



# **KONUŐMA SİNYALİNDEN ARKA PLAN MÜZİĐİNİN AYRILMASI**

**Sait Melih DOĐAN**

**YÜKSEK LİSANS TEZİ**

**ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ ANABİLİM DALI**

**GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**HAZİRAN 2018**

Sait Melih DOĞAN tarafından hazırlanan “KONUŞMA SİNYALİNDEN ARKA PLAN MÜZİĞİNİN AYRILMASI” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi Elektrik Elektronik Mühendisliği Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

**Danışman:** Prof. Dr. Özgül SALOR DURNA

Elektrik Elektronik Mühendisliği Anabilim Dalı, Gazi Üniversitesi

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

**Başkan:** Prof. Dr. Hakkı Gökhan İLK

Elektrik Elektronik Mühendisliği Anabilim Dalı, Gazi Üniversitesi

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

**Üye:** Prof. Dr. Cengiz TAPLAMACIOĞLU

Elektrik Elektronik Mühendisliği Anabilim Dalı, Gazi Üniversitesi

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

Tez Savunma Tarihi: 27/06/2018

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

.....

Prof. Dr. Sena YAŞYERLİ

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.

S.Melih DOĞAN

27/06/2018

KONUŐMA SİNYALİNDEN ARKA PLAN MÜZİĞİNİN AYRILMASI  
(Yüksek Lisans Tezi)

Sait Melih DOĞAN

GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

Haziran 2018

ÖZET

Bir müzikal parçanın vokal sesi ve arka plan kısımlarını ayırmak zor bir iştir. Literatürde, müzik parçalarından vokal sesi ve arkaplan parçalarını ayırma süreci genellikle müziğin tekrarlama özelliğini kullanır. Bu alandaki öncü çalışmalar arasında yer alan Yinelenen Desen Çıkarma Tekniğı (REPET) ve Sağlam Temel Bileşen Analizi (RPCA) yöntemlerinde arka plan müziğinin tekrarlama özelliğı kullanılarak müzikal bir parça vokal ve arka plan müziğı olarak ayrılabilir. Bu çalışmada REPET algoritmasının ayırma başarımını geliştirmek için REPET ve RPCA algoritmalarını birleştiren bir araştırma çalışması yürütülmüştür. Önerilen yöntemin başarımlarını REPET ve RPCA ile karşılaştırmak için MIR-1K veri setinden seçilen ses parçaları ile iki farklı test yapılmıştır. Her iki testte de önerilen yöntemin başarımının diğer iki yöntemle göre daha iyi olduğı görülmektedir.

Bilim Kodu : 90537  
Anahtar Kelimeler : Ses işleme, müzik-ses ayrımı, müzik yapı analizi, yinelenen model çıkarma tekniğı (REPET), sağlam temel bileşen analizi (RPCA), ortalama görüş puanı  
Sayfa Adedi : 73  
Danışman : Prof. Dr. Özgül SALOR DURNA

MUSIC /SINGING VOICE SEPARATION  
(M. Sc. Thesis)

Sait Melih DOĞAN

GAZİ UNIVERSITY  
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

June 2018

ABSTRACT

Separating the vocal and background parts of a piece of music is a very difficult task. In the literature, the process of separating vocal and background parts from musical pieces usually utilize music repetition feature. In both Repeating Pattern Extraction Technique (REPET) and Robust Principal Component Analysis (RPCA) methods, which are among the leading studies in this field, musical pieces are separated as vocal and background music by using repetition feature of the background music. In this paper, a research study is carried out combining REPET and RPCA algorithms in order to improve the separation performance of the REPET algorithm. In order to compare performances of the proposed method with REPET and RPCA, two different tests have been carried out with selected audio tracks from the MIR-1K dataset. It has been shown by both tests that the performance of the proposed method is much better than other two methods.

Science Code : 90537

Key Words : Audio processing, music-voice separation, music structure analysis, repeating pattern extraction technique(REPET), robust principal component analysis(RPCA), mean opinion score

Page Number : 73

Supervisor : Prof. Dr. Özgül SALOR DURNA

## TEŐEKKÜR

Çalıőmalarım boyunca yardım ve katkılarıyla deęerli yardım ve katkılarıyla beni yönlendiren, kıymetli tecrübelerinden faydalandığım danışmanım Prof. Dr. Özgöl SALOR DURNA'ya, manevi destekleriyle beni hiçbir zaman yalnız bırakmayan çok deęerli ailem; D. Mehmet DOĐAN, Hatice DOĐAN, Feyza DOĐAN, Ahmet Emre DOĐAN ve Safiye Bilge DOĐAN'a ve çalıőmalarımda bana destek olan iő arkadaşlarıma teőekkürü bir borç bilirim.



## İÇİNDEKİLER

	<b>Sayfa</b>
ÖZET .....	iv
ABSTRACT.....	v
İÇİNDEKİLER .....	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
2. KURAMSAL TEMELLER .....	5
2.1. REPET .....	5
2.1.1. REPET – 1. Adım: Vuruş spektrumunun hesaplanması .....	6
2.1.2. REPET – 2. Adım: Tekrarlanan segmentin bulunması .....	8
2.1.3. REPET – 3. Adım: Tekrarlanan desenlerin çıkarılması .....	9
2.2. RPCA .....	9
2.2.1. Normlar uzayı.....	9
2.2.2. Tekil değer ayrışımı (Singular Value Decomposition-SVD).....	12
2.2.3. Temel bileşen çözümleme (Principal component analysis - PCA) .....	14
2.2.4. Lagrange çarpanları.....	15
2.2.5. Artırılmış Lagrange çarpanı (Augmented Lagrange multiplier- ALM)..	16
2.2.6. RPCA Detaylı incelenmesi.....	17
2.3. İki Modelin Çıktılarının İncelenmesi .....	24
3. TEZDE ÖNERİLEN MODEL .....	25
4. ÖNERİLEN YÖNTEMİN BAŞARIMLARININ DEĞERLENDİRİLMESİ .....	31
4.1. MİR-1K Veri Seti.....	31
4.2. Ortalama Görüş Puanı (Mean Opinion Score).....	32

	<b>Sayfa</b>
4.2.1 Ortalama görüş puanı (MOS) kullanılarak sonuç değerlendirilmesi.....	32
4.3.Bss Eval Aracı.....	34
4.3.1. Bss Eval aracı kullanılarak pencereleme fonksiyonu seçimi .....	35
4.3.2. Bss Eval aracı kullanılarak başarımlar değerlendirilmesi .....	36
<b>5. SONUÇ VE ÖNERİLER .....</b>	<b>39</b>
<b>KAYNAKLAR .....</b>	<b>41</b>
<b>EKLER.....</b>	<b>45</b>
EK-1. Tezde Önerilen Metot, REPET Ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının Vokal ve Arka Plan Müziği Çıktılarının Spektrogramları .....	46
EK-2. 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanımlama Sonuçlarının İşlendiği Formlar.....	56
<b>ÖZGEÇMİŞ .....</b>	<b>73</b>

**ÇİZELGELERİN LİSTESİ**

<b>Çizelge</b>	<b>Sayfa</b>
Çizelge 4.1. Ortalama görüş puanı değerlendirme tablosu .....	32
Çizelge 4.2. Mos testi sonrasında elde edilen sonuçlar .....	33



## ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Repet algoritması akış şeması .....	5
Şekil 2.2. L1 ve L2 norm gösterimi.....	10
Şekil 2.3. Tekil Değer Ayrışımı (SVD) .....	13
Şekil 2.4. Tekil değer ayrışımı eksen azaltımı .....	15
Şekil 2.5. Lagrange çarpanı kullanımı ile maksimum noktanın tespiti .....	16
Şekil 2.6. Aykırı elemanların olması durumunda ana bileşen yönü .....	18
Şekil 2.7. Rpca kullanılarak düşük rank ve seyrek matrislerin oluşturulması .....	23
Şekil 2.8. Rpca ve zaman frekans maskesi ile vokal ve arka plan sesinin ayrılması.....	23
Şekil 3.1. Önerilen metodun akış diyagramı.....	25
Şekil 3.2. Vokal sesinin olmadığı seyrek kısımların ayrılması sonuçları .....	29
Şekil 3.3. Düşük frekansların ayrılması sonuçları .....	30
Şekil 4.1. Pencere fonksiyonlarına göre sdr değerleri .....	35
Şekil 4.2. Repet, rpca ve önerilen metot kullanılarak oluşan sonuçların sdr değerleri...	36
Şekil 4.3. Repet, rpca ve önerilen metot kullanılarak oluşan sonuçların sdr değerleri.....	37
Şekil 4.4. Repet, rpca ve önerilen metot kullanılarak oluşan sonuçların sdr değerleri..	37

## SİMGELER VE KISALTMALAR

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

<b>Simgeler</b>	<b>Açıklamalar</b>
<b>dB</b>	Desibel
<b>Hz</b>	Hertz
<b>KHz</b>	Kilo Hertz
<b>ms</b>	Mili saniye
<b>Kısaltmalar</b>	<b>Açıklamalar</b>
<b>ALM</b>	Augmented lagrangian mulitplier
<b>ASR</b>	Automatic speech recognition
<b>DFT</b>	Discerete fourier transformation
<b>FFT</b>	Fast fourier transformation
<b>HMM</b>	Hidden markov model
<b>LFPC</b>	Log frequency power coefficients
<b>MFCC</b>	Mel-Frequency cepstrum
<b>MMSE</b>	Minimum mean square error
<b>NMF</b>	Non-Negative matrix factorization
<b>PCA</b>	Principal component analysis
<b>PLP</b>	Perceptual linear predictive coefficients
<b>REPET</b>	Repeating pattern extraction technique
<b>RPCA</b>	Robust principal component analysis
<b>STFT</b>	Short time fast fourier transformation
<b>SVD</b>	Singular value decomposition
<b>YSA</b>	Yapay sinir ağları

## 1. GİRİŞ

Müziğin her bilim dalı açısından farklı anlamları olsa da genel tanımı, sesin biçim ve anlamlı titreşimler kazanmasıdır. Sesin oluşurken dinleyicilerde duygulara yönelik etkileşim yapması durumu melodi olarak kabul görür. Melodi, bir müzik parçasında baştan sona kadar belirli yerlerde tekrarlanır. Düzgün ve ardışık notaların bir müzik parçasında birbirini takip etmesi ritim olarak adlandırılır ve müziğin kulağa hoş gelen kısmı melodinin en önemli kısmını oluşturur. Ritim aslında hareketle, hareketin var oluş süresiyle, tekrarıyla ve kaç defa tekrar ettiği ile ilişkilidir. Ritmin olmazsa olmazı ise periyodikliğidir.

Ses ve sessizlik ile anlık olarak analiz ettiğimiz müziği bütünsel olarak analiz ederek müziği tekrar eden sesler ile (bunun karşıtı olarak) değişen seslerin dengesi olarak niteleyebiliriz. Müzik teorisyenlerine göre tekrarlama, müziğin yapısını analiz etmede büyük bir öneme sahiptir. Schenker, tekrarlama, müzikal parçanın içindeki en küçük yapısal unsur olarak tanımlamıştır [1]. Ruvet müzik parçasını küçük parçalara bölmek ve müzikte yer alan söz dizimini ortaya çıkarmak için tekrarlama kullanmıştır [2]. Ockelford çalışmasında müzikal bir parçada yer alan düzenin müziği estetik açıdan hoş hale getiren şeylerin başında geldiğini ve bu düzenin ise parçada yer alan tekrarlamaların bir sonucu olduğunu belirtmiştir [3].

Tekrar kavramı, müzik yapısı, ritim ve alan melodisinin tahmini için kullanılabilir. Müzik üzerine çalışan araştırmacılar müziğin tekrarlama özelliğini; ses segmentasyonu, ritim tahmini gibi alanlarda kullanırlar. Şarkıların vokal ve arka plan müziği olarak ayrılması, enstrüman-vokal teşhisi, melodi çıkarımı, ileri ses işleme gibi alanlarda kullanılabildiği için büyük ilgi görmektedir.

İnsanın bakış açısından, şarkı sesi dinleyicinin hemen dikkatini çeken etkenlerin en önemlisidir. İnsan işitme sistemi, çevredeki herhangi bir sese etkili bir şekilde odaklanma ve sesleri farklı kaynaklardan ayırmak için dikkate değer bir yeteneğine sahiptir. Bununla birlikte insan işitme sistemi, vokal seslerini arka plan müziği eşliğinden ayırma kabiliyetine sahiptir. Bu görev insanlar için zahmetsiz olmasına rağmen, makineler için daha zordur.

Bir şarkıyı müzik ve ses bileşenlerine verimli bir şekilde ayırma özelliği kullanılarak enstrüman / vokal tanımlaması, melodi çıkarma, ileri seviye ses işlenmesi, arka plan-gürültü çıkarımı , karaoke-oyun ve gibi geniş bir uygulama alanında kullanılabilir.

Konuşma ayrımı son yıllardan beri geniş bir biçimde incelenmiş olsa da arka plan müziğinden insan sesi ya da konuşmayı ayırma üzerindeki çalışmalar sayıca daha azdır.

### Bu konuda yapılan çalışmalar

Müzik / ses ayırma yöntemleri tipik olarak vokal / vokal olmayan parçaları tanımlamak için kullanılır. Vokallerin arka plan eşliğinden ayrılmasında matris çarpanlara ayırma (Nonnegative Matrix Factorization-NMF), Ampirik Dalgacık Dönüşümü (Empirical Wavelet Transform- EWT) ve Log Frekans Güç Katsayıları (Log Frequency Power Coefficients- LFPC) gibi çeşitli teknikler kullanılmaktadır [4-8,11].

Smaragdis ve arkadaşları çalışmalarında, negatif olmayan matris çarpanlara ayırma (Nonnegative Matrix Factorization-NMF) yöntemi kullanarak yeni bir denetimli gürültü ayırıcı metot geliştirmişlerdir. Anılan metot, konuşma geliştirme amaçlı Bayesian formülü ile NMF (BNMF) kullanımı temelinde geliştirilmiştir. Bu metot ile önceki çalışmalarda yer alan eğitim ve test aşamaları arasındaki uyumsuzluk problemi ortadan kaldırılmaya çalışılmıştır. İlk olarak Saklı Markov Modeli (Hidden Markov Model - HMM) ile BNMF kullanılarak asgari ortalama karesel hata tahmincisi (Minimum Mean Square Error-MMSE) arka planda yer alan gürültü hakkında bilgi olmadan türetilir. Sonrasında arka plandaki gürültüyü öğrenmek için BNMF modeli kullanılarak denetimsiz konuşma geliştirme sistemi geliştirilir [4].

Birçok müzik parçası üzerinde değişen vokallerin üst üste geldiği müzik zeminleri tekrarlanmaktadır. Rafii ve Pardo bu yöntemi, Tekrarlayan Desen Çıkarma Tekniği (Repeating Pattern Extraction Technique – REPET) olarak adlandırmışlardır. Bu metotta ilk önce tekrar eden tahmin periyodu bulunur. Daha sonra tekrarlama periyodu kullanılarak segmentasyon matrisi hesaplanır. Yinelenen segment arka plan modeli, tekrarlanan bir spektrogram modelini türetmek için kullanılır. Segmentasyon matrisi ile müzik karışımı karşılaştırılarak minimum elemanlar bulunur bu elemanlar arka plan müziğini ortaya çıkarır [5].

Kim ve arkadaşları, gürültüye dayalı Otomatik Konuşma Tanıma (Automatic Speech Recognition -ASR) gerçekleştirebilmek için NMF temelli bir gürültü azaltma yöntemi geliştirmişlerdir. Önerilen yöntem ile durağan ve durağan olmayan gürültünün azaltılması için Wiener filtreleme ile NMF tekniği birleştirilerek hibrit bir yaklaşım oluşturulmuştur. Bu sayede hedef arka plan gürültüsü zayıflatılıp vokal sesi ön plana çıkarılmıştır [6].

Erdoğan ve arkadaşları, tek kanallı konuşma-müzik ayırma algoritmalarında spektral maskeli NMF kullanmayı önermiştir. Algoritmada NMF kullanılarak konuşma ve müzik sinyallerinin eğitimi gerçekleştirilmekte, ardından maskeleyme kullanılarak müzik ve vokal sesi, karışım sinyalden ayrılmaktadır. Eğitim aşamasında NMF her kaynak için bir dizi temel vektör seti kullanmaktadır. Eğitim genlik spektrumu bölgesinde gerçekleşmektedir [7].

Nerkar ve arkadaşları çalışmalarında, Ampirik Dalgacık Dönüşümü (Empirical Wavelet Transform- EWT) yardımıyla sinyali etkin bir şekilde parçalayarak sinyalin içerisindeki vokal ile vokal olmayan (müzikal) kısmı tekrarlı arka planı bularak ayırmayı önermiştir. Nerkar ve arkadaşları, EWT metodunun durağan ve doğrusal olmayan sinyalleri bulmak için en uygun yol olduğu ve en son teknoloji yöntemlerine kıyasla daha iyi sonuçlar verdiğini belirtmişlerdir [8].

Huang ve arkadaşlarının çalışmalarında tekrarlı ve az değişen yapısı nedeniyle arka plan müzik eşleniğini düşük seviyeli bir alt uzayda olduğu diğer taraftan vokal seslerinin ise şarkılar içinde nispeten daha seyrek ve daha fazla değiştiği varsayımı üzerine kuruludur. Bu varsayımdan yola çıkarak şarkıları öncesinde herhangi bir eğitim kullanmadan Sağlam Temel Bileşen Analizi (Robust Principal Component Analysis – RPCA) kullanılmıştır. Bahse konu konveks problemin RPCA ile çözümü için Artırılmış Lagrange Çarpanı (Augmented Lagrange Multiplier- ALM) metodu kullanarak müzikal parçaları, seyrek ve düşük seviyeli olarak adlandırılan iki matrise ayırmışlar ardından frekans maskeleyme yöntemi kullanarak müzik eşleniği ile vokal sesine ulaşmışlardır [9].

Virtanen, zamansal süreklilik ve seyreklik kriterleri ile NMF kullanarak mono ses kaynağı ayırımı yapmayı önermiştir. Bu çalışmada negatif matris çarpanlarından temel matris ve kazanç matrisi / aktivasyon matrisi, yeniden yapılandırma hata terimi, zamansal süreklilik

terimi ve seyreklik terimi olan üç terimin ağırlıklı toplamı olan maliyet fonksiyonunu en aza indirgeyerek tahmin edilmektedir [10].

Vembu çalışmasında ilk olarak şarkıdaki vokal / vokal olmayan yerleri Mel Frekans Kepstrum Katsayıları (Mel-Frequency Cepstrum- MFCC), Algısal Lineer Öngörü Katsayıları (Perceptual Linear Predictive coefficients- PLP), Log Frekans Güç Katsayıları (Log Frequency Power Coefficients- LFPC) ve Yapay Sinir ağları (YSA) kullanarak belirlemeye çalışmıştır. Daha sonra, spektrogramı vokal ve vokal olmayan temel bileşenlerden ayırmak için NMF yöntemi kullanmıştır [11].

Raj ve Arkadaşları önceden bilinen vokal segmentleri kullanarak, Olasılıksal Gizli bileşen analizi (Probabilistic Latent Component Analysis - PLCA)'ne dayalı bir eşlik modeli geliştirmişlerdir. Daha sonra anılan eşlik modelini kullanarak vokal kısımları bulmayı amaçlamışlardır [12].

Yoshi ve arkadaşları geliştirdikleri davul sesi tanıma algoritmasında spektrogramdaki davul modellerine karşılık gelen zaman frekansı şablonlarını elemanların medyanını alarak bulmaya çalışmıştır [13].

Ozerov ve arkadaşları vokal ve vokal olmayan bölümlerin ayrımı yapmak için ilk olarak MFCC ve GMM kullanmışlardır. Bu işlemin sonucunda çıkan vokal olmayan kısmı ile Bayesian eşleşim modelinin eğitimini gerçekleştirmişlerdir [14].

Bu tez çalışmasında ise müziğin tekrarlama özelliğinden faydalanarak müziği vokal ve arka plan bileşenlerine ayrılması üzerine çalışılmıştır. Bu çalışmada Rafii ve Pardo'nun geliştirdiği REPET algoritması [5] ve RPCA metodu [9] hibrit olarak kullanılarak yeni bir metod geliştirilmiş olup, objektif ve sübjektif testler sonucunda, vokal ve arka plan müziğinin her iki yöntemden de daha başarılı ayrılabilirdiği gösterilmiştir.

## 2. KURAMSAL TEMELLER

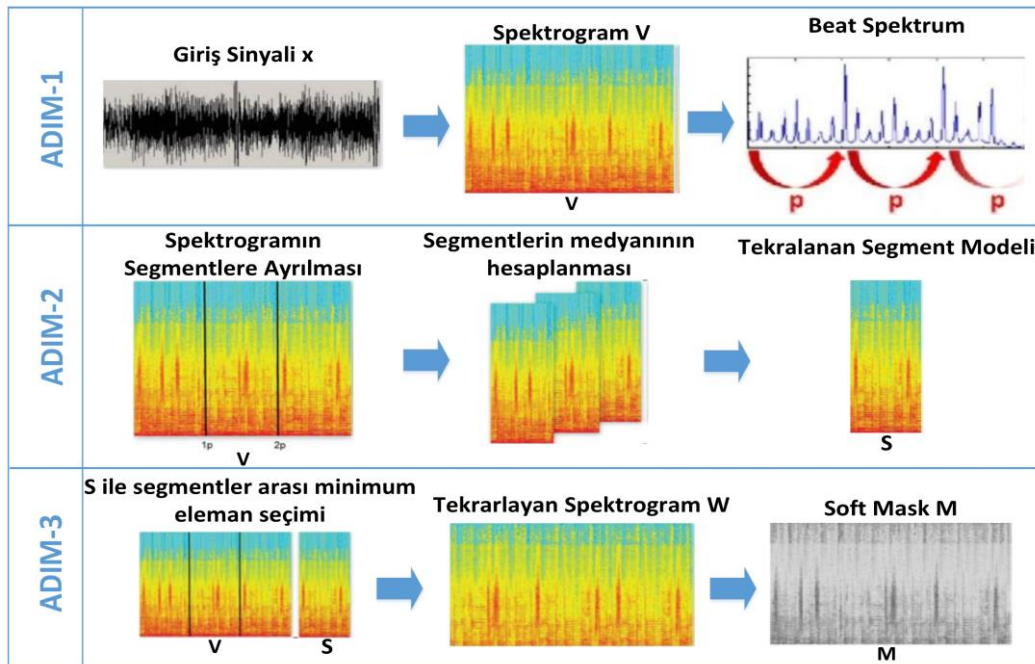
Bu bölümde tezde kullanılan yöntemleri açıklamak amacıyla bahse konu yöntemlerin kuramsal temelleri açıklanmıştır.

