu nedenle müzik eserlerinin dijitalleştirilmesi sonucu elde edilen ses sinyali, zaman serisi formunda olmakla birlikte varyansta durağan değildir.

### Durağanlığın önemi

- Stokastik sürecin (yığının) yapısı hakkında, süreçten gözlenen verilere dayalı olarak çıkarsama yapmak için, genellikle o prosesin yapısı hakkında bazı varsayımlar yapılması gerekir.
- En önemli varsayım durağanlıktır. Durağanlıkla ilgili temel düşünce, prosesin davranışını belirleyen olasılık kurallarının zaman boyunca değişmemesidir. Bir anlamda, proses istatistiksel olarak dengededir.
- Durağan olmayan proseslerde serinin bir süre ortalamadan uzaklaşsa bile tekrar döneceği uzun dönem ortalama söz konusu değildir. Bu durumda, serinin varyansı zaman serisindeki pozisyona bağlıdır ve zaman ilerledikçe varyans ta sonsuza gider. Diğer bir deyişle, otokorelasyonlar zaman içinde küçülmez ya da çok yavaş küçülürler.
- Durağan serilerde şoklar geçicidir ve zaman ilerledikçe enerjisini kaybeder. Bir zaman sonra yeni gözlemlere etkide bulunmaz (Bakır, 2019).

### **2.3.4. Durağanlık ölçüsü**

Bir zaman serisinin gecikmeli değerleri arasındaki birlikte değişimin ölçüsü olan kovaryansa otokovaryans, korelasyona ise otokorelasyon denmektedir. Otokorelasyon fonksiyonu (ACF), seride "ardışık korelasyon" olup olmadığını belirlemek için kullanılan bir istatistiktir. Dolayısıyla, zaman serilerinin durağanlığının incelenmesinde sıklıkla kullanılmaktadır (Bakır, 2019).

Korelasyon kavramının lineer ilişkinin bir ölçüsü olması, ardışık korelasyonların, zaman serilerinin davranışını anlamadaki önemini açıklamaktadır. Bu nedenle, rasgele değişken ve onun geçmiş değerleri arasındaki korelasyon (ardışık korelasyon), doğrusal zaman serileri analizinin temelini oluşturmaktadır. Bu korelasyon yapısına otokorelasyon denmektedir. Bir zaman serisinin otokorelasyon fonksiyonu, serinin otokovaryans fonksiyonu aracılığıyla elde edilebilir.

$\{Y_t : t \in T\}$  bir zaman serisi olmak üzere serinin otokovaryans fonksiyonu,

$$\begin{aligned}\gamma(k) &= \gamma_k = \text{Cov}(Y_t, Y_{t-k}) \\ &= E[(Y_t - \mu)(Y_{t-k} - \mu)] \\ &= \frac{1}{n-k} \sum_{t=1}^{T-k} (Y_t - \mu)(Y_{t-k} - \mu), k \in \mathbb{Z}\end{aligned}\quad (2.2)$$

Örnekleme otokovaryansı ise,

$$\begin{aligned}\hat{\gamma}_k &= \text{Cov}(Y_t, Y_{t-k}) \\ &= E[(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})] \\ &= \frac{1}{T-k} \sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}), k \in \mathbb{Z}\end{aligned}\quad (2.3)$$

biçiminde tanımlanır.

$\{Y_t : t \in T\}$  durağan bir zaman serisi olmak üzere, serinin otokovaryans fonksiyonu  $\gamma_k$  aşağıdaki özellikleri sağlar.

i.  $\gamma_k$  fonksiyonu simetriktir.  $\gamma_k = \gamma_{-k}$

ii. Her  $k$  için,  $|\gamma_k| \leq \gamma_0$

iii.  $\gamma_k$  fonksiyonu negatif olmayan tanımlıdır. Yani,  $\forall T, \forall a_j, j = 1, 2, \dots, T \in \mathbb{R}, \forall k_j, j = 1, 2, \dots, T \in \mathbb{Z}$  için,

$$\sum_{i=1}^T \sum_{j=1}^T a_i a_j \gamma(k_i - k_j) \geq 0$$

Otokovaryans fonksiyonu  $\gamma_k$  olmak üzere, serinin otokorelasyon fonksiyonu,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_{t-k})}}\quad (2.4)$$

olarak tanımlanır. Örnekleme otokorelasyonu ise  $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$  olarak ifade edilmektedir. Otokorelasyon fonksiyonu otokovaryans fonksiyonu aşağıdaki özellikleri sağlamaktadır.

i.  $\gamma_k$  fonksiyonu simetriktir.  $\rho_k = \rho_{-k}$

ii. Her  $k$  için,  $|\rho_k| \leq \rho_0$

iii.  $\rho_k$  fonksiyonu negatif olmayan tanımlıdır. Yani,  $\forall T, \forall a_j, j = 1, 2, \dots, T \in \mathbb{R}, \forall k_j, j = 1, 2, \dots, T \in \mathbb{Z}$  için,

$$\sum_{i=1}^T \sum_{j=1}^T a_i a_j \rho(k_i - k_j) \geq 0$$

Otokorelasyon fonksiyonu ( $\rho_k$ ) daha küçük dereceden otokorelasyonların da etkisini içerisinde barındırdığından  $X_t$  ile  $X_{t+k}$  arasındaki net korelasyonu vermemektedir. Ancak kısmi korelasyon katsayısı (PACF), kısa dönem gecikme değerleri sabitken  $X_t$  ile  $X_{t+k}$  arasındaki net korelasyonu vermektedir.  $X_t$  ile  $X_{t+k}$  arasındaki kısmi korelasyon katsayısı  $t$  ve  $t - k$  arasındaki gözlemler kümesi ( $\{X_{t-k+1}, \dots, X_{t-1}\}$ ) verilmişken  $X_t$  ile  $X_{t+k}$  arasındaki koşullu korelasyondur.  $k'$  ncı dereceden kısmi otokorelasyon katsayısı,

$$\hat{\rho}_{kk} = \begin{cases} \hat{\rho}_1 & , \quad k = 1 \\ \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\rho}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\rho}_{k-1,j} \hat{\rho}_j} & , \quad k = 2, 3, \dots \end{cases} \quad (2.5)$$

olarak tanımlanmaktadır (Bakır, 2019).





### 3. SES SINYALLERİ ARASINDAKİ BENZERLİĞİN TESPİTİNE İSTATİSTİKSEL YAKLAŞIM

Ses sinyallerinin durağan olmayan zaman serileri formunda olduğuna ilişkin kanıtlar çalışmanın önceki bölümlerinde sunulmuştur. Ses sinyalleri arasındaki benzerlik yapısının araştırılması esasında zaman serileri arasındaki benzerliğin araştırılmasına denk olacaktır. Ancak zaman serileri arasındaki benzerliğin belirlenmesinde kullanılacak olan uzaklıklar çoğunlukla durağanlık varsayımına dayanmaktadır. Ses sinyallerin durağan olmayışı, zaman serilerinde benzerliğe ilişkin literatürde önerilen uzaklıkların çoğunu kullanılamaz hale getirmektedir. Varyansta durağan olmama (heteroskedasticity) söz konusu olduğunda, akıllara ilk olarak çeşitli dönüşüm metodları ile (Örneğin, Box-Cox dönüşümü) varyansa durağan hale getirmek gelebilir. Ancak bu işlem, ses sinyalinin dijitalleştirilmesi sonucu elde edilen zaman-genlik verisine uygulandığı takdirde, sinyalin frekans karakteristiğini değiştirmektedir. Ses sinyalinin frekans dağılımının değişmesi, sesin değişmesi anlamına geldiğinden bu yöntemlerin uygulanması sakıncalıdır.

Durağanlık varsayımına dayanmayan benzerlik ölçüleri arasında en bilineni Dinamik Zaman Bükme (Dynamic Time Warping, DTW) algoritmasıdır. Bu algoritma, sinyalin zaman bağımlılığını ortadan kaldırarak sıradan uzaklıkları (Öklid vb.) kullanılabilir hale getirmektedir. Algoritmanın önemli özelliklerinden biriside farklı boyutlardaki sinyaller arası benzerliği de tespit edebilmesidir (Atasoy, 2011). Ancak, DTW "biçimsel" benzerliğin bir ölçüsüdür ve ses sinyalinin zaman-genlik verilerinin benzerliği bu iki sinyalin müzikal olarak benzer olduğu anlamına gelmediğinden (Örneğin, bir ritim kaydı ile bir eser icrası zaman-genlik bazında aynı olabilmektedir.) çalışmada kullanılmamıştır.

Çalışmanın ana odağını oluşturan Itakura-Saito uzaklığı, Fourier dönüşümü tabanlı oluşu sebebiyle sinyalin durağan olmasını gerektirse de, sinyalin frekans alanı üzerinde tanımlı bir uzaklık ölçüsü olduğundan büyük öneme sahiptir. Fourier dönüşümünde durağanlık varsayımı çeşitli uyarlamalar ile ortadan kaldırılabilirdiğinden, bu uzaklık ölçüsü kullanılabilir hale gelmektedir. Çalışmanın bu bölümünde durağan zaman serileri arasındaki benzerliğin bir ölçüsü olan Itakura-Saito uzaklığının temel prensipleri ve

metodolojisinden bahsedilerek, bu uzaklığın durağan olmayan zaman serilerinde kullanılabilmesine ilişkin önerilecek olan yöntem için zemin hazırlanacaktır.

### 3.1. Itakura-Saito Uzaklığı (ISU)

Itakura – Saito uzaklığı iki olasılık dağılım fonksiyonu arasındaki farkın simetrik olmayan bir ölçüsüdür. Spektrum, frekansa göre genliklerin dağılımı olduğundan, Itakura – Saito uzaklığı, original spectrum ( $P(\omega)$ ) ile bu spektrumun kestirimi ( $\hat{P}(\omega)$ ) arasındaki farkın bir ölçüsü olarak kullanılabilir ve tek bir sinyal için sinyalin, kestirilen sinyale uzaklığı olarak Eş. 3.1'deki gibi ifade edilir.

$$d_{IS}(P(\omega), \hat{P}(\omega)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[ \frac{P(\omega)}{\hat{P}(\omega)} - \log \left( \frac{P(\omega)}{\hat{P}(\omega)} \right) - 1 \right] d\omega \quad (3.1)$$

Burada  $d_{IS}$  Itakura-Saito uzaklığını,  $\omega$  frekansı,  $P(\omega)$  ise güç spektrumunu sembolize etmektedir (Itakura ve Saito, 1968). Zaman serilerinin kesikli zaman-frekans bölgesinde tanımlı olmaları sebebiyle, uygulamada Itakura – Saito uzaklığı sinyalin Fourier spektrumları üzerinden Eş. 3.2'deki gibi hesaplanmaktadır.

$$d_{IS}^{(a)}(Spec_1, Spec_2) = \sum \left\{ \frac{Spec_1}{Spec_2} - \log \left( \frac{Spec_1}{Spec_2} \right) - 1 \right\}$$

$$d_{IS}^{(b)}(Spec_2, Spec_1) = \sum \left\{ \frac{Spec_2}{Spec_1} - \log \left( \frac{Spec_2}{Spec_1} \right) - 1 \right\} \quad (3.2)$$

$$d_{IS}^* = \frac{1}{2} \left\{ d_{IS}^{(a)}(Spec_1, Spec_2) + d_{IS}^{(b)}(Spec_2, Spec_1) \right\}$$

Burada  $d_{IS}^{(a)}$  birinci sinyalin ikinci sinyale uzaklığını,  $d_{IS}^{(b)}$  ikinci sinyalin birinci sinyale uzaklığını,  $d_{IS}^*$  ortalama Itakura-Saito uzaklığını göstermektedir. Itakura-Saito uzaklığını uzaklığın iki yönlü hesaplanmasının sebebi, her bir sinyalin, diğerinin kestirimi olarak varsayılmasından ileri gelmektedir. Bu bakış açısı ile temelde kendi kestirimlerine olan uzaklıkları elde edilmektedir ve elde edilen uzaklıklar eşit olmak zorunda değildir. Bu nedenle Itakura-Saito uzaklığı benzerliğin simetrik olmayan bir ölçüsüdür.

Bu tanımlara ek olarak Itakura-Saito uzaklığı,  $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + e_t$  ve

$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + v_t$  biçimindeki iki otoregresif süreç arasındaki benzerliğin bir ölçüsü olarak özelleştirilebilmektedir. Bu durumda otoregresif süreçler için Itakura-Saito uzaklığı Eş. 3.3'de verildiği gibidir.

$$d_I(\alpha, \beta) = \log \left( \frac{MSE_{x,y}}{MSE_{x,x}} \right) = \log \left( \frac{\beta^T R_X \beta}{\alpha^T R_X \alpha} \right)$$

$$d_I(\beta, \alpha) = \log \left( \frac{MSE_{x,y}}{MSE_{x,x}} \right) = \log \left( \frac{\alpha^T R_Y \alpha}{\beta^T R_Y \beta} \right) \quad (3.3)$$

$$d_I^*(\alpha, \beta) = \frac{1}{2} [d_I(\alpha, \beta) + d_I(\beta, \alpha)]$$

Burada,  $e_t \sim WN(0, \sigma_1^2)$  ve  $v_t \sim WN(0, \sigma_2^2)$  serileri beyaz gürültü serisini,  $\alpha = (1, \alpha_1, \dots, \alpha_p)$  ve  $\beta = (1, \beta_1, \dots, \beta_p)$  modellerdeki bilinmeyen parametre vektörlerini olsun.  $R_X$  ve  $R_Y$  ise süreçlerin otokorelasyon matrislerini simgelemektedir (Itakura, 1975).

Eş 3.1 ve Eş. 3.2'de verilen Itakura-Saito uzaklıkları temel olarak klasik Fourier dönüşümü ile elde edilen spektrumlara dayanmaktadır.

