



Ses İşaretlerinin Ayrıştırılmasında Kör Kaynak Algoritmalarının Performans Analizleri*

Hüsamettin Çelik^{1**}, Nurhan Karaboğa²

¹ Erzincan Binali Yıldırım Üniversitesi, Tercan Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Erzincan, Türkiye (ORCID: 0000-0002-7662-0674)

² Erciyes Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Kayseri, Türkiye (ORCID: 0000-0002-4550-5251)

(Konferans Tarihi: 5-7 Mart 2020)

(DOI: 10.31590/ejosat.araconf52)

ATIF/REFERENCE: Çelik, H. & Karaboğa N. (2020). Ses İşaretlerinin Ayrıştırılmasında Kör Kaynak Algoritmalarının Performans Analizleri. *Avrupa Bilim ve Teknoloji Dergisi*, (Özel Sayı), 399-404.

Öz

İki veya daha fazla sesin karışımından birini elde etmek veya gürültülü ortamda kaydedilen seslerin gürültüden ayrıştırılması her zaman popülerliğini korumaktadır. Gürültü çeşidi, kaynağı ve seslerin karışım ortamları bilinmediği için ayrıştırma işleminde kullanılan algoritmaların işlem süresi ve performansları farklılık göstermektedir. Karışmış işaretleri ayrıştırmada kör kaynak algoritmaları kullanılmaktadır. Doğrusal karışmış işaret kaynaklarından oluşan veri kümesinden, işaretlerin veya gürültünün ayrı ayrı tahmin edilme işlemi kör kaynak ayrıştırma olarak ifade edilmektedir. Mühendislik uygulamalarında birçok yöntem olsa da tasarım kısıtlamaları ve gereksinimleri göz önüne alındığında hangi algoritmanın daha uygun olacağını analiz etmek gerekir. Bu çalışmada üç farklı yöntem olarak; Pearson bağımsız bileşen analizi (Pearson Independent Component Analysis -PICA), İkinci dereceden kör tanımlama (Second-Order Blind Identification-SOBI) algoritması ve Ortak yaklaşım özdeğerlerin köşegenleştirilmesi (Joint Approximation Diagonalization of Eigen-matrices-JADE) algoritmaları karşılaştırılmıştır. Algoritmaların performans analizleri dikkate alınarak başarı oranları ve işlem sürelerine göre değerlendirilmiştir.

Anahtar Kelimeler: Kör kaynak ayırma, ICA,PICA, SOBI, JADE.

Performance Analysis of Blind Source Algorithms in the Separation of Sound Signals

Abstract

Obtaining a mixture of two or more sounds or separating the sounds recorded in the noisy environment from noise is among the most studied topic. Since noise type, source and mixing environments of sounds are unknown, the processing time and performance of the algorithms used in the separation process differ. Blind source algorithms are used to decompose mixed signs. From the data set consisting of linear mixed signal sources, the process of estimating the signals or noise separately is expressed as blind source decomposition. Although there are many methods in engineering applications, it is necessary to analyze which algorithm will be more suitable considering the design constraints and requirements. In this study, as three different methods; Pearson independent component analysis (PICA), Second order blind identification (SOBI) algorithm and Joint approximate diagonalization of eigen matrices (JADE) algorithms were compared. It has been evaluated according to performance rates and processing times by considering the performance analysis of the algorithms.

Keywords: Blind Source Separation, ICA, PICA, SOBI, JADE.

* Bu makale *International Conference on Access to Recent Advances in Engineering and Digitalization (ARACONF 2020)* de sunulmuştur.

** Sorumlu Yazar: Erzincan Binali Yıldırım Üniversitesi, Tercan Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Erzincan, Türkiye, ORCID: 0000-0002-7662-0674, hcelik@erzincan.edu.tr

1. Giriş

Günümüzde birden fazla kişinin gürültülü ortamlarda konuşma sesinin kaydedilmesi ve gerektiğinde istenen kişi için sesinin ayırılması oldukça önemlidir. Tıp elektronigi, radar sistemleri, güvenlik, doğal afetler, ortam dinlemeleri ve yer tespitinde dinlenen sesin gürültüden arındırılması ve işlem süresinin kısıtlılığı hayati öneme sahip olabilir.

