





Frequency estimation for monophonical Turkish maqam music by using modified variational mode decomposition method

Berrak Öztürk Şimşek¹ , Aydın Akan² 

¹Istanbul University, Faculty of Engineering, Electrical and Electronics Engineering Department, Istanbul 34320, Turkey

²Chalabi Clerk Izmir University, Faculty of Engineering and Architecture, Biomedical Engineering Department, Izmir, 35620, Turkey

Highlights:

- Monophonic Turkish maqam music
- Fundamental frequency estimation
- Modified variational mode decomposition for music signals

Keywords:

- Variational mode decomposition
- Monophonic music
- Turkish maqam music
- Pitch frequency

Article Info:

Received: 26.04.2017

Accepted: 19.10.2017

DOI:

10.17341/gazimmfd.416457

Acknowledgement:

I would like to express my very great appreciation to Assoc. Prof.Dr. Baris Bozkurt for his valuable support and advise.

Correspondence:

Author: Berrak Öztürk Şimşek

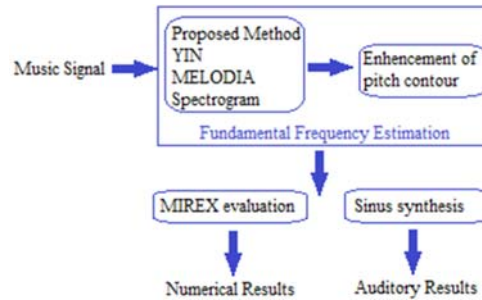
Şimşek

e-mail: bozturk@iku.edu.tr

phone: +90 2124984727

Graphical/Tabular Abstract

The aim of this study is to estimate the fundamental frequency of monophonic Turkish maqam music signals. We develop and use a modified Variational Mode Decomposition method to obtain the modes of input signal. In order to compare the proposed method with other common decomposition methods which are YIN, MELODIA, and original VMD algorithm, we utilized MIREX criteria for performance evaluation. In addition, the auditory results reveal that we can detect the correct melody of the music.



Purpose: In this study, a new modified Variational Mode Decomposition (VMD) method is presented for fundamental frequency estimation of Turkish maqam music records. The proposed method's performance is compared to the other common decomposition methods such as YIN, MELODIA, and original VMD methods.

Theory and Methods:

Monophonical Turkish maqam music signals are decomposed by using the modified VMD. VMD is a method to decompose a given input signal into some ensemble modes which has narrow-band properties. In this work, we propose Elasticnet regression for the optimization step in VMD algorithm. Besides, frequency estimation algorithm is applied to the autocorrelation of the input signal to provide further improvement.

Results:

We used four different maqam music pieces performed by 8 different instruments including wind, percussive and stringed to test the performance of our VMD algorithm. Proposed method has been evaluated according to the MIREX criteria. The comparison has been performed by means of Voicing Detection, Voicing False Alarm, Raw Pitch Accuracy, Raw Chroma Accuracy and Overall Accuracy, with encouraging results.

Conclusion:

We proposed a modified VMD based approach for fundamental frequency and pitch contour estimation in monophonic Turkish maqam music signals. It has been observed that the results obtained by the proposed method for monophonic Turkish maqam music are comparable with that of polyphonic music estimation algorithms such as YIN, and MELODIA in terms of MIREX criteria.



Monofonik Türk makam müziği eserlerinde iyileştirilmiş değişken kip ayrışım yöntemi ile temel frekans kestirimi

Berrak Öztürk Şimşek^{1*}, Aydın Akan²

¹Istanbul Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, İstanbul, 34320, Türkiye

²İzmir Kâtip Çelebi Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Biyomedikal Mühendisliği Bölümü, İzmir, 35620, Türkiye

Ö N E Ç I K A N L A R

- Monofonik Türk makam müziği
- Temel titreşim frekansı kestirimi
- Müzik sinyalleri için iyileştirilmiş değişken kip ayrışımı

Makale Bilgileri

Geliş: 26.04.2017

Kabul: 19.10.2017

DOI:

10.17341/gazimmfd.416457

Anahtar Kelimeler:

Değişken kip ayrışımı,
monofonik müzik,
Türk makam müziği,
temel titreşim frekansı

ÖZET

Bu çalışmada, Değişken Kip Ayrışım yöntemi tabanlı yeni bir yaklaşım kullanılarak monofonik Türk makam müziği kayıtları içerisindeki temel titreşim frekanslarının kestirimi önerilmektedir. Değişken Kip Ayrışım yöntemi, reel değerli bir sinyali sonlu sayıda alt sinyallere (kiplere) ayırmaktadır. Değişken Kip Ayrışım sinyali içerisindeki temel bandları, yinelemesiz ve uyarlanabilir olarak belirleyip eş zamanlı olarak uygun kiplerin kestirilmesini sağlar. Bu yöntemde sinyali en iyi temsil edebilecek kipler, Görgül Kip Ayrışım yönteminde kullanılan özgül kip fonksiyonları gibi dar bantlı olma özelliğine sahiptir. Bu çalışmada, klasik Değişken Kip Ayrışım algoritmasında en iyileşme aşamasında kullanılan Tikhonow düzenleme yerine Elastic Net yönteminin kullanılması önerilmekte ve böylece temel titreşim frekansı kestirim başarımları iyileştirilmektedir. Frekans kestirim algoritması, müzik sinyali yerine öz-iltinti fonksiyonuna uygulanarak iyileştirmeye katkıda bulunulmuştur. Yapılan benzetim çalışmalarında, müzik sinyalinde ayırıştırma aracı olarak sıkça kullanılan Spektrogram, MELODIA, YIN ve orijinal Değişken Kip Ayrışım yöntemleri ile karşılaştırıldığında, gerçek müzik ve sentetik test sinyalleri ile elde edilen temel titreşim frekansı kestirim sonuçlarının başarılı ve karşılaştırılabilir olduğu gözlenmiştir.

Frequency estimation for monophonical Turkish maqam music by using modified variational mode decomposition method

H I G H L I G H T S

- Monophonic Turkish maqam music
- Fundamental frequency estimation
- Modified variational mode decomposition for music signal

Article Info

Received: 26.04.2017

Accepted: 19.10.2017

DOI:

10.17341/gazimmfd.416457

Keywords:

Variational mode
decomposition,
monophonic music,
Turkish maqam music,
pitch frequency.

ABSTRACT

In this study, a new method is presented for the fundamental frequency estimation of Turkish maqam music recordings by using Variational Mode Decomposition. Variational Mode Decomposition is a method to decompose a real-valued signal into an ensemble of sub-signals (modes) which is entirely non-recursive, determines the relevant bands adaptively and estimates the corresponding modes concurrently. In order to decompose a given signal optimally, Variational Mode Decomposition obtains an ensemble of modes with narrow-band properties similar to the Intrinsic Mode Function definition used in Empirical Mode Decomposition. In this study, we propose using the ElasticNet Regression instead of the Tikhonow Regularization during the optimization stage of the classical Variational Mode Decomposition algorithm to improve the fundamental frequency estimation performance. Frequency estimation algorithm is applied on the autocorrelation function of the music signal to further improve the results. Simulation results on real music and synthetic test data show better or comparable performance to other common decomposition methods for music signals such as spectrogram, YIN, MELODIA and original Variational Mode Decomposition methods.

*Sorumlu Yazar/Corresponding Author: bozturk@iku.edu.tr / Tel: +90 212 498 4727

1. GİRİŞ (INTRODUCTION)

Kültürel mirasın temel öğelerinden biri olan müzik bilimi, bilişim teknolojilerinin gelişmesine ve yaygınlaşmasına bağlı olarak değişmektedir. Bu gelişmeler eşliğinde müziğin dinlenme ve üretilme biçimi hızla değişmekte ve bununla beraber hesaplamalı müzik bilimi ilerlemektedir. Müzik bilgi erişiminin büyük bir kısmının sembolik veriler yani notalar üzerinden yapılması, bilgi erişim uygulamalarının (Music Information Retrieval-MIR) yaygınlaşması ve teknolojinin gelişmesi otomatik olarak notaya dökme probleminin önemini artırmaktadır. Bu problem uluslararası alanda üzerinde son 15 yıldır çalışılan bir konu olmasına karşın, Türk makam müziği kayıtları için bu tür araştırmalar henüz yetersizdir [1]. Nota yazım programları, akort cihazları, sentezciler gibi müzik ile yakından ilgili olan teknolojilerin Türk müziği eserlerinin üretim ve araştırılması aşamasında verimli bir biçimde kullanılabilmesi için müzikoloji bilimi ve sinyal işleme araçlarının birlikte çalışarak otomatik notaya dökme problemine açıklık getirmesi gerekmektedir [2]. Batı müziği için hesaplamalı müzikoloji ve MIR uygulamaları özellikle son 15 yıldır hızla gelişmektedir. İlk kez 1980'lerin başlarına doğru polifonik müziklerin notaya dökülmesi için çoklu frekans algoritması tasarlanmıştır. Daha yeni notaya dökme sistemleri psikoakustik (insan ile dünyadaki sesler arasındaki psikolojik etkileşim) analiz prensipleri içermektedir ve yeni hesaplama yöntemleri sayesinde uygulama alanı genişlemiştir [3, 4].