### 3.2. Fourier Dönüşümü

İnsan kulağı, sesi frekans tayfına (farklı perdelerdeki ses miktarları) çevirerek dönüşümü gerçekleştirir. Beyin ise bu bilgiyi algılanmış sese çevirir. Benzer şekilde Fourier dönüşümü, her frekansa ait sinüs dalgası için genlik ve evre hesaplayan bir fonksiyon olarak düşünülmelidir. Fourier analizi, uzay-zaman alanındaki bir fonksiyonu frekans, genlik ve evresi değişen sinüzoidal elemanlara indirger. Fourier dönüşümü ise, her frekanstaki genlik ve evreyi gösteren bir fonksiyondur. Dönüşüm, sinyalin (ya da fonksiyonun) sürekli veya kesikli (ayrık) zaman aralıklarında olmasına bağlı olarak hesaplanır. Sinyal kesikli zaman aralıklarına ait değerlerden oluşuyorsa, ayrık frekanslardaki sinüzoidal fonksiyonların serisine, sürekli ise tüm frekanslardaki sinüslerin Fourier integraline indirgenebilir. Burada frekanslar, orijinal frekansın iki, üç ve daha büyük katlarıdır. Fourier dönüşümü, sinyalin sürekli veya kesikli zaman aralıklı olmasından bağımsız olarak, bir nesneyi ortogonal iki uzay arasında eşlemektedir (Cömert, 2015).

### 3.2.1. Sürekli zamanlı Fourier dönüşümü

Sürekli zamanlı Fourier dönüşümü Eş. 3.4'deki biçimde hesaplanmaktadır.

$$X(k) = \int_{-\infty}^{\infty} x(t) e^{-ikt} dt \quad (3.4)$$

Burada  $t$  zaman indisi,  $i = \sqrt{-1}$  ve  $X(k)$  ise  $x(t)$  sinyalinin sürekli zamanlı Fourier dönüşümünü göstermektedir.

### 3.2.2. Ayrık zamanlı Fourier dönüşümü

Ayrık zamanlı Fourier dönüşümü Eş. 3.5'deki biçimde hesaplanmaktadır.

$$X(k) = \sum_{t=0}^{T-1} x(t) e^{-i2\pi kt/T}, \quad k = 0, \dots, T-1 \quad (3.5)$$

Burada,  $T$  gözlem sayısı,  $t$  zaman indisi,  $i = \sqrt{-1}$  ve  $X(k)$  ise  $x(t)$  sinyalinin ayrık zamanlı Fourier dönüşümüdür.

Her iki durumda da Fourier dönüşümü kompleks sayılar kümesinde tanımlıdır ve sinyalin frekans bölgesinin elde edebilmesi için, bu kompleks değerli çıktının modülüsünün bulunması gerekir. Bu işlemin sonucunda elde edilen reel değerli fonksiyona ise Fourier Spektrum'u denir ve Eş. 3.6'deki biçimde gösterilir.

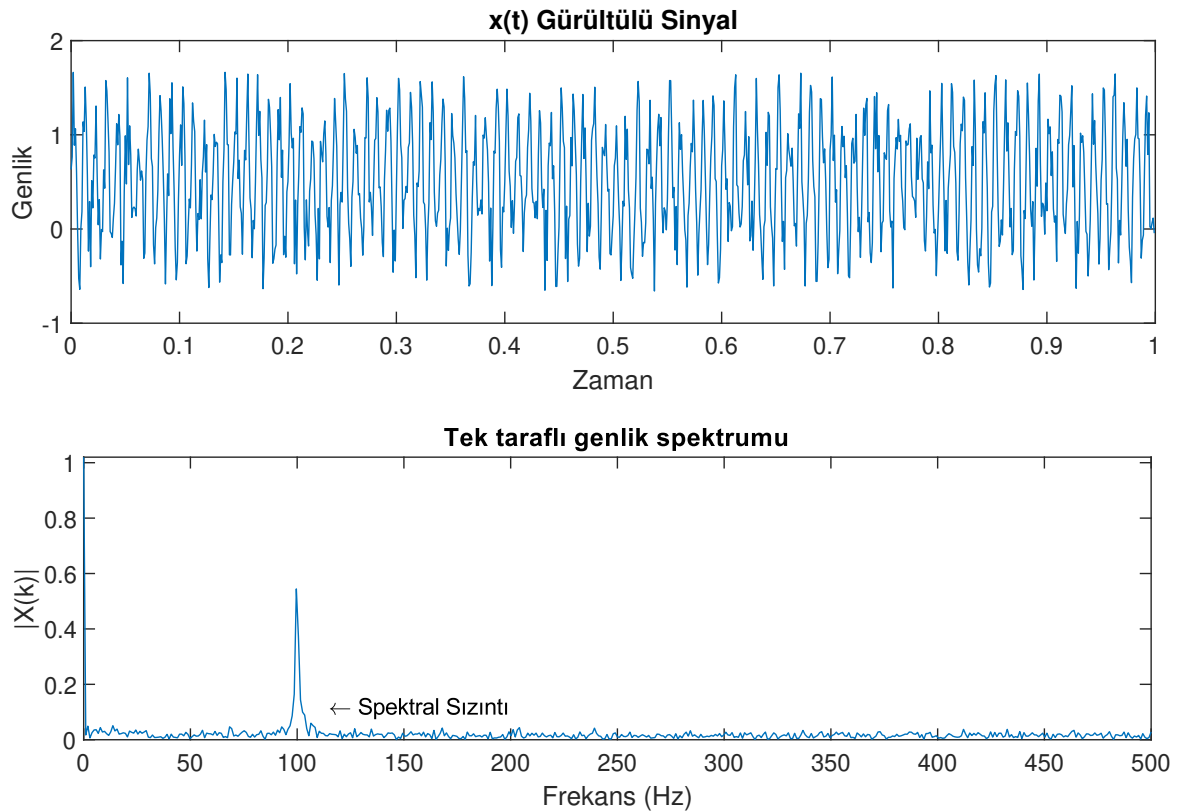
$$P_{xx}(k) = |X(k)|^2 \quad (3.6)$$

Fourier dönüşümü sinyalin içerdiği veri sayısına bağlı olarak maaliyetli bir işlemdir. Örneğin örnekleme frekansı 1000 Hz olan bir sinyalde birim zamanda 1000 veri okunmaktadır. Bu sinyalin sadece bir saniyesinin Fourier dönüşümü için bir milyon çarpma işlemi ve 1000 toplama işlemi yapılmaktadır. Sinyalin örnekleme frekansının ve süresinin artması ise bu işlemi güncel teknolojiye sahip bilgisayarlar için bile bir hayli zor olacaktır. Bu durumun önüne geçmek adına Hızlı Fourier dönüşümü (Fast Fourier Transform, FFT) geliştirilmiştir.

Hızlı Fourier dönüşümünün en önemli özelliği çarpımların sayısını azaltarak zamandan

kazanmasıdır. Bu işlemi örnek sayısını oranlayarak yapmaktadır. Örneğin 64 örnekten oluşan bir sinyalde  $64^2 = 4096$  adet çarpma işlemi gerekmektedir. Ancak, sinyal 32 birimlik iki parça ayrıldığında her bir parça için  $32^2 = 1024$  çarpma işlemi gerekecektir. Böylece, toplamda 2048 adet çarpma işlemi yapılacak olup, normalde gereken çarpma sayısı yarı yarıya düşürülmüş olacaktır (Cömert, 2015).

Fourier dönüşümü global bir dönüşümdür ve sinyal içerisindeki sabit dalgalanmaları belirlemek için yeterlidir. Bu özelliği dolayısıyla Fourier dönüşümü, sinyalin durağan olmasını gerektirmektedir. Fourier dönüşümünün durağan olmayan sinyallere uygulanması durumunda spektral sızıntı adı verilen problem ile karşılaşmaktadır. Spektral sızıntı, tek frekansa sahip bir sinyalin frekans spektrumunda, sahip olduğu frekans değerine yakın frekansları da örten bir eğri ile karşılaşılması durumudur. Spektral sızıntı problemi, sinüzoidal fonksiyonlar yardımıyla hipotetik olarak üretilen, tek frekansa sahip ve gürültü eklenmiş spektral ton üzerinden Şekil 3.1’de görselleştirilmiştir.



Şekil 3.1. Spektral sızıntı problemi

Daha önceki bölümde bahsedildiği üzere, ses sinyalleri durağan değildir. Bu nedenle klasik

Fourier dönüşümünün uygulanması bir çok problemi beraberinde getirecektir. Ancak sinyalin durağan olmaması durumunda da doğru frekans spektrumları elde etmek mümkündür.

### 3.3. Kısa Zamanlı Fourier Dönüşümü (STFT)

Kısa-zamanlı Fourier dönüşümü (Short-Time Fourier Transform, STFT), durağan olmayan bir sinyalin frekans içeriğinin zaman içinde nasıl değiştiğini analiz etmek için kullanılır. Başka bir deyişle bir sinyalin zaman bağımlı Fourier dönüşümü olarak da bilinir. Klasik Fourier dönüşümünün, sinyalin geçici (transitory) ya da durağan olmayan karaktere sahip olduğu durumdaki yetersizliğinin önüne geçmek adına Dennis Gabor (1947) Fourier dönüşümünü, sinyalin kısa-zamanlı bölümlerini inceleyecek şekilde uyarlamış ve yöntemi kısa-zamanlı Fourier dönüşümü olarak literatüre kazandırmıştır. Allen ve Rabiner (1977) sinyalin kısa-zamanlı bölümlerinin incelenmesinde pencere fonksiyonu kullanma fikrini önererek, kısa-zamanlı Fourier dönüşümüne kuvvet kazandırmıştır. Ayrıca, Allen (1982) tarafından yapılan çalışmada, kısa-zamanlı Fourier dönüşümü konuşma sinyalleri üzerinde sınanmış ve yöntemin başarısı tescil edilmiştir.

Allen ve Rabiner' in (1977) önerdiği yönteminin temeli, hareket eden pencere Fourier dönüşümüne dayanmaktadır. Pencere, sinyalin durağan olmamasından kaynaklı oluşan spektral sızıntının önüne geçmek için oluşturulan bir çeşit alt bölgedir. Pencere zaman alanında (time-domain) hareket eder ve sinyalin frekans içeriğinin incelenmesi için "spektrogram" denilen iki-boyutlu zaman-frekans dağılımını üretir. Fourier dönüşümüne benzer olarak, sinyalin sürekli veya ayrık zamanlı olmasına bağlı olarak hesaplanır. Sürekli zamanlı sinyalin STFT'si Eş. 3.7'deki biçimde tanımlanmaktadır.

$$X(\tau, f) = \int_{-\infty}^{\infty} x(t) h^*(t - \tau) e^{-i2\pi ft} dt \quad (3.7)$$

Burada  $h(\tau)$ ,  $\tau$  anında merkezlenmiş pencere fonksiyonudur (window function). Bu fonksiyon ile sinyal,  $\tau$  anına yakın olduğu anda kesilir ve Fourier dönüşümü ile bu zaman etrafında anlık bölgesel tahminler elde edilir.

STFT'nin hesaplanmasının en alışlagelmiş yolu sıfır etrafında merkezlenmiş ve  $\int_{-\infty}^{\infty} |h(\tau)|^2 d\tau = 1$  gücüne sahip olan sabitlenmiş, pozitif, çift yönlü pencere kullanmaktır.

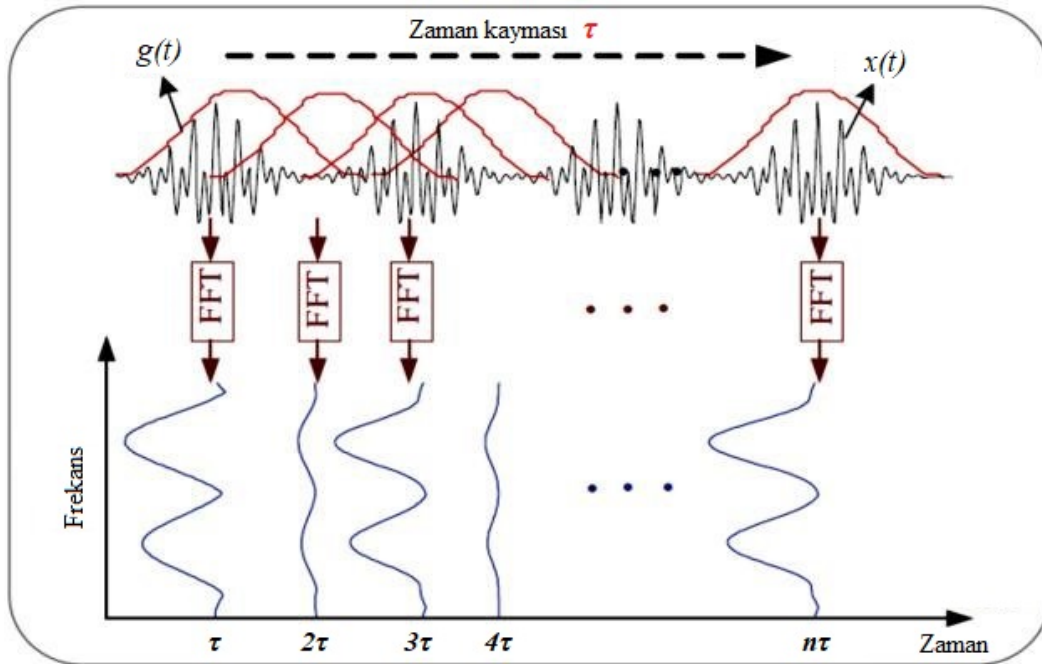
Ayrık zamanlı bir sinyalin STFT'si ise Eş. 3.8'deki biçimde tanımlanmaktadır.

$$X(\tau, f) = \sum_{-\infty}^{\infty} x(t) h^*(t - \tau) e^{-i2\pi ft} \quad (3.8)$$

Her iki durum için spektrogram, klasik Fourier dönüşümüne benzer olarak, Eş. 3.9'daki gibi elde edilmektedir.

$$S_X(\tau, f) = |X(\tau, f)|^2 \quad (3.9)$$

Eş. 3.9, zaman-değişken ve durağan olmayan sinyaller için sıklıkla kullanılır. Kısa-zamanlı Fourier dönüşümü ile bir sinyalin analizi Şekil 3.2 de gösterilmektedir.



