

İĞNECİKLI SİNİR AĞLARI İLE SES KAYNAĞI SINIFLANDIRMA

Peren J. CANATALAY¹

Osman Nuri UÇAN²

Özet

Bu çalışmada, mikrofon dizilerinden alınan verilerin işlenmesi ile ses kaynağının yer tayini üzerine çalışılmıştır. Ses kaynağı konumu belirlemede başarıyı en yüksek yöntemlerden biri olan faz dönüşümlü yönlendirilmiş cevap gücü algoritması üzerine yoğunlaşmıştır. Mikrofon dizileri ile yapılan kayıtlarda kaynakların ayrı ayrı dinlenebilmesi için odaklama yapılması gerekmektedir ve bunun için de ses kaynağı konumunun kesin bir şekilde bulunabilmesi gerekmektedir. Bu amaçla, konum bilgilerinin değerlendirilmesi ve yer tayini işlemi için iğnecikli sinir ağı ile sınıflandırma yapılmıştır. Elde edilen sonuçlar, oldukça yüksek bir başarıya işaret etmektedir. Belirlenen konumlar, bilinen gerçek konuşmacı konumları ile oldukça uyumludur ve hata oranları birkaç santimetreyi aşmamaktadır.

Anahtar Kelimeler: İğnecikli sinir ağı, Ses kaynağı yer bulma

¹ İstanbul Aydın Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü

² İstanbul Aydın Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü

I. GİRİŞ

Ses ve/veya konuşma kaydetmede ortam koşulları oldukça belirleyici olan olumsuz etkilere sahiptir. Bu etkilerin bertaraf edilmesi birden çok mikrofon kullanılması sayesinde sağlanabilmektedir. Birden çok mikrofonun kullanımı, mikrofonların yerleşim biçimlerine göre ortam içerisindeki tüm sesleri duyabilme olanağı vermektedir. Ancak, ortamdaki tüm seslerin duyulabilmesi, tüm seslere hakim olma anlamına gelmemektedir. Zira ortamda birden çok ses kaynağı bulunması ve/veya ortamın yankılama etkisinin yüksek olması gibi durumlarda sesler birbirine karışmakta, kimi sesler örtülmekte, geri planda kalmakta ve anlaşılammamaktadır. Bu durum genel olarak kokteyl partisi problemi olarak adlandırılmaktadır. Çözüm ise dizisel işaret işleme tekniklerinin çoklu mikrofon kayıtlarına uygulanmasıdır.

Bir diğer önemli problem de ses kaynağı ve/veya konuşmacının hareketli olması durumu veya bir başka ifadeyle kayıt esnasında yer değiştiriyor olması durumudur. Bu durumda da kayıt edilen ses ve/veya konuşmanın kalitesi sabit olamamakta ve zaman içerisinde çeşitli zayıflamalara maruz kalmaktadır. Nitekim her farklı konum, farklı gürültü kaynakları, bozucular ve/veya yansıtıcı nesnelere beraberinde getirmektedir. Bu tür bir problemin çözümü de yine dizisel işaret işleme yöntemlerine başvurarak çoklu mikrofon kayıtlarının işlenmesi olarak ortaya çıkmaktadır.

Dizisel işaret işleme, temel olarak birçok algılayıcıdan alınan verilerin, algılayıcıların uzaysal yerleşimlerine göre işlenmesi olarak