Batı müziği için geliştirilmiş olan otomatik notaya dökme teknikleri, iki müzik arasındaki farklar nedeniyle Türk makam müziğine doğrudan uygulanamamaktadır [1, 5]. Bu iki müzik arasında tavrı, tartım, icra ve kuram olarak birçok fark sıralanabilmektedir. Klasik Türk müziği icrası Batı müziğine oranla çok fazla süsleme içermektedir. Süslemelerin fazla olması özellikle temel titreşim frekansı analizini zorlaştıran sebeplerden biridir. Batı müziği ses sisteminde notalara karşılık gelen temel titreşim frekansları, frekans uzayının logaritmik olarak 12 eşit parçaya ayrılması ile oluşmaktadır. Bu parçaların dayanarak notaların karşılık geldiği frekanslar tanımlanmış ve bunlar standart haline getirilmiştir. Klasik Türk makam müziği için farklı ses sistemlerinden hangisinin icra ile kuramı en iyi şekilde örtüştürdüğü konusunda anlaşma yoktur. Frekans uzayını 53, 106, 159 eşit parçaya ayıran ses sistemlerinin yanı sıra tam sayı operatörlerinden faydalanarak notaların yerlerini ve frekanslarını tanımlayan sistemler de mevcuttur [1, 6]. Kullanılan sistemler notalar için belli bir merkez frekans değerini belirleyebilmekte iken, icracı tarafında belirli bir frekans aralığında icra edilmektedir. Bu nedenle nota isimlerinin belirli tek bir frekans yerine bir frekans aralığına atanması gerekmektedir. Türk müziğinde, Batı müziği için tanımlanmış olan standart akort frekansı ($L_{a4}=440\text{Hz}$) gibi geçerli tek bir akort veya perde frekansı bulunmamaktadır [1, 7]. Ney boyları ile ilişkilendirilen bir ahenk sistemi akort amacıyla kullanılmaktadır [8]. Bunun yanı sıra kullanılan mikrotonal perde yapısına sahip enstrümanlar analizi zorlaştırmaktadır. Türk makam müziğinin Batı müziğinden

çok farklı ve karmaşık tartım özellikleri vardır. Batı müziği için geliştirilmiş ritim analiz araçları, klasik Türk makam müziği kayıtlarında 4/4 ve 3/4'den farklı vuruşları olan ve güçlü vuruşlar içermeyen kayıtların analizinde düşük başarıya sahiptirler. Batı müziğinde ezginin ana karakterini veren tonalite genel olarak majör ve minör diziler ve bunlardan üretilen diziler ile karakterize edilse de, Türk müziğinde ezgi karakteri makam ile tanımlanır. Makam sayısı çok fazla olmakla birlikte, günümüzde sıkça kullanılan otuz adet makamdan söz etmek mümkündür. Türk makam müziği ile Batı müziği arasında bahsi geçen bu farklar nedeniyle, müzik analiz araçları ve uygulamalarının Türk makam müziği için geliştirilerek kullanılması gerekmektedir ve bu hususta bir açık vardır. Türk müziğinin ana karakteri olan makam kavramının, Ortadoğu, Asya ve Kuzey Afrika'nın önemli bir kısmında yaşamaya devam eden geleneksel müzik ile benzerlikleri bulunmaktadır. Bu nedenle Türk makam müziğinin notaya dökülmesi ve bilgi erişim uygulamalarının gelişmesi geniş bir coğrafyaya ait müziğin anlaşılması adına da katkılar sağlayacaktır [1, 9]. Otomatik notaya dökme; icracı tarafından performans edilmiş bir kayıttan, çeşitli sinyal işleme araçları ile işlenerek sembolik veriye dönüştürülmesi olarak tanımlanabilmektedir [10]. Bu problem uluslararası alanda da henüz tam olarak çözülmemiş fakat üzerinde yoğunlaşmış zor bir hedefdir. Bu konuda birçok yayın [11], doktora tezi [12, 13], ve açık kaynaklı uygulama [14, 15] bulunmaktadır.

Perde frekansının algısı, insan işitmesi ve sesin anlaşılması açısından önemli rol oynar. Dinleyiciler, birçok eş zamanlı sesin perdelerini algılayabilmektedir ve karışım içerisinde bir sesi akustik olarak ayırmak için perdeleri etkin bir biçimde kullanabilmektedir. Perde, sesin algısal bir özneliğidir ve fiziksel olarak işaretin saniyedeki tekrar sayısına, periyodikliğe karşılık gelmektedir [1, 16]. Perde periyodunun yerinin anlaşılması çok güçtür, bu sebeple ses analizinde daha çok gerçek f_0 ölçümü hedeflenir.

Klasik Türk müziği, Batı müziğinden farklı olarak çoğunlukla monofonik ve heterofonik olarak icra edilmektedir. Türk müziğinde kullanılan mikrotonal perdeleri enstrümanlar ve icracı tarafından yapılan süslemeler temel titreşim frekansı analizini zorlaştırmaktadır. Nota yazma uygulamalarında elde edilen temel titreşim frekansı değerinin belli bir nota ismine atfedilmesi gerekmektedir. Kayıt içerisinde vurmalı çalgı bulunması ritim analizini kolaylaştıran bir etken olmasına rağmen, temel titreşim frekans analizini zorlaştırmaktadır. Müziğin icrası sürekli zamanlı bir frekans uzayına karşılık gelirken notalar ayrık zamanlı bir frekans temsil ederler. Çalışmamızda DKA tabanlı önerilen yöntem kullanılarak monofonik Türk makam müziği kayıtları ayrıştırılmış ve temel titreşim frekans değerleri elde edilmiştir. Temel titreşim frekans analizi temel olarak sinyalin zaman-frekans çözümlemesine dayanmaktadır. Bu çerçeveden bakıldığında en yaygın kullanılan ve temel niteliğindeki analiz yöntemlerinden biri Kısa Zamanlı Fourier Dönüşümü (KZFD) temelli

Spektrogram yöntemidir [17]. Birçok temel frekans analiz yönteminin doğruluk karşılaştırılmasında kullandığı önemli bir araçtır. GKA yöntemi 1990'ların sonunda Huang tarafından önerilmiş ve Hilbert Huang Dönüşümünün (HHD) elde edilmesinde kullanılmıştır [1, 18]. Bu yöntemin amacı, verilen bir giriş sinyalini farklı izgesel bantlı, önceden bilinmeyen kiplere ayırmaktır [18]. Günümüzde birçok sinyal analizi ve sınıflandırma probleminin çözümünde HHD yaygın olarak kullanılmaktadır. Bu algoritmanın sağlam olmayan yanı analitik bir tanıma sahip olmamasıdır. Bu nedenle özellikleri ve yöntemin sınırlamaları net bir biçimde bilinmemektedir. Gürültü ve örnekleme konusunda hassas bir algoritmadır.

Spektrogram gibi sıkça kullanılan bir sinyal işleme aracı olan Dalgacık dönüşümü, Batı müziği uygulamalarında müzik sinyallerinin frekans kestirimi için kullanılmaktadır. Dalgacık dönüşümü, işitsel işaret işleme, görüntü işleme gibi bir ve iki boyutlu işaret işleme uygulamalarında yoğun olarak kullanılmaktadır. Yöntemin taban fonksiyonlarına "dalgacık" adı verilir. Bu yöntem dalgacık ölçeğinin seçimine dayanmaktadır. Bunlar önemli olmayan dalgacık katsayılarını, o kısımdaki ilgili sinyalin enerjisinin eşik değerini kullanarak yok ederler. Diğer bazı çalışmalarda [1], sinyali uyarlanabilir alt bantlarda uyarlanabilir dalgacıklara ayırıştırılan Görgül Dalgacık Dönüşümü (GDD) kullanılmıştır [20-22]. Bu yöntemde giriş işareti uygun taban fonksiyonları cinsinden yazılarak frekans bileşenlerinin ortaya çıkartılmasına dayanır.

Rezonatör Zaman Frekans İmgesi (RZFİ) isminden de anlaşılacağı gibi frekansa bağlı zaman-frekans gösterim yöntemlerinden biridir. Var olan diğer zaman-frekans analiz yöntemlerinin müzik sinyalleri için bazı önemli kısıtlamaları bulunurken, RZFİ yönteminde bazı parametrelerin düzenlenmesi ile farklı zaman-frekans çözünürlükleri kullanılabilir. RZFİ ise birinci dereceden karmaşık rezonatör bankası ile uygulanarak hesap yükü açısından avantaj sağlar [23]. Ancak monofonik kayıtlarda harmonik bileşenlerin frekans bölgesinde eşit bir biçimde dağılımları nedeniyle Eş dağılım analizi daha kullanışlı olmaktadır. RZFİ son yıllarda Batı müziği analizinde sıkça kullanılan yeni bir zaman-frekans gösterimi sunmaktadır [24]. YIN algoritması öz-tilinti tabanlı ve müzik sinyalleri için sıkça kullanılan algoritmalarından birisidir [2]. Verilen giriş işaretinin kendi üzerinde kaydırılması ile dalga şeklinin hangi sıklıkla tekrar ettiğinin ölçümüne dayanmaktadır. Burada kaydırılmış ve orijinal işaretin birbiri ile en çok örtüştüğü kaydırma miktarı temel periyodu vermektedir, temel titreşim frekansı f_0 bu değerden hesaplanır. Öz-tilinti tabanlı temel titreşim frekansı kestiriminde, sinyal belirli boyutta bir pencere ile pencerelenir ve bu pencere içerisindeki sinyal kendi üzerinde kaydırılarak her pencere için bir f_0 değeri elde edilir. Elde edilen f_0 değerleri müzik parçası boyunca elde edilerek temel melodiye ait frekans değerlerinin oluşturduğu perde konturu elde edilmiş olur. MELODIA son yıllarda Batı müziği analizinde sıkça kullanılan bir melodi kestirim algoritmasıdır [25]. Bu algoritmada amaç, verilen giriş sinyalinden sürekli zamanlı

diziler olarak bir perde konturu elde etmektir. Giriş sinyalinden elde edilen bu konturlardan bir dizi karakteristik belirlenerek melodik olan ve olmayan konturlar ayrıştırılır. Yöntem dört temel bloktan oluşur: sinüzoidalın çıkartılması, çıkıntı fonksiyonunun belirlenmesi ve perde konturlarının oluşturularak melodinin seçilmesi. Bu yöntemde insan işitmesinin algısal olarak hassas olduğu frekansları arttırarak algılanmayanları bastırarak bir filtreleme yapılmaktadır (Eş gürültü filtresi). Bu filtre, gürültü içeren çok tiz ve bas sesleri, yani düşük ve yüksek frekans banlarındaki enstrümanları temizlerken, orta-band frekansları arttırmaktadır. Uygulanan filtreleme sonrasında sinyale KZFD uygulanarak yerel uç noktalar ve izgesel tepeler elde edilir ve tepeler yardımıyla çıkıntı fonksiyona ulaşılır. Çıkıntı fonksiyonuna ait tepe noktaları, temel melodi için f_0 adaylarını, dolayısıyla perde konturu vermektedir.