Bu çalışmanın amacı konuşma seslerini veya karışmış olan ortam gürültüsünü ayırtmada en uygun algoritmayı bulmaktır. Yapılan çalışmada Pearson bağımsız bileşen analizi (Pearson Independent Component Analysis-PICA) (Karvanen, Eriksson, & Koivunen, 2000), İkinci dereceden kör tanımlama (Second-Order Blind Identification-SOBI) (Belouchrani, Abed-Meraim, Cardoso, & Moulines, 1993; 1997) algoritması ve ortak yaklaşım özdeğerlerin köşegenleştirilmesi (Joint Approximation Diagonalization of Eigenmatrices-JADE) (Cardoso & Souloumiac, 1993) algoritmaları kullanılarak iki konuşma sesi ile beyaz gauss gürültüsü, rastgele karışım matrisleriyle karıştırılıp sonrasında üç farklı algoritma ile ayırılmıştır. Yapılan bu işlemlerden sonra sinyal/gürültü oranı kullanılarak üç farklı algoritmanın performans analizleri yapılmıştır (Ilgin, Celik, & Sevim, 2012; Celik, Ilgin, & Sevim, 2019).

Kokteyl parti problemi (Cocktail Party Problem) olarak adlandırılan bu yöntemde, farklı mesafelerde bulunan üç mikrofon ve aynı ortamda bulunan iki ses kaynağını ve bir gürültü işaretinin ağırlıklı karışımlarını kaydettiği varsayılmıştır. Bu çalışmanın kapsamı mikrofon ile kaydedilen sinyalleri birbirinden ayırarak kaynak işaretlere en yakın işareti elde etmektir. Bu işlem literatürde "Kokteyl Parti Problemi" olarak geçmektedir (Bronkhorst, 2000). Ortam itibariyle kaynak işaretlerinin ve gürültünün nasıl karıştığını bilmediğimizden ayırma işlemi kör kaynak ayırma olarak adlandırılmaktadır.

Literatürde kullanılan Bağımsız Bileşen Analizi (ICA) algoritması bağımsız bileşenleri ayırmak için oldukça yaygın kullanılan bir yöntemdir (Hyvärinen, 1984). Kör kaynak ayırma işleminde kullanılan yöntemlerin temelinde ICA algoritması bulunmaktadır. Başka bir yöntem olan Kernel Bağımsız Bileşen Analizi (KICA) algoritması (Bach & Jordan, 2002) başarımlı performansı iyi olsa da işlem süresi bakımından çok yavaştır.

Yapılan bu çalışmanın adımları: 2. Bölümde JADE, SOBI ve PICA algoritmalarının çalışma prensipleri hakkında bilgi verilmektedir. 3. Bölümde karşılaştırılan algoritmaların başarımlı oranı ve işlem süreleri ile ilgili bulgular verilmektedir. 4. Bölümde ise sonuçlara yer verilmektedir.

2. Materyal ve Metod

2.1. Bağımsız Bileşen Analizi (ICA)

Bağımsız bileşen analizi klasik bir yöntem olup bir dizi doğrusal olarak karışan çok değişkenli sinyallerden orijinal sinyalleri analiz etme yöntemidir. PICA algoritmasının da temelini oluşturmuştur. ICA işlevselliğinin özü istatistiksel bağımsızlığa dayanır. Bu model Eşitlik (1) ile ifade edilirse;

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

Burada, $x(t)$ alınan sinyal vektörü, A karıştırma matrisi ve $s(t)$ orijinal kaynaklar vektörüdür. ICA algoritmasında, ayırma matrisi olarak W matrisi, A matrisinin tersi olarak kabul edilir. Bunun sonucunda ICA modeli:

$$Wx(t) = y(t) \quad (2)$$

$y(t)$ tahmini kaynak sinyallerini temsil eder (Hyvärinen & Oja, 2000).