Şekil 3.2. STFT ile sinyalin analizi (Jaber and Bicker, 2018)

Şekil 3.2 incelendiğinde pencere fonksiyonunun sinyalin kısa-zamanları üzerindeki etkisi ve sinyalin kısa-zamanlı Fourier dönüşümü ile elde edilen frekans spektrumunun yapısı açıkça görülmektedir. Özetle, sinyalin STFT ile elde edilen spektrumu zamanın her bir kesitindeki Fourier spektrumundan oluşmaktadır.

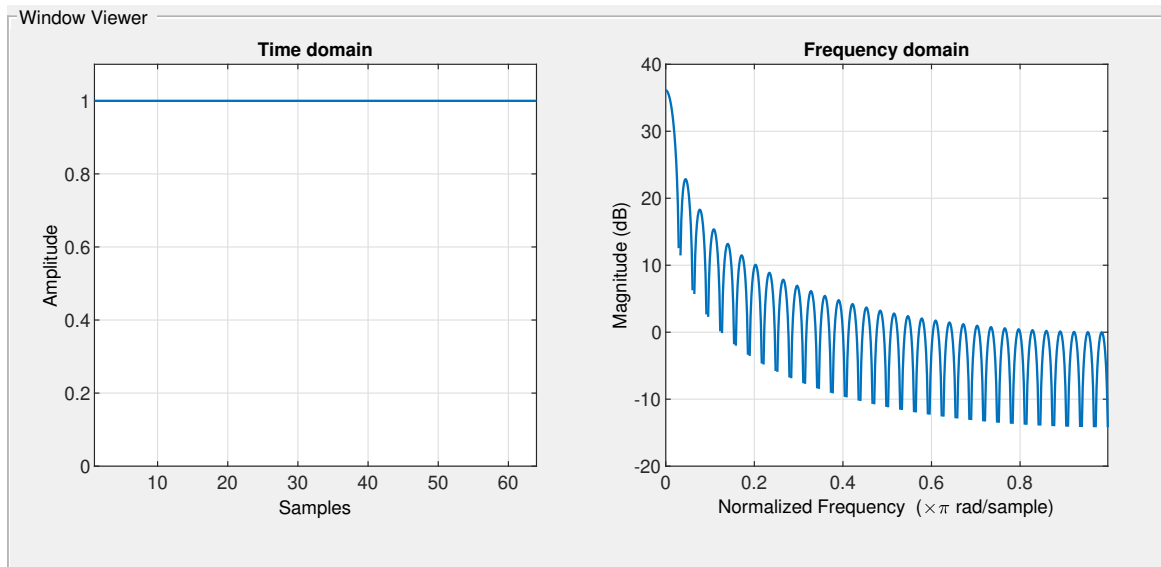
### 3.4. Pencere fonksiyonları

Sinyallerin alt karakteristik yapılarını ön plana çıkarmak için sinyal belli sayıda örnek içeren parçalara ayrılır. Bu parçaların herbirine pencere (window) denir. Pencere fonksiyonları, sinyal parçalarını ağırlıklandırarak, sinyal parçasının orta kısımlarının vurgulanması sağlayan fonksiyonlardır. Pencere fonksiyonu uygulandığı parçanın başlangıç ve bitiş bölgelerine yakın kısımları söndürmektedir. Bu özelliği sayesinde pencere fonksiyonları, spektral sızıntıyı zayıflatmak ve yüksek otokorelasyonu azaltmak için kullanılabilir bir araçtır. Pencere fonksiyonları "sabit" ve "ayarlanabilir" olarak iki gruba ayrılır.

#### 3.4.1. Sabit pencere fonksiyonları

Bu türden pencere fonksiyonları yaygın olarak sinyal işleme uygulamalarında tercih edilmektedir. Dörtgen, Hamming, Hann, Blackman, Bartlett yaygın biçimde kullanılan sabit pencere fonksiyonlarıdır. Sıklıkla kullanılan bu sabit pencere fonksiyonlarının, uygulandıkları sinyale zaman ve frekans alanlarında nasıl bir manipülasyon yaptıkları aşağıda açıklanmaktadır.

1. Dikdörtgen pencere fonksiyonu  $h(t) = \mathbf{1}_L$  biçimindedir. Burada  $L$  pencere uzunluğudur.



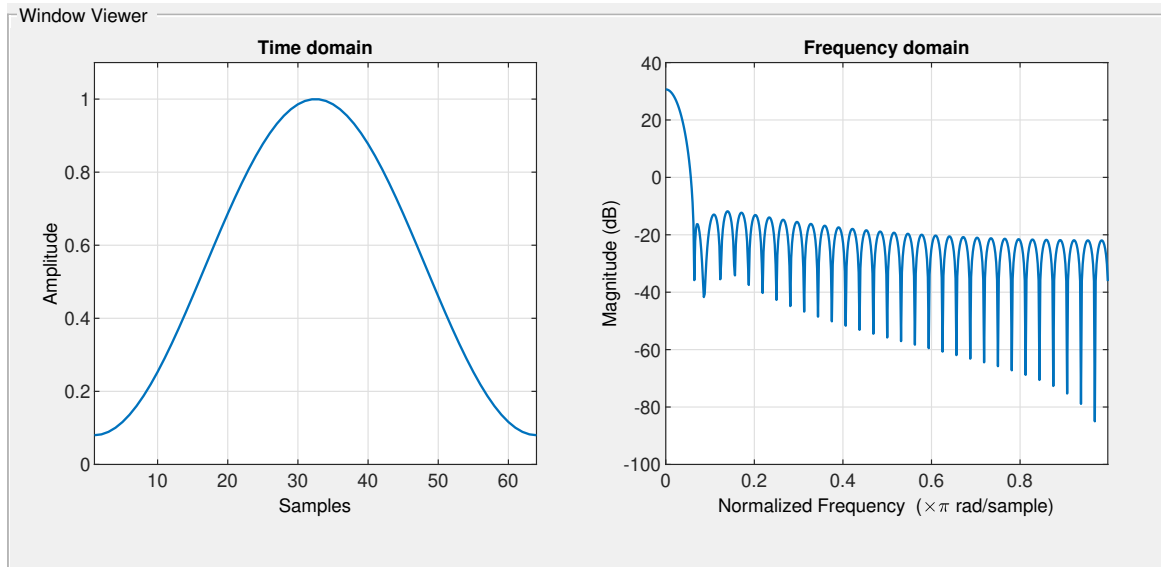
Şekil 3.3. Dikdörtgen pencere fonksiyonu ( $L = 64$ )



Sinyalin bir parçasını, dikdörtgen pencere fonksiyonu ile ağırlıklandırmak, Şekil 3.3'deki sol taraftaki grafikte, sinyalin belirli bir parçasının aynen alındığı (yani eşit ağırlıklandırıldığı), kalan kısmının ise sıfırlandığı bir sinyal görülmektedir. Sağ tarafında ise sinyali bu parçasına Dikdörtgen pencere fonksiyonunun uygulanması ile elde edilen pencere vektörüne ilişkin Fourier dönüşümünün karesinin desibele (dB) dönüştürülmüş hali görülmektedir.

2. Hamming pencere fonksiyonu: Konuşma sinyallerinin işlenmesi uygulamalarında sıklıkla kullanılan bir sabit pencere fonksiyonudur. Hamming pencere fonksiyonu, bağımsız değişken olarak sadece pencere uzunluğuna sahip olup, Eş. 3.10'daki gibi tanımlanır.

$$h(t) = 0.54 - 0.46 \cos\left(2\pi \frac{t}{L}\right), \quad 0 \leq t \leq L \quad (3.10)$$

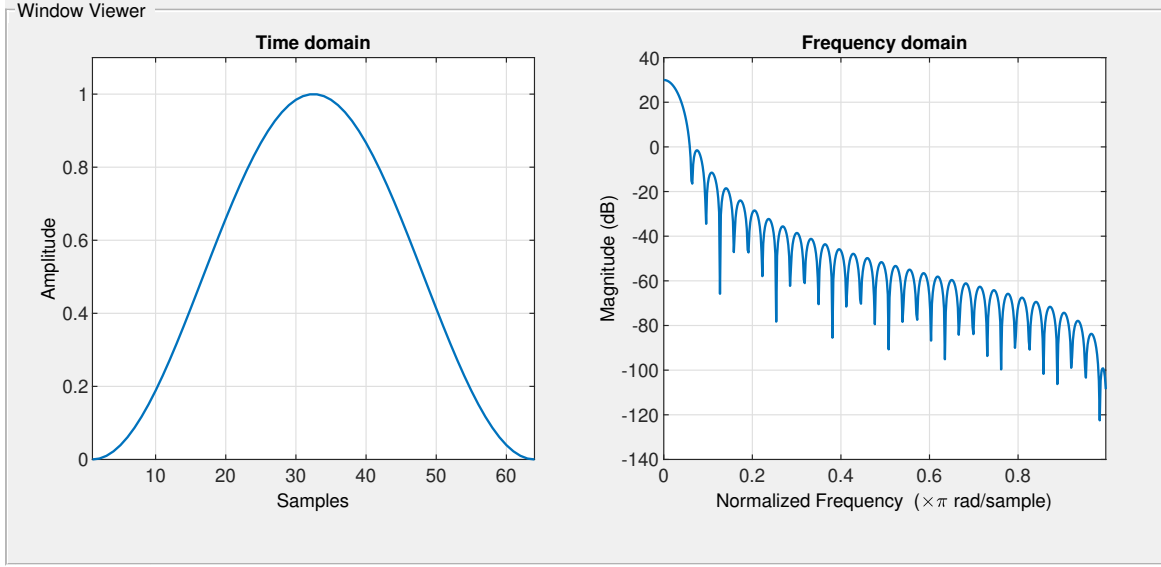


Şekil 3.4. Hamming pencere fonksiyonu ( $L = 64$ )

Şekil 3.4'in sol kısmında pencere genişliği 64 örnek olan bir Hamming penceresi, sağ kısmında ise bu pencere fonksiyonu ile sinyal parçasının ağırlıklandırılması ile elde edilen pencere vektörüne ilişkin Fourier dönüşümünün karesinin desibele (dB) dönüştürülmüş hali görülmektedir. Hamming pencere fonksiyonunda, Dikdörtgen pencere fonksiyonunun aksine sinyal parçasının orta kısımları daha fazla ağırlıklandırılmaktadır.

3. Hann pencere fonksiyonu, Hamming pencere fonksiyonu gibi konuşma sinyallerinin analizinde sıklıkla kullanılmaktadır. Hann pencere fonksiyonu Eş. 3.11'deki biçimde tanımlanır.

$$h(t) = 0.5 \left( 1 - \cos\left(2\pi\frac{t}{L}\right) \right), \quad 0 \leq t \leq L \quad (3.11)$$



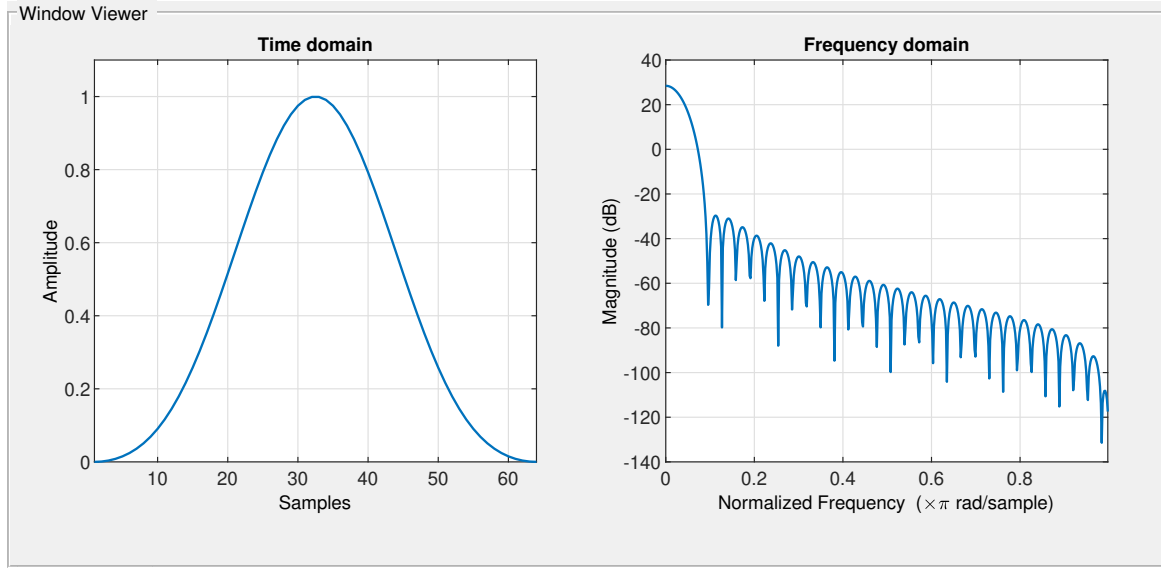
Şekil 3.5. Hann pencere fonksiyonu ( $L = 64$ )

Hann pencere fonksiyonu, Hamming pencere fonksiyonuna nazaran, sinyal parçasının orta kısımlarını daha fazla ağırlıklandırmaktadır. Başka bir deyişle, varyansı Hamming pencere vektörünün varyansından daha küçüktür.