2. DEĞİŞKEN KİP AYRIŞIMI TABANLI TEMEL FREKANS KESTİRİMİ (VARIATIONAL MODE DECOMPOSITION BASED FUNDAMENTAL FREQUENCY ESTIMATION)

Çalışmamızda monofonik Türk makam müziği kayıtları içerisindeki temel titreşim frekanslarının bulunması hedeflenmiştir. Türk makam müziğinde kullanılan mikrotonal perde yapısına sahip enstrümanlar ve icracı tarafından süslemelerin sıkça kullanılması, temel titreşim frekans analizini zorlaştırmaktadır. DKA, sinyallerin genliklerinin veya frekanslarının ne kadar birbirine yakın olduğuna bakmaksızın yüksek bir doğrulukla temel frekans kestirimi yapılmasına olanak sağlar [26]. DKA yönteminde verilen giriş sinyali sonlu sayıda alt sinyallere, kiplere ayrıştırılmaktadır. GKA ve DKA tanımında kip; yerel uç noktalar ve sıfır geçiş sayıları birden farklı olan sinyal olarak tanımlanır. Son yıllarda yapılan çalışmalarda, tanım yavaş yavaş ÖKF (b_k) olarak anılmaya başlanmıştır, bunlar Eş. 1'de gösterilen genlik modülasyonlu-frekans modülasyonlu (GM-FM) sinyallerdir.

$$b_k(t) = a_k(t) \cos(\Phi_k(t)) \quad (1)$$

$\Phi_k(t)$ azalmayan bir fonksiyon olan ($\Phi'_k(t) \geq 0$) fazı ifade etmektedir, zarf $a_k(t) \geq 0$ negatif olmayan bir fonksiyondur ve $\omega_k(t) := \Phi'_k(t)$ anlık frekansı göstermektedir [1, 19].

Gerçekte, ω_k bir modun ortalama frekansı ise, onun Eş. 2 ile gösterilen pratikte band genişliği, anlık frekansın merkezden en fazla sapması Δf ve bu sapmanın oranı f_{FM} ile artar. Dolayısı ile Carson kuralına göre band genişliği aşağıdaki şekilde hesaplanır.

$$BW_{GM-FM} = 2(\Delta f + f_{FM}) \quad (2)$$

Bir kipi toplam band genişliği (Eş. 3); frekans modülasyonunda maksimum frekans sapması Δf , $a_k(t)$ zarfının band genişliği ve FM sinyalin band genişliklerinin toplamı ile ifade edilir [1].

$$BW_{GM-FM} = 2(\Delta f + f_{FM} + f_{GM}) \quad (3)$$

Müzik sinyali analizinde kullanılan diğer yöntemlerin, gürültü ile başa çıkmada yetersiz olmaları, katı band sınırları ve filtre bankasının sınırlarının önceden belirlenmesinin gerekliliği gibi birçok kısıtlamaya sahip olmaları nedeniyle dezavantajları vardır [1, 19]. DKA yönteminde ise, ilgili bandların uyarlanabilir olarak belirlenmesi ve eşzamanlı olarak uygun kiplerin kestirilmesi için değişken model kullanılmaktadır. GKA'daki ÖKF tanımına karşılık gelen dar-band özellikleri bu yöntemde de kullanılır [1, 27]. Verilen giriş sinyalinden en iyi biçimde yeniden yapılandırılan bir dizi kip (b_k) aranır. Bu kipler ayrışım sırasında kestirilen merkez frekans (ω_k) civarında band-sınırlıdır. Değişken model özellikle giriş sinyali içindeki gürültünün varlığına çok dayanıklıdır. Wiener filtre ile sıkı ilişkisi, gerçekte gürültü ile ilgili bazı avantajlara sahip olduğunu belirtir. Gerçekte bu yöntem uyarlanabilir, çoklu band olarak klasik Wiener filtrenin genelleştirilmiş halidir. Kiplere ait band genişlikleri belirlenirken, kiplere ait analitik sinyal HHD yardımıyla hesaplanarak her bir kiplin frekans izgesi, ilgili kestirilmiş merkez frekansa ayarlı bir üstel ile çarpılarak temel banda kaydırılır. Böylece demodüle sinyalin ilgili band genişliği kestirilmiş olur. Her bir kip doğrudan Fourier bölgesinde yinelemeli olarak güncellenir, kiplerin merkez frekanslarının hali hazırdaki kestirimine uygun dar band Wiener Filtre tüm diğer kiplerin artık sinyallerinin kestirimi için uygulanır. Ardından merkez frekans, kiplerin güç izgesinin ağırlık merkezi olarak yeniden kestirilir.

DKA yönteminde bilinmeyen x sinyalini geri almak ve gürültüsüz sinyali kestirmek için Tikhonov düzenleme kullanılarak en küçükleme problemi yürütülmektedir. Eş. 4'ü göz önüne alalım; gözlenen sinyal $x_0(t)$, orijinal sinyale $x(t)$ sıfır ortalamalı Gauss gürültü η eklenmesiyle oluşmuş olsun;

$$x_0 = x + \eta \quad (4)$$

Tikhonov düzenleme problemi Eş. 5'deki gibi yazılabilir.

$$\min_x \{ \|x - x_0\|_2^2 + \alpha \|\partial_t x\|_2^2 \} \quad (5)$$

Burada en küçükleme ifadesinde yer alan α , denge parametresini (düzenleme parametresi) ifade eder ve $\alpha > 0$ pozitifdir. Bu en küçük ortalama karesel hata problemini çözerek bilinmeyen x sinyalini elde edebilmek için, Eş. 6'da görülen Euler-Lagrange eşitliği kullanılır.

$$x - x_0 = \alpha \partial_t^2 x \quad (6)$$

Eş. 6, Fourier alanında çözülerek Eş. 7 elde edilir.

$$\hat{x}(\omega) = \frac{x_0}{1 + \alpha \omega^2} \quad (7)$$

Burada $\hat{x}(\omega)$, $x(t)$ sinyaline ait $j^2 = -1$ için Fourier Dönüşümünü ifade etmektedir. Bu noktada x sinyalinin bir alçak geçiren filter yardımı ile tekrar elde edilebileceği net bir biçimde görülmektedir. Orijinal DKA yönteminde ait arttırılmış Lagrange (\mathcal{L}) Eş. 8 ile ifade edilmektedir;

$$L(b_k, \omega_k, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * b_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \|x - \sum b_k\|_2^2 + \langle \lambda, x - \sum b_k \rangle \quad (8)$$

Arttırılmış Lagrange eşitliğinde; b_k , sinyalin ayrıştırıldığı k . kipi, ω_k , k . kipe ait merkez frekansını, λ ise Lagrange çarpanını ifade etmektedir. Problemin çözümü Eş. 8'in boyun noktası olarak elde edilir.