### 2.1. REPET

REPET (Repeating Pattern Extraction Technique – Tekrarlayan Örüntüyü Ayırma Yöntemi) algoritması, ses karışımı içerisinde tekrarlanan yapıların periyodunu tahmin ederek, periyodik olarak tekrar eden arka plan modellenmesi yaparak ses karışımı içerisinde arka planı çıkarmayı amaçlamaktadır [5]. Diğer ayırma yaklaşımlarından farklı olarak REPET,

- Belirli istatistiklere (ör. MFCC veya renk özelliklerine) bağlı olmaması,
- Karmaşık çerçevelere (ör. kaynak / filtre modellenmesi) ihtiyaç duymaması,
- Önişleme (örn. Vokal / vokal olmayan bölütleme veya önceden eğitim) gerektirmemesi,
- Yalnızca kendine benzerliğe dayalı olması ve,
- Hızlı ve kolay uygulanabilir olması

gibi özelliklere sahiptir. Şekil 2.1’de özetlenmiş olan REPET yöntemi, üç ana adımdan oluşmaktadır.



Şekil 2.1. REPET algoritması akış şeması [5]

Sisteme girdi olarak verilen ses karışımı  $x$  (vokal ve arka plan müziği birlikte içeren) Fourier dönüşümü yapılarak frekans bölgesine taşınır. Fourier dönüşümü zaman bölgesindeki fonksiyonların frekans bölgesindeki fonksiyonlara dönüşmesini sağlayan bir orthogonal dönüşümdür. Birçok araştırmacı görüntü ve ses sinyalleri işleme uygulamalarında farklı dönüşüm yöntemleri kullanmaktadır (Ayrık Kosinüs Dönüşümü, Hermite Dönüşümü, Fourier Dönüşümü vb.). Fourier dönüşümü kullanılarak sinyal bileşenlerine ayrılır ve sinyali oluşturan frekans bileşenlerinin genlikleri belirlenir [15].

Ayrık Fourier Dönüşümü (Discrete Fourier Transform - DFT) sinyalin sonlu, kısa ve süreli olması durumunda düşük bir frekans çözünürlüğüne sahiptir. DFT’de frekans ekseninde lokal bir noktadan bahsederken bu zaman ekseninde tüm zaman aralığına karşı gelmektedir. Bu nedenle durağan olmayan sinyaller için DFT dönüşümü kullanışlı değildir.

Denis Gabor tarafından geliştirilen Kısa Zamanlı Fourier Dönüşümü (Short Time Fourier Transform- STFT), DFT dönüşüm sonuçlarının daha iyi hale gelmesi için geliştirilmiş standart yöntemlere göre birçok üstünlüğe sahip bir dönüşüm metodudur. Bu yöntemde sinyal bir pencere fonksiyonu ile çarpılarak pencere fonksiyonu ile seçilen kısma DFT uygulanır. STFT ile sinyalin küçük bir zaman aralığındaki durağan olarak kabul edilen birbirlerine bitişik veya örtüşen zaman aralıklarına ayrılarak pencere fonksiyonu (Dikdörtgen, Hamming, Hanning, Blackman pencerelerinden biri olabilir) ile çarpılır. Bu işlem sonucunda her parçanın DFT’si alınarak frekans bileşenleri bulunur ve STFT elde edilir [16]. STFT bir fonksiyonu, pencereleme kullanımı ile uygun küçük parçalara ayırıp inceleyerek anlık frekans içeriğini bulmakta kullanılmaktadır.

### **2.1.1.REPET – 1. Adım: Vuruş spektrumunun hesaplanması**

Sisteme girdi olarak verilen ses karışımı  $x$ ’ten STFT işlemi ile elde edilen fourier dönüşümünün mutlak değeri alınarak ve simetrik kısım atılarak genlik spektrogramı  $V$  oluşturulur. Dolayısıyla,  $V$  matrisinin dikey boyutu frekans, yatay boyutu zaman ekseni olur ve bu matrisin elemanları Fourier spektrumundaki genlikleri temsil etmektedir. Spektrogramlar çizilirken üç boyutlu çizim yerine genlikler renklerle ifade edilmektedir (Bkz. Şekil 2.1). Elde edilen spektrogramın elemanlarının karesi alınır ( $V^2$ ). Yapılan işlemde genlik spektrogramının karesinin kullanılmasının nedeni tepe noktaları

vurgulamak ve periyot tahminini kolaylaştırmaktır. Daha sonra  $V^2$  matrisinin her bir satırının özilintisi hesaplanarak  $B$  matrisi oluşturulur.

Özilinti fonksiyonu, Eş. 2.1'de tanımlandığı gibi bir serideki elemanların kendi aralarındaki benzerliğinin zamanın (gecikmenin) fonksiyonudur.

$$R(\tau) = \frac{1}{(T_{maks}-T_{min})} \int_{T_{min}}^{T_{maks}} s(t)s(t + \tau)dt; \quad T_{maks} > \tau > T_{min} \quad (2.1)$$

Eş. 2.1'deki ifadede  $R$ ,  $\tau$  gecikmesi fonksiyonu olarak özilintiye ifade eder. Bu kapsamda özilinti gözlemlenen eleman ile önceki eleman arasındaki benzerliğin, zamansal gecikme fonksiyonu ( $\tau$ ) olarak ifadesidir. Sinyal işleme alanında tekrar eden örüntülerin tanınması, bir sinyalin kayıp temel frekansının tespit edilmesi ve sinyaldeki periyodikliklerin bulunması gibi amaçlar için kullanılabilir. Bu tezde de ses sinyallerinin frekans spektrumu içerisinde yer alan periyodik yapıların tespitini yapmak ve dolayısıyla arka plan müziğinin tekrarlı yapısını ortaya çıkartmak amacıyla kullanılmıştır [17]. Özilinti matrisi  $B$ , Eş. 2.2'deki gibi hesaplanır. E

$$B(i,j) = \frac{1}{m-j+1} \sum_{k=1}^{m-j+1} V^2(i,k) V^2(i,k+j-1) \quad (2.2)$$

Eş. 2.2'de  $i$ , frekans boyutunu,  $j$  ise zaman boyutunu temsil etmektedir.  $m$  zaman boyutunun en yüksek elemanını, başka bir deyişle  $V^2$  matrisinin yatay boyutudur. Bu işlem sonrasında  $B$  matrisinin sütunlarının ortalamaları alınarak vuruş spektrumu (beat spectrum) olarak adlandırılan Eş. 2.3'te yer alan  $b$  matrisi oluşturulur [5].

$$b(j) = \frac{1}{n} \sum_{i=1}^n B(i,j) ; \quad b(j) = \frac{b(j)}{b(1)}$$

$$i = 1 \dots n \text{ (frekans)} ; \quad n = \frac{N}{2} + 1, j = 1 \dots m \text{ (zaman)} \quad (2.3)$$

Oluşan  $b$  vektörü (vuruş spektrumu), zamandaki tekrarlı yapıyı ortaya çıkartmaktadır.  $b$  vektörünün yinelenen periyodu olarak ifade edilen  $p$ 'yi, otomatik olarak tahmin etmek için basit bir prosedür kullanır (Bkz. Şekil 2.1). Temel fikir,  $b$  matrisinin hangi tam sayı katlarında en yüksek ortalama birikim enerjisine sahip olduğunu bulmaktır.

Vuruş spektrogramı oluşturulurken özilinti kullanılır. Fakat özilinti en uzun gecikme süresi çoğu zaman güvenilir değildir, çünkü hesaplama yapılırken zamanla ileriye gidildiğinde, benzerliği hesaplamak için daha az katsayı kullanılır. Bu yüzden  $b$  matrisin son 1/4 elemanı  $p$  periyodu kullanılırken ihmal edilir. Daha sonra geniş, ancak çok geniş olmayan pencere oluşturur. Pencere,  $b$  matrisinin son 1/4 elemanı atıldıktan sonraki değerinin en az 1/3 büyüklüğünde seçilmelidir. Penceredeki en büyük elemanın bulunduğu nokta  $p$  periyodu olarak kabul edilir. Tempo sapmaları nedeniyle, yinelenen zirveler  $p$ 'nin tamsayı katları olmayabilir. Bu durumda  $b$  matrisinin  $p$  periyodunun tam katı olması için  $b$  matrisine sıfır dolgulama (zero padding) işlemi yapılır.

### 2.1.2.REPET – 2. Adım: Tekrarlanan segmentin bulunması

Yinelenen periyot  $p$  bulunduktan sonra daha önce oluşturduğumuz genlik spektrogramı  $V$ 'den  $p$ 'nin tamsayı bir katı kadar uzunlukta matrisle işleme devam edilir.  $p$  uzunluğundaki  $V$  matrisinden ayrılan parçaların her birinin ortalama değerinin bulunması için medyan kullanılır. Bir dizideki elemanlar küçükten büyüğe doğru sıralandığında, ortada bulunan elemana ortanca değere medyan denir [18]. Dizideki elemanların sayısı tek ise ortanca değer en ortada bulunan değer; eleman sayısı çift ise ortanca değer, ortadaki iki değer toplamının yarısı olur. Bir dizideki olasılık dağılımı simetrik olamayıp, dizi beklenmedik büyük veya aykırı küçük değerler içeriyorsa medyan aritmetik ortalama daha iyi merkez değer vermektedir. Algoritmada yer alan merkez bulma işleminde medyan kullanılmasının nedeni, tekrarlanan ve tekrarlanmayan kısımları ayırmada daha iyi olduğunun düşünülmesidir. Ayrılan parçaların medyanı alınarak oluşturulan Eş. 2.4'te yer alan matris,  $S_m$  tekrarlanan segmentasyon modeli adı verilir.

$$S_m(i, l) = \underset{k=1..r}{\text{medyan}} \{V(i, l + (k - 1)p)\}$$

$$i = 1 \dots n \text{ (frekans)}, l = 1 \dots p \text{ (zaman)}, p = \text{periyot uzunluğu ve } r = \text{segment sayısı} \quad (2.4)$$

Böylece Şekil 2.1'deki 2. adımda gösterilen  $V$  matrisinden alınan  $p$  zaman uzunluğundaki parçaların her bir elemanının kendi içinde medyanından oluşan bir  $S_m$  matrisi oluşturulur.

### 2.1.3. REPET – 3. Adım: Tekrarlanan desenlerin çıkarılması

Tekrarlanan segmentasyon modeli,  $S_m$ , hesaplandıktan sonra,  $S_m$  ile daha önce ayrılan  $r$  adet bölümün elemanları karşılaştırılarak bunların minimum olanları bir matrise aktarılır. Bu matris  $W$  olarak adlandırılan, tekrarlanan spektrogram modelini Eş. 2.5'deki gibi oluşturulur.

$$W(i, l+(k-1)p) = \min \{S_m(i, l), V(i, l+(k-1)p)\}$$

$$i = 1 \dots n, (\text{frekans}) \text{ ve } j = 1 \dots m(\text{zaman}) \quad (2.5)$$

Bu işlemden sonra  $M$  olarak adlandırılan zaman-frekans maskesi,  $W$  matrisi elemanlarının  $V$  ile normalize edilmesi ile Eş. 2.6'deki gibi oluşturulur. Son olarak elde edilen  $M$  maskesi giriş sinyali spektrogramı olan  $V$  ile eleman eleman çarpılır. Elde edilen matristen ters STFT alınarak tahmin edilen arka plan müziği elde edilir. Müzikal parçada yer alan vokal sesi ise  $x$  giriş sinyalinden daha önce elde edilen arka plan müziğinin çıkarılması ile elde edilir.

$$M(i, j) = \frac{W(i, j)}{V(i, j)}, \quad M(i, j) \in [0, 1]$$

$$i = 1 \dots n (\text{frekans}) \text{ ve } j = 1 \dots m (\text{zaman}) \quad (2.6)$$

## 2.2. RPCA

RPCA (Sağlam Temel Bileşen Analizi - Robust Principal Component Analysis), günümüzde veri analizi ve boyut azaltımı için kullanılan popüler bir istatistiksel araçtır. Bu bölümde ilk olarak olarak RPCA problemi çözmek için gerekli olan kuramsal temeller anlatılmış olup bu kısmın sonrasında RPCA detaylı olarak incelenmiştir.

### 2.2.1. Normlar uzayı

Lineer cebirde  $\vec{v}$  vektörünün büyüklüğü  $\vec{v}$ 'nin normu olarak adlandırılmaktadır. Bu çalışmada kullanılan RPCA algoritmasında, birkaç farklı vektör normu kullanılmasından dolayı bu bölümde vektör normlarının tanımları yapılmıştır.

**L<sub>1</sub>-norm**

L<sub>1</sub> normu, Eş. 2.7’de görüldüğü üzere bir vektörün tüm bileşenlerinin mutlak değerlerinin toplamı olarak tanımlanır [19].

$$|x|_1 = \sum_{r=1}^n |x_r| \quad (2.7)$$

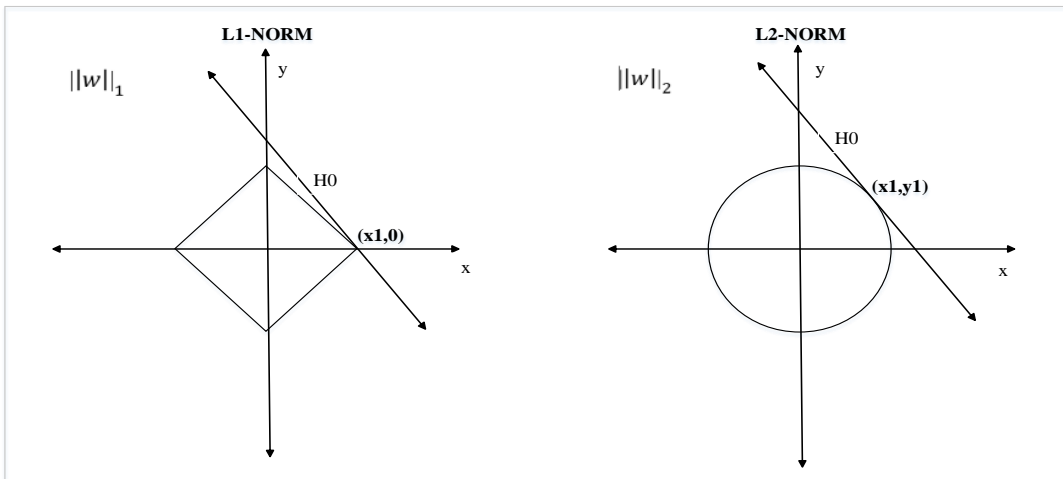
**L<sub>2</sub>- norm**

Gerçel ya da karmaşık sayılardan oluşan ve terimlerin mutlak değerlerinin karelerinin toplamı sonlu olan bütün dizilerin oluşturduğu uzaya L<sub>2</sub> uzayı denir. Eş. 2.8’da yer alan L<sub>2</sub> norm ayrıca Öklit norm olarak da adlandırılır [20].

$$|x|_2 = \sqrt{\sum_{k=1}^n |x_k|^2} \quad (2.8)$$

**L<sub>1</sub> norm ve L<sub>2</sub> norm karşılaştırması**

Genellikle L<sub>2</sub> norm kullanmak, L<sub>1</sub> norma göre tahmin hatasını asgariye indirmek için daha iyi olarak kabul edilmektedir. Bunun nedeni L<sub>1</sub> normun iki öngörücülerden birini seçerken, L<sub>2</sub> normun birbiriyle yüksek korelasyona sahip iki öngörücü kullanmasıdır. Fakat bu duruma karşın L<sub>1</sub> norm, sıfır değerli çok sayıda katsayı ya da az sayıda büyük katsayı içeren çok küçük değer üretme özelliğine sahiptir.



Şekil 2.2. L<sub>1</sub> ve L<sub>2</sub> norm gösterimi

Şekil-2.3’de iki boyutlu uzayda  $H_0$  hipotez eğrisi ve çözümde kullanılan  $L_1$  ve  $L_2$  norm çözümleri görülmektedir.  $L_1$  birim topları "elmas şeklinde" iken,  $L_2$  birim topları "döner simetrik" olarak tanımlanmıştır.

İki boyutlu uzayda yer alan sistemi  $L_2$  norm ile çözdüğümüzde  $H_0$  eğrisinin ve  $L_2$  norm çemberine teğet olduğu noktanın hem  $x$  ve hem de  $y$  eksenlerinde değerinin olduğunu, fakat sistemi  $L_1$  norm ile çözdüğümüzde  $L_1$  normun dörtgeninin doğası gereği  $H_0$  eğrisi ile kesiştiği noktanın sadece  $x$  eksenini kesiştiğinin görmekteyiz (Bkz. Şekil-2.3). Sistemin  $L_1$  norm çözümünde,  $y$ 'nin rolünün olmaması seyrek çözüm bulmada daha iyi olduğunu göstermektedir. Bu nedenle, bu tez kapsamında RPCA algoritmasında  $L_1$  norm tercih edilmiştir.

### Nükleer norm ( $L_*$ )

Bir matrisin rankını minimize etme, kontrol sistemleri, boyut indirgeme ve gerçekleştirme teorisi vb. alanlarında karşılaşılan oldukça zorlu bir sorundur. Birçok çarpanlara ayırma, kestirim veya yaklaşım metodunda, Eş. 2.9’da gösterildiği gibi bir matrisin rankını minimuma indirmek için nükleer norm kullanarak konveks problemin çözümü gerçekleştirilmektedir. Nükleer norm bir matrisin tekil değerlerinin toplamı olarak tanımlanmaktadır.

Minimizasyon problemi,  $R^{m \times n}$  matris-rank kısıtlamalarını içerecek düşük rank yaklaşımı şekilde Eş. 2.9’da verilmiştir.

$$\min_x = \text{rank}(X), \text{ öyle ki } L(X) = D \quad (2.9)$$

Bu ifadede  $L$ , matris alt örnekleme operatörü olmak ile birlikte  $L(X)$ ,  $X$ 'in tüm girişlerinin bir alt kümesidir.  $D$  ise oluşturulabilecek en düşük ranklı matrisi ifade etmektedir. Eş. 2.9’da yer alan ifade pratikte az kullanıldığı için bunun yerine " $\| \cdot \|_*$ " olarak temsil edilen matrisin nükleer normu Eş. 2.10’daki gibi gösterilebilir. Bu ve benzeri problemlerin çözümünde kullanılan en iyi konveks yaklaşımlardan biridir. Candès ve Recht çalışmalarında çoğu düşük-rank matrisin nükleer norm kullanarak yeniden oluşturulabildiğini Eş. 2.10’da gösterildiği şekilde kanıtlamışlardır [21].

$$\min_x \|X\|_* , \quad \text{öyle ki } L(X) = D \quad (2.10)$$

### Frobenius norm ( $L_F$ )

Frobenius norm,  $m \times n$  boyutlu bir matrisin elemanlarının mutlak değerlerinin karelerinin toplamının karekökü olarak hesaplanır ve *vektör normu* olarak da adlandırılır. Frobenius norm, Eş. 2.11'da gösterilmiştir [22].

$$\|X\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |x_{ij}|^2} \quad (2.11)$$

### Sonsuzluk normu ( $L_\infty$ )

Sonsuzluk normu Eş. 2.12'de ifade edildiği gibi;  $\vec{x}$  vektörünün bileşenlerinin mutlak değerlerinin maksimumu, sonsuzluk normu olarak adlandırılır ve " $\|\cdot\|_\infty$ " ifadesi ile gösterilir [23].

$$\|X\|_\infty = \max_i |x_k| \quad (2.12)$$

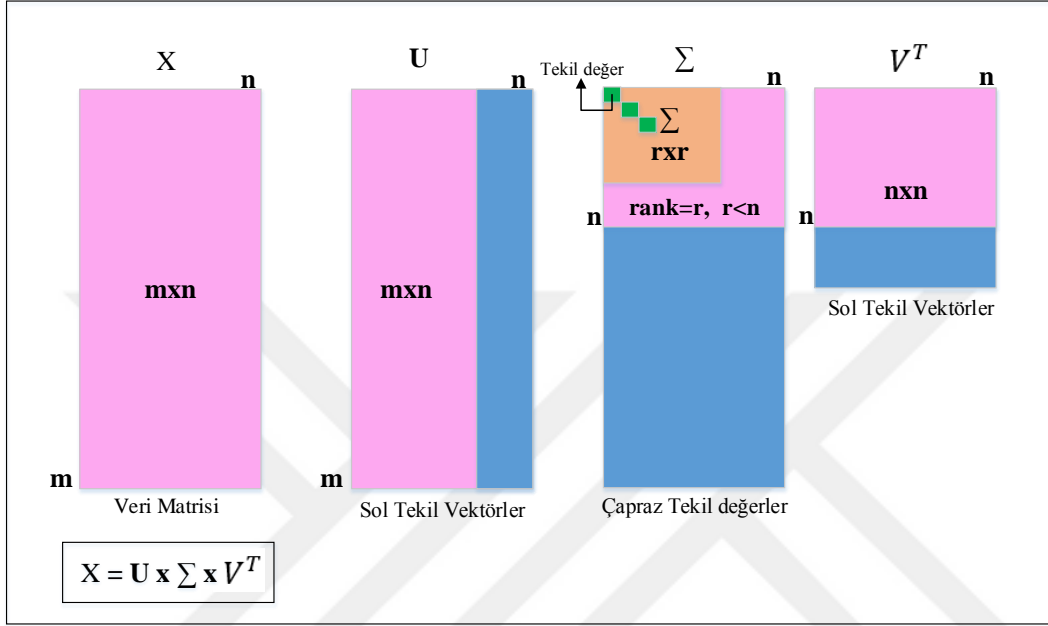
## **2.2.2. Tekil değer ayrışımı (Singular Value Decomposition-SVD)**

Tekil Değer Ayrışımı (Singular Value Decomposition – SVD) bir matrisi çarpanlara ayırmak, gen analizi, boyut azaltımı gibi geniş uygulama alanlarında kullanılmaktadır. SVD yönteminde amaç veriyi kovaryans değerlerine göre boyutlara ayırarak verinin boyutunu indirgemektir [24]. SVD ayrışımında elde edilen ilk bileşen en yüksek kovaryans değerine sahip olmakla birlikte daha sonraki bileşenlerde değer azalarak son elemanda en düşük kovaryans değerine sahip olmaktadır [25].