4. Blackman pencere fonksiyonu Eş. 3.12'de tanımlanmaktadır.

$$h(t) = 0.42 - 0.5\cos\left(\frac{2\pi t}{L-1}\right) + 0.08\cos\left(\frac{4\pi t}{L-1}\right), \quad 0 \leq t \leq M-1 \quad (3.12)$$

Burada  $L$  pencere uzunluğu,  $M$  ise  $T$  çift sayı iken  $\frac{L}{2}$ , tek sayı iken  $\frac{L+1}{2}$  dir.

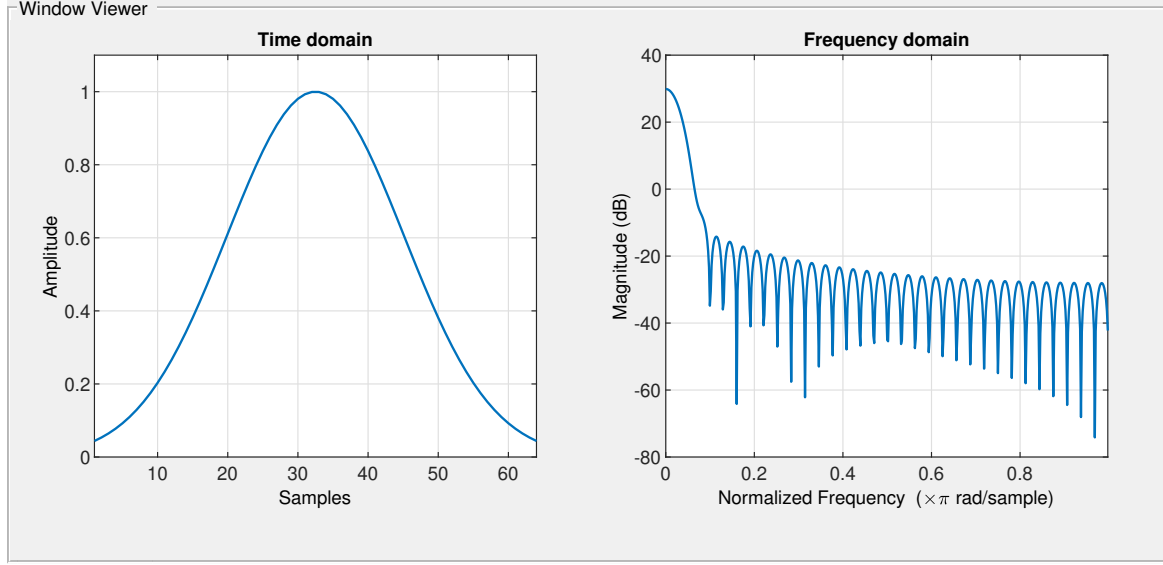


Şekil 3.6. Blackman pencere fonksiyonu ( $L = 64$ )

Blackman pencere fonksiyonu, Hann pencere fonksiyonunun uyarlanmış bir halidir. Blackman pencere vektörü, Hann pencere vektörüne nazaran daha düşük varyansa sahip olup, böylece orta kısımlara çok yakın örneklerin etkisi artırılırken, uzak örneklerin etkisi azaltılmıştır.

5. Gaussian pencere fonksiyonu: Bu pencere fonksiyonu normal dağılımdan türetilmiştir.  $L$  pencere genişliği ve  $\sigma$ , gaussian rastgele değişkenin standart sapması olması üzere Eş. 3.13'deki biçimde tanımlanmaktadır.

$$h(t) = e^{-t^2/2\sigma^2}, \quad -(L-1)/2 \leq t \leq (L-1)/2 \quad (3.13)$$



Şekil 3.7. Gaussian pencere fonksiyonu ( $L = 64$ )

Gaussian pencere fonksiyonu normal dağılım karakteristiğine sahip, iki parametrelidir (pencere genişliği ve varyans değeri) bir pencere fonksiyonudur. Bu yapısı ile diğer sabit pencere fonksiyonlarının aksine varyans kontrolüne sahiptir.

Sabit pencerelerin genel özelliği sinyal boyunca aynı kalmasıdır ve bu yapısı nedeniyle, sabit pencere fonksiyonlarının karmaşık sinyallerde etkin bir biçimde çalışmaz. Bu durumun üstesinden gelebilmek için ayarlanabilir (adaptive) pencere fonksiyonları önerilmiştir.

### 3.4.2. Ayarlanabilir pencere fonksiyonları

Mevcut literatürde önerilmiş çeşitli ayarlanabilir pencere fonksiyonları bulunmaktadır. Bunların en çok bilinenleri Dolph-Chebyshev, Kaiser, Saramaki ve ultraspherical ile son zamanlarda geliştirilen üstel, cosh, modifiye edilmiş cosh, modifiye edilmiş Kaiser pencere fonksiyonlarıdır. Bu çalışmada, desibel değerine göre ayarlanabilir parametreye sahip Kaiser pencere fonksiyonu, müzikal ses sinyalleri gibi değişken ve fazla alt karakteristik yapıya sahip sinyallerin, frekans bölgesindeki değişimleri etkili bir biçimde ortaya çıkarmasından dolayı tercih edilmiştir. Kaiser pencere fonksiyonu Eş. 3.14'deki gibi tanımlanır.

$$h_k(t) = \frac{I_0\left(\alpha_k \sqrt{1 - \left(\frac{2t}{L-1}\right)^2}\right)}{I_0(\alpha_k)}, \quad |t| \leq \frac{L-1}{2} \quad (3.14)$$

Burada  $L$  pencere genişliği,  $\alpha_k$  bağımsız ayarlanabilir parametreyi ve  $I_x$  ise sıfır-dereceli birinci tür geliştirilmiş Bessel fonksiyonunu göstermektedir (Avcı, 2014). Bessel fonksiyonu bir kuvvet serisidir ve Eş. 3.15'deki biçimde tanımlanmaktadır.

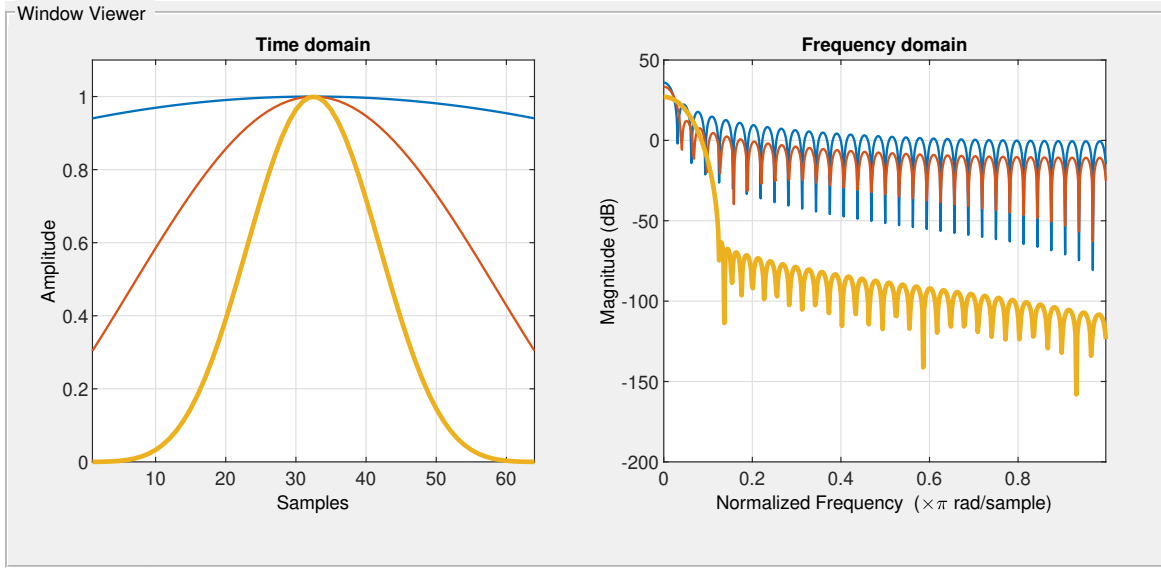
$$I_0(x) = 1 + \sum_{j=1}^{\infty} \left[ \frac{1}{j!} \left(\frac{x}{2}\right)^j \right]^2 \quad (3.15)$$

Sabit pencereler için pencere uzunluğunun artması analob (mainlob, sinyalin pencere fonksiyonu uygulanacak genişlikteki parçası) genişliğini artırırken, dalgalanma oranını sabit tutmaktadır. Ancak, Kaiser penceresindeki bağımsız ayarlanabilir parametredeki ( $\alpha_k$ ) artış, analob genişliğinde artışa, dalgalanmada ise azalmaya neden olur (Kaya ve İnce, 2011).

Ayarlanabilir parametre,  $k$  dB'lik yanlob (sidelob, sinyalin pencere fonksiyonu uygulanacak genişlikteki parçasının dışında kalan kısmı) azalmasına sahip FIR (Sonlu Dürtü Yanıtı, Finite Impulse Response) filtresini temsil eden bir Kaiser penceresi elde etmek için, yaygın olarak Eş. 3.16'daki biçimde kullanılmaktadır.

$$\alpha_k = \begin{cases} 0.1102(k-8.7) & , \quad k > 50 \\ 0.5842(k-21)^{0.4} + 0.07886(k-21) & , \quad 21 \leq k \leq 50 \\ 0 & , \quad k < 21 \end{cases} \quad (3.16)$$

Şekil 3.8'de Kaiser pencere fonksiyonunun, belirli ayarlanabilir parametreler için sinyalin ilgili bölümüne etkisi ve ağırlıklandırma sonucu elde edilen pencere vektörünün spektrumuna etkisi (desibel (dB) olarak) görülmektedir.



Şekil 3.8. Kaiser pencere fonksiyonu ( $L = 64, \alpha_k = (0.5, 2.5, 12)'$ )

Şekil 3.8'den görüleceği üzere, ayarlanabilir parametre  $\alpha_k$  arttıkça sinyalin orta kısımlarına daha fazla ağırlıklandırma uygulanmaktadır ve böylece sinyal parçasının frekans alanındaki etkisi sinyal boyunca sürekli olarak değişmektedir. Bunun bir sonucu olarak, fazla frekans yapısı içeren sinyaller etkili bir biçimde karakterize edilmektedir.

## 4. MÜZİKAL BENZERLİĞİN TESPİTİ İÇİN UYARLANMIŞ ITAKURA-SAITO UZAKLIĞI (UISU)

Itakura-Saito uzaklığı bilindiği gibi Fourier dönüşümü ile elde edilen spektrumlar üzerinden hesaplanmaktadır. Başka bir deyişle, Fourier dönüşümü tabanlı bir uzaklık hesaplama yöntemidir. Bu nedenle durağan olmayan sinyallerin benzerliğinde kullanılamaz. Bu bölümde, bahse konu problemin üstesinden gelebilmek için önerilen, kısa-zamanlı Fourier transformu tabanlı Itakura-Saito uzaklığı (UISU) açıklanmakta ve müzikal benzerliğin tespitindeki performansı gösterilmektedir.

Yöntem temel olarak kısa-zamanlı Fourier dönüşümüne dayanmaktadır. Klasik Itakura-Saito uzaklığında kullanılan Fourier spektrumları yerine kısa-zamanlı Fourier dönüşümü ile elde edilen spektrumlar kullanılarak sinyalin durağan olmamasından ileri gelen spektral sızıntı vb. problemlerin önüne geçilmiştir. Önerilen yöntem, klasik ISU'da kullanılan Fourier spektrumları yerine STFT spektrumlarının adapte edilmesinden dolayı "uyarlanmış Itakura-Saito uzaklığı" adı verilmiş ve Eş. 4.1'deki biçimde tanımlanmıştır.

$$d_{UIS}^{(a)}(S_1, S_2) = \sum \left\{ \frac{S_1}{S_2} - \log \left( \frac{S_1}{S_2} \right) - 1 \right\}$$

$$d_{UIS}^{(b)}(S_1, S_2) = \sum \left\{ \frac{S_2}{S_1} - \log \left( \frac{S_2}{S_1} \right) - 1 \right\} \quad (4.1)$$

$$d_{UIS}^* = \frac{1}{2} \left\{ d_{UIS}^{(a)}(S_1, S_2) + d_{UIS}^{(b)}(S_1, S_2) \right\}$$

Burada  $d_{UIS}^{(a)}$  birinci sinyalin ikinci sinyale uzaklığını,  $d_{UIS}^{(b)}$  ikinci sinyalin birinci sinyale uzaklığını,  $d_{UIS}^*$  ortalama uyarlanmış Itakura-Saito uzaklığını göstermektedir. Şekil 3.2' de gösterildiği gibi, kısa-zamanlı Fourier dönüşümünde, zamandaki kayma miktarı ( $\tau$ ) ve pencere genişliğine ( $L$ ) bağlı olarak,  $L\tau$  adet Fourier spektrumu elde edilmektedir. Bu sebeple, elde edilen  $L\tau$  adet Fourier spektrumu, toplam spektrumu elde etmek amacıyla toplanmış  $\left( S_i = \sum_j^{L\tau} Spec_{ij}, i = 1, 2 \right)$  ve bu genel spektruma Eş. 3.2'de tanımlanan Itakura-Saito uzaklığı uygulanarak, sinyallerin benzerliğini gösterecek şekilde Eş. 4.1'deki biçimde düzenlenmiştir. Uyarlanmış Itakura - Saito uzaklığı da bu nedenle toplam

spektruma dayalı toplam fonksiyonu biçiminde ifade edilmiştir. Uyarlanmış Itakura-Saito uzaklığı da, klasik Itakura-Saito uzaklığı gibi asimetrik bir uzaklık ölçüsüdür ve tanımı gereği alabileceği minimum değer sıfırdır. Yani Itakura-Saito uzaklığı sıfıra gittikçe ses sinyallerinin benzerliğinin arttığı biçiminde yorumlanır.