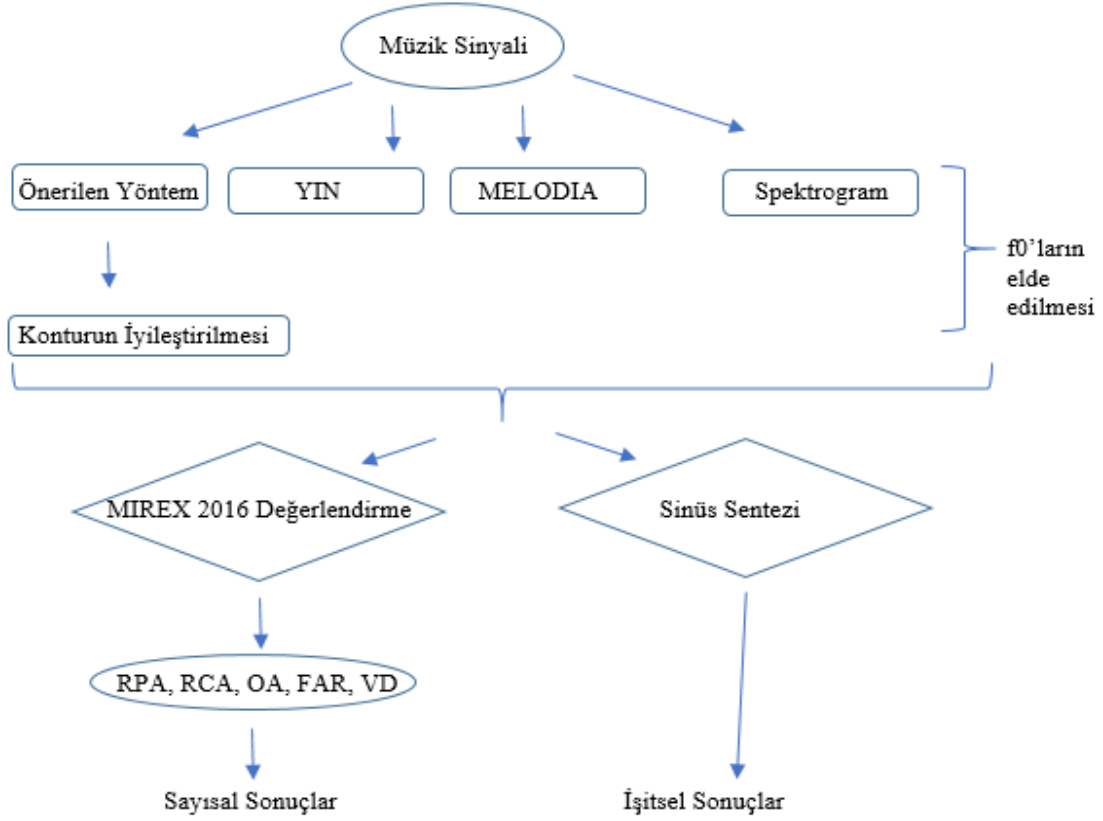
2.1. Önerilen Yöntem (The Proposed Method)

Çalışmamızda, müzik sinyallerinin kendisi yerine öz-iline fonksiyonu ile yapılan DKA tabanlı frekans kestiriminde daha başarılı sonuçlar alındığı görülmüştür [1]. Dolayısı ile önerilen yöntemde verilen monofonik Türk makam müziği sinyali $x(n)$ 'e ait öz-iline dizisi $R_{zz}(m)$, alt bileşenlerine ayrılarak frekans kestirimi yapılmıştır [26, 28].

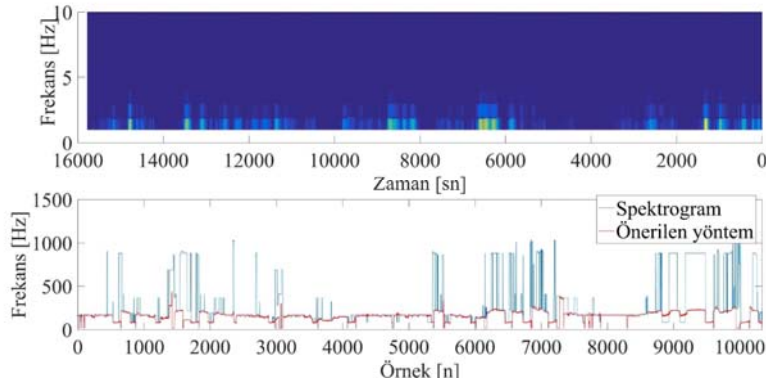
Şekil 1'de önerilen yöntemin temel adımları gösterilmektedir: müzik sinyali içerisinde temel titreşim frekansı değerleri önerilen yöntem, Spektrogram, YIN ve MELODIA yöntemleri kullanılarak elde edilir. Karşılaştırma aracı olarak MIREX2016 (Multiple Fundamental Frequency Estimation & Tracking- Çoklu Temel Titreşim Frekansı Kestirimi ve Takibi) kistasları [29] değerlendirilmekte ve yöntemler doğruluk bazında birbiri ile karşılaştırılarak sayısal sonuçlar elde edilmektedir. Sayısal sonuçların yanı sıra, yöntemlerden elde edilen temel titreşim frekans değerleri ile sinüs sentezi yapılarak işitsel sonuçlar da elde edilerek değerlendirmeye katılmaktadır.

Kullanılan yöntemlerin Matlab programında karşılaştırılabilir olması adına, MELODIA ve YIN yöntemleri Essentia [30] kütüphanesinden Matlab özütleyicileri çıkartılarak kullanılmıştır. Kütüphaneden çıkartılarak kullanılan yöntemlerin ön tanımlı çalışma protokolünde çalıştırılması sağlanmıştır. MELODIA yönteminde pencere boyutu 2048, atlama boyutu 128 olarak belirlenmiştir ve bu değer üzerinden çıkıntı fonksiyonunun hesaplanması sağlanmıştır. YIN yöntemi ise pencere boyutu 2048, atlama boyutu 256 olarak çalıştırılmıştır. Referans olarak kullanılacak Spektrogram çıktısı, yöntemlerle birebir karşılaştırma yapmayı kolaylaştırmak adına, karşılaştırılacak yöntemde kullanılan pencere boyutuna uygun bir biçimde düzenlenerek ayrı ayrı saklanır. Şekil 2'de hüzzam makamında bir parçaya ait Spektrogram görüntüsü zaman-frekans ekseninde verilmiştir. Altında ise Spektrogram tepelerinden elde edilen frekans konturu üzerine, önerilen yöntem ile elde edilen sonuçlar, n . pencere için kestirilmiş f_0 değerleri ile gösterilmektedir. Şekil 2'den de görüleceği gibi önerilen yöntem ile elde edilen frekans konturu, Spektrogram konturu ile net bir biçimde örtüşmektedir.

Gürültüsüz sinyali kestirmek için orijinal DKA yöntemi Tikhonov düzenleme olmadan faydalanırken, çalışmamızda Elastic Net regresyon [31, 32] kullanılarak problem yeniden ele alınmıştır. Önceki örnekte olduğu gibi Eş. 4'ten yola çıkılarak, en küçükleme problemi $\ell-1$ ve $\ell-2$ normlarının birlikte kullanılmasıyla Eş. 9'daki gibi yazılabilir.



Şekil 1. Önerilen yöntemin temel adımları (The basis steps of proposed method)



Şekil 2. Zaman-frekans izgesi ve Spektrogram konturu üzerine önerilen yöntemin sonucu (Time-frequency spectrum and the result of the proposed method in Spectrogram contour)

$$\min_x \{ \|x - x_0\|_2^2 + \alpha \|\partial_t x\|_2^2 + \beta \|\partial_f x\|_1 \} \quad (9)$$

Çözüm Eş. 10'daki gibi Fourier alanında gerçekleştirilir.

$$\hat{b}_k^{i+1} = \arg \min_{\hat{b}_k, \hat{b}_k = \hat{b}_k} \left\{ \alpha \left\| \int j\omega [(1 + \text{sgn}(\omega + \omega_k)) \hat{b}_k(\omega + \omega_k)] \right\|_2^2 + \beta \left\| \int j\omega [(1 + \text{sgn}(\omega + \omega_k)) \hat{b}_k(\omega + \omega_k)] \right\|_1 + \left\| \hat{x} - \sum \hat{b}_n + \frac{\hat{\lambda}}{2} \right\|_2^2 \right\} \quad (10)$$

ω değişkenini $\omega + \omega_k$ ile değiştirdiğimizde Eş. 11 elde edilmektedir.

$$\hat{b}_k^{i+1} = \arg \min_{\hat{b}_k, \hat{b}_k = \hat{b}_k} \left\{ \alpha \left\| \int j(\omega - \omega_k) [(1 + \text{sgn}(\omega)) \hat{b}_k(\omega)] \right\|_2^2 + \beta \left\| \int j(\omega - \omega_k) [(1 + \text{sgn}(\omega)) \hat{b}_k(\omega)] \right\|_1 + \left\| \hat{x} - \sum \hat{b}_n + \frac{\hat{\lambda}}{2} \right\|_2^2 \right\} \quad (11)$$

Problem Eş. 12 ve Eş 13'deki gibi yürütülür.

$$\hat{b}_k^{i+1} = \arg \min_{\hat{b}_k, \hat{b}_k = \hat{b}_k} \left\{ \int_0^\infty \alpha |j|^2 (\omega - \omega_k)^2 |2|^2 |\hat{b}_k(\omega)|^2 d\omega + \int_0^\infty \beta |j| (\omega - \omega_k) |2| |\hat{b}_k(\omega)| d\omega + 2 \int_0^\infty \left(\hat{x} - \sum \hat{b}_n + \frac{\hat{\lambda}}{2} \right)^2 d\omega \right\} \quad (12)$$

$$\hat{b}_k^{i+1} = 4\alpha(\omega - \omega_k)^2 |\hat{b}_k(\omega)|^2 + 2\beta(\omega - \omega_k) \hat{b}_k(\omega) + 2 \left(\hat{x} - \sum_{n \neq k} \hat{b}_n + \frac{\hat{\lambda}}{2} \right)^2 \quad (13)$$

Elde edilen ifadenin türevi sıfıra eşitlenerek (Eş. 14), en küçükleme problemin çözümü (Eş. 15, Eş. 16 ve Eş. 17) yürütülür ve sonuç Eş. 18’deki gibi elde edilir.