2.2. Pearson Bağımsız Bileşen Analizi (PICA)

Pearson sistemi, geniş kaynak dağılımlarını modellemek için kullanılacak parametrik bir kaynak dağılımıdır. Bu algoritma sabit doğrusal olmayan kontrast fonksiyonları ve maksimum olabilirlik yaklaşımını içermektedir. Doğrusal olmayan fonksiyon yaklaşımı tüm frekanslara eşit olarak uygulanır, ancak her frekans için gerçek konuşma sinyali farklı bir dağılıma sahiptir. Dağılım için diferansiyel denklem Eşitlik (3) ile tanımlanır:

$$f(x) = \frac{(x-a)f(x)}{b_0+b_1x+b_2x^2} \quad (3)$$

burada a , b_0 , b_1 ve b_2 dağılım parametreleridir ve moment yöntemi ile tahmin edilebilir. Bağımsız bileşen analizinde maksimum olabilirlik yaklaşımı, varsayılmış kaynak dağılımının sonuç fonksiyonunu:

$$\varphi(x) = -\frac{f'(x)}{f(x)} = -\frac{(x-a)}{b_0+b_1x+b_2x^2} \quad (4)$$

olarak bulunur. Pearson sisteminde payda iki gerçek köke sahip olduğunda genelleştirilmiş beta dağılımını temsil etmektedir. Momentler yönteminde, teorik momentler örnek momentlerle tahmin edilir.

$$\hat{\alpha}_1 = \bar{x} = \sum_{i=1}^n x_i/n \quad (5)$$

verilerden hesaplanır. Ortalama sıfır ve varyans bir olduğunda, pearson sistem parametrelerinin tahmin edicileri için aşağıdaki formül elde edilir.

$$\hat{b}_1 = \hat{a} = -\frac{\hat{a}_3(\hat{a}_4+3)}{c} \quad (6)$$

Burada $\hat{C} = 10\hat{a}_4 - 12\hat{a}_3^2 - 18$ olarak verilir. Pearson ICA algoritması hakkında daha fazla bilgi için kaynak (Karvanen, Eriksson, & Koivunen, 2000) ve (Kato, Nagahara, Araki, Sawada, & Makino, 2006) incelenmelidir.

2.3. İkinci Dereceden Kör Tanımlama Algoritması (SOBI)

İkinci dereceden kör tanımlama (SOBI) yöntemi, kaynak sinyali ve karıştırma matrisinin en iyi tahminini çözebilen ve ayrıca farklı kaynak sinyallerinden ayrılmış sinyal bileşenlerini elde edebilen orijinal gözlem verilerinin ikinci dereceden istatistiklerine dayanan kör kaynak ayırma yöntemidir (Belouchrani, Abed-Meraim, Cardoso, & Moulines, 1997). SOBI yöntemi, kaynak sinyal bileşenini tahmin edebilen ve göreceli olarak daha az veri noktası kullanarak birden fazla Gauss gürültü kaynağını ayırabilmektedir (Liu, Wang, Huang, & Yang, 2019). Bir dizi çalışmada modal parametre tanımlaması için umut verici bir alternatif olarak kabul edilmiştir (Rainieri, 2014). Son yıllarda, sinyal işleme ve mekanik arıza tespiti alanlarında SOBI yöntemi başarıyla uygulanmıştır.