Şekil 2.3 ve Eş. 2.13'te görüldüğü gibi her  $X \in \mathbb{C}^{m \times n}$ , matrisini SVD yöntemiyle 3 parçaya ayırarak çarpanlarına ayırabiliriz.

$$X = U \Sigma V^T \quad (2.13)$$

$U$  matrisi  $m \times n$  boyutlu,  $V$  matrisi  $n \times n$  boyutlu iki üniter matris ve  $\Sigma$ ,  $n \times n$  boyutlu köşegen matrisi olup,  $\Sigma$  matrisinde  $X$  matrisinin tekil değerleri tutulmaktadır. Eş. 2.14'de olduğu gibi  $\Sigma$  matrisinde en büyük tekil değer üst köşegenden başlamak üzere azalarak aşağıdaki gibi yazılmalıdır.



Şekil 2.3. Tekil değer ayrışımı (SVD)

$$\Sigma = \text{köşegen}(\sigma_1, \sigma_2, \dots, \sigma_{\min(i,j)}) \quad (\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(i,j)} \geq 0) \quad (2.14)$$

$U$ ,  $\Sigma$  ve  $V$  matrislerini bulmak için ilk olarak Eş. 2.15'te bulunan ifade kullanılır.

$$X^T X = U \Sigma^T \Sigma V^T \quad (2.15)$$

İlk olarak  $X$  matrisinin öz değerleri hesaplanır. Öz değerler kullanılarak anılan matrisin öz vektörleri bulunur.  $X$  matrisinin öz vektörlerinin birleşimi  $V$  matrisini oluşturur.  $X$  matrisinin öz değerlerinin köşegen matrisin köşegenine büyükten küçüğe sıralanması sonucunda tekil değerlerin tutulduğu  $\Sigma$  matrisine ulaşılır.  $U$  matrisine ulaşmak için Eş. 2.16'da ifade edilen denklem kullanılır.

$$XV = U\Sigma \quad (2.16)$$

SVD hesaplanırken  $V$ 'nin yalnızca  $r < n$  olmak üzere  $r$  adet sütunu kullanılarak dönüşüm yapıldığında, dönüştürülmüş verinin boyutu,  $n$ 'den  $r$ 'ye indirgenmiş olur. Bu özellik, tezin ilerleyen kısımlarında bir matrisin düşük rank (low-rank) bileşenini bulmak için kullanılacaktır.

### 2.2.3. Temel bileşen çözümleme (Principal component analysis - PCA)

PCA, dağılmış bir veri kümesinin temel özelliklerini toplam varyansa dayalı olarak bulmak için kullanılan istatistiksel bir yaklaşımdır [26]. Varyans, bir seri içindeki tüm sayıların, o serinin ortalaması ile farklarının, karelerinin, toplamlarının, ortalamasıdır. Eş. 2.17'de gösterilen eşitlikte  $N$  uzunluğundaki  $Z$  Serisi içindeki sayıların ortalamasını  $\bar{Z}$  olarak ifade edersek, varyans aşağıdaki gibi hesaplanır:

$$Z_X = \frac{1}{N} \sum_{i=1}^N (Z_i - \bar{Z})^2 \quad (2.17)$$

$X$ - $Y$  koordinat sisteminde çok değişkenli dağıtılmış veri seti göz önüne alındığında, PCA ilk önce orijinal veri kümelerinin maksimum varyanslarını bulur. Bu veri noktaları  $U$ - $V$  koordinat sistemi olarak adlandırılan yeni bir eksene aktarılır.  $U$  ve  $V$  ekseninin yönü ana bileşen olarak adlandırılır.

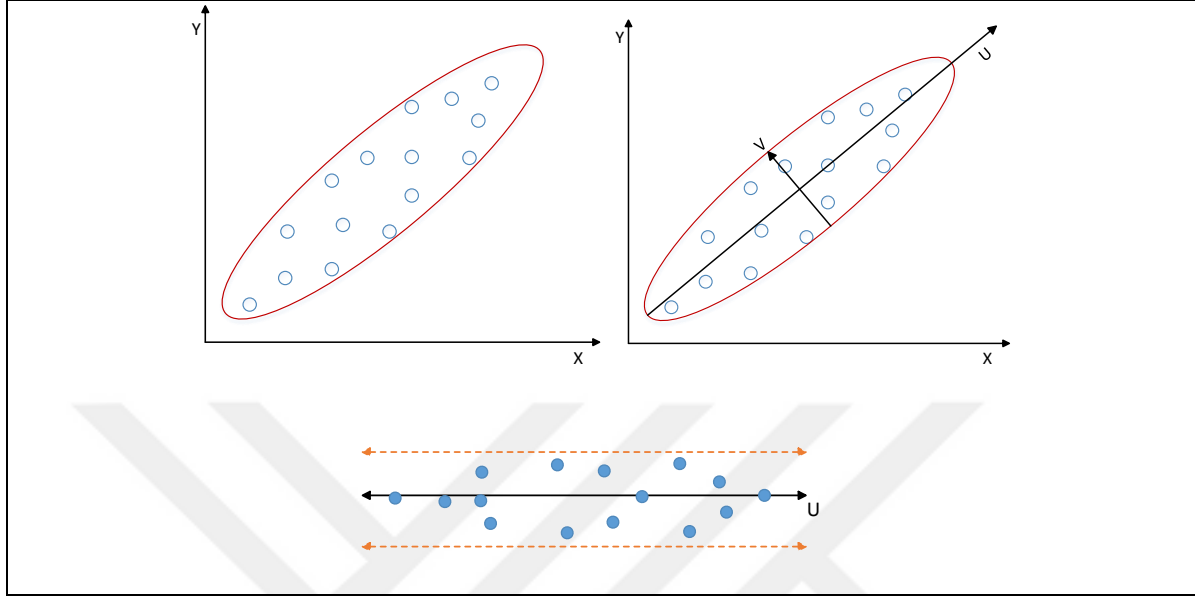
Ana bileşendeki elemanların  $V$  ekseninde sahip olduğu değerler sıfıra çok yakın olmasından dolayı (varyans değeri düşük),  $V$  değişkeni ihmal edilerek veri kümesi sadece  $U$  değişkeni ile temsil edilebilir. Bu işlemler esnasında verinin bazı özellikleri kaybolabilmektedir, fakat burada asıl amaç varyansı yüksek tutarak minimum derecede kayıpla çalışma yapabilmektir. Boyut indirgeme işlemi Şekil 2.4'te gösterildiği gibi; ilk olarak varyanslara göre eksenlerin çıkarılması ardından sıfır ortalama alınacak şekilde ayarlanıp eksen dönüştürülmesi şeklinde gerçekleştirilmektedir.

PCA yüksek boyutlu veri işleme ve analizi, resim sıkıştırma ve örüntü tanıma gibi bilimsel ve mühendislik alanlarında kullanılabilen popüler bir araçtır.

PCA'nın kullanılmasının temel amaçları:

- Verilerin boyutunu azaltmak,
- Düşük boyutlu alt uzayı etkili ve doğru bir şekilde tahmin etmek,

- Verileri analizler için görüntülemek, olarak sıralanabilir.



Şekil 2.4. Tekil değer ayrışımı eksen azaltımı

#### 2.2.4. Lagrange çarpanları

Lagrange çarpanları tekniği çok değişkenli fonksiyonların sınırlandırılmış aralıkta maksimum ve minimum değerlerini bulmak için kullanılır. Bu teknik aşağıdaki ifadeye benzeyen kısıtlamalar için geçerlidir. Eş. 2.18'de tanımlandığı gibi, Lagrange çarpanları metodu,  $g$  çok değişkenli bir fonksiyon ve  $k$  sabit sayı olmak üzere;

$$\nabla f(x, y) = \lambda \nabla g(x, y), \quad \lambda \neq 0$$

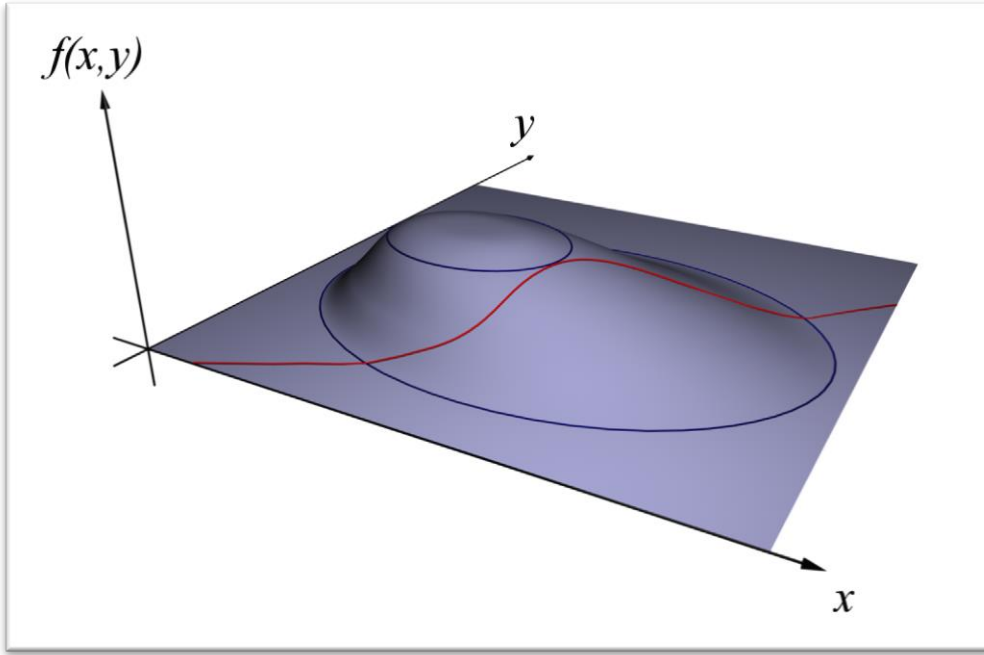
$$\nabla f // \nabla g \text{ maksimum ve minimum noktalar için } g(x, y) = c \quad (2.18)$$

olarak ifade edilebilir.

$\lambda$  değişkeni, "Lagrange çarpanı" olarak adlandırılır. Örnek olarak  $f$  fonksiyonunun girdi alanı iki boyutluysa,  $f$  fonksiyonu üzerine yansıtılan  $g(x, y) = c$  kısıtlaması Şekil 2.5'te yer alan grafikteki gibi kırmızı çizgi ile gösterilebilir. Lagrange çarpanı kullanımındaki amaç Şekil 2.5'te yer alan kırmızı çizgideki en yüksek noktayı bulmaktır. Temel fikir,  $f(x, y)$  ve  $g(x, y)$  fonksiyonlarının kontur çizgilerinin birbirine teğet olduğu noktaları aramaktır.

Maksimum ve minimum noktalarda  $f(x,y)$  fonksiyonun ve  $g(x,y)$  kısıtlama fonksiyonun gradyan (meyil) vektörleri bir birine paralel olmak zorundadır. Bu durum,  $f$  ve  $g$  gradyan vektörlerinin birbirine paralel olduğu noktaları bulma ile aynı mantığa sahiptir [27]. Lagrange eşitliği Eş. 2.17'deki ifade ile gösterilir.

$$L(x,y, \lambda) = f(x,y) - \lambda(g(x,y)-c) \quad (2.17)$$



Şekil 2.5. Lagrange çarpanı kullanımı ile maksimum noktanın tespiti [28]

Bu yöntemde ekstremum noktaları bulmak için  $f(x,y)$  ve sınır koşulu  $g(x,y)$  fonksiyonlarının fonksiyonlardaki değişkenlere göre kısmi türevi alınır.  $f$  fonksiyonun gradyan vektörü ( $\nabla f$ ) lambda ile çarpılmış  $g$  fonksiyonunun gradyan ( $\nabla g$ ) vektöründen çıkarılarak sıfıra eşitlenir. Bu denklem çözülerek ekstremum noktalara ulaşılır.

### 2.2.5. Artırılmış Lagrange çarpanı (Augmented Lagrange multiplier- ALM)

ALM kısıtlı optimizasyon problemlerinin çözmek için kullanılan Eş 2.19'da ifade edilen sınırlamalara sahip bir algoritmadır.

$$\min f(x) \text{ olmak üzere } h(x) = 0 \quad (2.19)$$

Sınırlandırılmış bir optimizasyon problemine penaltı teriminin eklenmesi ve problemin bir dizi kısıtlı problem haline getirilmesi nedeniyle penaltı metoduna benzemekle birlikte, penaltı terimine sahip olması ile Lagrange çarpanı tekniğinden ayrılmaktadır [29]. Penaltı metodu kısıtlı bir optimizasyon problemini, çözümleri asıl kısıtlı sorunun çözümüne ideal olarak yakınsayan bir dizi kısıtlanmamış problem ile değiştirir. Bu metotta genel mantık, kısıtlamaların ihlali için penaltı teriminin fonksiyona eklemesidir. ALM Fonksiyonu Eş. 2.20'deki gibi ifade edilir:

$$L(X, Y, \mu) = f(x) + \langle Y, h(x) \rangle + \frac{\mu}{2} \|h(x)\|_F^2 \quad (2.20)$$

Eş. 2.20'de  $\mu$  penaltı terimi pozitif bir skaler,  $Y$  Lagrange çarpanı ve  $h(x)$  problemin kısıtı olarak tanımlanır. RPCA problemini çözmek için Eş. 2.20'deki fonksiyon Eş. 2.21'de ifade edilen  $f(x)$  fonksiyonunu ve  $h(x)$  kısıtlama fonksiyonu kullanılmaktadır.

$$X = (L, S), f(x) = \|L\|_* + \lambda \|S\|_1, h(x) = X - L - S \quad (2.21)$$

RPCA probleminin çözmek için, ALM fonksiyonu Eş. 2.22'deki gibi ifade edilir:

$$L(L, S, Y, \mu) = \|L\|_* + \lambda \|S\|_1 + \langle Y, X - L - S \rangle + \frac{\mu}{2} \|X - L - S\|_F^2 \quad (2.22)$$

Eş. 2.22 de eşitliğin sağında kalan “ $\|L\|_* + \lambda \|S\|_1$ ” işlemi çözmek istenilen problemi, ikinci işlem “ $\langle Y, X - L - S \rangle$ ” Lagrange fonksiyonunu ve “ $\frac{\mu}{2} \|X - L - S\|_F^2$ ” işlemi ise penaltı fonksiyonunu ifade etmektedir

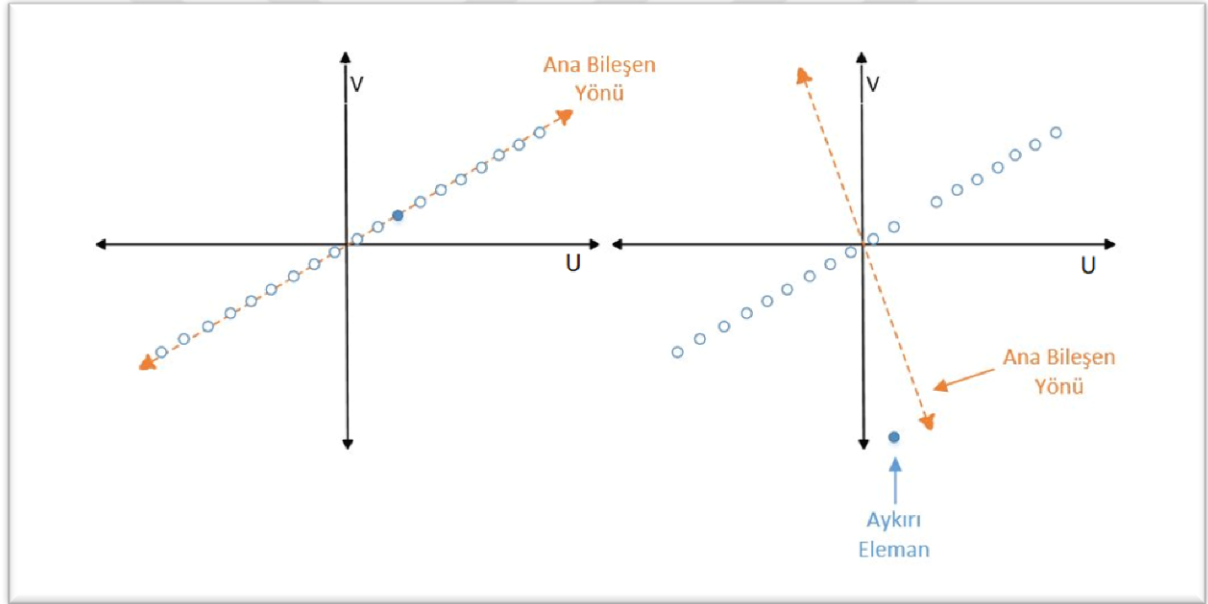
### 2.2.6. RPCA Detaylı incelenmesi

PCA tarafından oluşturulan ana bileşenlerin yönleri bazı aykırı elemanların olması durumunda Şekil 2.6'da gösterildiği gibi doğru sonuçları vermemektedir. Bunun nedeni PCA'nın maksimum varyanslar ile ana bileşenini bulmasıdır. Bu duruma en büyük varyansa sahip elemanın aykırı eleman olması neden olmuştur (Bkz. Şekil 2.6) [30].

Büyük boyutlu bir verinin  $D \in R^{m \times n}$  olacak şekilde matrisinin sütunları şeklinde yerleştirilmiş olarak düşünersek, düşük boyutlu alt uzayda düşük ranklı  $A$  matrisi bulmak için Eş. 2.23'teki kısıtlı optimizasyon kullanılır:

$$\begin{aligned}
& \min_{A,E} \|E\|_F, \\
& \text{rank}(A) \leq r, \\
& D = A + E
\end{aligned} \tag{2.23}$$

$r \ll \min(m,n)$ , olacak şekilde hedef alt uzayın boyutu, " $\|\cdot\|_F$ " verilerin Gauss gürültüsü tarafından bozulduğu varsayımına karşılık gelen Frobenius normudur. Bu problem,  $D$ 'nin ana sol tekil vektörleri tarafından kapsanan alt uzay üzerine yansıtılarak kolaylıkla çözülebilir. PCA, bozulmanın ilave Gauss gürültüsünden kaynaklandığı zaman yani gürültünün büyüklüğü küçük olduğu sürece olumlu sonuçlar vermektedir. Ancak büyük oranda bozulmanın olduğu şartlar altında düzgün sonuçlar vermemektedir.



Şekil 2.6. Aykırı elemanların olması durumunda ana bileşen yönü

Aslında,  $A$ 'nın yalnızca bir girişi aykırı olsa bile, klasik PCA ile elde edilen tahmini  $A$ , gerçek  $A$ 'dan uzak olabilir. Wright ve arkadaşları  $E$  matrisini yeterince seyrek olduğu sürece  $D = A + E$  eşitliğinden  $A$  düşük dereceli matrisini Eş. 2.24'de yer alan dışbükey optimizasyon problemlerini çözerek elde edilebildiğini kanıtlamışlardır [31].

$$\min \|L\|_* + \lambda \|S\|_1$$

$$L + S = |X| \tag{2.24}$$

$\|\cdot\|_*$  ifadesi matrisin nükleer normunu yani tekil değerlerinin toplamını,  $\|\cdot\|_1$  ifadesi bir matrisin elemanlarının mutlak değerlerinin toplamını,  $\lambda$  ise pozitif ağırlık elemanı olarak adlandırılır.

### RPCA probleminin çözümünde ALM kullanımı

RPCA problemini çözmek için tam SVD'nin hesaplanmasına gerek yoktur. Sadece belirli bir eşik değerinden daha büyük olan tekil değerlere ve bunların tekil vektörlerinin kullanımı yeterlidir. Belirli eşik değerinden büyük tekil değerleri PROPACK adlı yazılım paketi kullanılarak bulunabilmektedir [32]. RPCA probleminin ALM algoritması kullanılarak çözümünde aşağıdaki adımlar uygulanır [33]:

- 1- İlk olarak giriş değeri olan  $X$  matrisinin  $L_2$  normu, Eş. 2.25'deki gibi  $X$  matrisinin karelerinin toplamının karekökü alınarak hesaplanır.

$$\|X\|_2 = \sqrt{\sum_{k=1}^n |x_k|^2} \quad (2.25)$$

- 2- Daha sonra Eş. 2.26'da gösterildiği gibi  $X$  matrisinin sütunlarının  $L_\infty$  normu hesaplanır.

$$\|X\|_\infty = \max_i |x_k| \quad (2.26)$$

- 3- Eş. 2.27'de ifade edilen  $L_2$  norm ve  $L_\infty$  normun,  $\lambda$  simgesi ile gösterilen değişim parametresine bölümü karşılaştırılarak aralarında büyük olan seçilir ve ikili norm olarak adlandırılır.

$$\text{İkili norm} = \max(\|X\|_2, (\|X\|_\infty / \lambda)) \quad (2.27)$$

- 4-  $X$  matrisi ikili norma bölünerek iterasyonun başlangıcında kullanılacak  $Y$  matrisini oluşturulur.

- 5- Penaltı parametresi  $\mu$ , Eş. 2.28'deki gibi hesaplanır

$$\mu=1.25/\|X\|_2 \quad (2.28)$$

- 6- İterasyon aşamasına geçilmeden önce son olarak  $X$  matrisinin Frobenius normu Eş. 2.29'daki gösterildiği gibi hesaplanır.

$$\|X\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |x_{ij}|^2} \quad (2.29)$$

- 7- Düşük rank  $L$  ve seyrek (Sparse)  $E$  matrisleri boyutları  $M$  matrisinin boyutlarına eşit olacak şekilde Eş. 2.30'daki gibi sıfır matrisler oluşturulur.

$$X \in R^{n1 \times n2}, L \in R^{n1 \times n2}, E \in R^{n1 \times n2} \quad (2.30)$$

- 8- İterasyon ilk olarak; giriş matrisi  $X$ , iterasyon başlangıç matrisi  $Y$  ve penaltı parametresi kullanılarak Eş. 2.31'deki gibi geçici  $T$  olarak adlandırılan matrisin hesaplanması ile başlar.

$$T = X - L + (1/\mu) * Y \quad (2.31)$$

- 9-  $T$  matrisinden  $\lambda/\mu$  çıkarılması işleminden sonra,  $T$  matrisindeki elemanlardan sıfırdan büyük olanları seyrek  $E$  matrisine yazılır.