$$8\alpha(\omega - \omega_k)^2 \hat{b}_k(\omega) \hat{b}_k'(\omega) + 2\beta(\omega - \omega_k) \hat{b}_k'(\omega) = 4 \left(\hat{x} - \sum_{n \neq k} \hat{b}_n + \frac{\hat{\lambda}}{2} \right) \hat{b}_k'(\omega) \quad (14)$$

$$4\alpha(\omega - \omega_k)^2 \hat{b}_k(\omega) + \beta(\omega - \omega_k) = 2 \left(\hat{x} - \sum_{n \neq k} \hat{b}_n + \frac{\hat{\lambda}}{2} \right) - 2\hat{b}_k(\omega) \quad (15)$$

$$4\alpha(\omega - \omega_k)^2 \hat{b}_k(\omega) + 2\hat{b}_k(\omega) = 2 \left(\hat{x} - \sum_{n \neq k} \hat{b}_n + \frac{\hat{\lambda}}{2} \right) - \beta(\omega - \omega_k) \quad (16)$$

$$\hat{b}_k(\omega) [4\alpha(\omega - \omega_k)^2 + 2] = 2 \left(\hat{x} - \sum_{n \neq k} \hat{b}_n + \frac{\hat{\lambda}}{2} \right) - \beta(\omega - \omega_k) \quad (17)$$

$$\hat{b}_k^{i+1}(\omega) = \left[2 \left(\hat{x} - \sum_{n \neq k} \hat{b}_n + \frac{\hat{\lambda}}{2} \right) - \beta(\omega - \omega_k) \right] \frac{1}{4\alpha(\omega - \omega_k)^2 + 2} \quad (18)$$

Elastic Net regresyon; doğrusal ve sırt (Lasso ve Ridge) yöntemlerinin $\ell-1$ ve $\ell-2$ ceza terimlerinin doğrusal birleşiminden oluşmaktadır. Birbiri ile ilintili birçok özellik olduğu durumda kullanışlı bir yöntemdir. Bu yöntem, $\ell-1$ ceza terimine sahip Lasso yönteminin sınırlamalarının üstesinden gelebilmek için ikinci dereceden ceza terimini kullanmaktadır. İkinci dereceden ceza terimi, maliyet fonksiyonunu kesinlikle konveks hale getirmekte ve böylece tek bir minimuma sahip olması sağlanmaktadır. Daha önce de bahsettiğimiz gibi HHD yardımı ile kiplere ait analitik sinyal elde edilir ve bu analitik sinyal kullanılarak ilgili kipe ait band genişliği kestirilir. Kiplerin frekans izgesi, kestirilmiş merkez frekansına ayarlı bir üstel ile çarpılarak temel banda kaydırılır ve band genişliği elde edilir. HHD iki kısımdan oluşmaktadır; ilki eleme işlemi yapılarak verilen giriş sinyalinden ÖKF’lerin elde edildiği GKA yöntemi olup, ikinci kısım ise genlik, anlık faz ve anlık frekans değerlerinin elde edilmesini sağlayan HHD yöntemini içeren kısımdır. HHD yönteminde ilk olarak GKA yöntemi kullanılarak, giriş sinyali farklı frekanslardaki ÖKF bileşenlerine ayrıştırılır. Bu şekilde giriş sinyali artık, elde edilen ÖKF bileşenlerinin ve artık sinyalin toplanması ile Eş. 19’daki gibi elde edilir.

$$x(t) = r_n + \sum_{k=1}^n b_k \quad (19)$$

Burada; b_k ’lar k ’ncinci ölçekteki ÖKF’yi ifade ederken, r_n ise yinelemeler sonunda geriye kalan artık sinyali işaret eder. Genlik ve anlık frekans bilgilerine ulaşabilmek için, elde edilen ÖKF’lere HHD (9f) uygulanarak analitik sinyal elde

edilir. HHD’nin en sık kullanıldığı problem tamamen reel sinyalden bir analitik sinyalin oluşturulmasıdır. ÖKF’lere HHD uygulandıktan sonra Eş. 20’deki veri seti elde edilir.

$$\mathcal{H}x_k(t) = \frac{1}{\pi} p.v \int_{-\infty}^{\infty} \frac{b_k(v)}{t-v} dv \quad (20)$$

Analitik sinyal Eş. 21’deki gibi oluşturulur.

$$c_k(t) = b_k(t) + j\mathcal{H}x_k(t) = a_k(t) e^{j\phi_k(t)} \quad (21)$$

Burada, $\phi_k(t)$ fazı ve $a_k(t)$ gerçek zarf tarafından yönetilen genliği ifade etmektedir. Analitik sinyalin anlık genlik a_k ve anlık frekansı ω_k Eş. 22 ve Eş 23’deki gibi hesaplanır.

$$a_k(t) = \sqrt{b_k^2(t) + \mathcal{H}x_k^2(t)} \quad (22)$$

$$\omega_k(t) = \frac{d\phi_k(t)}{dt} \quad (23)$$

Zorlanmış problemin çözümü Eş. 24’deki gibi elde edilir.

$$\min_{b_k, \omega_k} \left\{ \begin{array}{l} \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * b_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \\ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * b_k(t) \right] e^{-j\omega_k t} \right\|_1 \end{array} \right\} \quad (24)$$

Giriş sinyali x , b_k ’ların toplamı olarak Eş. 25’deki gibi ifade edilir.

$$\sum b_k = x \quad (25)$$

DKA yönteminde yeniden yapılandırma karşılaşılan zorluklar, ikinci dereceden ceza terimi ve Lagrange çarpanı (λ) kullanılarak ortadan kaldırılmaktadır. Böylece, genişletilmiş Lagrange fonksiyonu Eş. 26 ile ifade edilir.

$$L(b_k, \omega_k, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * b_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \beta \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * b_k(t) \right] e^{-j\omega_k t} \right\|_1 + \|x - \sum b_k\|_2^2 + \langle \lambda, x - \sum b_k \rangle \quad (26)$$

Bu en küçükleme probleminin çözümü, Çarpanların Değişimli Yön Yöntemi (ÇDYM: Alternate Direction Method of Multipliers- ADMM) kullanılarak genişletilmiş Lagrange fonksiyonunun boyun noktası olarak elde edilir.

Sinyal kiplerinin merkez frekansları ω_k üzerinden en küçükleme problemi Eş. 27 ve Eş. 28’deki gibi yürütülür.

$$\omega_k^{i+1} = \operatorname{argmin}_{\omega_k} \left\{ \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * b_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (27)$$

$$\omega_k^{i+1} = \operatorname{argmin}_{\omega_k} \left\{ \int_0^\infty (\omega - \omega_k)^2 |\hat{b}_k(\omega)|^2 d\omega \right\} \quad (28)$$

İkinci dereceden denklem Eş. 29’daki gibi çözülür.

$$\omega_k^{i+1} = \frac{\int_0^\infty \omega |b_k(\omega)|^2 d\omega}{\int_0^\infty |b_k(\omega)|^2 d\omega} \quad (29)$$

Burada Lagrange çarpanının rolü gürültünün varlığına karşı sağlamlığın artırılmasıdır. İkinci dereceden hata terimi; yakınsamayı iyileştirirken, kısıtlamaları zorunlu tutar. Ayırıştırılmaya dâhil olmaması gereken güçlü gürültülerin varlığı durumunda, uygun seçim için Lagrange çarpanı düşürülürken yalnızca ikinci dereceden ceza terimi kullanılır. Düşük gürültü seviyesinde (gürültünün standart sapması ($\sigma=0.005$), tamamen rasgele bulunan merkez frekansına rağmen Lagrange çarpanı %100 başarı için yardımcı olur, bunun için güncelleme parametresi $\tau>0$ seçilir. Yüksek gürültü seviyelerinde Lagrange çarpanına ait güncelleme parametresi $\tau=0$ yapıldığında, gürültü $\sigma=10^{-3}$ düzeyinde iken başarı seviyesi 1 üzerinden 0.7-0.8 aralığındadır. Gürültünün artması ile ($\sigma=10^{-1}$) başarı seviyesi maksimum değeri olan 0.9'lara ulaşmakta, daha sonra gürültünün artışı ile birlikte hızla düşmektedir. Yüksek gürültüye sahip müzik kayıtları için bir diğer senaryo ise, Lagrange çarpanına ait güncelleme parametresinin ($\tau>0$) işleme katılmasıdır. Bu durumda düşük gürültü düzeyinde (0.001-0.5) başarı seviyesi maksimumda yani 1'dir. Artan gürültü ile birlikte başarı seviyesi hızla düşmektedir.

Önerilen yöntemde uygulamadan önce belirlenmesi gereken bir diğer giriş parametre ise data-uygunluk sınırlamasının denge parametresidir (α ve β). Denge parametresi band genişliği ile ters orantılı bir parametredir. Gürültünün varlığına bağlı olarak büyük seçildiğinde ($\alpha=1000$) band genişliğini azaltacak ve doğru merkez frekansının seçilmesini güçleştirecektir. Bunun tam tersi olan küçük denge parametresi seçimi ($\alpha=100$) durumunda ise band genişliğinin artması, dolayısıyla da kipi daha fazla gürültü içermesi anlamına gelmektedir. Her iki denge parametresinin eşit seçilmesi ($\alpha=\beta=1000$), en küçükleme probleminde ikinci parametrenin katkısı görünmez hale getirir. Dolayısıyla çalışmamız, $\alpha=1000$ ve $\beta=100000$ seçilerek yürütülmüştür.

Yöntemde belirtilen diğer bir parametre ise ω_k merkez frekansını başlangıç durumuna getirmek için kullanılan $init=0$ 'dır.

Yakınsama kistasının toleransını belirten ϵ parametresinin azalması, sinyalin frekansından bağımsız olarak hatayı da azaltmaktadır. GKA yönteminde ise hata, ayırıştırılacak sinyalin frekansı ile üstel bir biçimde artmaktadır. K parametresi, giriş sinyalinden elde edilecek kiplerin (b_k) sayısı olarak ifade edilir.

