

Kanonik Korelasyon Analiz Tabanlı Ses Ayırıştırma Algoritmalarının İşlem Süresi Azaltımı

Hüsamettin ÇELİK^{*1}, Fatih Yavuz ILGIN², Yusuf SEVİM³

¹ Erzincan Binali Yıldırım Üniversitesi, Tercan Meslek Yüksekokulu, Bilgisayar Programcılığı Bölümü, 24800, Erzincan

² Erzincan Binali Yıldırım Üniversitesi, Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, 24100, Erzincan

³ Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik -Elektronik Bölümü, 61080, Trabzon

(Alınış / Received: 05.06.2018, Kabul / Accepted: 23.05.2019)

Anahtar Kelimeler

Bağımsız bileşen analizi,
Kanonik korelasyon analizi,
Çekidek kanonik korelasyon analizi,
Ayrık dalgacık dönüşümü

Özet: Gürültülü ortamlarda ses işaretinin gürültüsüz olarak elde edilmesi oldukça önemlidir. Karışım ortamı ve gürültü kaynakları bilinemediğinden dolayı ayırıştırma işlemi oldukça zordur. En az iki sinyalin karışımını içeren bir veri kümesinden bu karışımı oluşturan her bir kaynağın tahmin edilmesi kör kaynak ayırıştırma olarak tanımlanmaktadır. Kör kaynak ayırıştırma yöntemlerinin başarımı, işlem süresi ve doğruluğu ile belirlenmektedir. Bu çalışmada doğrusal biçimde karışmış olan farklı ses işaretleri, Kanonik Korelasyon Analiz tabanlı algoritmaları kullanılarak ayırıştırılmıştır. Algoritmaların başarımlarıyla birlikte, işlemsel yükleri ile orantılı olan işlem süreleri karşılaştırılmış ve ayrıca algoritmaların işlem sürelerini düşürmek için Ayrık Dalgacık Dönüşümü yönteminden faydalanılmıştır. Bunlara ek olarak farklı veri uzunlukları kullanılarak, veri uzunluğunun algoritma performansları üzerindeki etkisi incelenmiştir.

Processing Time Reduction of Canonical Correlation Analysis Based Sound Separation Algorithms

Keywords

Independent component analysis,
Canonical correlation analysis,
Kernel canonical correlation analysis,
Discrete wavelet transform

Abstract: In noisy environments, it is very important to obtain the sound signal without noise. As the mixing environment and noise sources are not known, the decomposition process is quite difficult. The estimation of each source constituting this mixture from a data set containing a mixture of at least two signals is defined as blind source separation. The performance of the blind separation process is determined by the processing time and accuracy. In this study, the different audio signals which are mixed linearly are decomposed by using Canonical Correlation Analysis based algorithms. Along with the performance rates of the algorithms, the processing times which are proportional to the calculation cost are compared. In addition, Discrete Wavelet Transform was used to reduce the processing times of the algorithms. In addition, by using different data lengths, the effect of data length on algorithm performance was investigated.

1. Giriş

Literatürde ses ayırıştırma için kullanılan pek çok algoritma bulunmaktadır. Bu alanda yapılan ilk çalışmalar kaynak işaretlerin istatistiksel özelliklerini inceliyerek ayırıştırma işlemi yapılmaya çalışılmıştır. Fakat bu durum, kaynaklar hakkında bazı varsayım yapmayı gerektirmektedir. Bu varsayımlar da kaynak işaretlerin istatistiksel olarak bağımsız olmasını

gerektirmektedir. Çoğu algoritmada, algoritmaların işlemsel yükü ile ayırıştırma performansları doğru ilişki içerisindedir [1]. Bu çalışmada kullanılan Kanonik Korelasyon Analiz (KKA) tabanlı algoritmaların işlemsel yükünü azaltmak için Ayrık Dalgacık Dönüşümü (ADD) kullanılmıştır. Ses ayırıştırma senaryosu olarak herhangi bir ortamda bir konuşma sesi ve iki tane farklı ses kaynakları eşzamanlı olarak ses çıkarmakta olduğu ve aynı