- 10-  $T$  matrisi elemanlarına  $\lambda/\mu$  eklenmesi işleminden sonra  $T$  matrisindeki elemanlardan sıfırdan küçük olan elemanların oluşturduğu matris  $E$  matrisi ile Eş. 2.32'deki gibi toplanır.

$$E_{k+1} = \arg \min_k L(L, E_k, Y_k, \mu_k) \quad (2.32)$$

- 11- SVD kullanılarak  $sv$  olarak adlandırılan başlangıçta belirlenen tahmini boyut kadar tekil değer Eş. 2.33'teki gibi hesaplanır.

$$(U, S, V) = \text{svd}(X - L_k + \mu_k^{-1} * Y_k) \quad (2.33)$$

- 12- SVD işlemi sonucu çıkan tekil değerlerin  $1/\mu$  değerinden büyük olanlarının sayısı tespit edilir ve iterasyonda kullanılan tahmini boyut ve kullanılan tekil değerler Eş. 2.34'teki ifadeye göre güncellenir.

$d=\min(m,n)$  olmak üzere;

$$sv_{k+1} = \begin{cases} svp_k + 1 & , svp_k < sv_k \\ \min(svp_k + \text{round}(0.05d), d), & svp_k = sv_k \end{cases} \quad (2.34)$$

Eş. 2.34'te  $sv_k$  , tahminin boyut ilk iterasyonda on olarak alınır,  $svp_k$  ise SVD işlemi sonucu çıkan tekil değerlerin  $1/\mu$  değerinden büyük olanlarının sayısıdır.

- 13- Düşük-ranklı  $L$  matrisi Eş. 2.35'deki gibi  $sv$  kadar tekil değer kullanılarak oluşturulur.

$$L_{k+1} = US_{\mu_k-1} [S] V^T \quad (2.35)$$

- 14- İterasyonda oluşturulan  $L$  ve  $E$  matrisleri  $X$  Giriş matrisinden çıkarılır ve  $Z$  matrisi Eş. 2.36'daki gibi oluşturulur.

$$Z_{k+1} = X - L_{k+1} - E_{k+1} \quad (2.36)$$

- 15-  $Y$  matrisi Eş. 2.37'deki gibi işlem yapılarak bir sonraki iterasyonda kullanılmak üzere güncellenir.

$$Y_{k+1} = Y_{k+1} + \mu_k (X - L_{k+1} - E_{k+1}) \quad (2.37)$$

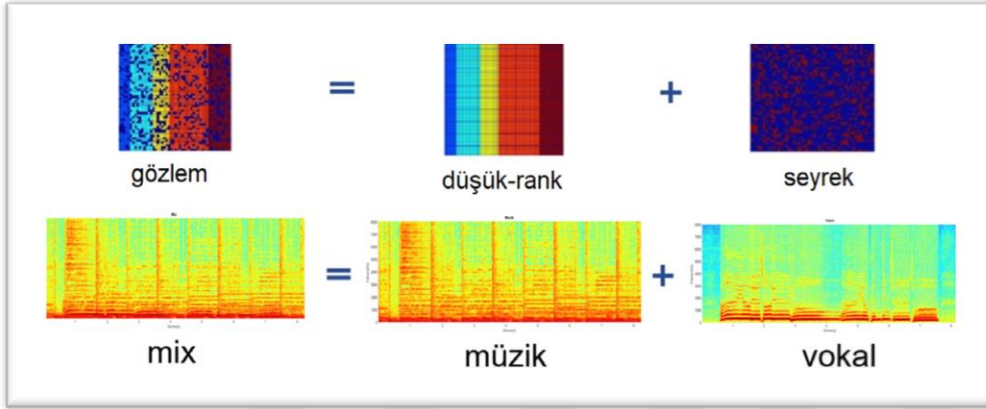
- 16- İterasyonun bitmesi için iki koşul vardır. İlk koşul  $Z$  matrisinin Frobenius normunun  $X$  matrisinin Frobenius normuna bölümünün Eş. 2.38'deki gibi sifıra yakınsamasıdır (bu çalışmada varsayılan değer  $10^{-7}$  alınmıştır). Diğer koşul ise maksimum iterasyon sayısına ulaşılmasıdır (varsayılan 1000 iterasyon).

$$\|Z\|_F / \|X\|_F < 10^{-7} \quad (2.38)$$

- 17- İterasyonlar boyunca sürekli güncellenen  $L$  matrisinin rankı bir süre sonra sabit kalacaktır.

### RPCA Kullanılarak Düşük Rank Ve Seyrek Matrislerin Oluşturulması

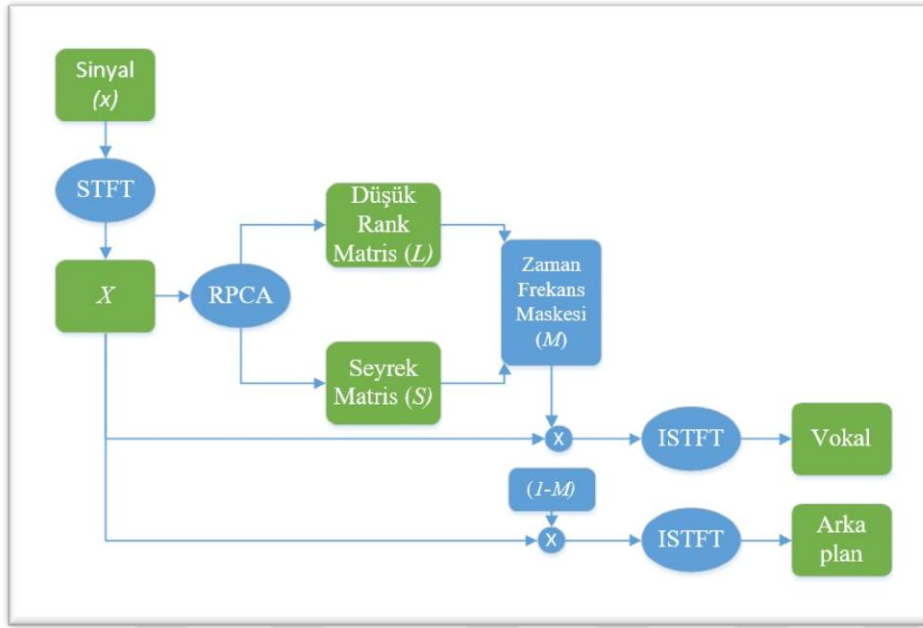
Bu algoritma tekrarlamamanın müzikte temel ilke olduğu fikrine dayalı olarak geliştirilmiştir. Bu nedenle, Şekil 2.7’de görüldüğü gibi müzik eşliklerinin düşük ranklı ve daha az değişen, vokalin ise daha fazla değişen ama ses karışımında seyrek olduğu varsayımından yola çıkılmıştır. Bu varsayımlara dayanarak, düşük ranklı ve seyrek matrisleri çözmek için bir matris çarpanlara ayırma algoritması olan RPCA kullanılmaktadır [9].



Şekil 2.7. RPCA kullanılarak düşük rank ve seyrek matrislerin oluşturulması

Müzik enstrümanlarının çıkardığı ses, genel olarak müziğin altta yatan tekrarlı yapısını oluşturduğu için, arka plan müziğini düşük ranklı bir sinyal (low-rank) olarak kabul ederiz. Vokal sesi ise aksine parça içinde daha çok değişime uğrar ve frekans alanlarında daha seyreklerdir. Bu gözlemden hareketle RPCA kullanımı ile Şekil 2.8’de olduğu gibi  $L$  düşük rank matrisini,  $S$  arka plan müzik eşleniğini ve  $M$  ses karışımını ifade ederek, problemin Eş. 2.39’deki denklem üzerinden çözümünü gerçekleştirir.

$$\text{minimize } \|L\|_* + \lambda \|S\|_1 \text{ ve } L + S = M \quad (2.39)$$



Şekil 2.8. RPCA kullanımı ardından zaman frekans maskesi kullanarak vokal ve arka plan sesinin ayrılması

RPCA kullanılarak arka plan müziği ve vokal sesinin ayrılması işlemi kısaca aşağıdaki adımların uygulanması ile gerçekleştirilir (Bkz. Şekil 2.8).

- 1- İlk olarak ses karışımından STFT hesaplanarak  $M \in R^{n_1 \times n_2}$  olacak şekilde  $M$  matrisi oluşturulur.
- 2- Bu işlemin sonrasında, RPCA problemini çözmek için etkili bir algoritma olan ALM yöntemi kullanılır. Bu işlemin sonucu olarak  $L \in R^{n_1 \times n_2}$  ve  $S \in R^{n_1 \times n_2}$  olacak şekilde iki çıktı elde edilir, seyrek yapıların olduğunu ve vokal etkinliği belirten  $S$  matrisi ve düşük ranklı müzik eşleniğinin yer aldığı  $L$  matrisi elde edilir.
- 3- Elde edilen  $L$  ve  $S$  matrisinin elemanları karşılaştırılarak Eş. 2.40'da gösterildiği gibi  $M_b$  olarak adlandırılan zaman frekans maskesi oluşturulur.

$$M_b(m, n) = \begin{cases} 1 & , |S(m, n)| > |L(m, n)| \\ 0 & , \text{diğer} \end{cases}$$

$$m = 1 \dots n_1 \text{ ve } n = 1 \dots n_2 \quad (2.40)$$

- 4- Eş.2.41'de gösterildiği gibi  $M_b$  maskesinin  $M$  sinyaline uygulanması sonrasında vokal sesini ifade eden  $X_{vokal}$ ,  $(1-M_b)$  maskesinin  $M$  sinyaline uygulanması ile arka plan müziğini ifade eden  $X_{vokal}$  yapıları ortaya çıkar.

$$\begin{cases} X_{vokal} \\ X_{arkaplan} \end{cases} = \begin{cases} M_b(m, n)M(m, n) \\ (1 - M_b(m, n))M(m, n) \end{cases} \\ m = 1 \dots n_1 \text{ ve } n = 1 \dots n_2 \quad (2.41)$$

- 5- Elde edilen  $X_{vokal}$  ve  $X_{arkaplan}$  matrislerin ters STFT alınarak müzik ve arka plan değerlerine ulaşılır.

### 2.3. İki Modelin Çıktılarının İncelenmesi

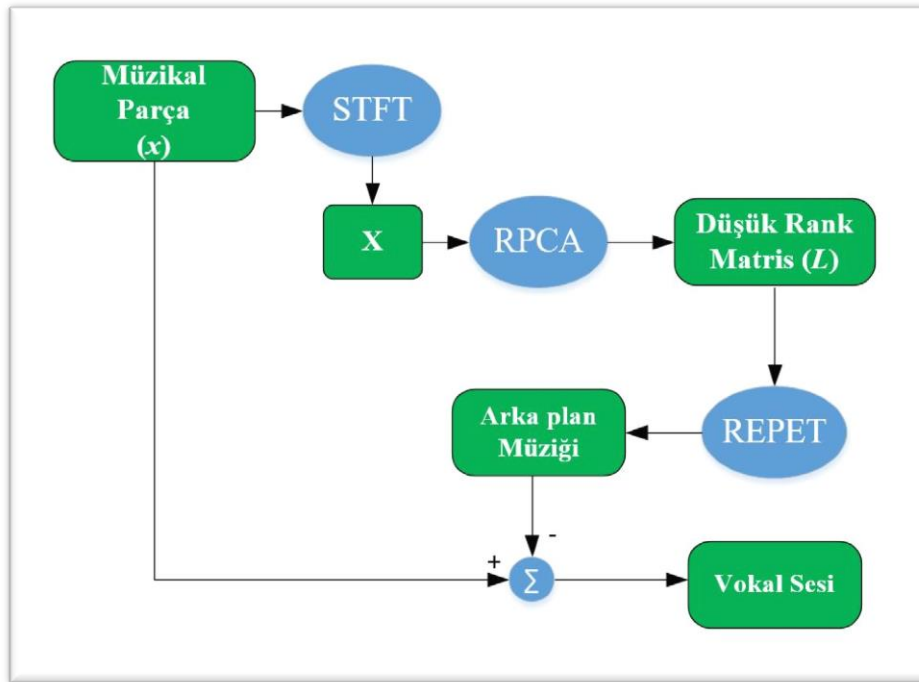
Ses karışımından vokal ve arka plan müziğinin ayrılmasında kullanılan bu iki algoritmanın sonuçları incelendiğinde;

- 1- REPET algoritmasının vokal olarak ayırdığı kısımda RPCA algoritmasına göre daha fazla arka plan müziğinin yer aldığı,
- 2- RPCA algoritmasının vokal olarak ayırdığı kısım dinlendiğinde REPET'e göre daha fazla yapay ses içerdiği,
- 3- RPCA algoritmasının davul sesini düşük-rank olarak görmediği, bu yüzden vokal sesin içerisine karıştığı, REPET'in ise beklendiği gibi davul sesinin tekrarlı yapısından dolayı vokal sesi içerisinde yer vermediği,
- 4- REPET ve RPCA algoritmasının spektogramları incelendiğinde RPCA algoritmasının düşük frekanslarda vokal arka plan ayrımını daha iyi yaptığı ve REPET algoritmasının da yüksek frekanslarda ayrımı daha iyi yaptığı,
- 5- Her iki algoritmanın da ön işlem gerekmeksizin çalıştığı, gözlemlenmiştir.

Yukarıdaki gözlemlerden hareketle, bu çalışmadaki ana fikir, iki algoritmanın kombinasyonunu kullanarak hem düşük hem de yüksek frekanslarda vokal ve arka plan müziğinin daha iyi ayrılmasını sağlamak için kullanılabileceğidir.

### 3. TEZDE ÖNERİLEN MODEL

Bu tezde sunulan model, daha önce bahsedilen RPCA ve REPET algoritmalarının birleştirilmesiyle geliştirilen hibrit bir modeldir [5,9,31]. İşlem kısaca giriş sinyalinin RPCA algoritması ile işlenerek düşük rankının bulunması ve seyrek kısımlarının çıkarılması ile başlamaktadır. Daha sonra RPCA kullanılarak ön işlem geçirilmiş ve seyrek kısımları çıkarılmış sinyal REPET yöntemiyle işlenerek arka plan müziği elde edilir.



Şekil 3.1. Önerilen metodun akış diyagramı

Sistemin işleyişi Şekil 3.1’de gösterilmek ile birlikte aşağıda ayrıntılı olarak listelenmiştir.

- 1- Orijinal karışım sinyalinin STFT dönüşümü, yarı örtüşen pencere kullanılarak elde edilir. Pencereleme fonksiyonu olarak Hamming pencereleme fonksiyonu kullanılmış olup bunun nedeni ayrıntılı olarak Bölüm 4.3.2’de anlatılmıştır. Eş 3.1’de gösterilen STFT alınmış  $x$  sinyali,  $X$  ile gösterilir:

$$X = \text{STFT}(x) \quad (3.1)$$

- 2- STFT alınan  $X$  matrisi RPCA kullanılarak seyrek ve düşük rank olarak adlandırılan Eş. 3.2’de ifade ile gösterilen iki ayrı matrise ayrılır. Ortaya çıkan matrislerin boyutu  $X$

matrisinin boyutu ile aynı olmalıdır. ( $X \in R^{n1x n2}$ ,  $L \in R^{n1x n2}$ ,  $S \in R^{n1x n2}$ ). RPCA ile Eş. 3.2'de yer alan problemleri çözmek için ALM metodu kullanılmaktadır. Yapılan işlemler sonrasında  $L$  ve  $S$  matrisleri elde edilir.  $L$  matrisi düşük ranklı matris arka plan müziğini ve  $S$  matrisi seyrek matris vokal sesini ifade eder.

$$\text{minimize } \|L\|_* + \lambda \|S\|_1 \quad \text{ve} \quad L+S = |X| \quad (3.2)$$

- 3-  $\|\cdot\|_*$  ibaresi tekil değerlerin toplamını,  $\|\cdot\|_1$  ibaresi ise matris elemanlarının mutlak değerlerinin toplamını ifade etmektedir. Denklemden yer alan  $\lambda$  ise  $L$  ve  $S$  matrisleri arasındaki değişim parametresi olarak kullanılmaktadır.  $\lambda$ , değişim parametresi Eş. 3.3'te belirtildiği gibi hesaplanmak ile birlikte sıfırdan büyük olmalıdır.

$$\lambda_k = k / \sqrt{\max(m, n)} \quad (3.3)$$

- 4- Elde edilen  $L$  matrisinin, ayna frekansı bileşenlerini çıkarılır ve genlik spektrumu alınarak  $V$  ile gösterilir.
- 5- Genlik spektrogramı  $V$  matrisinin karesinin her satırının özilintisi hesaplanarak Eş. 3.4'te ifade edilen,  $B$  özilinti matrisi elde edilir.  $V$ 'nin karesini kullanılmasının nedeni tepe noktalarını vurgulamak ve tahmin periyodunu daha iyi kestirmektir.

$$B(i, j) = \frac{1}{m-j+1} \sum_{k=1}^{m-j+1} V^2(i, k) V^2(i, k+j-1) \quad (3.4)$$

- 6-  $b$  olarak adlandırılan vuruş spektrum,  $B$  matrisinin satırlarının ortalaması alınarak Eş. 3.5'deki gibi hesaplanır daha sonra ilk oluşan  $b(0)$  ile normalize edilir.

$$b(j) = \frac{1}{n} \sum_{i=1}^n B(i, j); \quad b(j) = \frac{b(j)}{b(0)}$$

$$i = 1 \dots n(\text{frekans}), \quad n = \frac{N}{2} + 1, \quad i = 1 \dots m(\text{zaman}) \quad (3.5)$$

- 7- Ses karışımı  $x$ 'de tekrarlayan kalıplar varsa, vuruş spektrumunda ( $b$ ) periyodik olarak farklı seviyelerde tekrarlanan tepeler oluşturacaktır. Vuruş Spektrumunun periyodu, kendi harmonikleri üzerinde en büyük ortalama birikmiş enerjiye sahip olduğu periyot tahmin edilerek bulunur ve  $p$  ile adlandırılır.

- 8- Matrislerin tekrarlama periyodu  $p$  bulunduktan sonra, genlik spektrogramı  $V$ , hesaplanan  $p$  periyodunun boyuna göre  $r$  adet eşit parçaya Eş. 3.6'da gösterildiği gibi bölünür. Ayrılan parçaların medyanı alınarak oluşturulan matrise  $S_m$ , tekrarlanan segmentasyon modeli adı verilir.

$$S_m(i,l) = \underset{k=1..r}{\text{medyan}} \{V(i,l+(k-1)p)\} \quad (3.6)$$

- 9-  $V$  tekrarlanan segmentasyon modeli  $S_m$  hesaplandıktan sonra,  $S_m$  ile daha önce ayrılan  $r$  adet segmentin elemanları karşılaştırılarak bunların minimum olanları bir matrise aktarılır ve  $W$  olarak adlandırılan Eş. 3.7'de belirtilen tekrarlanan spektrogram modeli oluşturulur.

$$W(i,l+(k-1)p) = \min\{S_m(i,l), V(i,l+(k-1)p)\} \\ i = 1 \dots n, (\text{frekans}) \text{ ve } j = 1 \dots m(\text{zaman}) \quad (3.7)$$

- 10-  $M$  olarak adlandırılan zaman-frekans maskesi,  $W$ 'nin elemanlarının  $V$  ile normalize edilmesi ile Eş. 3.8'de olduğu gibi oluşturulur.

$$M(i,j) = \frac{W(i,j)}{V(i,j)}, \quad M(i,j) \in [0,1] \\ i = 1 \dots n(\text{frekans}) ; j = 1 \dots m(\text{zaman}) \quad (3.8)$$

- 11-  $M$  maskesi düşük rank  $L$  matrisine uygulanır ve ters STFT alınarak zaman eksenine geçilir ve arka plan eşleniğine ulaşılır.
- 12- Elde edilen arka plan eşleniği sinyali başlangıçta kullandığımız  $x$ , ses karışımından çıkartılarak vokal sesine ulaşılır.