2.1.1. Önerilen yönteme basit bir örnek (Simple example for proposed method)

Önerilen yöntem, sinyal içerisindeki frekansların birbirine ne kadar yakın olduğuna bakmaksızın sinyalleri başarı ile ayırıştırabilmektedir. Müzik sinyalleri genellikle tek bir frekans yerine bu frekansa ait harmonikleri de içermektedir. Bu nedenle örnek ayırıştırımlar harmonikli sentetik ve gerçek müzik sinyalleri üzerinde gerçekleştirilmiştir.

Sentetik sinyal; Ana bileşeni 20Hz ve ikinci harmoniği 40Hz olan bir sentetik sinyalin önerilen yöntem ile kiplerine ayırıştırılması, genlik-zaman ekseninde, Şekil 3 ve 4'te gösterilmektedir.

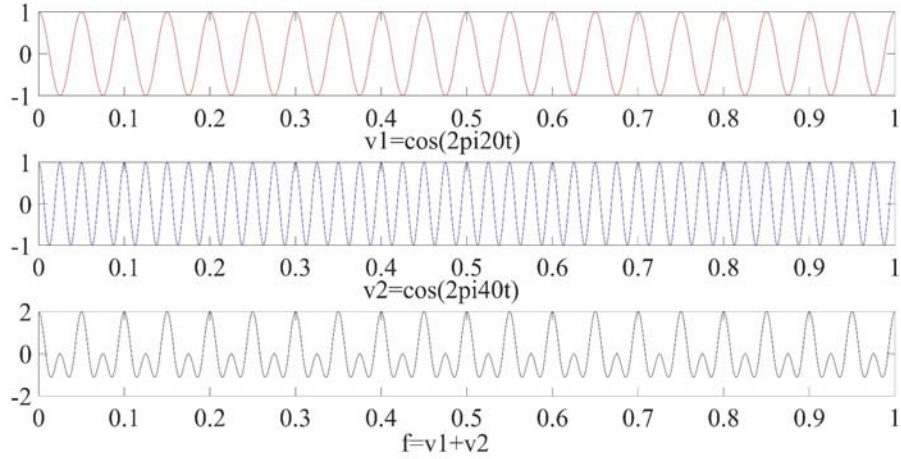
Geçek müzik sinyali; Piyano tarafından 2s boyunca çalınarak kayıt edilmiş, harmonik içeren bir notanın önerilen yöntem kullanılarak kiplerine ayırıştırılması, n; pencere sayısını ifade etmek üzere, Şekil 5'te gösterilmektedir. Notanın temel titreşim frekansı 516Hz olmakla birlikte Spektrogram görüntüsü ve yöntem sonucundan da görüldüğü gibi 1055Hz ve 1593Hz frekanslarında harmoniklere sahiptir. Önerilen yöntem sonucu ve Spektrogram görüntüsü net bir biçimde örtüşmektedir.

2.2. Perde Konturunun İyileştirilmesi (Enhancement of Pitch Contour)

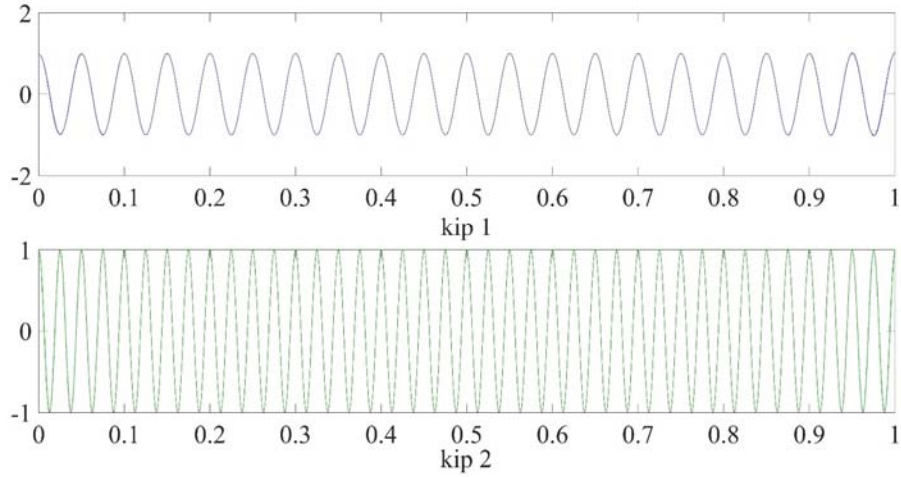
Temel titreşim frekansı kestirim yöntemleri bazı durumlarda sonuç olarak hatalı frekans değerleri verebilmektedir. Bu hatalı çıkışların bir kısmının yöntem uygulandıktan sonra filtreleme ile düzeltilmesi mümkündür. Filtre kullanarak, frekans değerlerinde devamlılığın yitirildiği noktalarda oluşan oktav hataları ve hatalı değerlerin düzeltilmesi sağlanmış olur [26, 33]. Sonradan iyileştirme olarak kullanılan bu filtre ile süreklilik arz eden bir perde konturu elde edebilmek için; frekans paketlerinin boyutu beşten büyük seçilir, kısa süreli frekans paketleri, tam sayı katlarındaki daha uzun süreli paketlerin başına veya sonuna eklenir, melodik dinamik aralık ise 4 olarak kabul edilir. Şekil 6'da örnek bir filtreleme gösterilmiştir. Önerilen yöntem tarafından elde edilmiş olan mavi düz çizgi ile gösterilmiş olan perde kontur, filtrelenerek düzenlenmiş, oktav hataları veya hatalı değerlerin olduğu noktalar sıfır frekansa atanmış ve filtrelenmiş kontur kırmızı yıldız ile gösterilmiştir.

3. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

Önerilen yöntem ile Türk makam müziği eserlerinin temel titreşim frekansı kestiriminde, üflemler, vurmalar ve yaylı enstrümanlar kullanılarak icra edilen çeşitli makamlardan 31 kayıt kullanılmıştır. Kullanılan eserler canlı performanslardan elde edilmiş tek kanallı kayıtlardan oluşmaktadır. Örneğin; Hüzzam makamında 3.17dk'lık bir parça için; çello, kanun, keman, kemençe, ney, tambur, ud ve ritim enstrümanları ayrı kanallardan kayıt edilmiş ve her bir enstrüman bir monofonik kayıt oluşturmuştur. Önerilen yöntem ile elde edilen monofonik kayıtlar 30s'lik parçalara ayrılarak alt setler oluşturulmuştur. Alt setler üzerinde önerilen yöntem uygulanarak, sonuçlar MIREX2016 kistasları kullanılarak değerlendirilmiştir. DKA yönteminde olduğu gibi önerilen yöntem için de sinyalin ayırıştırılacağı kip sayısının önceden belirlenmesi gerekmektedir, bu çalışmada kip sayısı 7 olarak belirlenmiş ve $init=0$ seçilerek merkez frekans sıfırdan başlatılmıştır. α ve β parametreleri ise sırasıyla 100,000 ve 1000 değerinde tutulmuştur, böylelikle band genişliği azaltılmış ve merkez frekans



Şekil 3. 20Hz ve 40Hz frekanslı sinüsoidal sinyaller ve toplamı (Sinusoids with 20Hz and 40Hz frequencies and sum of them)



Şekil 4. Önerilen yöntem sonucu elde edilen kipler (Modes obtained by the proposed method)

değerinin doğru seçilme olasılığı arttırılmıştır. $\tau=0$ seçilerek gürültünün varlığında başarı değerini arttırmak hedeflenmiştir. Tolerans değeri 1.10^{-7} seçilmiştir. Karşılaştırmada kullanılacak yöntemler ve devam eden çalışmalar dikkate alınarak önerdiğimiz yöntem için 46ms'lik pencere ve 2.9ms'lik atlama boyutu kullanılmıştır.

Değerlendirme; ses belirleme, yanlış ses alarmı, ham perde doğruluğu, ham renk doğruluğu ve tam doğruluk kıstasları üzerine yapılmaktadır. Önerdiğimiz yöntemde ise ses ile ilgili bir ayırım yapılmamakta, kullanılan eserler söz içermektedir. Doğru titreşim frekansları önceden bilinmeyen kayıtlar, Spektrogramdan elde edilen değerler ile karşılaştırılmıştır.

Ses Belirleme (SB): Kayıt içerisinde ses olan bölgelerde doğru veya yanlış olarak yakaladığı f_0 örneklerinin toplamının sessiz bölge dışında kalan örnek sayısına oranı ile ifade edilir [26, 34]. Yanlış Ses Alarmı (YSA): Referans

kayıt içerisindeki titreşimsiz bölgelerin titreşimli olarak işaretlenmiş olmasının ölçüsünü vermektedir.