*:fyilgin@erzincan.edu.tr

ortamda yerleştirilen üç tane mikrofonun da bu ses ve gürültü işaretlerinin ağırlıklı karışımlarını kaydetmekte olduğu varsayılmıştır. Buradaki amaç sadece kaydedilen mikrofon sinyallerini kullanarak insan sesini ve diğer ses kaynaklarını elde etmektir. Kullanılan kaynak işaretlerinin ve de bunların nasıl karıştıklarını o an tahmin edemediğimizden dolayı ayırıştırma kör kaynak ayırıştırma olarak tanımlanmaktadır. Bu işlemin literatürdeki en iyi örneksel karşılığı kokteyl parti problemi olarak tanımlanmaktadır [2]. Karışmış sinyallerden orijinal sinyalleri elde etmek için KKA, Çekirdek Kanonik Korelasyon Analiz (ÇKKA), Ayrık Dalgacık Dönüşümlü (ADD) KKA ve ADD'li ÇKKA algoritmaları kullanılmıştır. Bu algoritmaların ayırıştırma performanslarını karşılaştırmak için SNR oranları hesaplanmış, algoritmaların çalışma süreleri kayıt altına alınarak işlemsel yükleri incelenmiştir.

2. Materyal ve Metot

Verilen bir veri dizisi içerisinde istatistiksel olarak bağımsız sinyalleri bulmak için Bağımsız bileşen analizi kullanılmıştır. Bu yöntem ile bağımsız bileşenler bilinmeyen bir karıştırma matrisi ile karıştırılarak analiz için gözlem verilerini meydana getirirler.

Kanonik korelasyon analizinde iki veri kümesi arasındaki ilişkiyi analiz etmek için geliştirilmiştir. Burada amaç değişken veri kümeleri arasındaki maksimum korelasyonu bulmaktır.

Çekirdek kanonik korelasyon analizinde veriler yüksek boyutlu öznitelik uzayına taşınarak veri analizi için çeşitli çözümler sunar.

Verilerin zaman-frekans analizinde yüksek frekanslar için dar, düşük frekanslar için ise geniş pencere boyutlarında incelenmesi için ayrık dalgacık dönüşümünden yararlanılır.

2.1. Bağımsız bileşen analizi

Bağımsız bileşen analizi (BBA); çok değişkenli verinin istatistiksel özellikleri kullanılarak, bağımsız bileşenler oluşturan bir analiz yöntemidir. BBA yönteminde bağımsız bileşenler doğrusal veya doğrusal olmayan bir biçimde bilinmeyen bir karıştırma işleminden geçerek gözlem verilerini oluşturdukları varsayılmaktadır [3]. Doğrusal karışım için matematiksel gösterim eşitlik (1)'deki gibidir:

$$x = As \quad (1)$$

burada x vektörü gözlem verilerini (mikrofon), s vektörü kaynak verilerini, A matrisi de karıştırma matrisini (ortamın karıştırma etkisi) temsil etmektedir. BBA algoritmalarında ayırıştırma işlemi basitleştirmek için, A matrisi kare kabul edilerek kaynak sayısı ile bağımsız bileşen sayısının eşit olması

sağlanır [4]. BBA algoritmalarının amacı, bilinmeyen s kaynak değerlerinin yaklaşımı olan \hat{s} değerlerini, sadece x gözlem vektörünü kullanarak kestirilmesidir. Bunun için, bağımsız bileşenler olan s kaynaklarının non-Gaussian dağılıma sahip olduğu varsayılır iken, merkezi limit teoreminden dolayı gözlem verisi x 'in dağılımının Gauss dağılıma daha yakın olduğu varsayılmaktadır. BBA algoritması birden fazla gaussian dağılıma sahip veri üzerinde ayırıştırma performansı iyi değildir, bundan dolayı birden fazla Gauss dağılıma sahip veri var ise farklı yöntemler uygulanmalıdır. BBA işlemi sonucunda bulunan ayırıştırma matrisi (W); A matrisinin tersi olduğu varsayılır. W matrisini BBA ile bulmak için amaç fonksiyonları tanımlanır ve bu amaç fonksiyonların optimizasyonu ile W elde edilir. Elde edilen bu matris kullanılarak kaynak işaretlerinin tahmini aşağıdaki gibi hesaplanır:

$$\hat{s} = Wx \quad (2)$$

Çalışma metni içerisinde denklem yer alacak ise bunlar kenarlıklar gizlenmiş tek satırlık bir tablo içerisinde iki hücre ile verilmelidir. Sol hücreye denklem eklenmeli ve bu hücre ortali olmalıdır. Sağ hücrede ise denklem numarası yer almalı ve bu hücre sadece denklem numarasının sığacağı büyüklükte, sağa dayalı olmalıdır. Aşağıda denklem örnekleri verilmiştir.

2.2. Kanonik korelasyon analizi

KKA iki veya daha fazla değişken kümesi arasındaki ilişkileri ve bu değişken kümelerinin doğrusal fonksiyonları arasındaki maksimum korelasyonları bulmaya çalışarak, veri analizi yapan bir yöntemdir [5]. Verilerin doğrusallığı sağlayamadığı durumlar için doğrusal olmayan kanonik korelasyon analizi geliştirilmiştir [6].

İki değişken kümesi arasındaki korelasyonda, X değişken kümesi için p adet, Y değişken kümesi için ise q adet ($p \leq q$) değişken olduğunda, bu iki kümedeki değişkenlerin doğrusal olarak kombinasyonları hesaplanarak bunlar arasındaki korelasyon bulunabilir. KKA yöntemi, değişken kümelerinin doğrusal bileşenleri olan $U = a^T X$ ve $V = b^T Y$ değerleri arasındaki korelasyonu maksimum yapacak şekilde a ve b vektörlerini bulmaya çalışmaktadır [7]. Amaç U ve V birim varyanslı değişkenler arasındaki korelasyonu maksimum yapacak a ve b vektörlerinin seçilmesidir. Kanonik değişkenler arasındaki korelasyon, yani kanonik korelasyon:

$$p = \frac{Kov(a^T X, b^T Y)}{\sqrt{Var(a^T X)Var(b^T Y)}} = \frac{a^T C_{xy} b}{\sqrt{a^T C_{xx} a b^T C_{yy} b}} \quad (3)$$

olarak ifade edilir. Her bir doğrusal bileşenin aşağıdaki denklemlerdeki gibi birim varyansa sahip oldukları varsayılırsa:

$$\text{Var}(U) = a^T C_{xx} a = 1 \quad (4)$$

$$\text{Var}(V) = b^T C_{yy} b = 1 \quad (5)$$

Varyans değerleri bir olacak, şekilde a ve b vektörlerinin seçilmesinden dolayı korelasyon katsayısı, U ve V arasındaki kovaryans değerine eşit olmaktadır. Burada yapılması gereken bu ifadenin maksimum olmasıdır [7].

$$\max_{a,b} a^T C_{xy} b \quad (6)$$

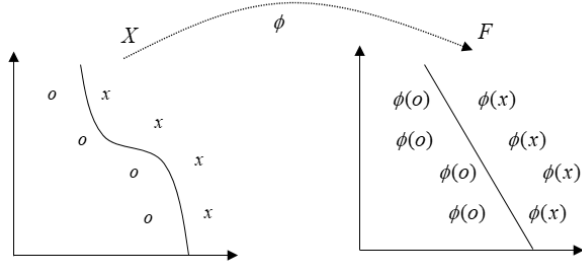
Bundan dolayı katsayıların maksimizasyon problemi olduğu düşünülüp, λ_1 ve λ_2 Lagrange çarpanları olmak üzere Lagrange fonksiyonu kullanılarak (6) denklemi aşağıdaki gibi şekillendirilir [8].

$$L = a^T C_{xy} b - 0.5\lambda_1 (a^T C_{xx} a - 1) - 0.5\lambda_2 (b^T C_{yy} b - 1) \quad (7)$$

Üsteki denklemde a ve b 'ye göre türev alınarak kanonik korelasyon için gerekli olan değerler bulunur.