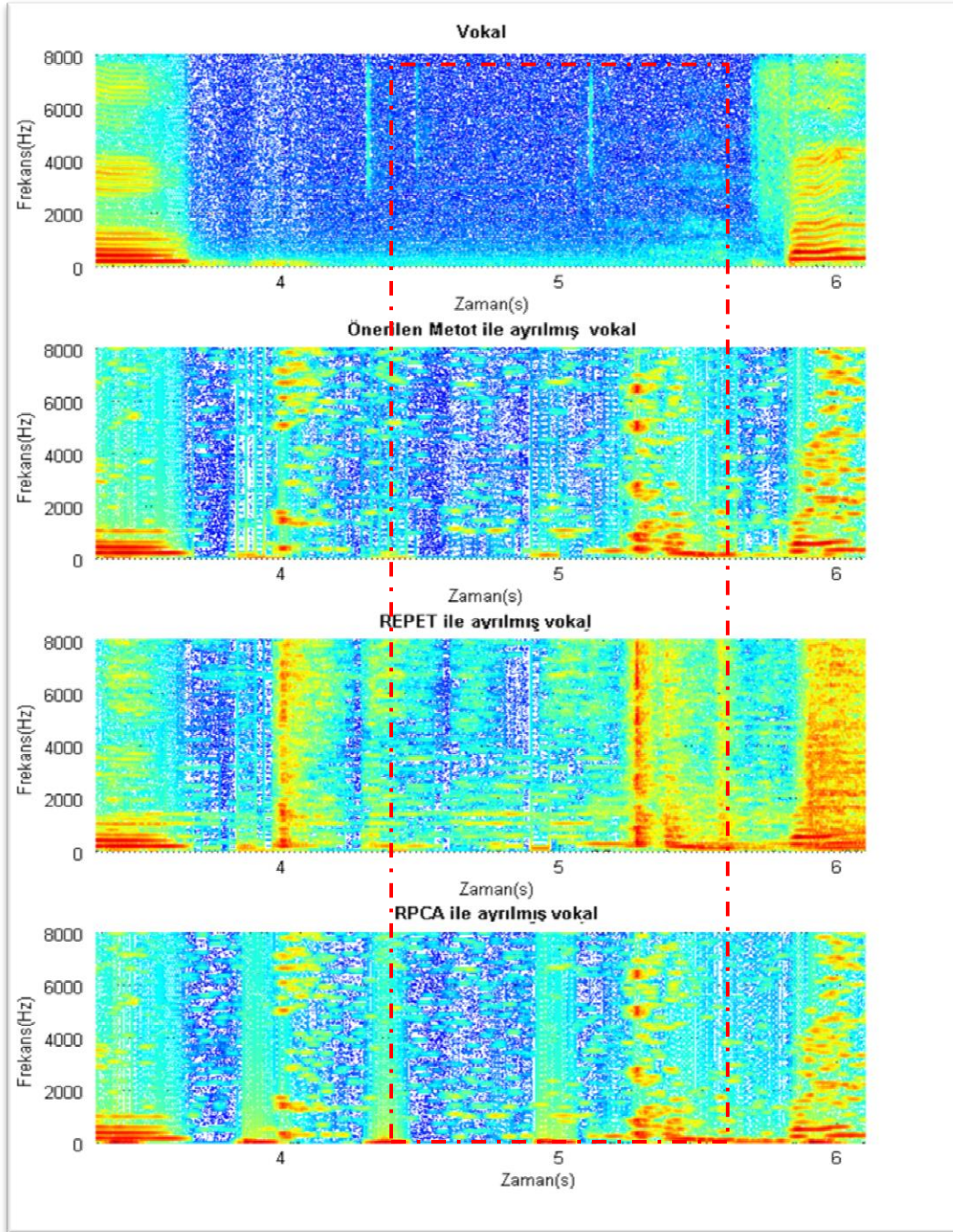
Tez çalışmasında, Bölüm 4'de performans ölçümü için kullanılan beş ayrı müzikal parçanın arka plan müziği ve vokal sesi REPET, RPCA ve önerilen metot ile ayrılmış ve çıktılarının spektrogramları çıkarılıp EK-1'de sunulmuştur. Söz konusu spektrogramlar incelendiğinde:

- REPET yöntemi ile ayrılan parçaların vokal sesi çıktılarında diğer yöntemlerden daha fazla arka plan müziğinin yer aldığı,

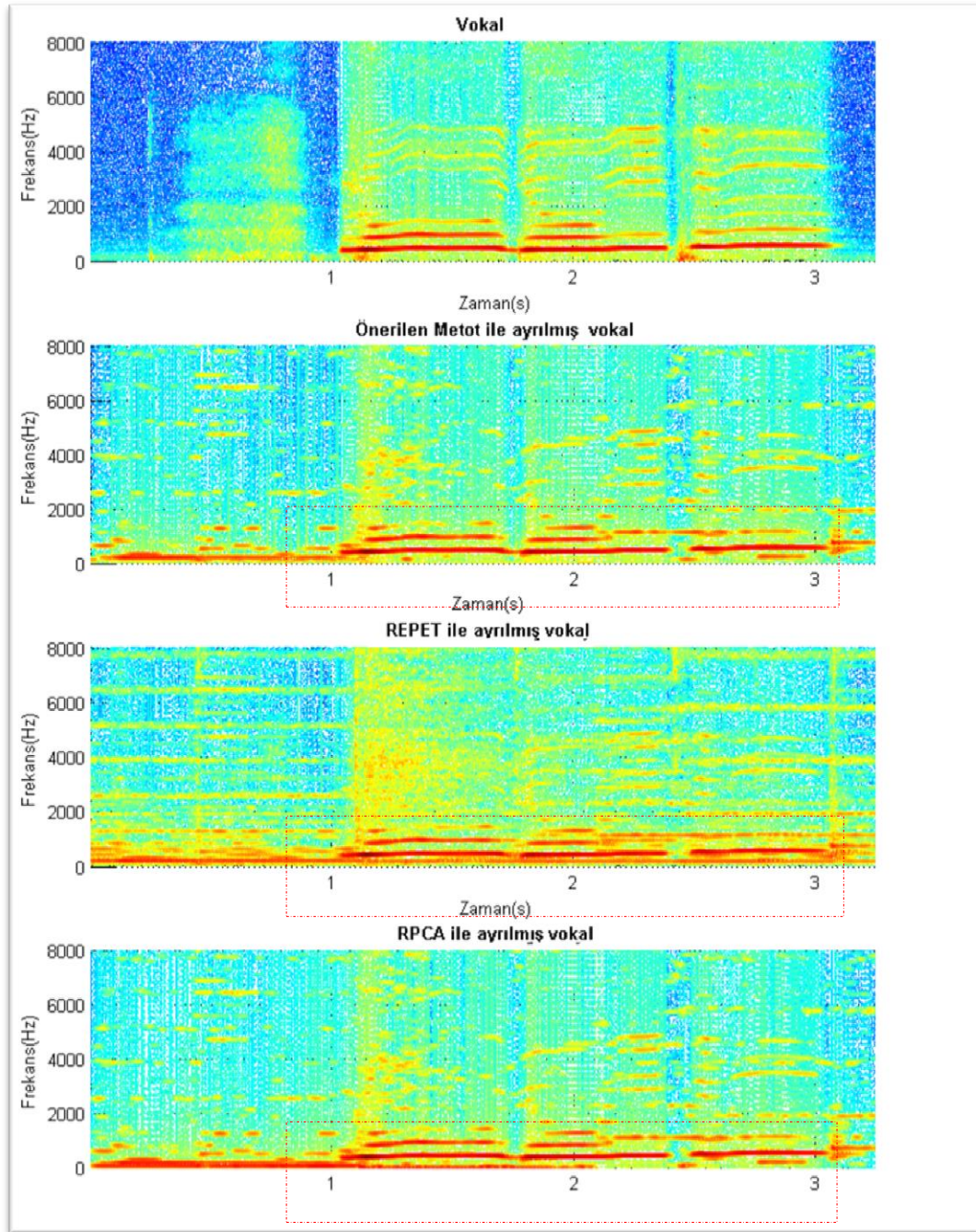
- REPET ve RPCA yöntemi ile ayrılan parçaların vokal sesi çıktıları incelendiğinde; düşük ve yüksek frekanslarda ve parçanın vokal sesi olmayan bölümlerinde Şekil 3.3'te olduğu gibi düzgün ayırım yapılamadığı,
- Parçada vokal sesi olmayan bölümlerinde önerilen metotta RPCA algoritmasının ön işlem olarak kullanılması sayesinde vokal sesi spektogramının Şekil 3.2'de olduğu gibi daha seyrek olarak görüldüğü, ayrılması zor hem müzikal enstrümanların hem de insan sesinin yer aldığı düşük frekans bandında diğer iki metoda göre daha iyi ayırım yaptığı ve önerilen metodun vokal sesi spektogramlarının orijinal vokal sesi spektogramlarına daha çok benzediği,

gözlemlenmiştir.





Şekil 3.2. Vokal sesinin olmadığı seyrek kısımların ayrılması sonuçları



Şekil 3.3. Düşük frekansların ayrılması sonuçları

## 4. ÖNERİLEN YÖNTEMİN BAŞARIMLARININ DEĞERLENDİRİLMESİ

Geliştirilen yöntemin başarımlarını doğrulamak için iki ayrı doğrulama yöntemi kullanılmıştır. Bunlardan ilki, Bss\_Eval aracı kullanılarak sistemin başarımlarının matematiksel olarak ölçümüdür [34]. Diğer yöntem ise öznel bir başarımlar ölçümü sağlayan ortalama görüş puanı (Mean Opinion Score-MOS) testidir. Testlerde, MIR-1K veri seti kullanılmıştır [35]. Testlerde geliştirilen hibrit yöntem, RPCA ve REPET algoritmaları karşılaştırarak başarımları karşılaştırılmıştır

### 4.1. MIR-1K Veri Seti

Tayvan Ulusal Tsing Hua Üniversitesi bünyesinde 1999 yılında kurulan Multimedya Bilgi Erişim Laboratuvarı (Multimedia Information Retrieval-MIR), müzik, ses ve görüntü işleme ve sentezleme amacıyla kurulmuştur [36]. MIR-1K veri seti, bahse konu laboratuvar tarafından müzikal parçadan vokal sesi ve arka plan müziği ayrıştırması vb. konularda çalışan araştırmacılar tarafından kullanılması amacıyla tasarlanmıştır [35]. MIR-1K veri setinde:

- 1000 adet vokal sesi ve arka plan müziği sırasıyla sağ ve sol kanallara kaydedilmiş müzikal parça yer almaktadır. Parçalar sağ ve sol kanalların toplanması ile mix hale getirilmiştir.
- Şarkıları söyleyen kişi tarafından şarkının sözlerinin ezgisiz kaydedilmiş hali ayrıca yer almaktadır.
- Veri setinde yer alan şarkılar, 5000 adet Çin pop şarkısından herhangi bir nedenden bağımsız olarak rastgele seçilmiş olup, sekiz kadın ve on bir erkekten oluşan laboratuvar çalışanları tarafından seslendirilmiştir. Şarkıları seslendiren kişilerin çoğu amatör olup profesyonel müzik eğitimi almamıştır.
- Veri setinde yer alan parçalarının süresi dört ile on üç saniye arasında değişmektedir.
- Veri setindeki parçaların örnekleme frekansı 16 kHz'dir.

Müzikal parçadan vokal sesi ve arka plan müziği ayrıştırması vb. konularda geçmişte yapılan çalışmalarda sıklıkla MIR-1k veri setinin kullanılması dolayısıyla, bu tez çalışmasında geçmişte yapılan çalışmalar ile önerilen metodun karşılaştırılması amacıyla

MIR-1K veri setinden faydalanılmıştır. Anılan veri seti içerisinde yer alan 1000 adet müzikal parça arasından sekiz saniyeden uzun süreli 438 adet müzikal parça, BSS\_ aracı kullanılarak, geliştirilen sistemin ve diğer yöntemlerin başarımları ölçülmüştür. Ayrıca veri setinden seçilen 438 adet parça içerisinde rastgele seçilen beş adet parça ile de algısal başarımlar ölçümü (MOS test) gerçekleştirilmiştir.

## 4.2. Ortalama Görüş Puanı (Mean Opinion Score)

Sistemin sonuçlarını değerlendirmek ve doğrulamak için kullanılan yöntemlerden birisi uzun yıllardır ses ve video haberleşmesinin kalitesini ölçümünde kullanılan Ortalama Görüş Puanı (Mean Opinion Score) olarak adlandırılan subjektif bir yöntemdir.

Bu yöntemde katılımcılardan sessiz bir odada özel kulaklıklar kullanarak duydukları sesin kalitesini değerlendirmeleri istenir. Katılımcıların testte kullanılan parçaları istedikleri kadar tekrar dinleme izinleri vardır. Katılımcılardan her sese Çizelge-4.1’de yer alan derecelendirme şeması kullanılarak birden beşe kadar puan vermeleri istenmiştir [37].

Çizelge 4.1. Ortalama görüş puanı değerlendirme tablosu

Ortalama görüş puanı (MOS)		
MOS	Kalite	Değer düşüklüğü
5	Mükemmel	Tam olarak algılanabilir
4	İyi	Algılanabilir ama can sıkıcı değil
3	Makul	Biraz can sıkıcı
2	Zavallı	Can sıkıcı
1	Kötü	Çok sinir bozucu

### 4.2.1 Ortalama görüş puanı (MOS) kullanılarak sonuç değerlendirmesi

Geliştirilen yöntemin MOS testi ile değerlendirilmesi, MIR-1K veri setinden beş ayrı müzikal parça ile on yedi farklı dinleyici ile gerçekleştirilmiştir. Katılımcılardan anılan müzikal parçaların REPET, RPCA ve geliştirdiğimiz model kullanılarak elde edilen sonuçlara puan vermeleri istenmiştir. Puanlama işleminden sonra her parçaya farklı yöntemler için verilen puanların aritmetik ortalaması alınmıştır.

Eş. 4.1’de gösterildiği gibi;  $N$  katılımcıların sayısını,  $Rn$  parçalara verilen puanı temsil etmektedir. MOS testi sonrasında elde edilen sonuçlar Çizelge-4.2’de gösterilmiştir.

$$MOS = \frac{\sum_{n=0}^N Rn}{N} \quad (4.1)$$

MOS sonuçları incelendiğinde sunulan modelin, RPCA ve REPET’e göre hem vokal sesi hem de arka plan müziği sonuçlarının her parça için diğerlerinden daha fazla puan aldığı görülmektedir. Testte her bir parça için verilen puanlara göre önerilen metot puanlarının birçok parça için üçün üstünde veya üçe yakın olduğu görülmüştür. Diğer yöntemlerden sadece RPCA metodu ile ayrılmış üç adet arka plan müziğinin üç civarında puan aldığı, REPET metodu ile ayrılan parçaların puanlarının ise üçün üzerine çıkamadığı gözlemlenmiştir. Testte kullanılan parçalar üzerinden inceleme yapıldığında vokal sesi ayırımında en çok puan alan parçanın Bobon\_5\_06 olduğu, önerilen metot kullanılarak vokal sesi ayrıştırılması yapıldığı ve 3,24 puan aldığı ayrıca arka plan müziği ayırımında en çok puan alan parçanın Titon\_2\_09 olduğu, önerilen metot kullanılarak arka plan müziği ayrıştırılması yapıldığı ve 3,59 puan aldığı görülmektedir. Teste katılan katılımcıların her bir parça için verdiği puanlar EK-2’de yer almaktadır.

Çizelge 4.2. MOS testi ile elde edilen sonuçlar

Parça Adı	REPET		RPCA		Önerilen Yöntem	
	Vokal	Müzik	Vokal	Müzik	Vokal	Müzik
Ani_1_07	2,41	2,53	2,12	3,06	2,88	3,24
Kenshin_2_05	2,47	2,59	2,24	3,00	2,59	3,06
Leon_8_10	2,47	2,47	2,18	2,76	2,94	3
Titon_2_09	2,53	2,76	2,88	3,06	3	3,59
Bobon_5_06	2,41	2	2,12	2,35	3,24	2,53
<i>ORTALAMA</i>	2.45	2.47	2.3	2.84	2.93	3.08

### 4.3. Bss Eval Aracı

BSS Eval aracı orijinal kaynak sinyali ile kaynak ayırıştırma algoritmalarının başarımlarını ölçmek üzere geliştirilen bir Matlab aracıdır [34]. Ölçümlerde, kaynak sinyal ile ayırıştırma yapıldıktan sonra elde edilen sinyaller karşılaştırılarak; sinyalde yer alan bozulma, parazit ve yapaylığı matematiksel olarak değerlendirmektedir. Bahse konu algoritmayı kullanmak için orijinal sinyale ihtiyaç duyulmaktadır (vokal ve arka plan müziğinin ayrı ayrı kaydedilmiş hali). Bu değerlendirme yöntemi birçok veri türü ile (ses, biyomedikal vs.) ve algoritmada (beamforming, ICA vb) kullanılmakla birlikte tezde de bahsedilen REPET, RPCA ve önerilen metodun sonuçlarının değerlendirilmesinde de kullanılmıştır.

Sinyal ayırma başarımı tahmin edilen sinyal  $\hat{s}_j$  ile orijinal sinyalin karşılaştırılmasıyla  $s_j$  vasıtasıyla ölçülür. İlk olarak tahmin edilen sinyal; gerçek sinyal, sinyaldeki bozulma, gürültü ve yapaylık hata terimleri olarak ayrıştırılır [34].

$$\hat{s}(t) = s_{target}(t) + e_{interf}(t) + e_{noise}(t) + e_{artif}(t) \quad (4.2)$$

$s_{target}$  izin verilen ölçüde gürültü içeren hedef sinyal;  $e_{interf}$ ,  $e_{noise}$  ve  $e_{artif}$  sırasıyla kaynak sinyaldeki karışma, gürültü, ayrılmadan sonra oluşan yapaylık olarak ifade edilmektedir. Bu dört eleman toplandığında tahmin edilen sinyali oluşturmaktadır.

İkinci aşamada, bu dört terimden her birinin sinyal süresince veya yerel çerçevelerde görelî miktarını değerlendirmek için enerji oranlarını hesaplanır.  $\hat{s}_j$  'nin parçalanmasından yola çıkarak, sayısal performans kriterlerini, desibel(dB) cinsinden ifade eden enerji oranları hesaplanırken Eş. 4.3-5'de yer alan formülleri kullanılır.

- Kaynak Bozulma Oranı (Source Distortion Ratio),

$$SDR = 10 \log_{10} \left( \frac{\|s_{target}\|^2}{\|e_{interf} + e_{artif}\|^2} \right) \quad (4.3)$$

- Kaynak Girişim Oranı (Source to Interferences Ratio),

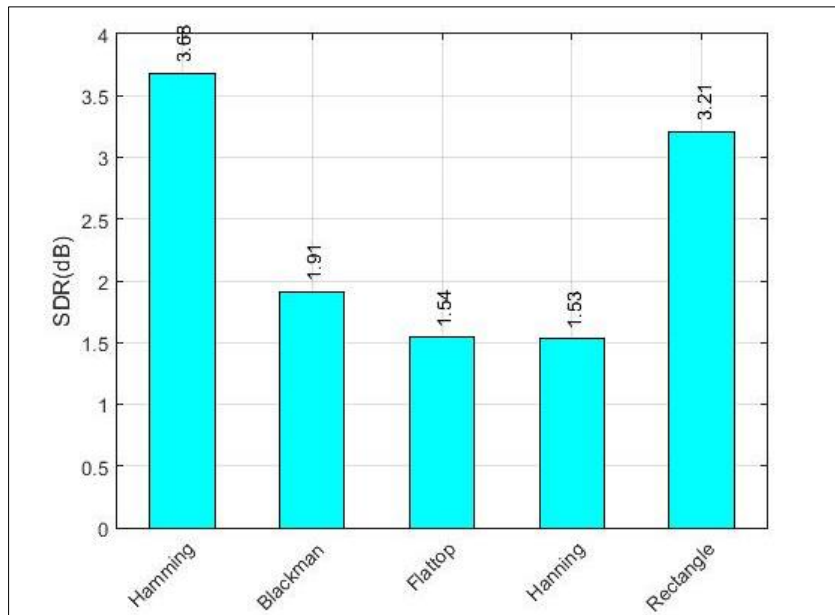
$$SIR=10 \log_{10} \left( \frac{\|s_{target}\|^2}{\|e_{interf}\|^2} \right) \quad (4.4)$$

- Kaynak Yapaylık Oranı (Sources to Artifacts Ratio),

$$SAR=10 \log_{10} \left( \frac{\|s_{target}+e_{artif}\|^2}{\|e_{artif}\|^2} \right) \quad (4.5)$$

#### 4.3.1. Bss Eval aracı kullanılarak pencereleme fonksiyonu seçimi

Fourier Dönüşümü kullanılarak bir sinyalin bütün frekans bileşenlerini çıkarılabilir. Ancak bu dönüşüm kullanılarak sinyal içerisinde yer alan frekans bileşenlerinin zaman içerisindeki konum bilgileri belirlenemez. Bu sebep ile STFT vasıtasıyla belirli zaman aralıklarında, pencereleme fonksiyonu kullanılarak frekans bileşenlerinin zaman-konum bilgisi de elde edilebilmektedir. Bu nedenle STFT kullanımı ile sinyalden hem frekans bilgisi hem de frekansların zaman bilgisi çıkarılabilir [38]. Sinyal işlemede, bir pencere fonksiyonu, seçilen aralıklar dışında sıfır değerli olan matematiksel bir fonksiyondur. STFT kullanırken seçilen pencereleme fonksiyonun şekli, darlığı veya genişliği ortaya çıkacak sonucun doğruluğu hususunda büyük önem taşımaktadır.

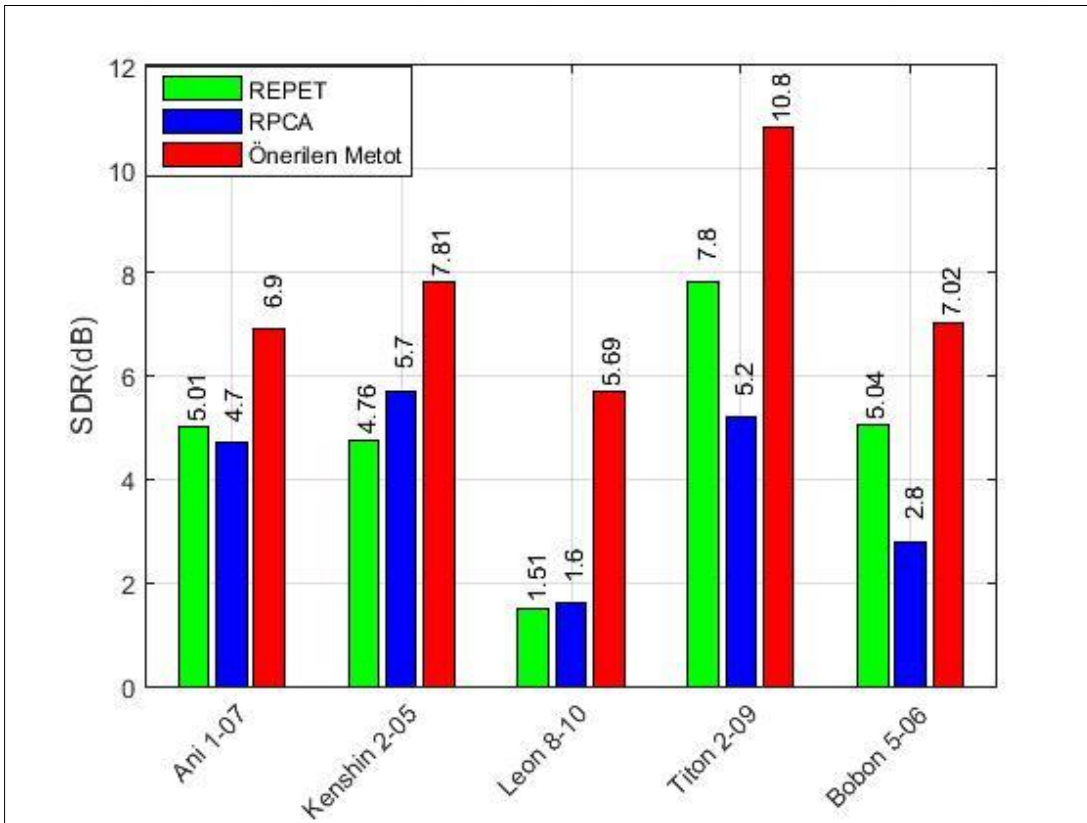


Şekil 4.1. Pencere fonksiyonlarına göre SDR değerleri

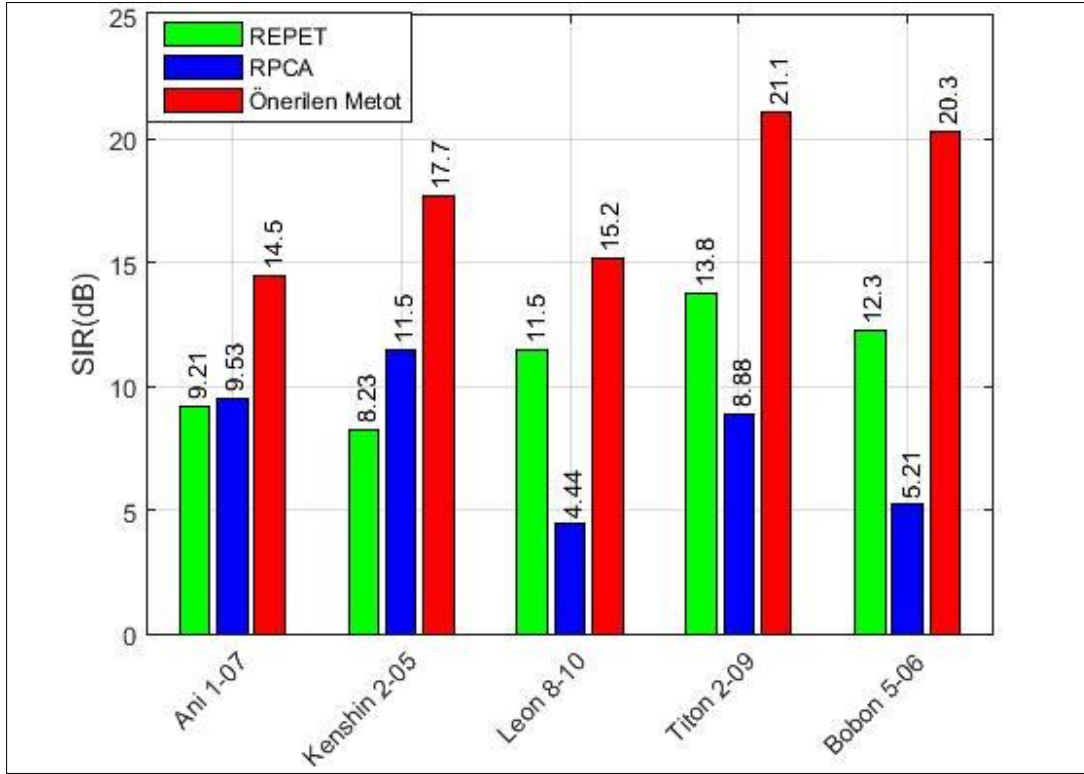
Geliştirilen metotta STFT yaparken kullanılacak pencerenin seçimi için muhtelif pencereler ile testler gerçekleştirilmiş olup, BSS eval aracı kullanılarak elde edilen sonuçların kaynak bozulma oranları incelenmiştir (SDR). Bu kapsamda Blackman, Flattop, Rectangle, Hanning ve Hamming pencereleri kullanılmıştır. Testte Tamy Macedo adlı şarkıcının tanto faz adlı şarkısı kullanılmıştır. Şekil 4.1 de görüldüğü gibi yapılan test sonucunda en iyi sonucu Hamming pencerenin verdiği görülmüş olup bu nedenle bu pencereleme fonksiyonu geliştirilen metotta kullanılmıştır.