Bu algoritma, sinyallerin zaman-korelasyon yapısı varsayımını çürütmek için ikinci sıra istatistiklere dayanır ve bunun için aşağıdaki adımların hesaplanmasını gerektirir:

- Beyazlatma
- Gecikmeli Korelasyon Matrislerinin Hesaplanması
- Ortak Köşegenleştirme

SOBI'nin ana konsepti, aşağıdaki ifadeye göre duran gecikmeli korelasyon matrislerinin diyagonal formu hakkında varsayımdır:

$$\begin{aligned} R_x(\tau) &= E\{x(t)x(t+\tau)^T\} \\ &= AR_s(\tau)A^T \end{aligned} \quad (7)$$

Burada R_s , kaynak sinyallerinin korelasyon matrisi ve R_x ise, gecikmeli korelasyon matrisidir. $\tau \in \{\tau_j | j = 1, 2, \dots, k\}$ ve $x(t)$ beyazlatılmış kovaryan matrisidir.

Eşitlik (7) dikkate alındığında tüm τ değerleri için, tüm korelasyon matrislerini birlikte köşegenleştiren üniter bir $U=BA$ matrisini oluşturur:

$$U^T R_x(\tau) U = R_s(\tau) \quad (8)$$

Bu nedenle, yaklaşık orijinal sinyaller $U^# Bx(t)$ ile hesaplanır (Sahonero & Calderon, 2017).

2.4. Ortak Yaklaşım Özdeğerlerin Köşegenleştirilmesi (JADE)

JADE algoritması, dördüncü dereceden kümülatörlerin köşegenleştirilmesine dayanan kör kaynak ayırma yöntemidir. Köşegenleştirilecek matris sayısının artırılmasıyla kaynak ayırma işleminde iyileşme elde edilir (Cardoso & Souloumiac, 1993; Miettinen, Nordhausen, & Taskinen, 2017). Herhangi bir $p \times p$ boyutunda M matris için dördüncü dereceden kümülatör matrisi şu şekilde tanımlanır:

$$C(M) = E[(x_{st} M x_{st}^T) x_{st}^T x_{st}] - M - M^T \text{tr}(M) I_p \quad (9)$$

burada x_{st} standartlaştırılmış bir değişkendir. JADE'de C matrisleri yaklaşık olarak dikey bir matris tarafından çapraz olarak köşegenleştirilir. Dördüncü dereceden çapraz kümülatörlerin toplamı;

$$J_{JADE}(W) = \sum_i \|diag(WF(M)_i W^T)\|^2 \quad (10)$$

şeklinde dir. Burada $\|diag()\|^2$ köşegen karelerin toplamı anlamına gelir. $F(M)$, M matrisin sabit bir sayıyla çarpımıdır. W ise kestirilecek parametre sayısını azaltmak için beyazlatılmış matristir (Hyvärinen, 1984).