2.3. Çekirdek kanonik korelasyon analizi

Çekirdek gösterimler, doğrusal öğrenme makinelerinin hesaplama yeteneğini arttırmak için verileri yüksek boyutlu öznitelik uzayına eşleyerek alternatif bir çözüm sunan yöntemlerdir.



Şekil 1. Öznitelik eşlemesi

Şekil 1.'de iki boyutlu giriş uzayından iki boyutlu öznitelik uzayına bir öznitelik eşleme örneği göstermektedir. Bu örnekte veriler doğrusal bir fonksiyonla giriş uzayında ayrılmazlarken öznitelik uzayında ayrılabilir hale getirilebilmektedirler.

Kernel KKA'de orijinal verinin doğrusal olmayan işlevler aracılığıyla bir özellik alanına eşlendiği varsayılır. Ardından, özellik alanında doğrusal KKA uygulanır. Eğer üzerinde durulan gerçek fonksiyonlar sadece sonlu norma sahip olanlarla sınırlandırılırsa, ortaya çıkan uzay Hilbert uzayı olarak bilinir [9]. Hilbert uzaylarının en temel özelliklerinden birisi, bu uzayın herhangi bir elemanı, sayılabilir sayıdaki ortonormal fonksiyonların ağırlıklı bir toplamı olarak ifade edilebilir olmasıdır [10,11]. Burada kullanılan çekirdek, iki girişe karşı düşen öznitelik vektörlerinin iç çarpımını gerçekleştirir.

$$K(x, z) = \langle \phi(x), \phi(z) \rangle \quad (8)$$

$$K(x, z) = \langle Ax, Az \rangle = x^T A^T A z = x^T B z \quad (9)$$

Eğer bu yaklaşım kullanılmak istenirse önce karmaşıklaştırılmış öznitelik uzayının oluşturulması, sonra bu uzayda iç çarpımın ne olacağı çözülür ve son olarak da o değeri orijinal girişler cinsinden hesaplamak için direkt bir yöntemin bulunması gerekir. Bu yüzden dolayı öznitelik uzayı tanımlanır [12]. Konu hakkında daha detaylı bilgi [11,12] kaynaktan elde edilebilir.

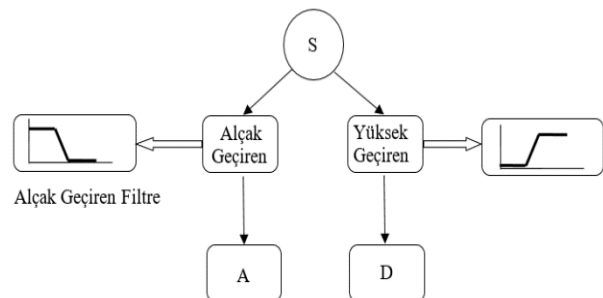
2.4. Ayrık dalgacık dönüşümü

Durağan olmayan sinyallerde Fourier Dönüşümü işaretin sadece frekans bileşenlerini gösterdiği için bu işaretlerin zaman düzlemindeki davranışı hakkında yeterli bilgi vermemektedir. Fakat durağan olmayan sinyallerde frekans bileşeni zamanda da değişkenlik göstermektedir. Bu nedenle, ayrık dalgacık dönüşümü (ADD) kullanılarak sinyalin zaman ve frekanstaki davranışı aynı anda gözlemlenebilmektedir. ADD'nin bir diğer faydası ise, algoritmaların hesapsal yükünü azaltmasıdır. ADD işleminin matematiksel ifadesi aşağıdaki denklemdeki gibi ifade edilmektedir:

$$ADD(m, n) = 2^{-m/2} \int f(t) \psi(2^{-m}t - n) dt \quad (10)$$

burada $f(t)$ işlenecek sinyali, m parametresi ölçeklemeyi, n parametresi ötelemeyi ve $\psi(z)$ ise uygun dalgacığı temsil etmektedir. ADD' de ölçekleme ve öteleme değerleri 2 ve 2'nin katlarını örnek alacak şekilde seçilir [13,14].