#### 4.3.2. Bss Eval aracı kullanılarak başarımların değerlendirilmesi

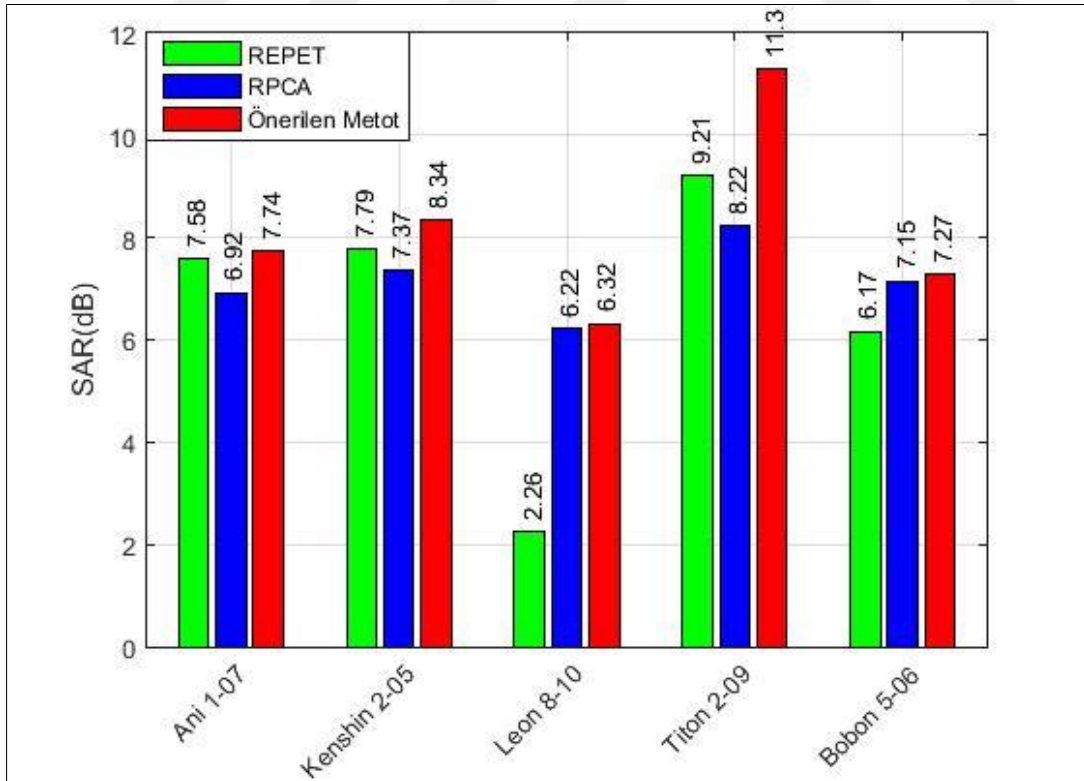
Geliştirilen yöntemin BSS\_Eval ile değerlendirilmesinde MIR-1K veri setinden beş ayrı parça ile gerçekleştirilmiştir. Her parça için ayrı ayrı SDR, SIR ve SAR değerleri hesaplanmıştır. Başarımların değerlendirilmesinin yapılabilmesi için müzikal parçanın arka plan eşlenişinin ve vokal kısmının ayrı ayrı kaydedilmiş örneklerinin olması gerekmektedir. Hesaplanan SDR, SIR ve SAR başarımlarının yüksek olması algoritmanın daha iyi ayrım yaptığını göstermektedir.



Şekil 4.2. REPET, RPCA ve önerilen metot kullanılarak oluşan sonuçların SDR değerleri



Şekil 4.3. REPET, RPCA ve önerilen metot kullanılarak oluşan sonuçların SIR değerleri



Şekil 4.4. REPET, RPCA ve önerilen metot kullanılarak oluşan sonuçların SAR değerleri

Elde edilen sonuçlardan SDR değerleri Şekil-4.2’de, SIR değerleri Şekil-4.3’te ve SAR değerleri Şekil-4.4’te gösterilmektedir. Test sonuçları incelendiğinde önerilen metodun, SDR ve SIR değerlerinin testte kullanılan bütün müzikal parçalar için RPCA ve REPET metotlarına göre oldukça yüksek değere sahip görülmektedir. Bununla birlikte SAR değerleri yine diğer yöntemlerden yüksek olmasına rağmen SDR ve SIR sonuçlarında daha az artış sağlamıştır. SAR başarımının SDR ve SIR sonuçlarına göre daha az artma sebebi RPCA metodunun sonuçlarının fazla yapaylık oluşturmasıdır. Bu durum RPCA’nın önerilen metotta da kullanılması sebebiyle SAR sonuçlarına tesir etmektedir. Testte kullanılan müzikal parçalar üzerinden inceleme yapıldığında en yüksek değere sahip parçanın titon\_2\_09, en düşük alan parçanın ise leon\_8\_10 olduğu görülmektedir. leon\_8\_10 parçasının diğer parçalara göre az puan almasının nedeninin, parçada yoğun miktarda bas ses yer alması olduğu, bu durumun tersi olarak titon\_2\_09 parçasında bas sesi yer alamaması nedeniyle yüksek puan aldığı düşünülmektedir. BSS aracı ile yapılan başarım ölçüm sonuçlarının ile MOS testi karşılaştırıldığında iki başarım ölçüm metodunda da en yüksek puanı alan parçanın titon\_2\_09 olduğu ve iki testte de benzer parçaların daha puan almaları nedeniyle iki testin sonucunun da paralellik gösterdiği söylenebilmektedir.

## 5. SONUÇ VE ÖNERİLER

Bu tezde, şarkıları vokal ve arka plan müziği olarak otomatik olarak ayrılması için geliştirilen yeni bir yöntem anlatılmaktadır. Önerilen yöntemde daha önce bahse konu uygulamalarda kullanılan REPET ve RPCA yöntemlerini birleştirilerek hibrit bir yöntem oluşturulmuştur. Her iki yöntemde de, vokal sesi ve arka plan müziğinin ayrılması, müzik parçasının tekrarlama prensibine dayanmaktadır. Önerilen metot ile bu iki metodun avantajlarından yararlanarak daha yüksek bir ayırma başarımı gerçekleştirilmesi amaçlanmıştır.

Önerilen yöntemde REPET' metodunun vokal sesi ayırım sonuçlarındaki arka plan müziğini azaltmayı amaçlanmış olup, bu durumun önüne geçilebilmesi için RPCA algoritması ön işlem parçası olarak kullanılmıştır.

Önerilen yöntemin başarımını test etmek için, iki ayrı test gerçekleştirilmiş olup bunlardan ilki MATLAB ortamı için geliştirilen BSS\_Eval\_Tool'u kullanılmıştır. Anılan test ile önerilen yöntem ve REPET metodunu nesnel olarak ölçümü gerçekleştirilmiştir. İkinci test, ise sonuçların öznel olarak değerlendirilebileceği MOS testidir. Söz konusu test de onyediyedi adet katılımcıya beş adet şarkının REPET ve önerilen metodun sonuçları dinletilerek puan vermeleri istenmiştir. Her iki teste de MIR-1K veri setinden parçalar kullanılmıştır.

Gerçekleştirilen testler incelendiğinde her iki teste göre önerilen yöntemin REPET ve RPCA metotlarına göre daha iyi vokal sesi ve arka plan müziği ayırımı yaptığı görülmektedir. Tezde geliştirilen metottun müzikal bir parçanın tekrarlayan arka plan müziği ve tekrarlanmayan vokal kısmını ayırdığından bahsedilmesine rağmen bu metot genelleştirilerek özellikle periyodik gürültü ve girişimleri gidermek amaçlı (güç hattı paraziti) ve konuşma sinyali (radyo iletişimi) uygulamalarında da kullanılabilir. İleriki çalışma olarak yöntemin başarımının hangi durumlarda arttığı araştırılması ve buna paralel olarak yöntemin geliştirilmesi planlanmaktadır.



## KAYNAKLAR

1. Schenker, H. (1954). *Harmony*. Chicago, Illinois: University of Chicago Press.
2. Ruwet, N., Everist, M. (1987). Methods of analysis in musicology. *Music Analysis*, 6(1), 3–911.
3. Ockelford, A. (2005). *Repetition in music: Theoretical and metatheoretical perspectives*. U.K., Farnham: Royal Musical Association Monographs. Ashgate, 13
4. Mohammadiha, N., Smaragdis, P., and Leijon, A. (2013). Supervised and unsupervised speech enhancement using nonnegative matrix factorization. *IEEE Translational on Audio, Speech, and Language Processing*, 21(10), 2140-2151.
5. Rafii, Z., and Pardo, B. (2013). Repeating pattern extraction technique (REPET): a simple method for music/voice separation. *IEEE Translational on Audio, Speech, and Language Processing*, 21(1), 73-84.
6. Kim, S. M., Park, J. H., Kim, H. K., Lee, S. J., and Lee, Y. K. (2012). *Non-negative matrix factorization based noise reduction for noise robust automatic speech recognition. in latent variable analysis and signal separation*. Berlin, Heidelberg: Springer, 338-346.
7. Grais, E. M., and Erdogan, H. (2011). *Single channel speech music separation using nonnegative matrix factorization and spectral masks*. Paper presented at the 17th International Conference on Digital Signal Processing, Corfu, Greece.
8. Nerkar, A. R., and Joshi, M. A. (2016). *Singing-voice separation from monaural recordings using empirical wavelet transform*. Paper presented at the 2016 International Conference on Ramanathapuram, India, 795-800.
9. Huang, P. Chen, S. D. Smaragdis, P. Hasegawa-Johnson, M. (2012). Singing-voice separation from monaural recordings using robust principal component analysis. *International Conference on Acoustic, Speech, and Signal Processing*, 57-60.
10. Virtanen, T. (2007). Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. *The Institute of Electrical and Electronics Engineers Transactions on Audio, Speech, and Language Processing*, 15(3), 1066-1074.
11. Vembu, S., and Baumann, S. (2005). Separation of vocals from polyphonic audio recordings. URL: <http://www.webcitation.org/query?url=http%3A%2F%2Fciteseerx.ist.psu.edu%2Fviewdoc%2Fdownload%3Bjsessionid%3D19E33B7CD0202D003378AEEF38B38B47%3Fdoi%3D10.1.1.139.5510%26rep%3Drep1%26type%3Dpdf&date=2018-06-28>, Son Eriřim Tarihi: 21.05.2018.
12. Raj, B., Smaragdis, P., Shashanka, M., and Singh, R. (2007). *Separating a foreground singer from background music*. India, Mysore: International Symposium Frontiers of Research in Speech & Music, pp. 3-5. URL:

- <http://www.webcitation.org/query?url=http%3A%2F%2Fparis.smaragd.is%2Fpubs%2FFRSM.pdf&date=2018-06-28>, Son Erişim Tarihi: 21.05.2018
13. Yoshii, K., Goto, M., and Okuno, H. G. (2004- Oct). Adamast: A drum sound recognizer based on adaptation and matching of spectrogram templates. Spain, Barcelona, *5th International Conference on Music Information Retrieval*, 184–191. URL:  
<http://www.webcitation.org/query?url=https%3A%2F%2Fpdfs.semanticscholar.org%2Ff809%2Ff31876181c93d212defed9212d11b1af40e4.pdf&date=2018-06-28> Son Erişim Tarihi: 21.05.2018
  14. Ozerov, A., Philippe, P., Bimbot, F., and Gribonval, R. (2007). *Adaptation of Bayesian models for single-channel source separation and its application to voice/music separation in popular songs*. Paper presented at the Translational of Audio, Speech, and Language Processing, Signal Processing Society
  15. Köktürk, G. (2006). Kısa süreli fourier dönüşümü için başarımlı karşılaştırma. *Dokuz Eylül Üniversitesi Fen ve Mühendislik Dergisi*, 8(3), 11-18.
  16. Sezer, E. (2008). *Epilepsi teşhisi için EEG sinyal analizi*, Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya, 10-12.
  17. İnternet: Sarl, I. Autocorrelation. URL:  
<http://www.webcitation.org/query?url=http%3A%2F%2Fautocorrelation.net%2F&date=2018-06-28> , Son Erişim Tarihi: 08/01/2018.
  18. İnternet: Medyan. *Wikipedia*. URL:  
<http://www.webcitation.org/query?url=https%3A%2F%2Ftr.wikipedia.org%2Fwiki%2FMedyan&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
  19. İnternet: Weisstein, E. W. L1 Norm. *MathWorld*. URL:  
<http://www.webcitation.org/query?url=http%3A%2F%2Fmathworld.wolfram.com%2FL1-Norm.html&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
  20. Karaçay, T. (2009). *Genel Topoloji*. Ankara: Başkent Üniversitesi, 220-229.
  21. İnternet: Nuclear norm and rank minimization. *Textons*. URL:  
<http://www.webcitation.org/query?url=https%3A%2F%2Ftextons.wordpress.com%2F2012%2F10%2F26%2Fnuclear-norm-and-rank-minimization%2F&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
  22. İnternet: Weisstein, E. W. Frobenius Norm. *MathWorld*. URL:  
<http://www.webcitation.org/query?url=http%3A%2F%2Fmathworld.wolfram.com%2FFrobeniusNorm.html&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
  23. İnternet: Cowlshaw, M. Fillmore, N. Lecture 14:Linear Algebra. *University of Wisconsin-Madison*. URL:  
<http://www.webcitation.org/query?url=http%3A%2F%2Fpages.cs.wisc.edu%2F%7Eamos%2F412%2Flecture-notes%2Flecture14.pdf&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.

24. İnternet: Fazel, M. Matrix rank minimization with applications. *Stanford University Department of Electrical Engineering*. URL: <http://www.webcitation.org/query?url=https%3A%2F%2Ffaculty.washington.edu%2Fmfazel%2Fthesis-final.pdf&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
25. Hopcroft, J., and Kannan, R. (2013) *Foundations of data science*. Cornell University, 130-135
26. Bao, B. K., Liu, G., Xu, C., and Yan, S. (2012) Inductive robust principal component analysis. *The Institute of Electrical and Electronics Engineers Transactions on Image Processing*, 21(8), 3794-3800.
27. İnternet: Lagrange multipliers, introduction. *Khan Academy*. URL: <http://www.webcitation.org/query?url=https%3A%2F%2Fwww.khanacademy.org%2Fmath%2Fmultivariable-calculus%2Fapplications-of-multivariable-derivatives%2Fconstrained-optimization%2Fa%2Flagrange-multipliers-single-constraint&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
28. İnternet: Lagrange multipliers, examples. *Khan Academy* examples. URL: <http://www.webcitation.org/query?url=https%3A%2F%2Fwww.khanacademy.org%2Fmath%2Fmultivariable-calculus%2Fapplications-of-multivariable-derivatives%2Fconstrained-optimization%2Fa%2Flagrange-multipliers-examples&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
29. İnternet: Gould, N. Penalty and augmented Lagrangian methods for equality constrained optimization. *Research Councils U.K.* URL: <http://www.webcitation.org/query?url=http%3A%2F%2Fwww.numerical.rl.ac.uk%2Fpeople%2Fnmimg%2Foupartc%2Flectures%2Fparts%2Fpart5.2.pdf&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
30. Yazdani, S., Shanbehzadeh, J., and Taghi Manzuri Shalmani, M. (2012). *RPCA: A Novel Preprocessing Method for PCA*. *Advances in Artificial Intelligence*. 2012, 7
31. Wright, J., Ganesh, A., Rao, S., and Ma, Y. (2009). Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. *Journal of the Association for Computing Machinery*.
32. Larsen, R.M. (1998) Lanczos bidiagonalization with partial reorthogonalization. *Department of Computer Science, Aarhus University, Technical report*, DAIMI PB-357,
33. İnternet: Lin, Z., Lin, M. Chen, and Y. Ma. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv*. URL: <http://www.webcitation.org/query?url=https%3A%2F%2Farxiv.org%2Fpdf%2F1009.5055.pdf&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
34. F'évotte, C., Gribonval R., and Vincent, E. (2005). BSS EVAL toolbox user guide. *France, Rennes: IRISA Technical Report 1706*.

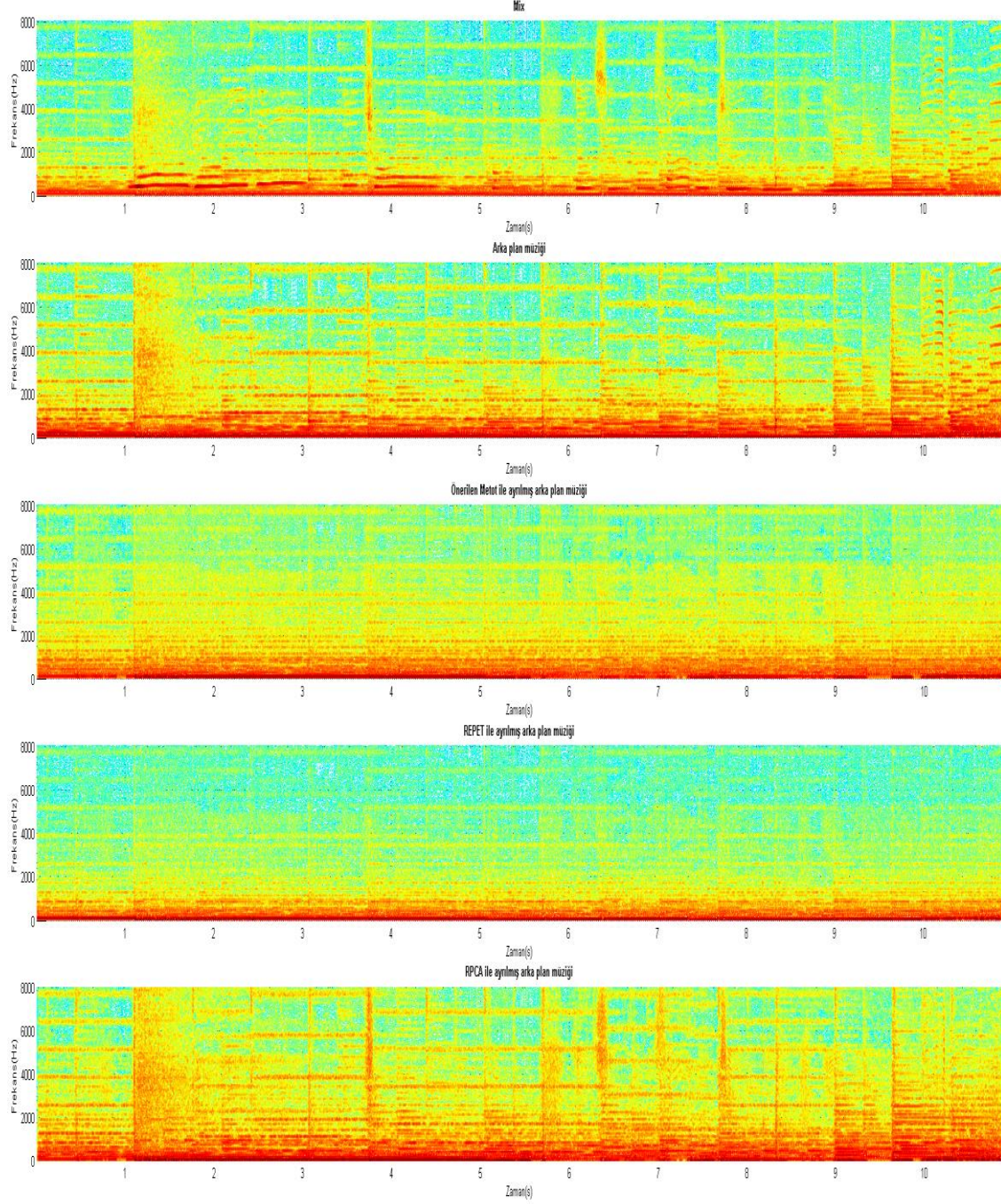
35. İnternet: MIR-1K Dataset. *Mirlab* URL: <http://www.webcitation.org/query?url=https%3A%2F%2Fsites.google.com%2Fsite%2Funvoicedsoundseparation%2Fmir-1k&date=2018-05-21>, Son Erişim Tarihi: 21.05.2018.
36. İnternet: About Mirlab. *Mirlab*. URL: [http://www.webcitation.org/query?url=http%3A%2F%2Fmirlab.org%2Fabout\\_intro.asp&date=2018-05-21](http://www.webcitation.org/query?url=http%3A%2F%2Fmirlab.org%2Fabout_intro.asp&date=2018-05-21), Son Erişim Tarihi: 21.05.2018.
37. Fitzek, F.H., Rein, S., Seeling, P., Reisslein, M. (2005) Robust header compression (rohc) performance for multimedia transmission over 3g/4g wireless networks. *Wireless Personal Communications* 32(1), 23–41.
38. Taşaner M, (2009). *Çoklu silindirik hedeflerin sınıflandırılması*, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 8-12.