Takip eden kısımda, Uyarlanmış Itakura-Saito uzaklığının başarısı, hipotetik olarak üretilen sinyaller ve gerçek müzik eserlerinin dijitalleştirilmesi ile elde edilen ses sinyalleri üzerinde test edilmektedir.

#### 4.1. Simülasyon Çalışması

Çalışmanın bu kısmında kullanılacak olan ses sinyalleri, bilgisayar ortamında sinüzoidal fonksiyonlar aracılığıyla üretilen sabit genliğe sahip spektral tonların (notaların) farklı zaman, frekanslarda karıştırılması ile elde edilmiştir. Ayrıca, daha detaylı bir karşılaştırma yapabilmek için, bazı sinyallere orijinal yapısını bozmayacak genlikte gürültü eklenmiştir.

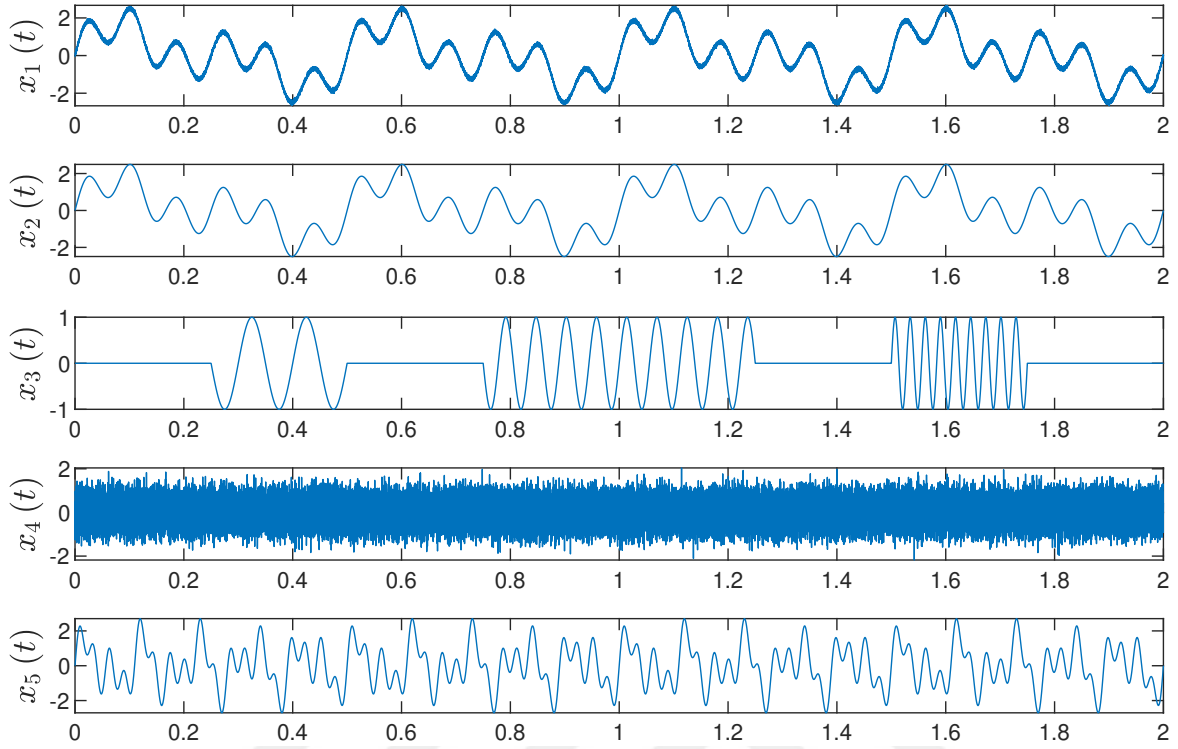
Ses sinyallerinin alt karakteristik yapısını oluşturacak spektral tonlar, MATLAB® programında  $f_1 = 2 \text{ Hz}$ ,  $f_2 = 4 \text{ Hz}$  ve  $f_3 = 12 \text{ Hz}$  frekansları ile 2'şer saniye,  $f_4 = 10 \text{ Hz}$ ,  $f_5 = 18 \text{ Hz}$  ve  $f_6 = 36 \text{ Hz}$  frekansları ile 0.25 ve 2'şer saniye olacak şekilde üretilmiş olup, CD kalitesinde ses elde etmek amacıyla örnekleme frekansı  $44100 \text{ Hz}$  alınmıştır. Her bir spektral tonun üretilmesi için kullanılan sinüzoidal fonksiyon Eş. 4.2' de verilmektedir.

$$x_j = \sin(2\pi f_j t), \quad j = 1, 2, 3, \quad t \in [0, 2] \quad (4.2)$$

Üretilen spektral tonlar, farklı kuvvetlerde benzerlikler yaratmak adına belirli formlarda karıştırılmış ve bazı karışımlara  $WN(0, \sigma^2)$  dağılımlı gürültü sinyali eklenmiştir. Dolayısıyla, birinci ses sinyali ( $x_1(t)$ ), üretilen tüm spektral tonların karışımına gürültü eklenmesi ile, ikinci ses sinyali ( $x_2(t)$ ), birinci sinyale kuvvetli benzerlik içermesi için, tüm spektral tonların karışımı ile, üçüncü ses sinyali ( $x_3(t)$ ), diğer sinyallere düşük benzerliğe sahip olması için, farklı frekanslarda ( $f_4, f_5$  ve  $f_6$ ) ve farklı zamanlarda (2 saniyelik zaman zarfı içerisinde, farklı zaman dilimlerinde meydana gelen 0.25 saniyelik spektral tonlar) bir araya getirilmesi ile ve son olarak dördüncü ses sinyali ( $x_4(t)$ ), düşük genlikte gerçekleşen seslerinde incelenebilmesi adına, birinci sinyale eklenen gürültü sinyalidir.



Böylece üretilen ses sinyalleri Şekil 4.1'deki gibi elde edilmiştir.



Şekil 4.1. Hipotetik ses sinyalleri

Şekil 4.1'deki hipotetik ses sinyalleri, ikili kombinasyonlar biçiminde klasik ve uyarlanmış Itakura-Saito uzaklıkları ile karşılaştırılmış olup, karşılaştırma sonuçları Çizelge 4.1'te verilmektedir.

Çizelge 4.1. Hipotetik sinyallerin ISU ve UISU karşılaştırması

	<i>ISU</i>	<i>UISU</i>
Ses 1 & Ses 2	2.9978e+09	0.01398
Ses 1 & Ses 3	5991.79267	85.66928
Ses 1 & Ses 4	2.27999	0.05775
Ses 1 & Ses 5	1.5813e+04	913.04485
Ses 2 & Ses 3	0.84598	23.42529
Ses 2 & Ses 4	14.50828	1762.01254
Ses 2 & Ses 5	0.82849	601.11536
Ses 3 & Ses 4	12.90019	563.51494
Ses 3 & Ses 5	7.87502	181.82409
Ses 4 & Ses 5	1.1509e+10	8443.50612

Çizelge 4.1'deki sonuçlar incelendiğinde, önerilen uzaklık UISU'nun, tasarlanan tüm benzerlik yapılarını ortaya çıkardığı görülmektedir.

Dolayısıyla, uyarlanmış Itakura-Saito uzaklığı, klasik Itakura-Saito uzaklığının sinyalin durağan olmamasına bağlı yanılgılarını aştığı ve son derece anlamlı sonuçlar verdiği gözlemlenmiştir. Itakura-Saito uzaklığı uygulanmış olduğunda, normal olarak sinyal 1, sinyal 2 ve sinyal 1 sinyal 4 arasında kuvvetli benzerlik, diğer sinyaller arasında ise benzer olmama beklenir. Halbuki, en çok benzerlik olan sinyal 1 ve sinyal 2 arasında sıfıra doğru giden bir uzaklık beklenirken, ISU diğerlerine nazaran en yüksek değeri alarak bu sinyallerin en benzemez sinyaller olduğunu ortaya koymuştur. Bunun aksine UISU beklenti doğrultusunda davranarak en düşük uzaklık ölçüsünü vermiş ve 0.01398 gibi hemen hemen sıfır olan bir değerle bu iki sinyalin birbirine çok yakın olduğunu göstermiştir. Ayrıca, kuvvetli benzerliğin beklendiği diğer iki sinyal (sinyal 1 ve sinyal 4) için, her iki uzaklık ölçüsünde (UISU ile elde edilen uzaklık değeri sıfıra daha yakın olmak üzere) anlamlı değerler vermiştir. Ancak, ISU ile elde edilen 2.27999 değeri, ISU ölçüsünün, sinyalin durağan olmama durumundan kaynaklı tutarsızlığının bir sonucudur ve güvenilir değildir. En yüksek benzemezliğin beklendiği sinyaller arasında ise UISU'ya nazaran çok daha küçük değerler vererek nispeten bir benzerlik durumunu ortaya koymuştur. UISU ise bu sinyaller için benzer sinyallerin ölçüsüne nazaran çok daha yüksek benzerlik ölçülerini vererek benzemezliği başarılı bir biçimde yakalamıştır. Ancak benzerliği düşük olan sinyal 4, sinyal 5 arasındaki uzaklık incelendiğinde, her iki uzaklık ölçüsünün de başarılı sonuç verdiği görülmektedir. Bu durum, ilk bakışta sinyal 4 ün beyaz gürültü süreci olmasından kaynaklı gibi görünse de, benzer duruma sahip sinyal 2 ve sinyal 4 arasındaki uzaklığı dikkate alarak, ISU'nun yanıltıcı olduğu kanısına varılmaktadır.

#### **4.2. Gerçek Müzik Sinyallerinin Benzerliği**

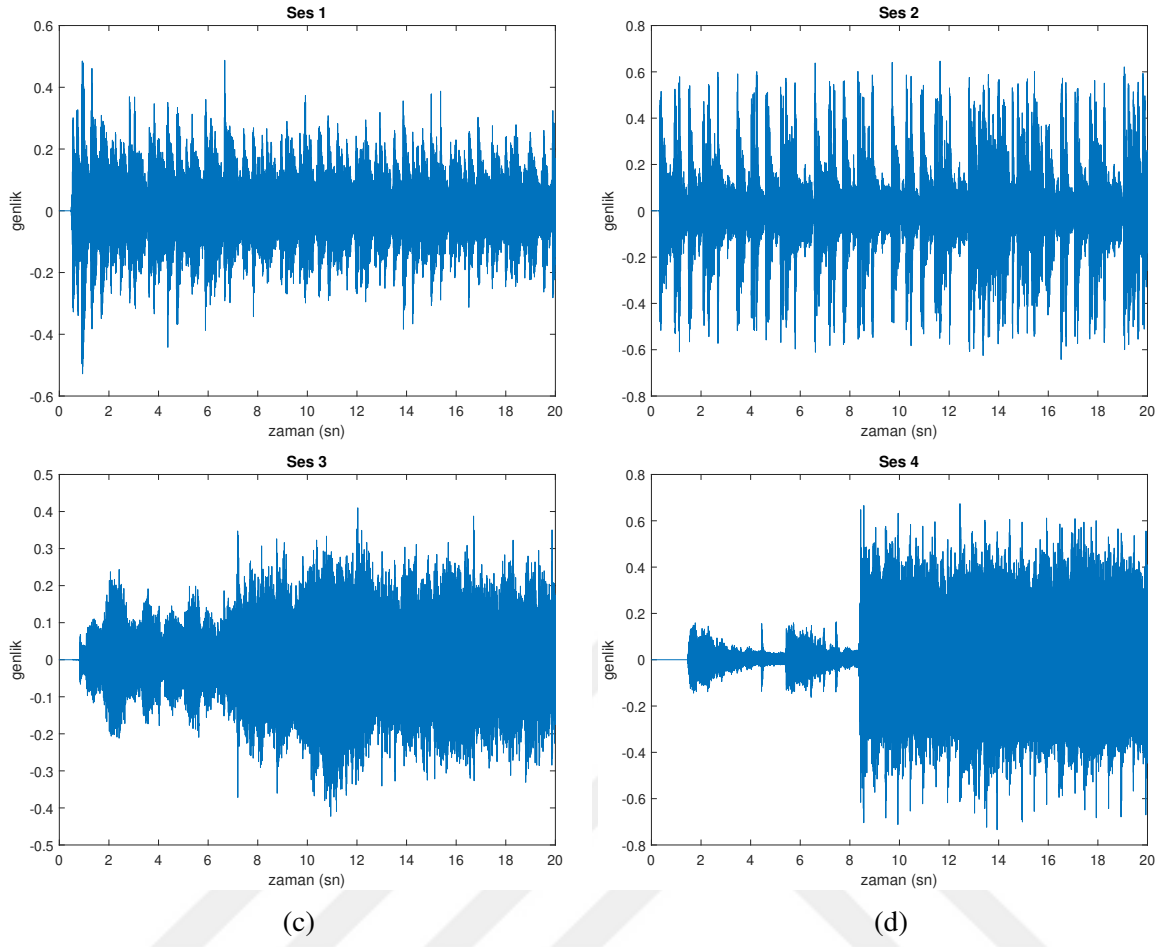
Çalışmanın bu bölümünde ses sinyallerinin benzerliğin tespitine yönelik başarısını ölçmek adına klasik Itakura-Saito uzaklığı ile önerilen uyarlanmış Itakura-Saito uzaklığı karşılaştırılacaktır. Uygulama çalışmasında kullanılacak eserler sırası ile,

- Eser 1: Muhlis Akarsu - Allah Allah desem gelsem
- Eser 2: Nelly Furtado - Wait for you
- Eser 3: Vassilis Saleas - Weeping eyes
- Eser 4: Michael Angelo Batio - Tribute to Metallica Medley

olarak belirlenmiştir. Buradaki ilk iki eser üzerinde bir telif hakkı davası olması nedeniyle özellikle seçilmiştir. Söz konusu davada eserlerin benzerliği uzman müzisyenlerden oluşturulan bir heyet tarafından kanıtlanmış olup, Muhlis Akarsu nezdinde Kalan Müzik® lehine sonuçlanmıştır. Diğer iki eser ise, bu eserlere "müzikal" benzerliği olmayan eserlerden seçilmiştir.