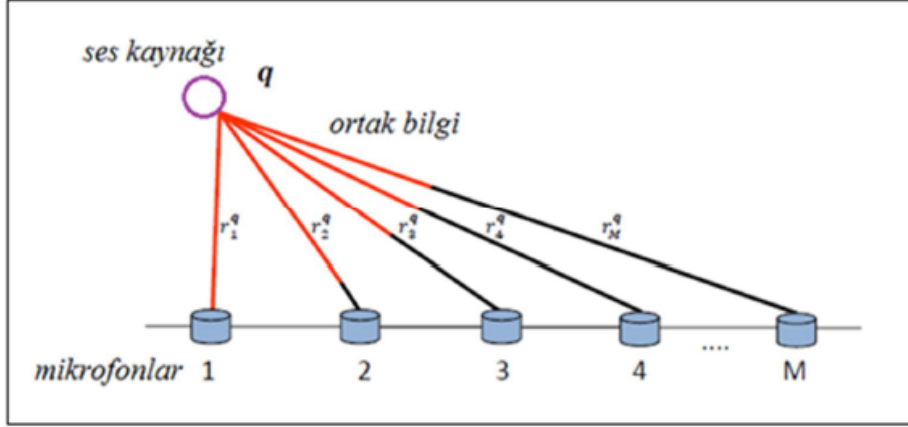
tanımlanabilir. Sonuçta hedeflenenler ise genel olarak kaynakların sayı ve konumlarının belirlenmesi, işaret gürültü oranının iyileştirilmesi ve birden fazla hareketli kaynağın takibi olarak özetlenebilir. Radar, sonar, sismik araştırmalar, tomografi ve kablosuz iletişim gibi alanlarda kullanılan dizisel işaret işleme, hem alıcı hem de verici dizileri için uygulanabilmektedir. Bu durumda yayılmak istenen işaretler belirlenen yönelimdeki belirlenen hedefe ulaştırılabilmektedir.

Dizisel işaret işlemenin gerçekleştirilebilmesi için bazı ön kabuller yapılmaktadır. Bunlardan ilki, kaynaktan çıkan enerjinin tüm yönlere düzenli bir şekilde yayıldığı ve ortamın enerjisi dağıtmadığı varsayılmıştır. Bir diğer ön kabul ise, işaretle birlikte ortalaması sıfır olan bir beyaz gürültünün varlığıdır ki bu beyaz gürültü işaretle ilintili değildir.

II. YÖNTEM

Bir mikrofon dizisi, birden çok mikrofondan oluşan ve kayıt edilen akustik işaretlerden uzaysal bilginin elde edilmesini sağlayacak şekilde konumlandırılan bir çoklu algılayıcı sistemidir [13]. Akustik kayıt alma sistemi olarak kullanılan bir mikrofon dizisi ile gelen ses dalgalarının uzaysal özellikleri verimli bir şekilde ortaya çıkarılabilir ve bu işlem tek bir mikrofon ile mümkün olmayan bir işlemdir. Bu işlemin gerçekleşebilmesi mikrofon dizisinin her bir elemanına gelen ses dalgaları arasındaki zaman farkına dayanmaktadır. Varıştaki zaman farkı olarak adlandırılabilen bu olgu, mikrofonlara ulaşan bilgi işaretleri arasında meydana gelen farklı zaman gecikmeleri olarak

açıklanmaktadır. Şekil 1’de de gösterilen bu olgu, mikrofonlar arasındaki fiziksel uzaklıkların bir sonucudur ve dizisel işaret işleme algoritmalarının üzerine bina edildiği, algılayıcı dizilerinin pek çoğu için geçerli olan en temel özelliktir.



Şekil 1 Ses dalgasının mikrofon dizisine farklı zamanlarda varışı

Sesin yayılımını incelemek ve ses kaynağı yer tespiti gibi uygulamalarda ses dalgasının modellenmesi gerekmektedir. Pratik uygulamalar için makul olabilecek ve geçerliliği gösterilmiş bir model ile yola çıkmak temel teşkil edecektir.