Birçok işaret için, işaretin düşük frekanslı içeriği en önemli kısımdır. Düşük frekans bileşenleri, işaretler için genellikle tanıma özelliklerini içerirler ve diğer taraftan yüksek frekanslı bileşenler ise işaretin kendisiyle düşük frekanslı kısım arasındaki farkı oluştururlar. Sinyallerin düşük frekanslı bileşenlerini incelemek için büyük öteleme değerleri, yüksek frekanslı bileşenlerini analiz etmek için ise küçük öteleme değerleri kullanılır [15]. Sinyalin alçak geçiren (AGS) ve yüksek geçiren (YGS) süzgeçlerle katlaması alınarak çıkışlar elde edilir [16]. İncelenen bir S işareti Şekil 2.'de görüldüğü gibi birbirini tümleyen alçak ve yüksek geçiren filtrelerden geçirilerek alçak olarak A ve yüksek olarak D frekanslı bileşenlerine ayrıştırılır. Konu hakkında daha detaylı bilgi [17] kaynaktan elde edilebilir.



Şekil 2. ADD filtre algoritması

2.5. Performans ölçümü

Algoritmaların performans analizleri belirlenirken aşağıda formülü verilen işaret/gürültü (SNR) oranından yararlanılmıştır [18].

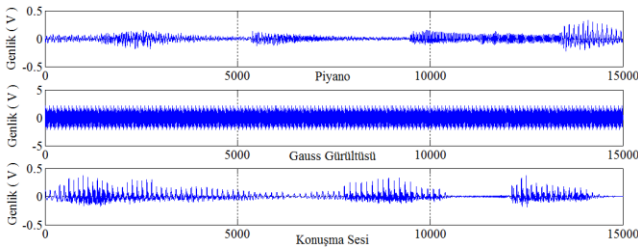
$$SNR = \frac{\sum_{n=0}^{N-1} x^2(n)}{\sum_{n=0}^{N-1} (x_q(n) - x(n))^2} \quad (11)$$

$$SNR \text{ dB} = 10 \cdot \log_{10}(SNR) \text{ dB} \quad (12)$$

3. Bulgular

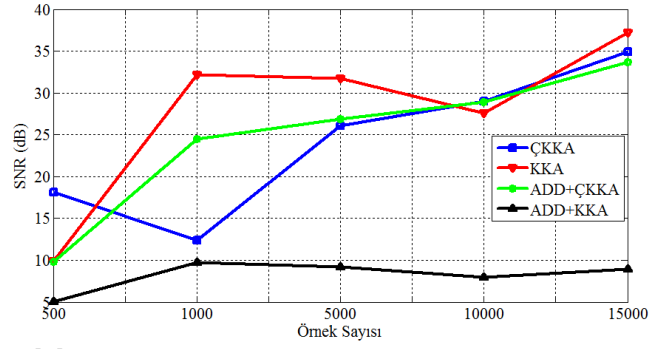
Yapılan uygulamada piyano, gauss gürültüsü ve insan konuşma sesi verilerinin üzerinde çalışma yapılmıştır. Bu sinyaller 100'er defa rastgele seçilen karıştırma matrisleriyle karıştırıldıktan sonra ÇKKA, KKA ve ADD algoritmalarına uygulanmış olup sonuçlar SNR(dB) cinsinden kaydedilmiştir. Performans indeksini hesaplamak için rastgele karışım oranları başlatılarak hesaplanan 100 adet SNR değerleri eşitlik (12)'ye göre hesaplanarak ortalamaları alınmıştır. Bu işlemler periyodik olarak veri uzunluğu 500, 1000, 5000, 10000 ve 15000 örnek alınarak tekrarlanmıştır. Böylece algoritmaların değişen veri uzunluklarına göre performans analizleri de yapılmıştır.

Şekil 3. 'de bu üç sinyale ait 15000 örnek sayısı için kaynak sinyalleri gösterilmiştir.