Burada müzikal benzerlikten anlaşılması gereken, eserlerin makamsal yapısı değildir. Aynı makama sahip olan ama birbirinden tamamen farklı olan binlerce ve hatta milyonlarca eser bulunabilmektedir. Çünkü makam bir nota dizgesidir ve müzikal benzerlik ise benzer zamanlarda, benzer nota süreleri ve dizilişi ile gerçekleşmektedir. Bu karakteristikler, sesin dijitalleştirilmesi ile elde edilen genlik-zaman verisinde her zaman kolay bir biçimde görülememektedir. Bunun sebebi, genlik-zaman grafiğinin eserdeki anlamlı şiddetleri veren bir grafik olmasıdır. Örneğin bir bateri solosu ile keman konçertosu benzer genlik-zaman verisine sahip olmakla birlikte, tamamiyle farklı frekans alanlarına sahiptir. Eserlerde benzerliği veren genlik-zaman verisinin frekans tayfidır. Dolayısıyla eserlerdeki benzerlik bu frekans tayflarının benzerliği üzerinden yapılmalıdır.

Öncelikle bu ses sinyallerinin analiz edilebilmesi için analog formdan dijital forma dönüştürülmesi A/D dönüşüm süreci ile yapılmıştır. İlk iki eserin benzerliği özellikle ilk 20 saniyede gerçekleştiğinden tüm eserler için bu süre içerisindeki ses sinyalleri dizisi esas alınmıştır. Böylece,  $44100 \times 20 = 882000$  uzunluğundaki zaman serileri ile çalışılmıştır. Uygulama temel olarak, bu dört eserin ikili kombinasyonlar biçiminde, klasik ve uyarlanmış Itakura-Saito uzaklıkları kullanarak, birbirleri ile karşılaştırılması ve sonuçlar üzerinden yöntemlerin birbirine üstünlüğünün sınanmasından oluşmaktadır. Söz konusu eserlerin dijital ortama aktarılması sonucu elde edilen zaman-genlik grafikleri Şekil 4.2 de verildiği gibidir.



Şekil 4.2. Karşılaştırma için seçilen ezgilerin zaman-genlik grafikleri

Şekil 4.2’de görüleceği üzere Eser 1 ve Eser 2’nin sinyallerinde sezgisel olarak diğerlerine nazaran bir benzerlik sezilenmektedir. Yukarıda tanımlanan eserler için hesaplanan ISU ve UISU, ikili karşılaştırmalar bazında Çizelge 4.2 de yer almaktadır.

Çizelge 4.2. Gerçek ses sinyallerinin ISU ve UISU karşılaştırılması

	<i>ISU</i>	<i>UISU</i>
Ses 1 & Ses 2	11.755	1.396
Ses 1 & Ses 3	22.097	15.371
Ses 1 & Ses 4	15.624	1036.390
Ses 2 & Ses 3	24.151	29.510
Ses 2 & Ses 4	17.811	321.170
Ses 3 & Ses 4	14.973	310.740

Çizelge 4.2 deki sonuçlar incelendiğinde çalışmada önerilen uyarlanmış Itakura-Saito uzaklığının daha anlamlı sonuçlar verdiği açıkça görülmektedir.

Ayrıca ses sinyallerinin durağan olmamasının, klasik Itakura-Saito uzaklığında hatalı sonuçlara sebep olacağı iddiasını da destekleyen bulgulara ulaşılmıştır.

Her iki uzaklık ölçüsü de gerçekte çok benzer olan Eser 1 ile Eser 2 arasındaki uzaklık değerini, diğerlerine nazaran en düşük olarak hesaplayarak, bunların diğerlerine göre daha benzer olduğunu ortaya koymuştur. Her iki uzaklık ölçüsünün de Fourier tabanlı olmasından dolayı, aynı ölçek tabanına dayanması bu iki uzaklık ölçüsünün karşılıklı olarak değerlerinde doğrudan karşılaştırılmasında bir sakınca ortaya koymaz. Bu nedenle UISU'nun ISU'ya nazaran (her ikisi de bu eserleri diğerlerine nazaran daha benzer bulmakla birlikte) UISU sıfıra çok daha yakın değer vererek gerçeklikteki yüksek benzerlik durumunu daha iyi yansıtmıştır. Diğer bir dikkat çekici husus ise, Eser 1 ve Eser 4 arasında diğerlerine nazaran benzemezliğin en yüksek derecede olması beklenirken, ISU'nun 15.624 değeri ile üç kombinasyondan daha düşük bir değere sahip olmasıdır. UISU ise bunun aksine, 1036.390 değeri ile en yüksek değere sahip olmuş ve sıfırdan oldukça uzağa düşmüştür. Böylece, simülasyon ile üretilen sinyallerde olduğu gibi gerçek ses sinyallerinde de UISU benzerliği doğru bir biçimde yakalamıştır.



## 5. AMPİRİK MOD AYRIŞIMININ UYARLANMIŞ ITAKURA-SAITO UZAKLIĞINA ETKİSİ

Çalışmanın bu bölümünde lineer ve durağan olmayan sinyallerin ayrıştırılmasında kullanılan Ampirik Mod Ayrışımı (Empirical Mode Decomposition, EMD) kullanılarak, ses sinyallerinin alt karakteristik yapılara ayrıştırılması ve bu alt karakteristik yapılar üzerinde uyarlanmış Itakura-Saito uzaklığının test edilmesi amaçlanmıştır.

### 5.1. Ampirik Mod Ayrışımı (EMD)

EMD lineer ve durağan olmayan zaman serisi sinyallerinin ayrıştırılması, analiz edilmesi ve güvenilir öngörülerin yapılabilmesi için son yıllarda kullanım sıklığı artan bir yöntemdir. EMD yönteminin önemli özelliklerinden bir tanesi de adaptif (uyarlanabilir) olmasıdır. Uyarlanabilir bir yöntem olmasının sebebi, herhangi bir spesifik olarak tanımlanmış matematiksel bir fonksiyona sahip olmayıp, veri-bazlı (data-driven) olmasından ileri gelmektedir. EMD yöntemi, sinyalin ayrıştırılmasından türetilen sonradan-tanımlı (post-definite) bir temele sahiptir (Huang ve diğerleri, 1998).

EMD yöntemi her verinin çok sayıda içsel mod fonksiyonlarından (IMF) oluştuğu varsayımına dayanmaktadır. Bu sayede lineer ve durağan olmayan serilerin alt karakteristik yapılarını içsel mod fonksiyonları aracılığıyla ortaya çıkarmaktadır. EMD yönteminin uygulanması sonucunda içsel mod fonksiyonları ve bu fonksiyonlar ile ifade edilemeyen artık (residual) serisi elde edilmektedir. Bu yöntemin en önemli özelliklerinden birisi, orijinal sinyalin, ayrışım sonucu elde edilen IMF'ler ve artık serisinin toplamı ile tekrardan elde ediliyor olmasıdır. İçsel mod fonksiyonu aşağıdaki iki şartı sağlamalıdır.

- i. Bir IMF ardışık iki çapraz düğüm (zero-crossing, node) arasında sadece bir ekstremuma (yerel minimum ya da maksimum) sahiptir. Başka bir deyişle, yerel maksimum sayısı ile yerel minimum sayısı arasındaki fark en fazla 1 (bir) olmalıdır.
- ii. Bir IMF sıfır ortalamalı olmalıdır.

Burada dikkat edilmesi gereken husus, ikinci şartın IMF'nin ortalamada durağan olduğunu ima etmesidir. Bu da söz konusu fonksiyonun analizini kolaylaştıran bir özelliktir. Bununla

birlikte, IMF'in durağanlığı genlik ve frekanstaki değişimi önlememektedir. Başka bir deyişle, IMF yapısal anlamda durağan gibi gözükse de genlik ve frekans dağılımı açısından durağan olmamaktadır.

IMF'ler, veriden eleme (sifting) yöntemi olarak bilinen yinelemeli bir işlemle elde edilir. Bu işlemin sonucunda  $x(t)$  sinyali, IMF'lere  $(x_n(t))$  ve IMF olarak ayrıştırılamayan kısım  $(r_n(t))$  ile Eş. 5.1 teki gibi ifade edilebilmektedir.

$$x(t) = \sum_n x_n(t) + r_n(t) \quad (5.1)$$

Eleme prosesi aşağıdaki algoritma ile yapılır.

A0. Başlangıç olarak  $n = 1, r_0(t) = x(t)$  alınır.

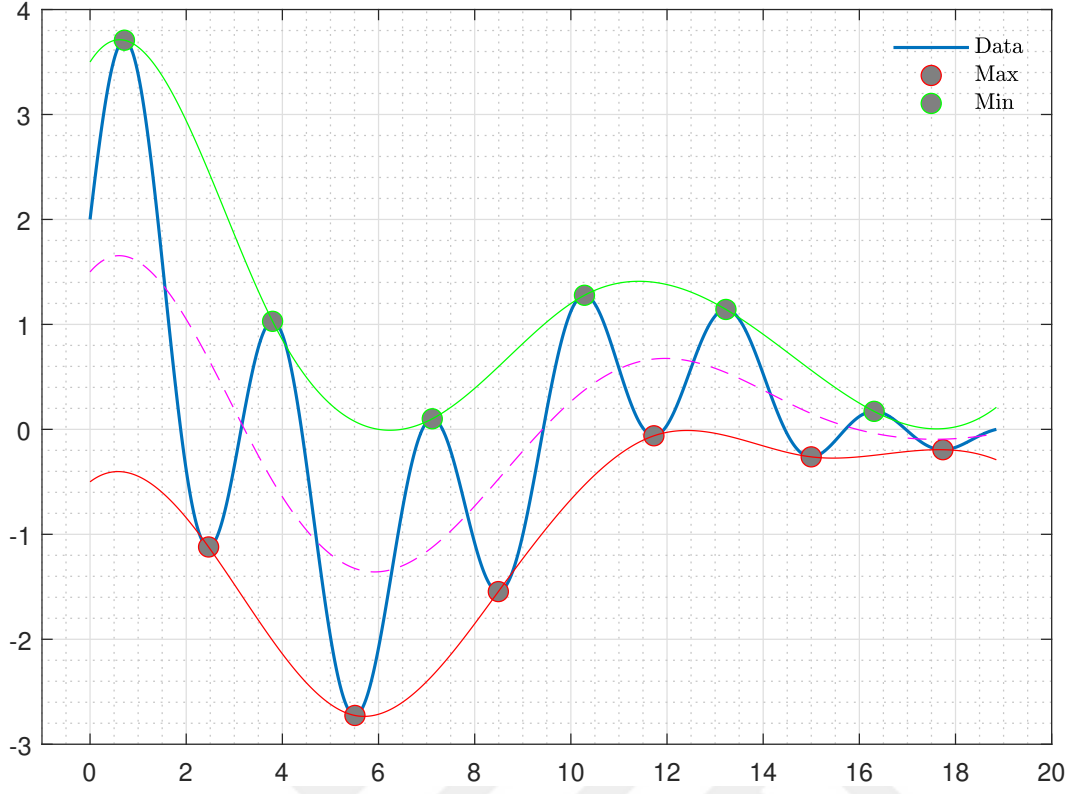
A1.  $n$  inci IMF aşağıdaki gibi elde edilir.

- i.  $k = 1$  iken  $h_0(t) = r_{n-1}(t)$
- ii.  $h_{k-1}(t)$  nin tüm yerel maksimum ve minimumları belirlenir.
- iii. yerel maksimum ve minimum noktaları kendi aralarında kübik spline interpolasyonu ile birleştirilerek üst  $U_{k-1}(t)$  ve alt  $L_{k-1}(t)$  zarf fonksiyonları (envelope) elde edilir.
- iv.  $h_{k-1}(t)$  nin ortalama zarf fonksiyonu  $m_{k-1}(t) = \frac{1}{2}(U_{k-1}(t) - L_{k-1}(t))$  bulunur.
- v.  $k$  inci bileşen  $h_k(t) = h_{k-1}(t) - m_{k-1}(t)$  biçiminde elde edilir.
  - Eğer  $h_k(t)$ , IMF olabilme şartlarının hiçbirini sağlamıyorsa,  $k \rightarrow k + 1$  alınır ve proses "ii." adımdan itibaren tekrar edilir.
  - Eğer  $h_k(t)$ , IMF olabilme şartlarının tamamını sağlıyorsa,  $x_n(t) = h_k(t)$  ve  $r_n(t) = r_{n-1}(t) - x_n(t)$  olacak şekilde ayarlanır.

A2. Eğer  $r_n(t)$  kalanı temsil ediyorsa eleme prosesi durdurulur. Etmeyorsa  $n \rightarrow n + 1$  olarak ayarlanır ve "A1" den itibaren süreç tekrar edilir.