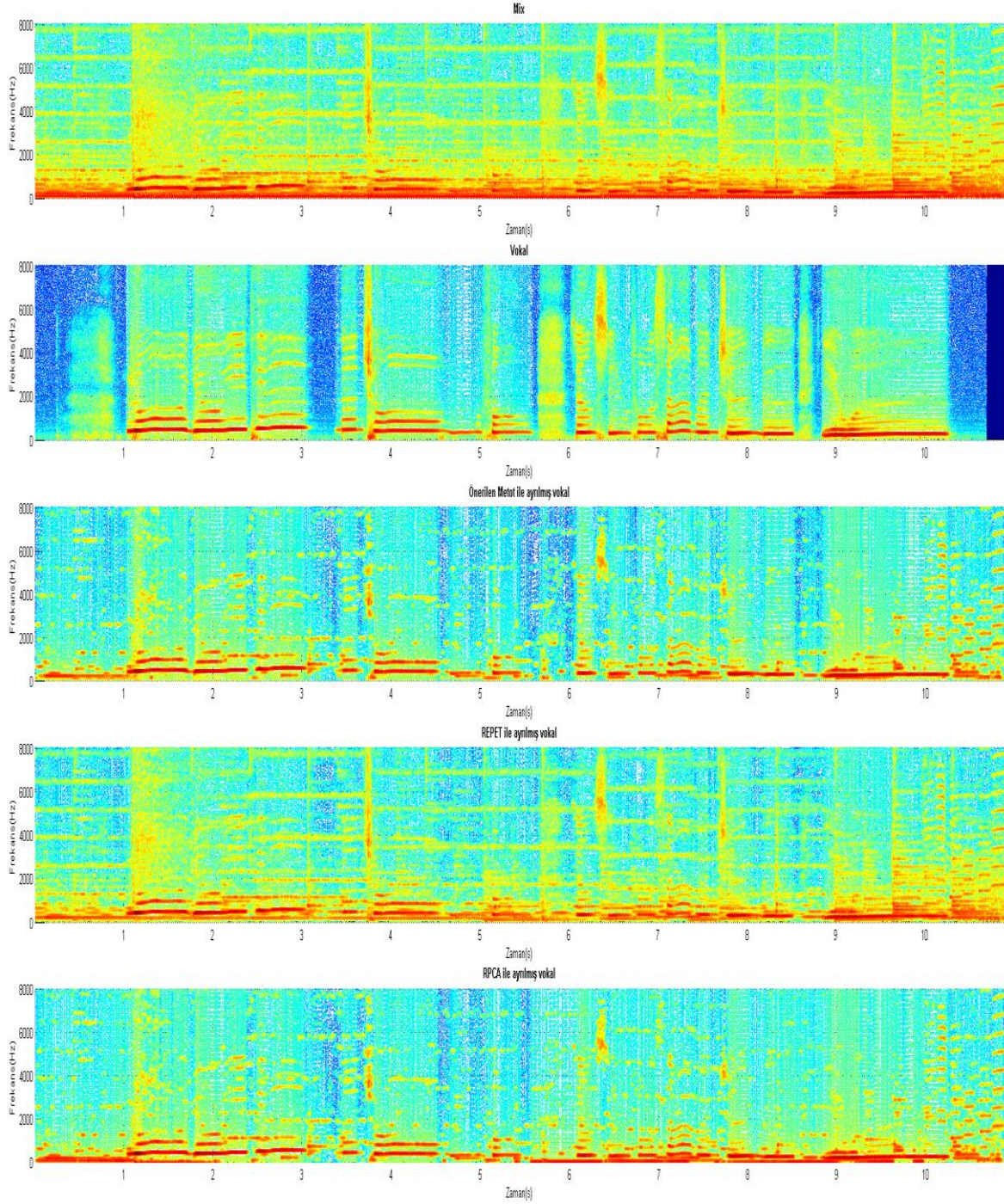
**EKLER**

EK-1. Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının Vokal Ve Arka Plan Müziği Çıktılarının Spektrogramları



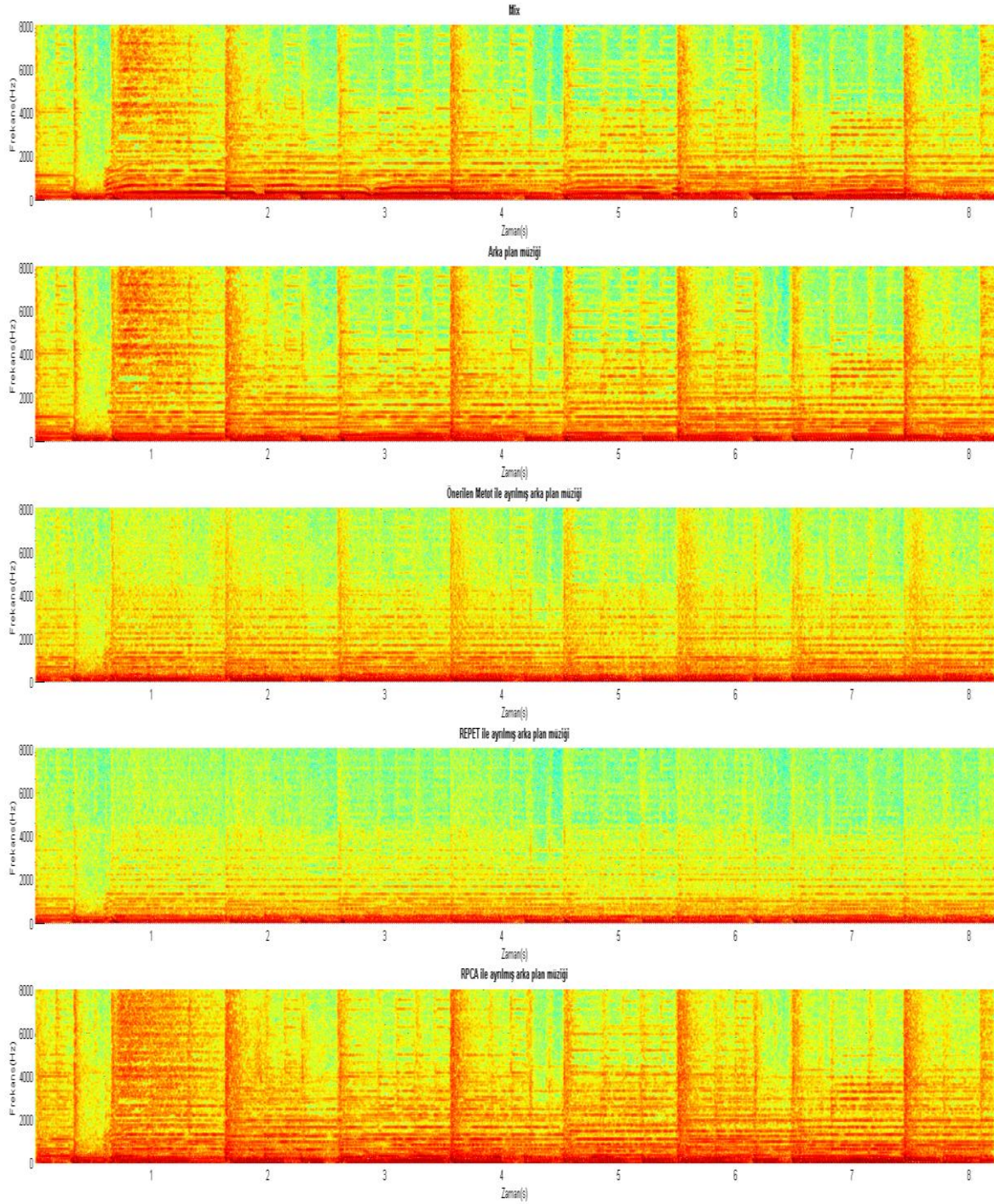
Şekil 1.1. Titon\_2\_09 parçası arka plan müziği spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



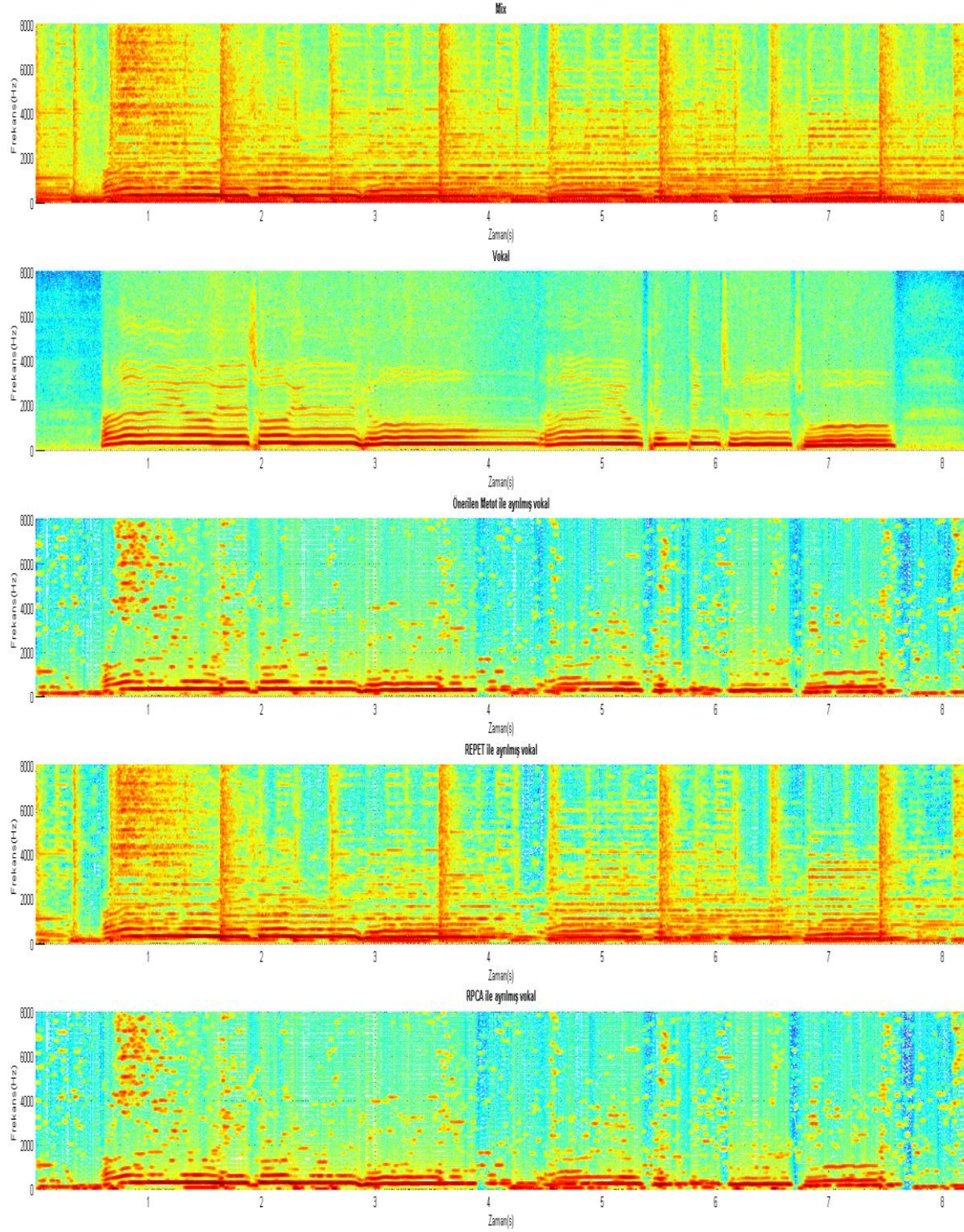
Şekil 1.2. Titon\_2\_09 parçası vokal spektogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



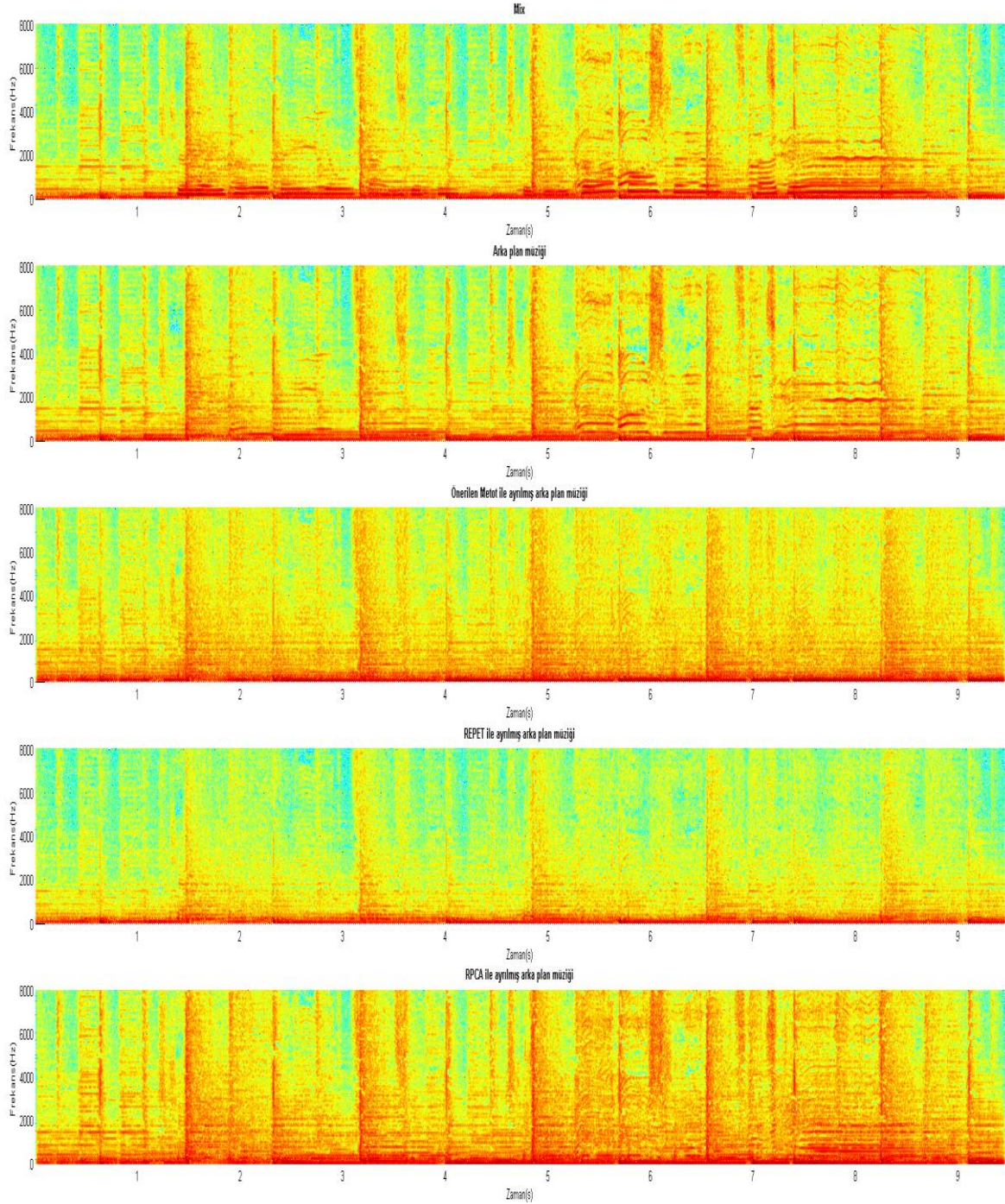
Şekil 1.3. Kenshin\_2\_05 parçası arka plan müziği spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



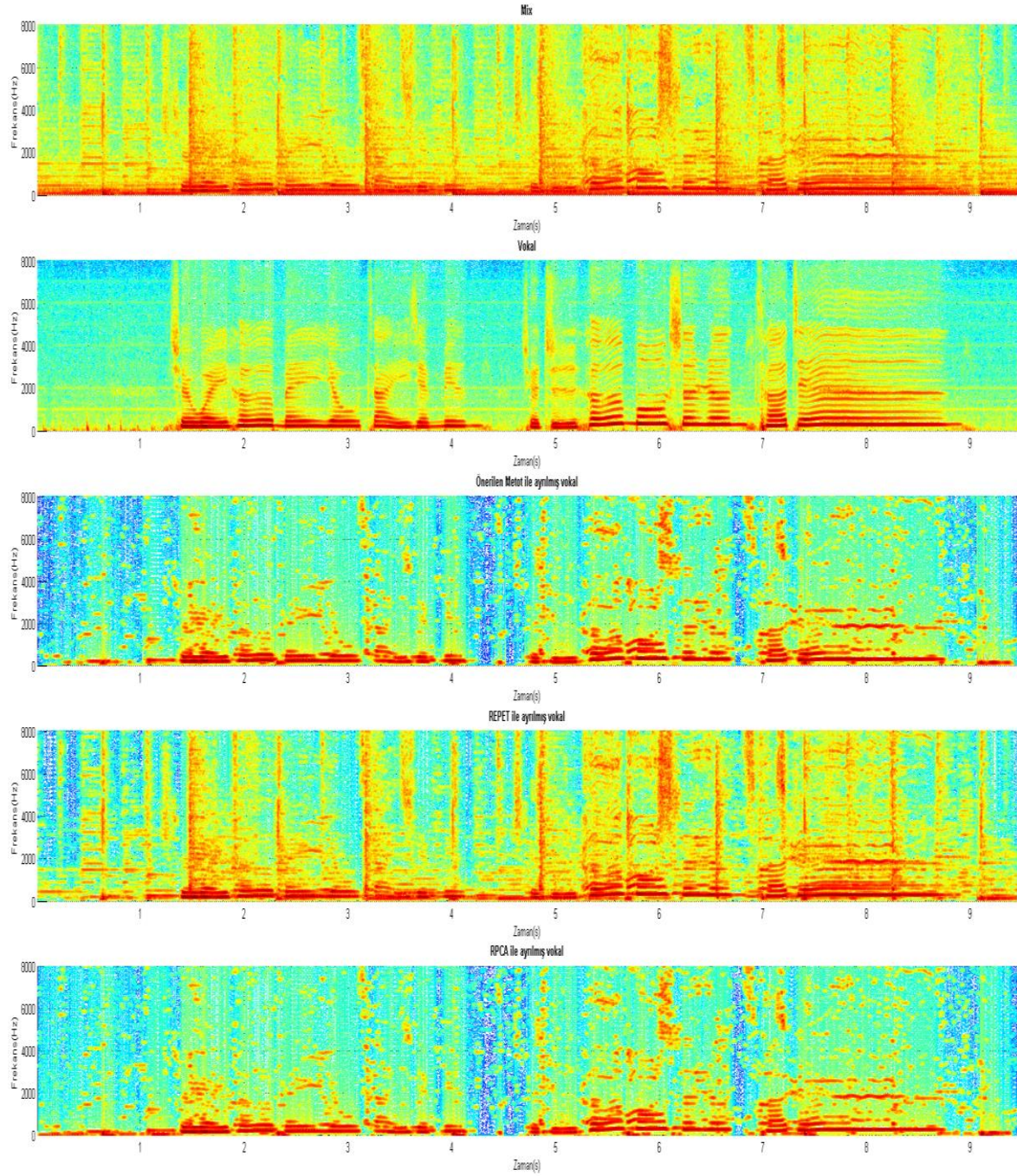
Şekil 1.4. Kenshin\_2\_05 parçası vokal spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



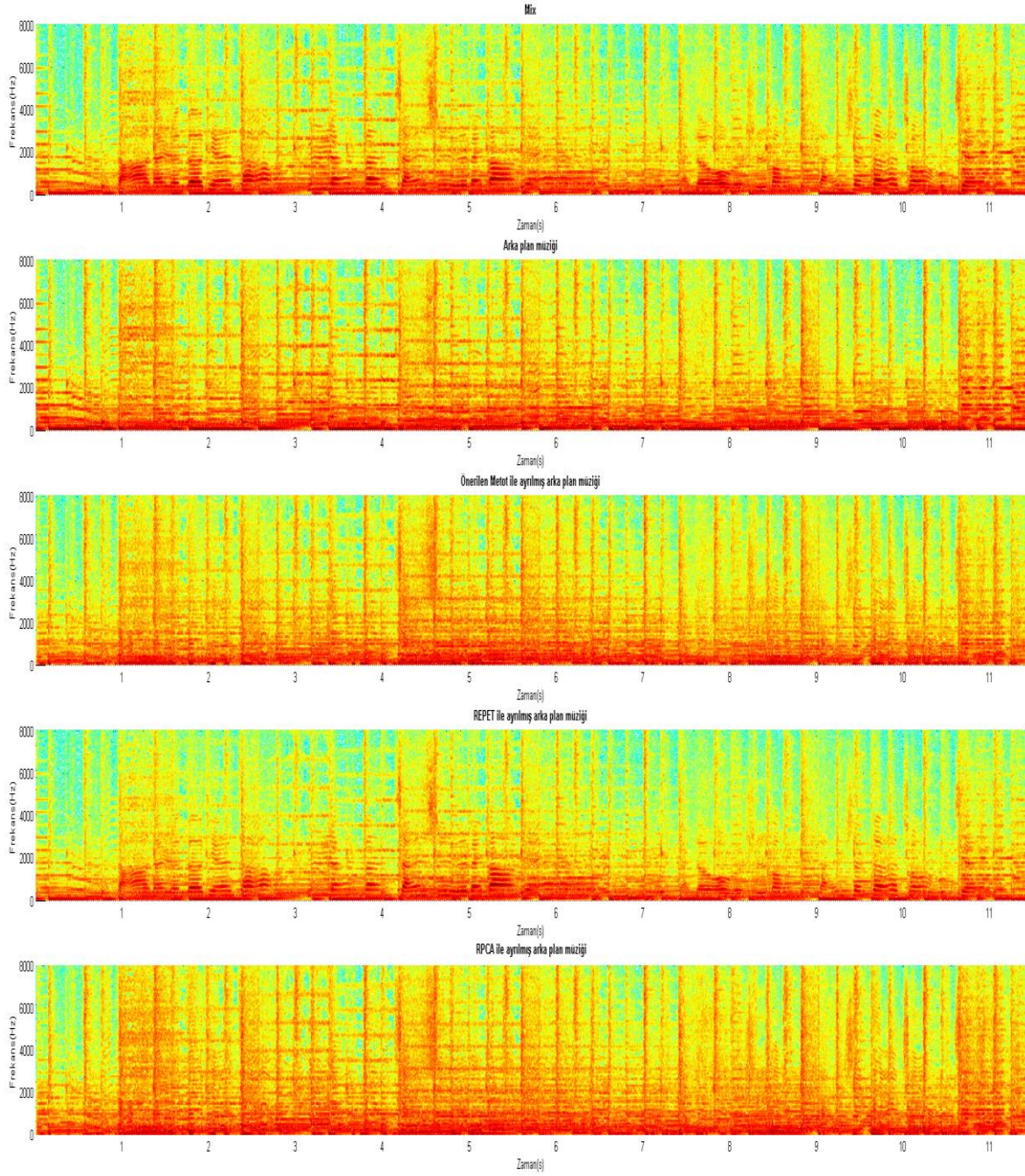
Şekil 1.5. Ani\_1\_07 parçası arka plan müziği spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