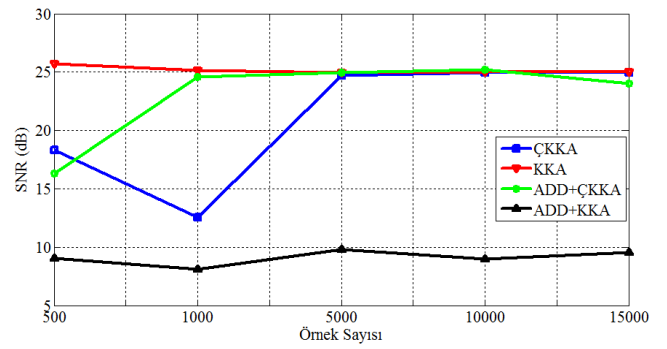
Şekil 3. Kaynak sinyaller

Algoritmaların uygulanması sonucunda SNR değerleri elde edilmiştir. Her bir kaynak ses işareti için sonuçlar ayrı ayrı grafiklerde gösterilmiştir. Şekil 4. 'de Piyano sesi için elde edilen algoritmaların SNR oranları verilmiştir. Şekil 4. 'de ki sonuçlara bakıldığında veri boyutlarının artmasıyla SNR oranının arttığı ve algoritmalar arasında veri boyutu düşükken en iyi sonucu KKA'nın verdiği gözlemlenmiştir.



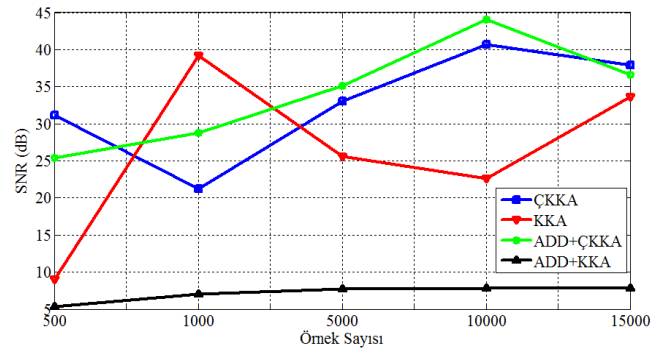
Şekil 4. Piyano sesi

Şekil 5. 'de Gauss gürültüsü için algoritma sonuçları verilmiştir. Şekilden de görüleceği üzere veri boyutu arttıkça tüm algoritmalar aynı SNR oranlarına sahip olduğu gözlemlenmiştir.



Şekil 5. Gauss gürültüsü

Şekil 6. 'da ise insan sesi için algoritmaların SNR oranları verilmiştir. Genel olarak bakıldığında ADD+ÇKKA algoritması diğer algoritmalarla karşılaştırıldığında daha iyi sonuç vermektedir. Bu yöntem ile ölçüm sonuçlarında ortalama olarak ADD'nin ÇKKA'ya göre SNR değerlerinde iyileştirme yapılmıştır.



Şekil 6. Konuşma sesi işareti

Aşağıda görülen tabloda ise bu algoritmaların çalışma süreleri karşılaştırılmıştır. Çalışma süresi aynı zamanda algoritmanın işlem yükü ile doğrudan orantılıdır. KKA algoritmasının diğer algoritmalarla karşılaştırıldığında çok hızlı sonuç verdiği gözlemlenmiştir. ADD+ÇKKA'da ise çalışma süresinin ÇKKA'ya göre yarı yarıya düştüğü ölçülmüştür. ADD algoritmasının süre bakımından performansını iyi anlamak için KKA'ya uygulanıp çalıştırılmıştır. ADD+KKA algoritmasının veri boyutu arttıkça çalışma süresinin de azaldığı Tablo 1. 'de gösterilmiştir.