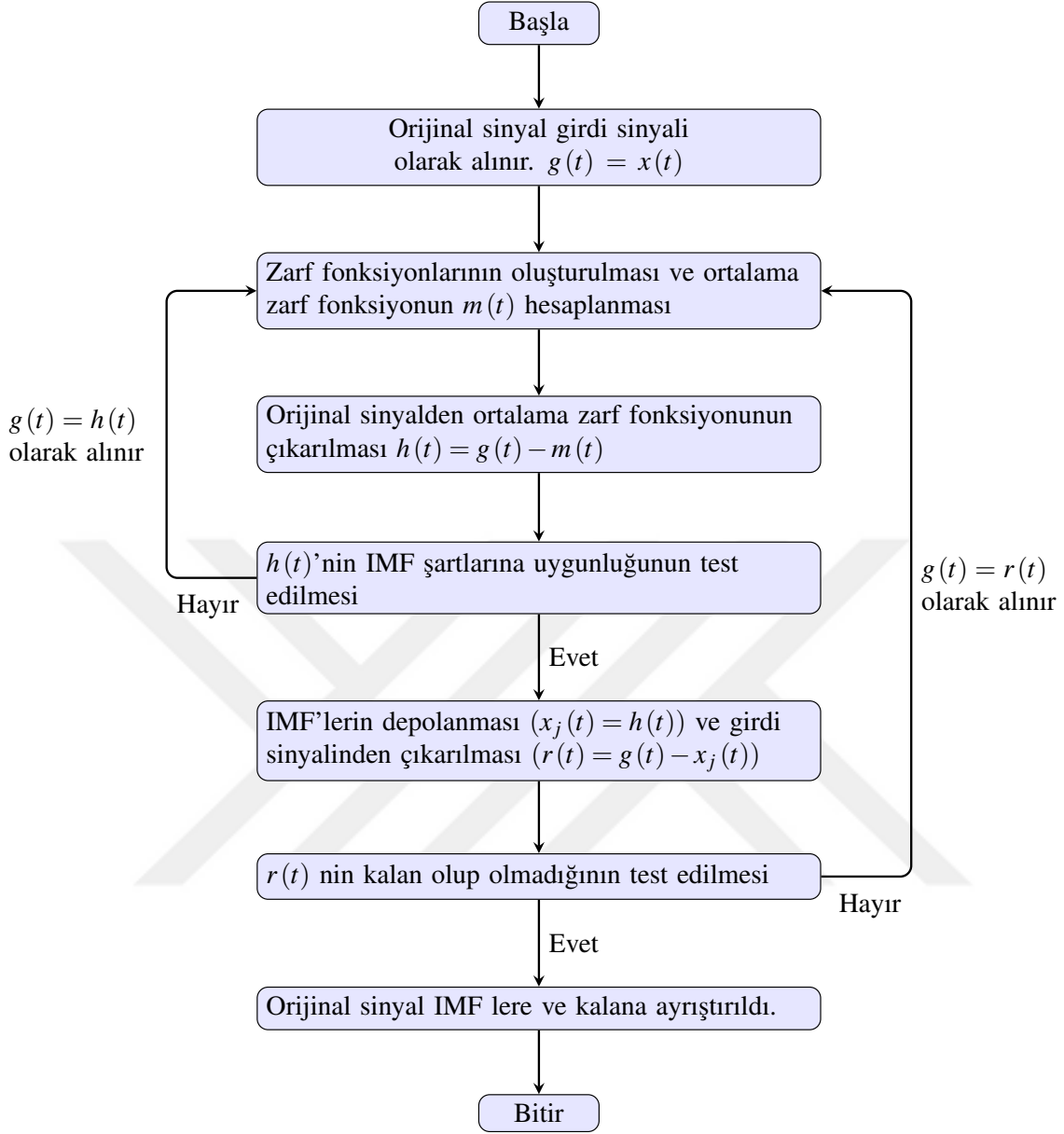
Şekil 5.1 de, yerel minimum ve yerel maksimum değerlerinin kübik eğriler ile birleştirilmesi sonucu elde edilen zarf fonksiyonları ve bu zarf fonksiyonlarının ortalaması gösterilmektedir. Ayrıca, bu şekil IMF'lerin elde edilmişinin bir temsilini göstermektedir.





Şekil 5.1. IMF lerin elde edilşinin basit diyagramatik gösterimi

EMD yönteminin algoritması akış diyagramı biçiminde Şekil 5.2 deki biçimde ifade edilebilmektedir.

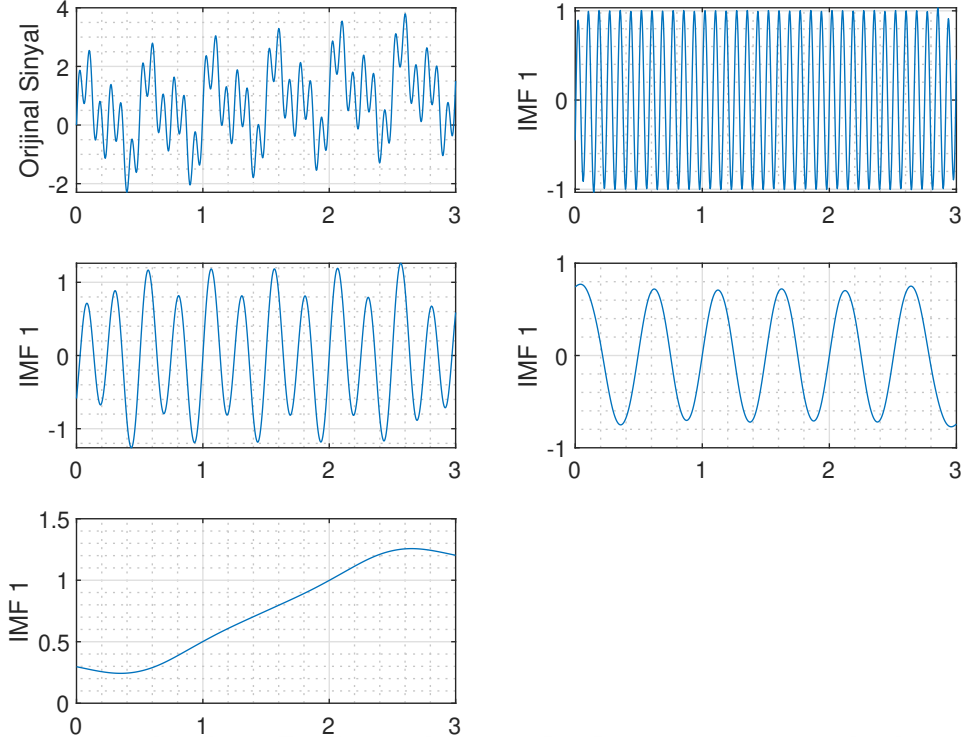


Şekil 5.2. EMD algoritmasının akış diyagramı

Eleme prosesi, sinyali durağan olmayan orijinal sinyal ile bölgesel olarak çakışmayan, benzersiz IMF lere ayrıştırır. Her bir IMF'nin ortalama periyodu örneklem sayısının iki katının düğüm sayısına bölünmesi ile bulunur. Ayrıca, bir zaman serisinden elde edilen IMF'lerin sayısı,  $T$  sinyalin uzunluğu olmak üzere, kabaca  $\log_2(T)$ 'ye eşit olacaktır.

EMD, dalgacıklar gibi önceden tanımlanmış dalga formlarına sahip bir alt-grup filtreleme tekniği değildir. Aksine bu yöntemde modların seçimi, otomatik ve adaptif zaman-değişen (time-varying) bir filtremeye karşılık gelir.

Şekil 5.1 de sinüzoidal fonksiyonlar yardımıyla hipotetik olarak oluşturulan  $2Hz$ ,  $4Hz$  ve  $12Hz$  frekanslara sahip spektral tonların karması ile elde edilen örnek bir ses sinyalinin EMD analizinin sonucu gösterilmektedir.



Şekil 5.3. Ampirik Mod Ayrışımı (EMD)

Şekil 5.1 incelendiğinde sinyalden ayrıştırılan her bir içsel mod fonksiyonunun benzersiz olduğu görülecektir. Ayrıca orijinal sinyalin kendisini oluşturan alt karakteristik yapılar (2 Hz, 4 Hz ve 12 Hz frekanslarına sahip spektral tonlar) mükemmel biçimde ayrıştığı görülmektedir.

EMD yönteminin başarısı, elde edilen IMF'lerin benzersizliği ile ölçülmektedir. Ancak, bazı sinyallerde benzersiz IMF'lerin elde edilmesi bir sorun olarak ortaya çıkmaktadır. IMF'lerin benzersiz olmaması, sinyali oluşturan alt karakteristik yapıların yüksek ilişki halinde olduğu ve dolayısıyla sinyalin tam anlamıyla ayrıştırılamadığı anlamına gelmektedir. Bu durum literatüre "mod karışımı problemi" olarak geçmiştir (Xu, Zhengxiang and Wang, 2016). Mod karışımı problemine, sinyali oluşturan alt karakteristik yapıların yakın frekanslı olması ya da zaman alanında çakışık olması durumu sebep olabilmektedir. Başka bir deyişle, yakın frekanslı dalgaların karması biçiminde oluşan sinyal büyük ölçüde önceki karakteristiğine

benzer bir davranış sergilemeye devam eder ve bu durum sinyalin kendisini oluşturan alt karakteristik yapılara ayrıştırılmasını güçleştirir.

Mod karışımı probleminin önüne geçmek adına EMD algoritmasının uyarlanmış bir hali olan Grupsal Ampirik Mod Ayrışımı (Ensemble Empirical Mode Decomposition, EEMD) yöntemi geliştirilmiştir.

## 5.2. Grupsal Ampirik Mod Ayrışımı (EEMD)

EEMD (Ensemble Empirical Mode Decomposition) yöntemi, IMF bileşenlerini orijinal sinyale ek olarak sonlu genliğe sahip anlamlı beyaz gürültünün olduğu her denemenin grupsal ortalaması olacak şekilde düşünerek mod karışımı problemini azaltmayı hedefler. Bu sürecin sonunda beyaz gürültünün yaratacağı etki giderilir (Gupta ve diğerleri, 2016). Beyaz gürültünün de eklenmesi ile birlikte  $i$ -nci denemedeki sinyal Eş. 5.2 deki gibi ifade edilir.

$$X^i(t) = X(t) + a_0 \omega^i(n), i = 1, 2, \dots, l \quad (5.2)$$

Burada  $\omega^i(n)$   $i$ -nci denemedeki birim varyansa ve  $a_0$  genliğine sahip beyaz gürültüyü simgelemektedir. Her bir  $i = 1, 2, \dots, l$  değeri için  $IMF_k^i$  beyaz gürültünün farklı gerçekleştirmeleri için hesaplanır.  $k$ -ncı içsel mod fonksiyonu  $IMF_k$  nın ortalaması Eş. 5.3 deki gibi ifade hesaplanır.

$$\overline{IMF_k} = 1/l \sum_{i=1}^l IMF_k^i \quad (5.3)$$

EEMD yöntemi, mod karışımı probleminin varlığı durumunda çoğunlukla IMF'leri başarılı bir şekilde ayırtmakla birlikte, bazen sinyale eklenen gürültü, sürecin sonunda sinyalden tamamen ayrıştırılmadığından, yanlış sonuçlara yol açabilmektedir. Bu durumun önüne geçmek adına, Torres ve diğerleri (2011) tarafından Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) yöntemi önerilmiştir.

### 5.3. Ses Sinyallerinin Benzerliğinin IMF'ler Arasındaki Benzerlik Üzerinden Tespiti

Bilindiği üzere bir sinyalin içsel mod fonksiyonları, sinyali oluşturan alt karakteristik yapıları göstermektedir. Dolayısıyla, iki sinyalin benzerliğini, sinyalleri oluşturan alt karakteristik yapılar üzerinden ele almanın daha hassas sonuç çıkarımı yapmayı önemli ölçüde etkilemesi beklenir.

Müzik eserleri çoğunlukla birbirine yakın ve uzak frekanslı, çakışık zamanlı spektral tonların bir karması biçimindedir. Bu yapısı nedeniyle doğal olarak mod karışımı problemine sahiptir. Bu nedenle ses sinyallerin içsel mod fonksiyonları EEMD yöntemi ile elde edilmelidir. Çalışmanın bu bölümünde, önerilen uyarlanmış Itakura-Saito uzaklığı önceki bölümde tanımlanan dört ezgiden EEMD yöntemi ile elde edilen içsel mod fonksiyonları üzerinde sınanmış olup ilgili sonuçlar Çizelge 5.1 de listelenmiştir.

Çizelge 5.1. IMF'ler arasındaki uyarlanmış Itakura-Saito uzaklıkları

IMF	$UISU_{1-2}$	$UISU_{1-3}$	$UISU_{1-4}$	$UISU_{2-3}$	$UISU_{2-4}$	$UISU_{3-4}$
1	0.19	0.90	26.37	5.74	34.65	2.62
2	1.99	5.44	366.59	6.95	335.80	52.27
3	0.23	79.52	346.70	139.83	257.09	127.79
4	0.32	255.23	17.85	249.80	16.73	28.81
5	1.19	3.10	1.32	22.95	5.20	1.35
6	1.61	60.14	1.25	39.22	2.79	4.09
7	0.87	11.85	1.05	10.80	1.98	2.41
8	0.73	1.27	2.14	1.58	1.88	0.81
9	0.93	5.11	8.10	2.28	7.26	1.12
10	1.09	2.05	1.60	0.94	1.27	0.72
11	1.81	1.07	1.44	1.38	2.33	0.99
12	0.73	1.11	1.24	1.97	0.91	3.15
Toplam	11.67	426.80	775.66	483.43	667.87	226.12

Burada,  $UISU_{i-j}$ ,  $i \neq j$ ,  $i$ -nci eser ile  $j$ -nci eser arasındaki UISU'yu göstermektedir.

Ampirik Mod Ayrışımı yönteminin önemli bir özelliği olan ve Eş. 5.1'de verilen, orijinal sinyalin, ayrışım sonucu elde edilen IMF'ler ve artık serisinin toplamı biçiminde ifade edilebilmesi özelliği, IMF'ler arasındaki UISU uzaklıklarının toplamalarının yorumlanmasının gerekliliğini ortaya koymaktadır. Bu nedenle Çizelge 5.1'de IMF'ler arası uzaklıklar toplamaları ile birlikte verilmektedir.