3. Araştırma Sonuçları ve Tartışma

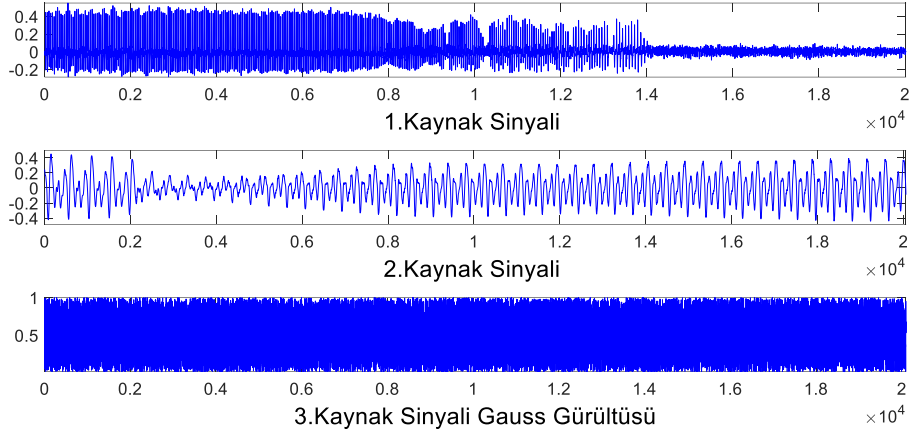
3.1. Bulgular

Bu çalışmada iki konuşma sesi ve gauss gürültüsü kullanılarak bir çalışma yapılmıştır. Rastgele seçilen karıştırma matrisleri ile bu üç işaret karıştırılarak PICA, SOBI ve JADE algoritmalarına uygulanmıştır. Algoritmaların stabil çalıştığını gözlemek için bu işlemler 100 kere tekrarlanmış ve sonuçlar dB cinsinden işaret/gürültü yani SNR oranı ile değerlendirilmiştir. Elde edilen sonuçların her bir algoritma için ortalamaları alınmıştır. SNR (dB) değerleri için Eşitlik (11) denklemi kullanılmıştır (Tan, 2007).

$$SNR (dB) = 10 \cdot \log_{10} \left(\frac{\sum_{n=0}^{N-1} x^2(n)}{\sum_{n=0}^{N-1} (x_q(n) - x(n))^2} \right) \quad (11)$$

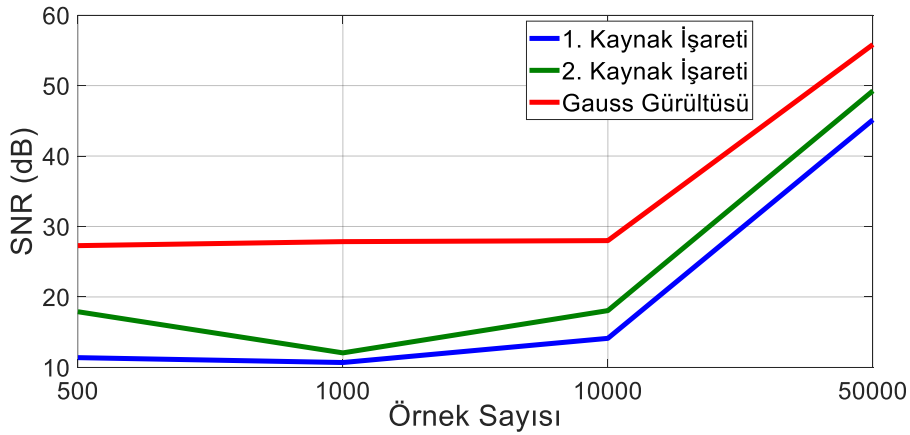
Kaynak işaretleri için 500, 1000, 10000 ve 50000 örnek sayısı alınarak sonuçlar incelenmiştir. Dört farklı sayıda veri uzunlukları dikkate alınarak algoritmaların performans analizleri yapılmıştır. Aynı zamanda algoritmaların işlem süreleri de kaydedilerek karşılaştırma yapılmıştır.

Şekil 1'de konuşma sesleri ve gauss gürültüsüne ait 20000 örnek sayısı için sinyaller gösterilmiştir. Burada kullanılan konuşma sesleri gürültüsüz ortamda kaydedilmiştir.

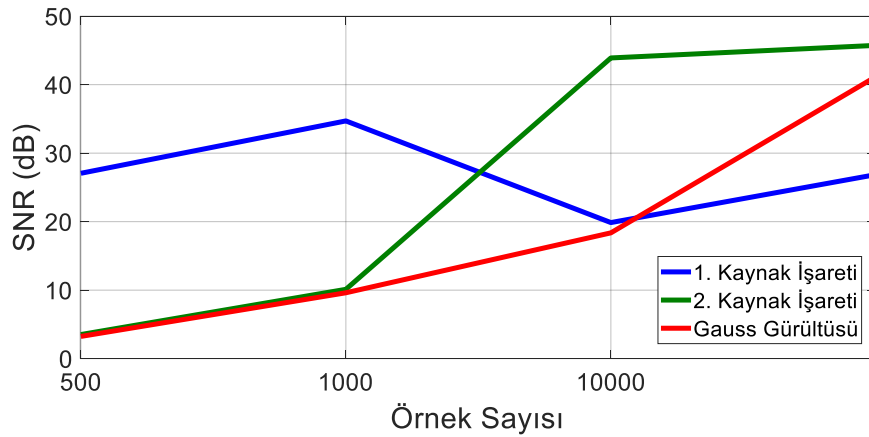


Şekil 1. Kaynak İşaretleri (1. ve 2. Sinyal Konuşma, 3. Sinyal Gauss Gürültüsü)