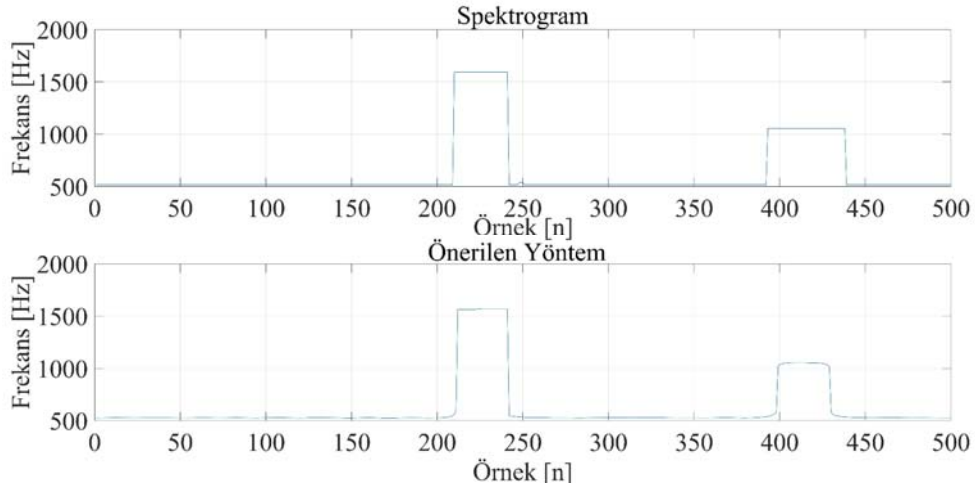
Ham Perde Doğruluğu (HPD): Algoritma tarafından kestirilen f_0 değeri, referans olan f_0 değerinin $\pm 1/4$ ton ($\pm 3\%$) içerisinde ise, referans içindeki titreşimli örneklerin miktarı olarak tanımlanır. Algoritma titreşimin olmadığı yerlerde de bir f_0 kestirebilir, bu nedenle HPD doğru olmayan titreşim kestirimlerinden etkilenmez.

Ham Renk Doğruluğu (HRD): HPD ile aynı anlamdadır, tek fark, referans ve kestirilen f_0 değerleri tek bir oktava karşılık gelmelidir. Yani oktav hatası ihmal edilir.

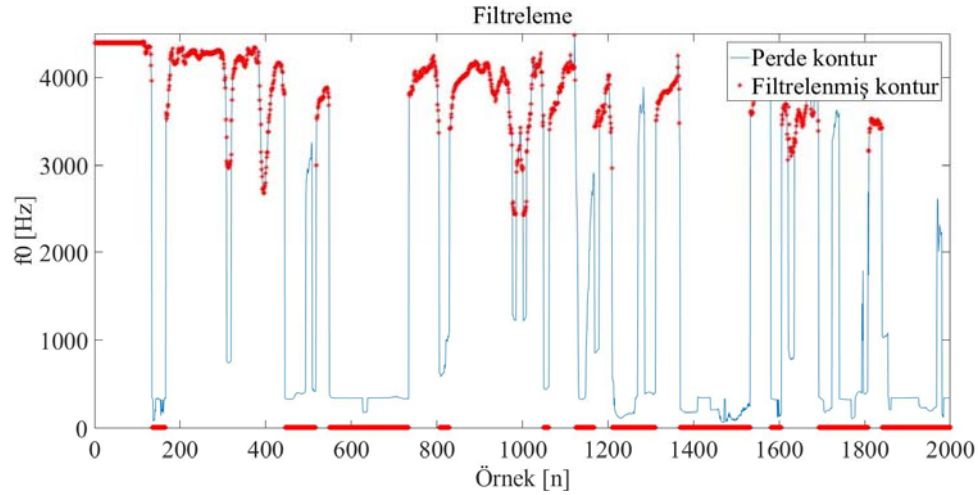
Tam Doğruluk (TD): Bu ölçüm, sistemin tüm performansını puanlandırmak adına perde titreşim kestirimi ve ses belirleme performanslarını birleştirir. Algoritma tarafından doğru biçimde kestirilen, yani melodi içerisindeki titreşimli bölgeleri doğru f_0 değerinde, titreşimsiz bölgeleri de

titreşimsiz olarak etiketleyen örnek miktarını ifade eder. Tablo 1’de MIREX kıstasları kullanılarak önerdiğimiz yöntem, orijinal DKA, MELODIA ve YIN yöntemleri ile yüz üzerinden puanlandırılarak karşılaştırılmıştır. Tablodan da görüldüğü gibi, ses belirleme kıstasında önerilen yöntemin başarısının düşük olmasına karşın, HPD, HRD ve YSA kıstaslarında en yüksek başarıya sahiptir. Bu da bize, önerdiğimiz yöntemin daha az hata ile (düşük YSA), daha yüksek doğrulukta (yüksek HPD) ve daha az oktav hatası yaparak (yüksek HRD) temel titreşim frekanslarını

kestirdiğini ifade etmektedir. YIN yönteminin SB kıstasında yakaladığı başarı (%99.56323), YSA kıstasından da görüldüğü gibi yüksek oranda yanlış olarak (%92.11158) kestirilmiş seslerden oluşmaktadır. Bu nedenle MIREX kıstaslarını tek başına değerlendirmek yerine birlikte gözlemlemek, sonucu daha sağlıklı görmemize yardımcı olacaktır. Tam doğruluk olarak tanımlanan büyüklük ise, yöntem tarafından sessiz bölgelerin de, sesli olan bölgeler gibi doğru etiketlenerek bulunmasının ölçüsünü vermektedir. Şekil 7’de ise yukarıdaki tablonun bir özeti,



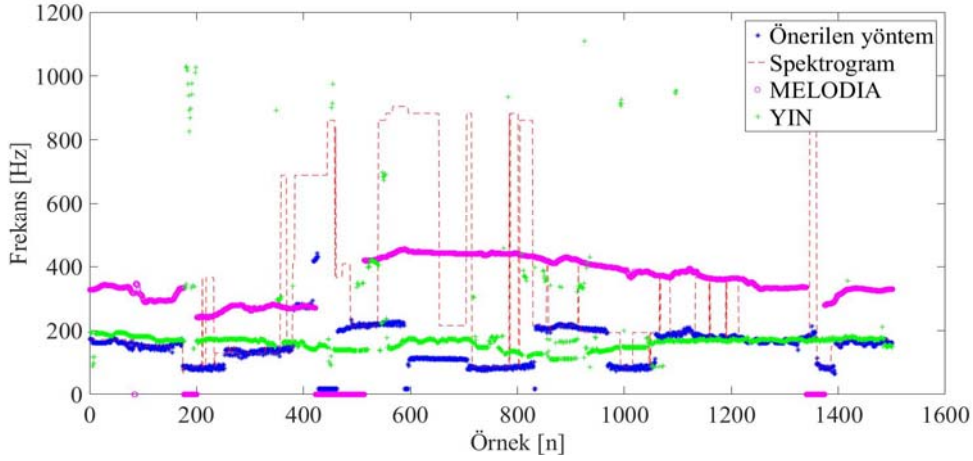
Şekil 5. Spektrogram ve önerilen yöntem sonucu elde edilen frekans değerleri
(Frequencies obtained by spectrogram and proposed method)



Şekil 6. Temel titreşim frekans konturunun iyileştirilmesi (Enhancement of fundamental frequency contour)

Tablo 1. Önerilen yöntemin Orijinal DKA, YIN ve MELODIA ile karşılaştırılması
(Comparison of the proposed method with original VMD, YIN methods)

Algoritma	Orijinal DKA	Önerilen Yöntem	MELODIA	YIN
Ses Belirleme	64,37331	59,21458	74,58969	99,56323
Yanlış Ses Alarmı	48,32537	27,05325	34,67062	92,11158
Ham Perde Doğruluğu	53,25328	61,12405	57,26712	39,45762
Ham Renk Doğruluğu	38,14621	42,42055	42,40832	23,32288
Tam Doğruluk	18,48432	22,08341	33,12235	38,34012



Şekil 7. Temel titreşim frekans konturları (Fundamental frequency contours)

şekil olarak verilmiştir. Hüzzam makamına ait heterofonik bir kayıttın analizi yapılarak, YIN ve MELODIA algoritmaları ile karşılaştırmalı olarak gösterilmiştir. Şekil 7’de, önerilen yöntemin Spektrogramı büyük bir doğrulukla takip ettiği görülmektedir. Farklı oktavlara çok fazla kaymadan, uygun bir oktavda ilerleyerek ve kesintisiz olarak konturun devam etmesi, yöntemden sonra uygulanan filtre yardımıyla sağlanmaktadır. MELODIA algoritması ise bir üst perdeden temel titreşim frekans konturunu elde etmektedir. Bu nedenle HRD oktav hataları nedeniyle düşmektedir. YIN yöntemi ile elde edilmiş olan konturun ise yüksek frekans değerlerine sıçramalar yaptığı ve kontur düzenini aksattığı şekilden görülmektedir.