Çizelge 5.1'de, çok benzer olan Eser 1 ile Eser 2'nin IMF'leri arasındaki toplam uzaklık, 11.67 değeri ile diğerlerine nazaran daha düşük olarak hesaplanmış ve bu iki eserin diğerlerine göre daha benzer olduğu görülmüştür. Ayrıca, eserler arasındaki en benzemez iki eser olan Eser 1 ile Eser 4 arasındaki uzaklık değeri, 775.66 olarak hesaplanmış ve beklenildiği gibi diğerlerine göre en benzemez olarak bulunmuştur. Elde edilen bu sonuçların, daha önceki bulguları destekler nitelikte olması yöntemin tutarlılığına ilişkin bir kanıt ortaya koymaktadır.

## 6. TARTIŞMA

Itakura-Saito uzaklığı minimumu sıfır olan ancak maksimumuna ilişkin herhangi bir sınır bulunmayan, veri bazlı bir uzaklıktır. Maksimum uzaklığın bilinmemesi, ölçekleme problemini beraberinde getirmekte ve iki sinyalin benzerliği yüzde ile ifade edilememektedir. Sıfıra yakınlık derecelendirilemediği sürece yorumlar subjektif olacaktır. Bu durumu nedeniyle Itakura-Saito uzaklığı ve çalışmada önerilen uyarlanmış Itakura-Saito uzaklığı, "görelî benzerliğin" bir ölçüsüdür. İzafî yanı nedeniyle, ses sinyallerinin benzerliğinin belirlenmesinde kullanılan sezgisel metotlara, nicel destek sağlanması adına kullanımı çok daha anlamlı olacaktır.







## 7. SONUÇ VE ÖNERİLER

Uyarlanmış Itakura-Saito uzaklığı (UISU), simülasyon ve gerçek ses sinyallerinin benzerliklerinin araştırılmasında başarılı sonuçlar elde etmiş olup, klasik Itakura-Saito uzaklığının, sinyalin durağan olmaması kaynaklı yanılgılarından etkilenmemiştir. Özellikle, çalışmada aralarında açılmış bir telif davası bulunan iki müzik eserinin benzerliğini son derece hassas bir biçimde kanıtlamış olması ve bunun yanı sıra en benzemez iki eser arasındaki uzaklığı benzerlik noktası olan sıfırdan bir hayli uzak olarak hesaplamış olması, yöntemin nicel anlamda verebileceği desteği ortaya çıkarmaktadır. Ayrıca uyarlanmış Itakura-Saito uzaklığı ile, ezgilerin EEMD yöntemi ile elde edilen IMF leri arasındaki benzerlik yapıları incelenmiş olup, elde edilen sonuçların direkt olarak ezgiler arasındaki uyarlanmış Itakura-Saito uzaklıkları ile tutarlı sonuçlara sahip olduğu gözlemlenmiştir. IMF'ler ilgili sinyalin alt karakteristik yapılarını gösterdiğinden, iki sinyalin IMF lerinin karşılaştırılması yüksek önem arz etmektedir. Karşılaştırma sonucu elde edilen sonuçların iddiaları doğrular nitelikte oluşu çalışmanın anlamlılığını arttırmaktadır.

Çalışmada önerilen yöntem, mod karışımı olarak bilinen ve sinyalin EMD'si sonucu elde edilen IMF lerin benzersiz olmaması anlamına gelen problemin tespitine de ışık tutmaktadır. Problemin tespitine ilişkin literatürde bazı çalışmalar mevcuttur. Örneğin, Terrien ve diğerleri (2011) tarafından, IMF'lere yapılan durağanlık testleri ile mod karışımı probleminin belirlenmesine ilkin bir yöntem önerilmiştir. Fosso ve Molinas (2017) tarafından ise, bu problemin tespitine ilişkin, sinyal maskeleye tekniğine dayanan bir yöntem önerilmiştir. Sinyalden EMD yöntemi ile elde edilen IMF'lerin ikili kombinasyonları arasındaki uyarlanmış Itakura-Saito uzaklıklarına bakılarak, IMF'lerin benzerlik yapısının ortaya çıkarılması, mod karışımı probleminin varlığına ilişkin kanıt niteliğinde olacaktır.



## KAYNAKLAR

- Ahonen T., E., Lemstr'Om K. and Linkola S. (2011). *Compression- Based Similarity Measures In Symbolic, Polyphonic Music*, International Society For Music Information Retrieval Conference, ISIMR.
- Akdi, Y. (2010). *Zaman Serileri Analizi (Birim Kökler ve Kointegrasyon*. (2. Baskı). Yayın Yeri: Gazi Kitabevi.
- Allen J., B. (1982). Application of the short-time Fourier transform to speech processing and spectral analysis. *Proc. IEEE ICASSP-82*, 1012-1015.
- Allen J., B. and Rabiner L., R. (Nov. 1977). A unified approach to short-time Fourier analysis and synthesis. *Proc. IEEE*, 65, 1558-1564.
- Avcı, K. (2014). Kaiser-Hamming Penceresi ve Yinelemesiz Sayısal Süzgeç Tasarımı İçin Performans Analizi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 29(4).
- Birgül, S. ve Çoğulu, T. (2014). *Temel Müzik Eğitimi*. İstanbul: BGST.
- Cambouropoulos, E., (2001). Melodic Cue Abstraction, Similarity and Category Formation: A Formal Model, *Music Perception An Interdisciplinary Journal*, Spring, 18(3).
- Cambouropoulos, E., Crochemore, M., Iliopoulos, C., Mohamed, M., Sagot, M. F. (2005). A Pattern Extraction Algorithm for Abstract Melodic Representations that Allow Partial Overlapping of Intervallic Categories. *ISMIR 2005 - 6th International Conference on Music Information Retrieval*, 167-174.
- Fosso, O. B. and Molinas, M. (2017). Method for Mode Mixing Separation in Empirical Mode Decomposition, *eprints in Physics, Mathematics, Statistics 1709.05547*,1(1), 1-7.
- Gabor D. (1947). Theory of Communication. *Journal of Institution of Electrical Engineers*, 93(3), 429-457.
- Gupta, A., Kumar, D., Chakraborti, A. and Sharma, K. (2016). Performance Evaluation of Empirical Eode Decomposition Algorithms For Mental Task Classification. *BioRxiv*, 1(1), 1-25.
- Haas, W., Frans, W., F. and Veltkamp, R., C. (2013). A geometrical distance measure for determining the similarity of musical harmony, *International Journal of Multimedia Information Retrieval*, 2(3), 189-202.
- Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q. and Liu, H. (1998). The Empirical Mode Decomposition and the Hilbert Spectrum For Nonlinear and Non-

stationary Time Series analysis. *Proceedings of The Royal Society*, 454(1971), 903-995.

Itakura, F. (1975). Minimum Prediction Residual Principle Applied to Speech Recognition, *IEEE Transactions on Acoustics, Speech, And Signal Processing*, 23(1), 67-72.

Itakura, F. and Saito, S. (1968). Analysis synthesis telephony based on the maximum likelihood method. *In Proc. 6th of the International Congress on Acoustics* (pp. C-17-C-20). Los Alamitos, CA: IEEE.

İnternet: Atasoy, H. (2011). Dinamik Zaman Bükme Algoritması. URL: <http://www.atasoyweb.net/Dinamik-Zaman-Bukme-Algoritmasi>, Son Erişim Tarihi: 24.11.2019.

İnternet: BAKIR, M., A. (2019). Time Series Analysis. Prof. Dr. M. Akif BAKIR - Personal Web Page. URL: <https://makifbakir.wordpress.com/time-series-analysis-2/>, Son Erişim Tarihi: 24.11.2019.

İnternet: Cömert, Z. (2015). Fourier Dönüşümü. Zafer Cömert Kişisel Web Sayfası. URL: <http://www.zafercomert.com/IcerikDetay.aspx?zcms=157>, Son Erişim Tarihi: 24.11.2019.

İnternet: Ses. (2018). URL: <http://key.com.tr/ses-dalgalar.html>, Son Erişim Tarihi: 12.07.2017.

İnternet: Suits, B. H. (n.d.). Frequencies of Musical Notes. URL: <https://pages.mtu.edu/suits/notefreqs.html>, Son Erişim Tarihi: 24.11.2019.

Jaber, A. and Bicker, R. (2018). Development of a Condition Monitoring Algorithm for Industrial Robots Based on Artificial Intelligence and Signal Processing Techniques. *International Journal of Electrical and Computer Engineering*. 81, 996-1009.

Kaya, T. ve İnce, M. (2010). Pencere Fonksiyonu Aileleri ve Uygulama Alanları. *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi*, 26(3), 291-306.

Kaya, T , İnce, M . (2011). *Kaiser Pencere Fonksiyonu Spektral Parametrelerinin Yapay Sinir Ağları Yardımıyla Hesaplanması*, Elektrik-Elektronik Bilgisayar Sempozyumu (FEEB 2011) Sempozyum Kitapçığı Cilt 2.

Logan, B. and Salomon, A. (June 2001). *A Content-Based Music Similarity Function*, Cambridge Research Laboratory, MIT.

Mongeau M. and Sankoff D, (1990). Comparison of Musical Sequences, *Computers and the Humanities*, 24(3), 161-175.

- Park, M. W. and Lee, E. C., (2013). Similarity Measurement Method between Two Songs by Using the Conditional Euclidean Distance, *Wseas Transaction On Information Science and Applications*, 12(10).
- Shin Mi-Hae, Kim Eui-Jeong, Seo Su-Seok and Kim Young-Chul, (2014). Implementation of Plagiarism Analysis System through Digital Conversion Processing of Sound Source, *International Journal of Multimedia and Ubiquitous Engineering*, 9(10), 33-34.
- Slaney, M., Rhodes, C. and Casey, M. (2008). Analysis of minimum distances in high – dimensional musical spaces. *IEEE Trans. On Audio, Speech, and Language Processing*, 16(5), 1015 – 1028.
- Suneja, K. and Bansal, M. (2015). *Comparison of time series similarity measures for plagiarism detection in music*. 2015 Annual IEEE India Conference (INDICON).
- Terrien, J., Marque, C. and Karlsson, B. (2011). Automatic Detection of Mode Mixing in Empirical Mode Decomposition Using Non-stationarity Detection: Application to Selecting IMFs of Interest and Denoising. *EURASIP Journal on Advances in Signal Processing*, 1(1), 1–8.
- Toiviainen P. and Eerola T. (2002). *A Computational Model of Melodic Similarity Based on Multiple Representations and Self-Organizing Maps*, Proceedings of the International Conference on Music Perception and Cognition, Sydney.
- Torres, M. E., Colominas, M. A., Schlotthauer, G. and Flandrin, P. (2011). *A Complete Ensemble Empirical Mode Decomposition With Adaptive Noise*. In ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 1(1), 4144–4147.
- Urbano J., Lloréns J., Morato J. and Cuadrado S., S. (2012). Symbolic Melodic Similarity: Hybrid Sequence Alignment with Geometric Representations, *MIREX*.
- Velankar M.R., Dr. Sahasrabuddhe H.V. and Dr. Kulkarni P.A., (2015). Modeling Melody Similarity Using Music Synthesis and Perception, *Procedia Computer Science*, 45, 728-738.
- Vempala, N., N. and Russo F., A. (2012). *A Melodic Similarity Measure Based on Human Similarity Judgements*, 12th International Conference on Music Perception and Cognition. Greece
- Weihls C., Ligges U., Mörchen F. and Müllensiefen D. (2007). Classification in Music Research, *Advances in Data Analysis and Classification*, 200(3), 255-291.
- Xiao L., Zheng Y., Tang W., Yao G. and Ruan L. (2014). GPU Acceleration of Melody Accurate Matching in Query-by-Humming, *The Scientific World Journal*.

Xu, G., Zhengxiang, Y. and Wang, S. (2016). *Study on Mode Mixing Problem of Empirical Mode Decomposition*. Joint International Information Technology, Mechanical and Electronic Engineering Conference, 1(1), 389–394.



## ÖZGEÇMİŞ

### Kişisel Bilgiler

Soyadı, Adı : ERDİŞ, Atacan  
 Uyuğu : T.C.  
 Doğum tarihi ve yeri : 24.06.1991, Ankara  
 Medeni hâli : Bekâr  
 e-mail : atacanerdis@gazi.edu.tr



### Eđitim

Derece	Eđitim Birimi	Mezuniyet Tarihi
Yüksek Lisans	Gazi Üniversitesi, İstatistik Bölümü	Devam Ediyor
Yüksek Lisans	Ankara Üniversitesi, İstatistik Bölümü	2017
Lisans	Ankara Üniversitesi, İstatistik Bölümü	2013

### İř Deneyimi

Yıl	Yer	Görev
2019 – Hâlen	Gazi Üniversitesi, İstatistik Bölümü	Arařtırma Görevlisi

### Yabancı Dil

İngilizce

### Yayınlar

1. Erdiř, A., Kardiyen, F. ve Bakır, M., A. (2017, 5-7 Ekim). *Müzik Eseri Hırsızlıđı Sorununa İstatistiksel Yaklařım*. 18. Uluslararası Ekonometri Yöneyem Arařtırması ve İstatistik Sempozyumu, KTÜ, Trabzon, Türkiye.
2. Jaiteh, M., I., Bakır M., A. , Erdiř, A. (2018, 28-30 April). *Mode Mixing In Empirical Mode Decomposition (EMD)*. 4th International Researchers, Statisticians And Young Statisticians Congress, Çeřme, İzmir, Turkey.
3. Erdiř, A., Bakır, M., A., Jaiteh, M. I. (2019, 4-8 October). *Detection of Mode-Mixing Problem in EMD of Time Series Signal Data*. 11th International Statistics Congress, Bodrum, Turkey.

### Hobiler

Müzik, Bilgisayar







*GAZİ GELECEKTİR...*