Her bir algoritma için kaynak işaretleri ayrı ayrı gösterilmiştir. Şekil 2'de PICA algoritmasının çalıştırılması ile karışmış kaynak işaretlerinin SNR değerleri gösterilmiştir. Sonuçlar incelendiğinde kaynak işaretlerinin belirli bir veri boyutuna kadar SNR değerinin ortalama olarak sabit olduğu, sonrasında ise veri boyutu arttığında SNR değerinin de arttığı gözlemlenmiştir. Özellikle gauss gürültüsünü seslere nazaran daha iyi ayırtmıştır.



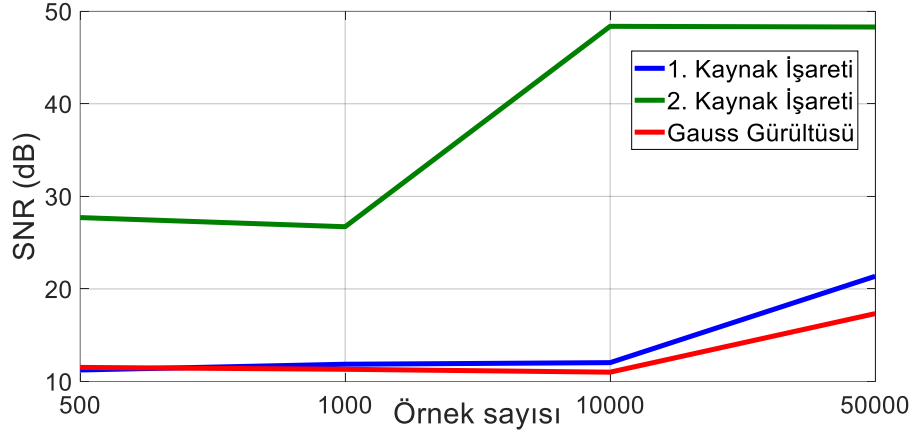
Şekil 2. PICA Algoritması SNR Değerleri



Şekil 3. SOBI Algoritması SNR Değerleri

Şekil 3 incelendiğinde SOBI algoritmasında Gauss gürültüsünü ve 2. kaynak işaretini iyi ayırttığı gözlemlenirken 1. kaynak işaretinde düşüş gözlemlenmiştir. Bu düşüş sesin genliğine göre değişiklik göstermektedir. Genel itibarıyla bakıldığında tüm işaretleri en iyi ayırtıran SOBI algoritmasının olduğu görülmektedir.

Son olarak JADE algoritmasının performansı Şekil 4’de incelendiğinde ise 2. kaynak işaretinin ayrıştırılmasında oldukça başarılı sonuçlar ortaya çıkmıştır. Genel olarak üç algortmada da veri boyutuna bağlı olarak SNR değerlerinin arttığı görülmektedir.



Şekil 4. JADE Algoritması SNR Değerleri

Tablo 1’ de algoritmaların işlem süreleri verilmiştir. Örnek sayısı ile işlem süresi doğru orantılı olarak artmıştır. SOBI algoritmasının işlem süresi diğerlerine göre oldukça başarılı olduğu görülmektedir.