Tablo 1. Algoritmaların çalışma süreleri

Adım Sayısı	ÇKKA	KKA	ADD+ÇKKA	ADD+KKA
500	2.61 sn	0.12 sn	3.01 sn	0.18 sn
1000	4.69 sn	0.13 sn	3.49 sn	0.13 sn
5000	16.55 sn	0.13 sn	5.73 sn	0.16 sn
10000	34.81 sn	0.15 sn	11.5 sn	0.17 sn
15000	63.19 sn	0.18 sn	24.02 sn	0.2 sn

4. Tartışma ve Sonuç

Bu çalışmada ses ayırma algoritması olarak kullanılan algoritmaların işlem hacminin ADD algoritması kullanılarak düşürülebileceği gösterilmiştir. Yapılan çalışmalar sonucunda ADD algoritmasının işlem hacmini düşürmenin yanında ses sinyalinin sonuçlarını iyileştirdiği gözlemlenmiştir. Pişano ve gauss gürültü için ADD algoritma sonuçlarında iyileşme gözlemlenmemiştir. Şekil 4., 5. ve 6.'ya bakıldığında, hangi sinyal olursa olsun veri boyutunun artırılmasının bu sinyallerin SNR oranlarını artırdığı görülmüştür.

Kaynakça

- [1] Serrano, M.C. 2009. Application of sound source separation methods to advanced spatial audio systems. Technical University of Valencia, Ph. D. Thesis , 212, Spain
- [2] Sezer, O., G. Erçil, A., Keskinöz, M. 2005. Destek Vektör Makinesi Kullanarak Bağımsız Bileşen Tabanlı 3B Nesne Tanıma. IEEE 13th Signal Processing and Communications Applications Conference, 16-18 May, Kayseri, 99-102.
- [3] Hyvarinen, A., Oja, E. 2000. Independent Component Analysis: Algorithms and Applications. Neural Networks 13, 411-430.
- [4] Hyvarinen, A. 1999. Survey On Independent Component Analysis. Neural Computing Surveys 2, 99, 94-128.
- [5] Hotelling, H. 1936. Relation Between Two Sets of Variates. Biometrika, 28, 321-377.
- [6] Kettenring, J. R. 1971. Canonical Analysis of Several Sets of Variables. Biometrika, 58, 433-451.
- [7] Anderson, T. W. 2003. An introduction to multivariate statistical analysis. 3rd edition, WILEY Series In Probability and Atatistics, 752s.
- [8] Tatlıdil, H. 1996. Uygulamalı çok değişkenli istatistiksel analiz. Cem Web Ofset Ltd. Şti. Ankara, 424s.

- [9] Keser, İ. K., Ertaş, K. 2007. Düzgünleştirilmiş Fonksiyonel Ana Bileşenler Analizi ve Bir Uygulama. Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 22, 1-26.
- [10] Ramsay, J. O. 1982. When The Data Are Functions. Psychometrica, 47(4), 379-396.
- [11] Bach, F. R., Jordan, M. I. 2002. Kernel Independent Component Analysis. Journal of Machine Learning Research, 3, 1-48.
- [12] Aykut, M. 2007. Örüntü tanıma amaçlı doğrusal olmayan yöntemlerin geliştirilmesi ve uygulanması. Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 111s, Trabzon.
- [13] Strang, G., Nguyen, T. 1996. Wavelets and Filter Banks. 2nd edition, Wellesley-Cambridge Press, 520s.
- [14] Bilgin, S. 2008. Kalp hızı değişkenliğinin dalgacık dönüşümü ve yapay sinir ağları kullanılarak analizi. Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 156s, Sakarya.
- [15] Ustundağ, M. 2013. ECG Sinyallerinde Gürültü Gidermek için Dalgacık Dönüşümünün FPGA Tabanlı Donanımsal Gerçekleşmesi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 25(2), 63-68.
- [16] Tepe, C., Sezgin, H. 2007. EKG Sinyallerinde Gürültü Gidermede Ayrık Dalgacık Dönüşümünde Farklı Ana Dalgacıkların Ve Ayırma Seviyelerinin Karşılaştırılması. Elektrik – Elektronik –Bilgisayar ve Biyomedikal Müh. 12. Ulusal Kongresi, 14-18 Kasım, Eskişehir.
- [17] Olkkonen, J. 2011. Discrete Wavelet Transforms. Theory And Applications, Mart 2011, İndia, 1-268.
- [18] Tan, L. 2007. Digital Signal Processing. Fundamentals and applications. Academic Press, 840s.