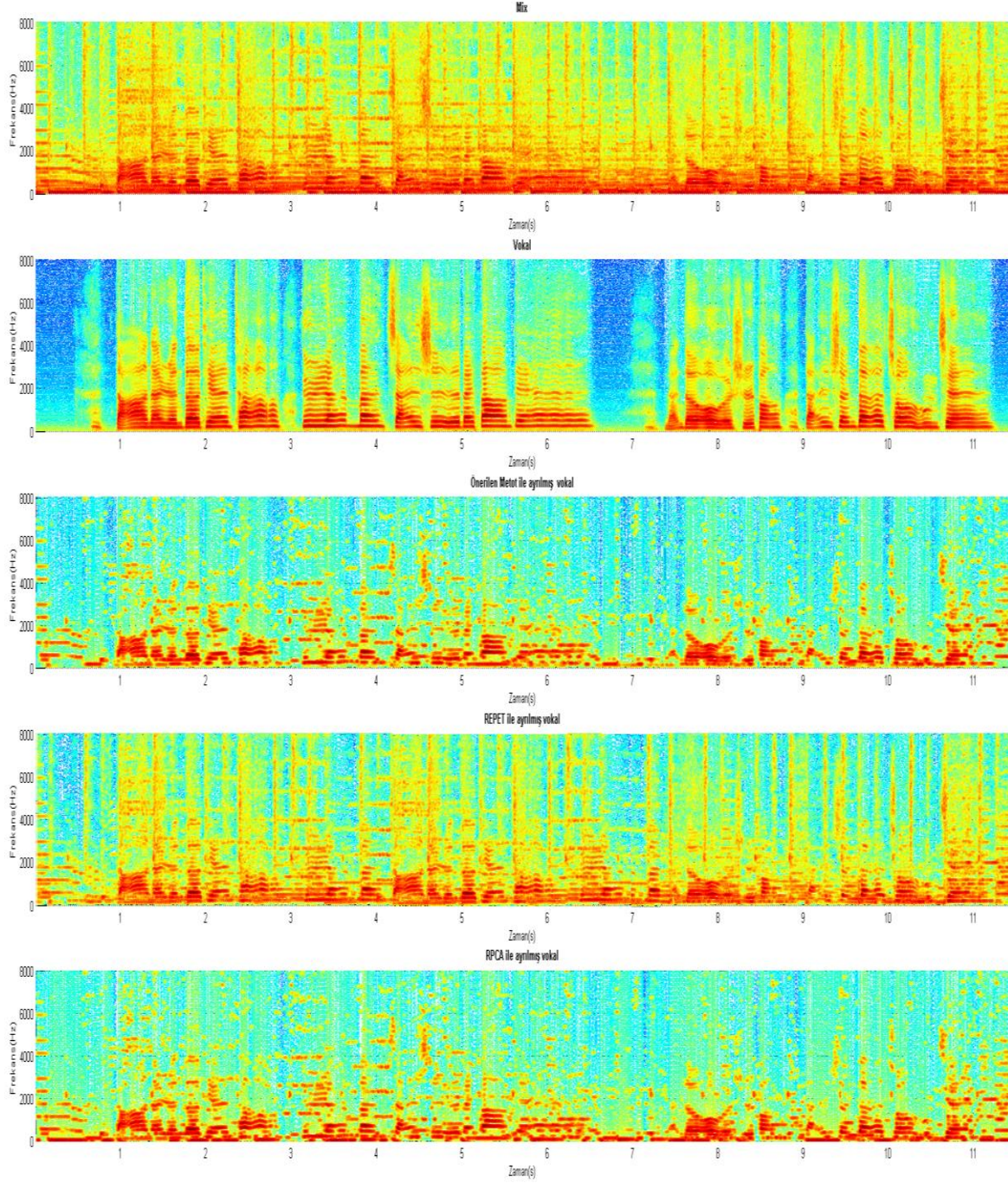
Şekil 1.6. Ani\_1\_07 parçası vokal spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



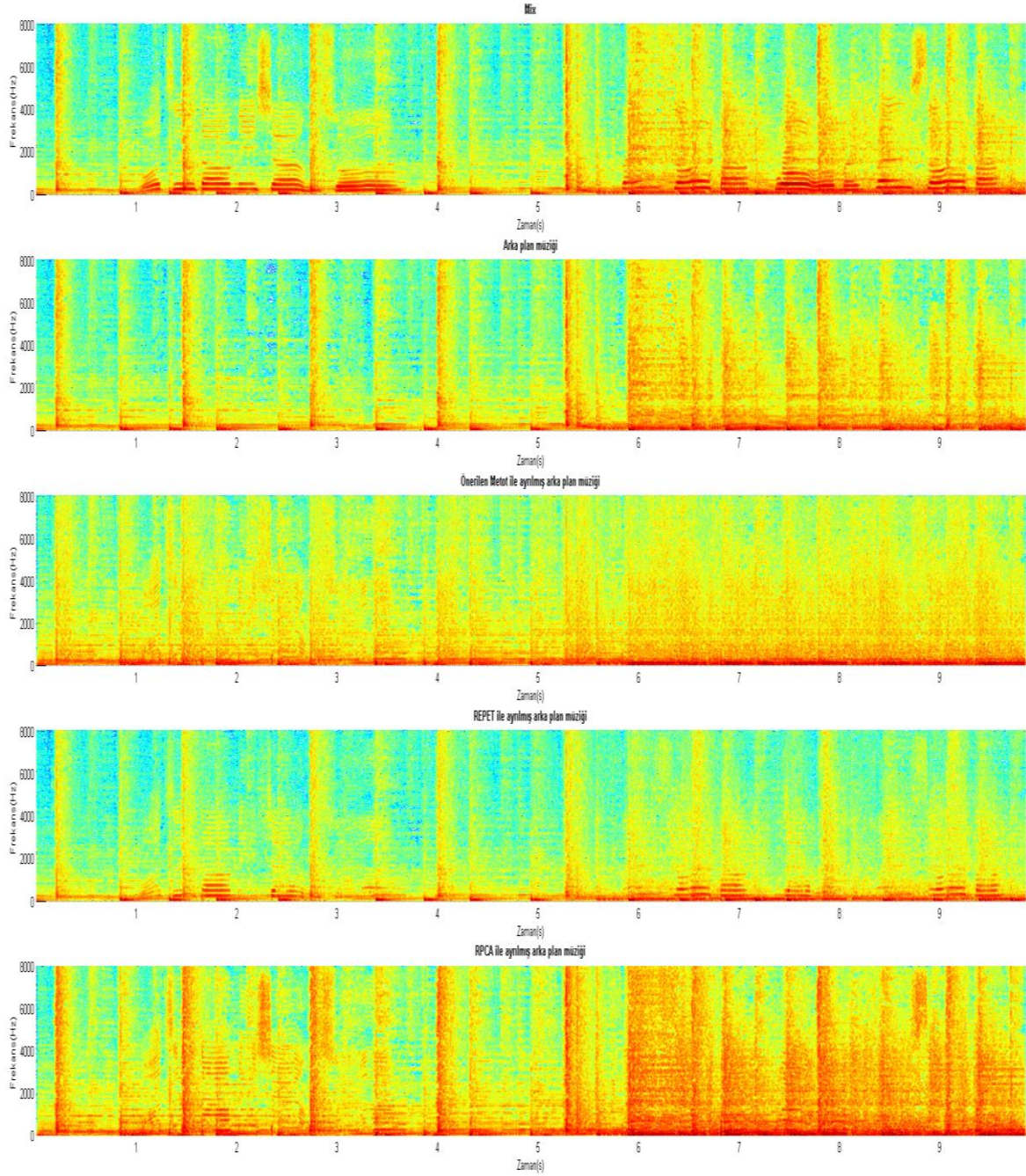
Şekil 1.7. Leon\_8\_10 parçası arka plan müziği spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



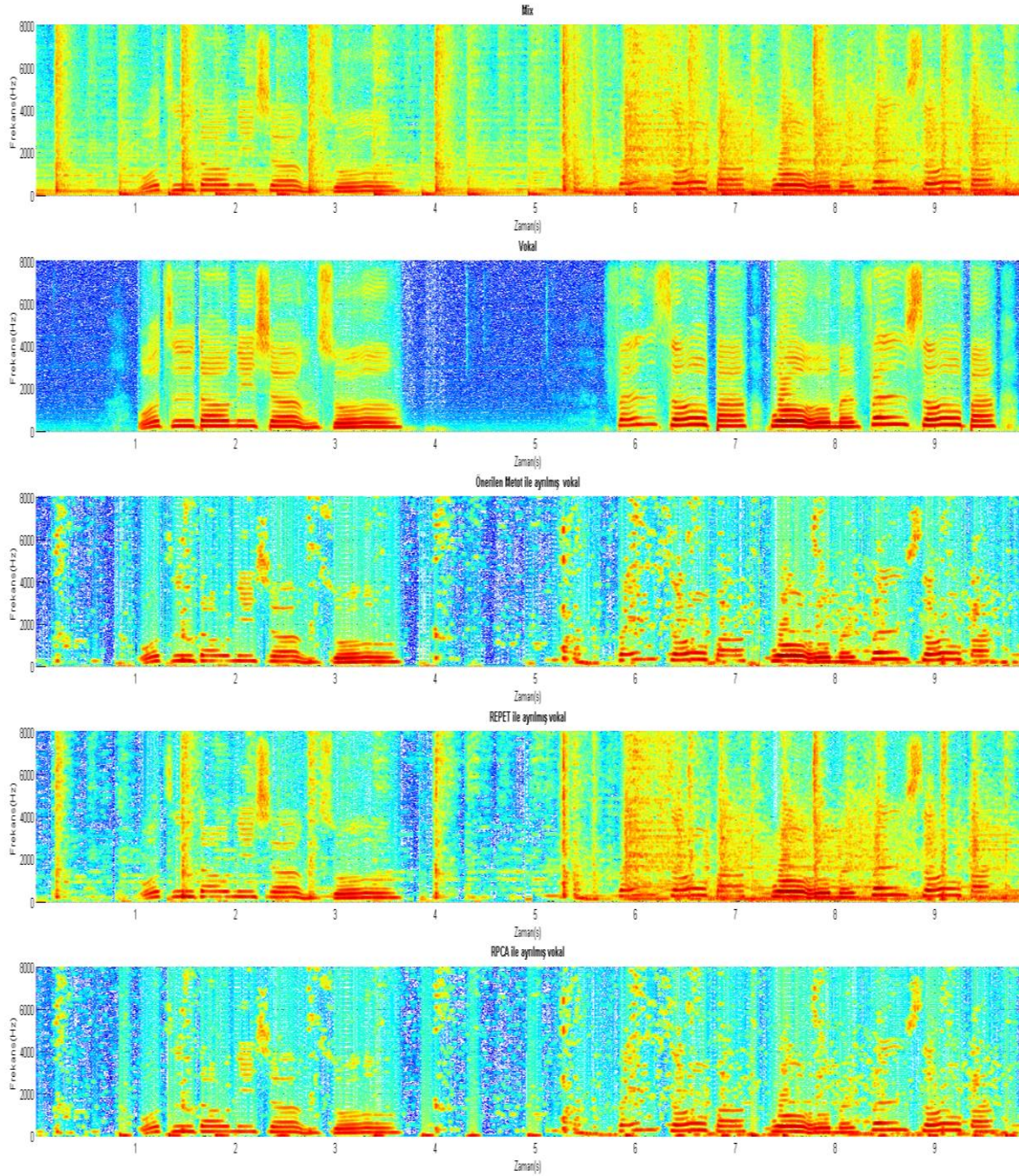
Şekil 1.8. Leon\_8\_10 parçası vokal spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



Şekil 1.9. Bobon\_5\_06 parçası arka plan müziği spektrogramları

EK-1. (devam) Tezde Önerilen Metot, REPET ve RPCA Yöntemlerinin Karşılaştırılması Amacıyla Kullanılan 5 Müzik Parçasının vokal ve Arka Plan Müziği Çıktılarının Spektrogramları



Şekil 1.10. Bobon\_5\_06 parçası vokal spektrogramları

## EK-2. 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 1. Katılımcı

**Yaşı** : 20

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

### Çizelge 1.1. Katılımcı-1'e ait test sonuçları

Katılımcı-1		1.Parça(Ani 1 07)	2.Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	3	2	3	2	2
	Müzik	3	3	2	5	2
RPCA	Vokal	2	3	1	3	4
	Müzik	4	4	2	3	2
Önerilen Metot	Vocal	5	4	4	5	5
	Music	2	4	4	3	4

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 2. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.2. Katılımcı-2'e ait test sonuçları

Katılımcı-2		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	1	3	2	3	3
	Müzik	2	1	4	1	2
RPCA	Vokal	3	4	2	1	2
	Müzik	3	3	2	2	2
Önerilen Metot	Vocal	1	1	1	1	1
	Music	2	2	1	2	2

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 3. Katılımcı

**Yaşı** : 31

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.3. Katılımcı-3'e ait test sonuçları

Katılımcı-3		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	4	5	3	3	3
	Müzik	4	4	3	5	5
RPCA	Vokal	3	3	4	4	4
	Müzik	3	2	3	3	3
Önerilen Metot	Vocal	5	4	5	5	5
	Music	5	5	4	4	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 4. Katılımcı

**Yaşı** : 29

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.4. Katılımcı-4'e ait test sonuçları

Katılımcı-4		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	2	2	1	2	2
	Müzik	3	3	2	2	1
RPCA	Vokal	1	1	1	4	1
	Müzik	4	3	3	3	2
Önerilen Metot	Vocal	3	3	3	3	3
	Music	4	3	4	2	2

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 5. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.5. Katılımcı-5'e ait test sonuçları

Katılımcı-5		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	1	1	3	1	1
	Müzik	3	3	3	3	2
RPCA	Vokal	4	5	5	5	3
	Müzik	3	3	3	3	2
Önerilen Metot	Vocal	3	3	5	3	2
	Music	4	4	4	5	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 6. Katılımcı

**Yaşı** : 28

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.6. Katılımcı-6'e ait test sonuçları

Katılımcı-6		1.Parça(Ani_1_07)	2. Parça(Kenshin_2_05)	3.Parça(Leon_8_10)	4.Parça(Titon_2_09)	5.Parça(Bobon_5_06)
REPET	Vokal	1	1	1	1	3
	Müzik	3	3	1	3	1
RPCA	Vokal	1	2	3	2	1
	Müzik	3	3	3	3	1
Önerilen Metot	Vocal	3	3	3	2	3
	Music	3	1	3	3	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 7. Katılımcı

**Yaşı** : 28

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.7. Katılımcı-7'e ait test sonuçları

Katılımcı-7		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	2	2	2	3	2
	Müzik	2	2	2	4	3
RPCA	Vokal	3	2	2	4	3
	Müzik	3	3	3	3	2
Önerilen Metot	Vocal	5	4	4	5	5
	Music	5	4	3	5	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 8. Katılımcı

**Yaşı** : 26

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.8. Katılımcı-8'e ait test sonuçları

Katılımcı-8		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	3	3	4	2	3
	Müzik	3	4	3	3	2
RPCA	Vokal	4	3	2	2	2
	Müzik	3	3	3	1	3
Önerilen Metot	Vocal	2	2	2	2	2
	Music	1	1	1	1	1

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 9. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.9. Katılımcı-9'e ait test sonuçları

Katılımcı-9		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	2	2	2	2	2
	Müzik	3	2	1	2	2
RPCA	Vokal	2	2	2	1	1
	Müzik	3	3	3	3	3
Önerilen Metot	Vocal	1	1	1	1	1
	Music	2	3	3	3	2

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 10. Katılımcı

**Yaşı** : 35

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.10. Katılımcı-10'e ait test sonuçları

Katılımcı-10		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	2	2	3	3	2
	Müzik	2	3	3	1	1
RPCA	Vokal	2	1	1	4	1
	Müzik	4	3	3	3	3
Önerilen Metot	Vocal	4	3	4	4	4
	Music	4	4	4	4	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 11. Katılımcı

**Yaşı** : 29

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.11. Katılımcı-11'e ait test sonuçları

Katılımcı-11		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	4	4	4	4	3
	Müzik	3	4	3	4	2
RPCA	Vokal	3	2	1	3	2
	Müzik	2	4	3	3	3
Önerilen Metot	Vocal	2	2	2	3	4
	Music	2	3	2	4	2

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 12. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.12. Katılımcı-12'e ait test sonuçları

Katılımcı-12		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	2	2	2	2	2
	Müzik	2	2	2	3	2
RPCA	Vokal	1	2	2	4	3
	Müzik	3	3	3	4	3
Önerilen Metot	Vocal	4	4	4	4	5
	Music	4	3	3	4	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 13. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.13. Katılımcı-13'e ait test sonuçları

Katılımcı-13		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	1	2	1	2	2
	Müzik	2	2	1	3	2
RPCA	Vokal	1	1	2	4	1
	Müzik	3	3	1	3	2
Önerilen Metot	Vocal	4	4	3	4	4
	Music	4	3	2	5	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 14. Katılımcı

**Yaşı** : 27

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.14. Katılımcı-14'e ait test sonuçları

Katılımcı-14		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	4	3	4	3	3
	Müzik	1	2	4	3	2
RPCA	Vokal	2	3	3	1	3
	Müzik	3	2	3	3	2
Önerilen Metot	Vocal	1	1	1	1	1
	Music	2	2	2	3	1

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 15. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.15. Katılımcı-15'e ait test sonuçları

Katılımcı-15		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	4	3	2	3	3
	Müzik	4	3	1	2	1
RPCA	Vokal	1	1	1	2	1
	Müzik	2	3	3	4	3
Önerilen Metot	Vocal	2	2	3	2	4
	Music	3	4	4	5	2

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 16. Katılımcı

**Yaşı** : 30

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.16. Katılımcı-16'e ait test sonuçları

Katılımcı-16		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	2	2	3	3	2
	Müzik	1	1	3	1	2
RPCA	Vokal	1	1	2	4	2
	Müzik	3	3	3	4	2
Önerilen Metot	Vocal	3	2	4	4	3
	Music	4	3	5	5	3

EK-2. (devam) 3'ü Kadın 14'ü Erkek 17 Katılımcı İle Yapılan Test Aşamasından Sonra Tanıma Sonuçlarının İşlendiği Formlar

Sait Melih DOĞAN in tez çalışması kapsamında gerçekleştirdiği test çalışmaları ile ilgili gerekli bilgilendirme tarafıma yapılmıştır. Aşağıda sonuçları verilen test çalışmasında, tarafımda gerçekleştirilmiştir.

**Adı Soyadı** : 17. Katılımcı

**Yaşı** : 27

**Cinsiyeti** : Bay  Bayan

**İletişim Bilgisi** : .....

**İmza** :

Çizelge 1.17. Katılımcı-17'e ait test sonuçları

Katılımcı-17		1.Parça(Ani 1 07)	2. Parça(Kenshin 2 05)	3.Parça(Leon 8 10)	4.Parça(Titon 2 09)	5.Parça(Bobon 5 06)
REPET	Vokal	3	3	2	4	3
	Müzik	2	2	4	2	2
RPCA	Vokal	2	2	3	1	2
	Müzik	3	3	3	4	2
Önerilen Metot	Vocal	1	1	1	2	3
	Music	4	3	2	3	3

## ÖZGEÇMİŞ

### Kişisel Bilgiler

Soyadı, adı : Sait Melih DOĞAN  
 Uyuğu : T.C.  
 Doğum tarihi ve yeri : 03.06.1987, Ankara  
 Medeni hali : Evli  
 Telefon : 0 (555) 7063990  
 E-Posta : saitmelihdogan@hotmail.com



### Eğitim

Derece	Okul/Program	Mezuniyet tarihi
Yüksek Lisans	Gazi Üniversitesi/Elektrik - Elektronik	Devam ediyor
Lisans	Gazi Üniversitesi/Elektrik - Elektronik	2011
Lise	Kocatepe Mimar Kemal Lisesi	2005

### İş Deneyimi

Yıl	Çalıştığı Yer	Görev
2011 - 2012	Netsite İletişim ve Elektronik Sis. A.Ş.	Proje Mühendisi
2012 - 2014	Milli Piyango İdaresi Başkanlığı	Mühendis
2014 – 2015	Türkiye Kalkınma Bankası A.Ş.	Bilişim Uzmanı
2015 -	T.C Başbakanlık	Bilişim Uzmanı

### Yabancı Dil

İngilizce

### Yayınlar

Doğan, S.M. & Salor, O. (2018). *Music/singing voice separation based on repeating pattern extraction technique and robust principal component analysis*. Paper presented at the 2018 5th International Conference on Electrical and Electronic Engineering (ICEEE), Istanbul, Turkey.



*GAZİ GELECEKTİR..*