Tablo 1. Algoritmaların Çalışma Süreleri

Örnek Sayısı	PICA	SOBI	JADE
500	2.05 sn	0.45 sn	0.91 sn
1000	2.86 sn	0.56 sn	1.22 sn
10000	5.24 sn	1,29 sn	2.17 sn
50000	7.13 sn	1.53 sn	3.08 sn

4. Sonuç

Kör kaynak ayırmada birçok yöntem olmakla beraber bu çalışmada üç yöntem karşılaştırılmış ve başarı oranları ile birlikte işlem süreleri incelenmiştir. Karışmış kaynak işaretlerinin ayrıştırılmasında kullanılan algoritmalar içerisinde örnek sayısına oranla PICA ve SOBI algoritmasının daha iyi sonuçlar verdiği gözlemlenmiştir. Diğer iki algoritmaya göre SOBI algoritmasındaki değişimler daha keskindir. Dikkat edilirse kaynak sinyallerin birbirinden ayrılması her algoritma için çok farklı sonuçlar çıkarmıştır. Birinde gürültüyü çok iyi ayrıştırırken diğerinde ikinci kaynağı çok iyi ayrıştırmıştır. İşlem süresi bakımından incelendiğinde ise SOBI algoritması, JADE algoritmasına göre iki kat, PICA algoritmasına göre ise dört kat daha kısa sürede işlem sonucu verdiği kanıtlanmıştır.

Kaynaklar

- Bach, F. R., & Jordan, M. I. (2002). Kernel independent component analysis. (s. 1-48). Journal of machine learning research.
- Belouchrani, A., Abed-Meraim, K., Cardoso, J. -F., & Moulines, E. (1993). Second-order blind separation of temporally correlated sources. *In Proc. Int. Conf. Digital Signal Processing*, (s. 346-351). Citeseer.
- Belouchrani, A., Abed-Meraim, K., Cardoso, J. -F., & Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, (s. 434-444).
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review on speech intelligibility in multiple-talker conditions. (s. 117-128). Acta Acustica united with Acustica.
- Cardoso, J. F., & Souloumiac, A. (1993). Blind beamforming for non-Gaussian signals. *IEE proceedings F (radar and signal processing)* (s. 362-370). IET Digital Library.

- Celik, H., Ilgin, F. Y., & Sevim, Y. (2019). Müzik işaretlerin tek kanal kör kaynak ayrıştırma ile ayrıştırılması. (s. 26-38). NWSA Engineering Sciences.
- Hyvärinen, A. (1984). Independent component analysis: recent advances. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, (s. 371).
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. (s. 411-430). Neural networks.
- Ilgin, F. Y., Celik, H., & Sevim, Y. (2012). Stable signal separation algorithm for ECG signals. *2012 International Symposium on Innovations in Intelligent Systems and Applications* (s. 1-4). IEEE.
- Karvanen, J., Eriksson, J., & Koivunen, V. (2000). Pearson system based method for blind separation. *Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation (ICA2000)*, (s. 585-590). Helsinki, Finland.
- Kato, H., Nagahara, Y., Araki, S., Sawada, H., & Makino, S. (2006). Parametric-Pearson-based independent component analysis for frequency-domain blind speech separation. *14th European Signal Processing Conference*, (s. 1-5). Florence.
- Liu, X., Wang, H., Huang, M., & Yang, W. (2019). An improved second-order blind identification (SOBI) signal de-noising method for dynamic deflection measurements of bridges using ground-based synthetic aperture radar (GBSAR). *Applied Sciences*, (s. 3561-3561).
- Miettinen, J., Nordhausen, K., & Taskinen, S. (2017). Blind source separation based on joint diagonalization in R: The packages JADE and BSSasymp. *Journal of Statistical Software*.
- Rainieri, C. (2014). Perspectives of second-order blind identification for operational modal analysis of civil structures. (s. 1-9). Shock and Vibration.
- Sahonero, G., & Calderon, H. (2017). A comparison of SOBI, FastICA, JADE and infomax algorithms. *Proceedings of The 8th International Multi-Conference on Complexity*.
- Tan, L. (2007). *Digital Signal Processing: Fundamentals and Applications*. Academic Press.