4. SİMGELER (SYMBOLS)

a_k	: ÖKF'nin zarfı
$argmin$: Argüman minimizasyon
b_k	: Özgül kip fonksiyonu
BW_{GM-FM}	: Band genişliği
c_k	: Analitik sinyal
f_{FM}	: FM sinyalin band genişliği
f_{GM}	: Zarfın band genişliği
f_0	: Temel titreşim frekansı
\mathcal{H}	: HHD
Hz	: Hertz
$init$: Başlangıç durumuna getirme parametresi
\mathcal{L}	: Lagrange simgesi
min	: Minimizasyon
r_n	: Artık sinyal
R_{zz}	: Öz ilinti fonksiyonu
sgn	: Signum fonksiyonu
x	: Orjinal sinyal
x_0	: Gözlenen sinyal
$\hat{x}(\omega)$: x sinyalinin fourier dönüşümü
ω_k	: merkez frekans
α ve β	: Denge parametreleri
ϕ_k	: ÖKF'nin fazı
Δf	: Maksimum frekans sapması
∂_t	: Zamana göre türev

η	: Sıfır ortalamalı gauss gürültü
λ	: Lagrange çarpanı
δ	: Birim dürtü fonksiyonu
σ	: Standart sapma
τ	: Güncelleme parametresi

5. SONUÇLAR (CONCLUSIONS)

Bu çalışmada monofonik Türk makam müziği eserlerinde temel frekans kestirimi için DKA tabanlı bir yöntem önerilmiştir. Önerilen yöntem ile elde edilen kiplere ait frekans değerleri eser içerisindeki temel frekanslar ile iyi bir biçimde örtüşmektedir. Önerilen yöntem müzik sinyali analizi için sıkça kullanılan YIN ve MELODIA algoritmaları ile karşılaştırmalı olarak değerlendirilmiştir. Değerlendirmeler MIREX2016 kıstaslarına uygun olarak gerçekleştirilmiştir. Sonuçlar, önerilen yöntem ile monofonik Türk makam müziği kayıtlarında temel frekans değerlerinin başarı ile elde edildiğini göstermiştir. Ses belirleme kıstasında önerilen yöntemin başarısı %64 civarında iken YIN %99 ses belirleme yeteneğine sahiptir. Önerilen yöntemde belirlenen ses düzeyi YIN'e göre düşük olmasına karşın bu ses dizisi içerisinde %53 başarı ile doğru perde frekansını yakalayabilmektedir. Toplam doğrulukta başarının düşmesinin sebeplerinden biri de; sonradan-işleme sessizlik olarak işaretlenmesidir, bu aynı zamanda SB kıstasının başarısını da azaltmaktadır. Fakat uygulanan sonradan-işleme sonucunda perde konturunun sürekliliği ve oktav hatalarının ortadan kaldırılması sağlanmış olur.

Bu çalışma, otomatik notaya dökme sistemi içerisinde önemli bir adımı oluşturmaktadır. Bu anlamda önerilen yöntem Türk makam müziği için bilgi erişim alanında; melodik analiz, ritim analizi, çoklu frekans kestirimi gibi birçok çalışma için önemli bir adım olacaktır [35]. Bundan sonra yapılacak çalışmalarda, heterofonik Türk makam müziği eserlerinin önerilen yöntem ile çoklu temel frekans değerlerinin kestirilmesi incelenecektir. Bu çalışma, makam karakterini taşıyan birçok müzik türü için bir adım oluşturmaktadır.

TEŞEKKÜR (ACKNOWLEDGEMENT)

Çalışmamızda, verilerin sağlanması, incelenmesi ve uygulanması konusunda desteklerini hiçbir zaman esirgemeyen sayın Doç. Dr. Barış Bozkurt'a teşekkür ederiz.

KAYNAKLAR (REFERENCES)

- Ozturk Simsek B., Akan A., Bozkurt B., Fundamental frequency estimation for monophonical Turkish music by using VMD, 23th Signal Processing and Communications Applications Conference (SIU), Malatya, 1022-1025,16-19 May 2015.
- Gedik A.C., Bozkurt B., Evaluation of the Makam Scale Theory of Arel for Music Information Retrieval on Traditional Turkish Art Music, Journal of New Music Research, 38 (2), 103-116, 2009.
- Klapuri A.P., Multiple Fundamental Frequency Estimation Based on Harmonicity and Spectral Smoothness, IEEE Transactions on Speech and Audio Processing, 11 (6), 804-816, 2003.
- Scheirer E.D., Tempo and Beat Analysis of Acoustic Music Signals, J. Acoust. Soc. Amer, 103 (1), 588-601, 1998.
- Benetos E., Holzapfel A., Automatic Transcription of Turkish Makam Music, Proceedings of ISMIR - International Conference on Music Information Retrieval, Curitiba, Brazil, November. 4-8th, 2013.
- Bozkurt B., Gedik A.C., Karaosmanoglu M.K., Türk Müziği için Müzik Bilgi Erisimi: Problemleri Çözüm Önerileri ve Araçlar, 17th Signal Processing and Communications Applications Conference (SIU), Antalya, 804-807, 09-11 Nisan 2009.
- Yarman O., 79-tone Tuning & Theory for Turkish Maqam Music, PhD Thesis, Istanbul Technical University, Social Sciences Inst., Istanbul, 2008.
- Ergüner S., Ney "metod", Ergüner Müzik, istanbul, 2007.
- Bozkurt B., Gedik A. C., Karaosmanoğlu M. K., Klasik Türk Müziği için Otomatik Notaya Dökme Sistemi, IEEE 19th Signal Processing and Communications Applications Conference (SIU), Antalya, 17-20, 20-22 Nisan 2011.
- Abidin D., Öztürk Ö., Özacar Öztürk T., Using Data Mining for Makam Recognition in Turkish Traditional Art Music, Journal of the Faculty of Engineering and Architecture of Gazi University, 32 (4), 1221-1232, 2017.
- Klapuri A., Automatic transcription of music, MSc Thesis, Tampere University of Technology, 1998.
- Martin K. D., Sound-Source Recognition: A Theory and Computational Model, Phd Thessis, Massachusetts Institute of Technology, Doctor of Philosophy in Electrical Engineering and Computer Science, Cambridge, 1999.
- Salamon J., Melody Extraction from Polyphonic Music Signals. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2013.
- Mauch M., Cannam C., Bittner R., Fazekas G., Salamon J., Dai J., Bello J.,and Dixon S., Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency, Proceedings of the First International Conference on Technologies for Music Notation and Representation, 2015.
- Cannam C., Landone C., Sandler M.,and Bello J., The Sonic Visualiser: A Visualisation Platform For Semantic Descriptors From Musical Signals. Proceedings of ISMIR, Victoria, Canada, 324-327, 2006.
- Klapuri A., Automatic Music Transcription as We Know it Today, Journal of New Music Research, 33 (3), 269-282, 2004.
- Cohen L., Time-frequency analysis, Editor: Englewood Cliffs, NJ: Prentice Hall PTR, A.B.D, 778, 1995.
- Huang N. E., Samuel S. S., Hilbert-Huang transform and its applications, World Scientific, A.B.D, 5, 2005.
- Dragomiretskiy K., Dominique Z., Variational Mode Decomposition, IEEE Transactions on Signal Processing, 62 (3), 531-544, 2014.
- Gilles J., Empirical wavelet transform, Signal Processing, IEEE Transactions on, 61 (16), 3999-4010, 2013.
- Aydın F., Aslan Z., Diagnosis of neuro degenerative diseases using machine learning methods and wavelet, Journal of the Faculty of Engineering and Architecture of Gazi University, 32 (3), 749-766, 2017.
- Albostan A., Önöz B., Wavelet application approach on the chaotic analysis of daily river discharge, Journal of the Faculty of Engineering and Architecture of Gazi University, 30 (1), 39-48, 2015.
- Zhou R., Marco M., A new time-frequency representation for music signal analysis: Resonator time-frequency image, Signal Processing and Its Applications, . ISSPA 2007. 9th International Symposium on. IEEE, Sharjah, 1-4, 12-17 February 2007.
- Zhou R., Mattavelli M., Zoia G., Music onset detection based on resonator time frequency image, IEEE Transactions on Audio, Speech, and Language Processing 16 (8), 1685-1695, 8 November 2008.
- Salamon J., Emilia G., Melody extraction from polyphonic music signals using pitch contour characteristics, Audio, Speech, and Language Processing, IEEE Transactions, 20 (6), 1759-1770, 2012.
- Ozturk. Simsek B., Bozkurt B., Akan A., Fundamental Frequency Estimation for Heterophonical Turkish Music by using VMD, 2016 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, 1625-1628, 2016.
- Mert A., Akan A., Eğimden Arındırılmış Dalgalanma Analizi İle Görgül Kip Ayırışımı Tabanlı Eeg Analizi, Tıp Teknolojileri Ulusal Kongresi, Kapadokya, 187-190, 25-27 Eylül 2014.
- Stanek M., Smatana T., Comparison of Fundamental Frequency Detection Methods and Introducing Simple Self-Repairing Algorithm for Musical Applications, 25th International Conference Radioelektronika

- (RADIOELEKTRONIKA), Pardubice, 217-221, 21-22 April 2015.
29. MIREX (Music, Information Retrieval Exchange), Multiple Fundamental Frequency Estimation & Tracking, http://www.music-ir.org/mirex/wiki/2016:Multiple_Fundamental_Frequency_Estimation_%26_Tracking, Yayın tarihi 17 February 2016, Erişim tarihi 30 March 2017.
 30. Essentia, Open-source C++ library, <http://essentia.upf.edu/>, Erişim tarihi February 2017.
 31. Soomro B.N., Xiao L., Soomro S. H., Molaie M., Bilayer Elastic Net Regression Model for Supervised Spectral-Spatial Hyperspectral Image Classification, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9 (9), 4102-4116, 2016.
 32. Zhang Z., Lai Z., Xu Y., Discriminative Elastic-Net Regularized Linear Regression, IEEE Transactions on Image Processing, 26 (3), 1466-1481, 2017.
 33. Bozkurt B., An Automatic Pitch Analysis Method for Turkish Maqam Music, Journal of New Music Research, 37 (1), 1-13. 2008.
 34. Poliner G.E., Ellis D.P.W., Ehmann A.F., Gomez E., Streich S., Ong B., Melody Transcription From Music Audio: Approaches and Evaluation, IEEE Transactions on Audio, Speech, and Language Processing, 15 (4), 1247-1256, May 2007.
 35. Bozkurt B., Gedik A.C., Karaosmanoğlu M.K., An automatic transcription system for Turkish music, Signal Processing and Communications Applications (SIU), Antalya, 17-20, 20-22 Nisan 2011.