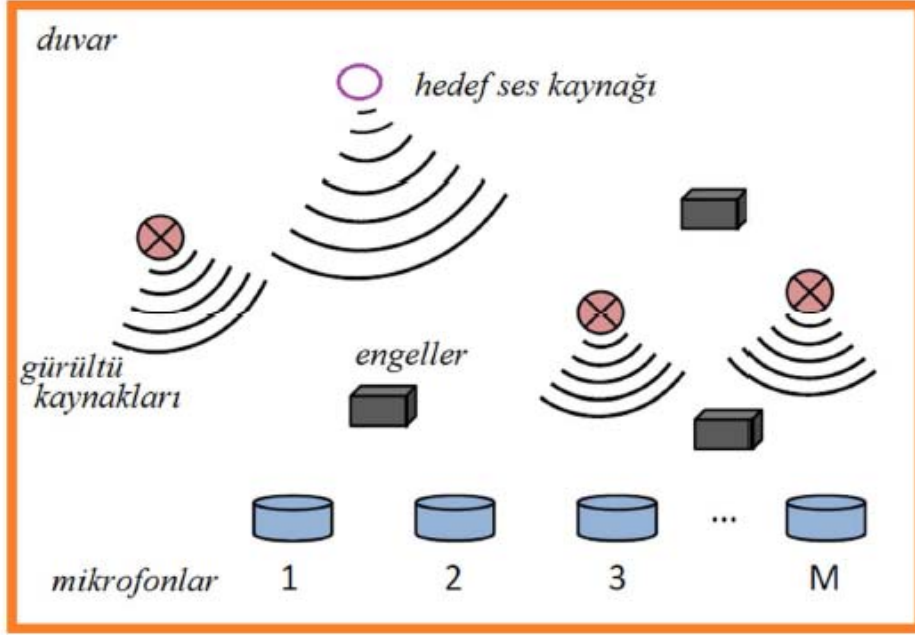
Ses kaynağının yerini bulma ve/veya konuşma iyileştirme gibi uygulamalar söz konusu ise, çoğunlukla kullanılacak yöntem demet oluşturma ve/veya türevleri olan algoritmalarıdır. Demet oluşturma teknikleri temelde mikrofon dizisinin bir hedef noktasına odaklanması ve dolayısı ile hedeflenen bu kaynak noktasından gelen işaretleri almasıdır. Hedef dışındaki tüm uzaydan gelen işaretler zayıflatmaya ve/veya

sıfırlanmaya tabi tutularak girişim ve bozulmalar engellenir. Bu yöntem uzaysal filtreleme olarak da adlandırılır ve sonuçta elde edilen işaret demet oluşturma işleminin yönlendirilmiş cevabı olarak nitelendirilir [13].

A. Faz Dönüştümlü Yönlendirilmiş Cevap Gücü

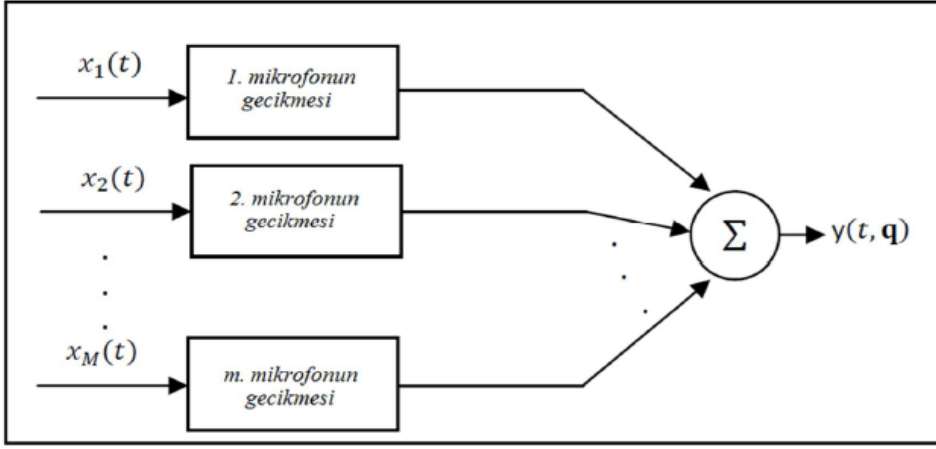
Şekil 2’de örnek bir kayıt ortamında hedef ses kaynağı, engeller, gürültü kaynakları ve mikrofon dizisi görülmektedir. Bu ortam oldukça gerçekçi bir şekilde betimlenmiştir. Ses kaynağından yayılan akustik dalgalar duvarlardan ve yayılıma engel nesnelere yankılanacak, çeşitli gürültü kaynaklarından çıkan seslerden dolayı girişime maruz kalacak ve mikrofon dizisine oldukça bozulmuş ve/veya karışmış bir halde ulaşacaktır.

Tüm bu olumsuz koşulların varlığına karşın, demet oluşturma yöntemlerinin kullanımı ile hedef ses kaynağının yeri belirlenebilmekte ve/veya konuşma üzerinde işlem yapılabilir [15]. Elbette ki burada hangi demet oluşturma tekniğinin kullanılacağı da önem kazanmaktadır, nitekim bu çalışmada bu tekniklerden başarı oranı en yüksek olan yöntemlerden biri olan faz dönüştümlü yönlendirilmiş cevap gücü yöntemi üzerinde durulacaktır.



Şekil 2 Gerçekçi bir ses kayıt ortamı

Demet oluşturma, kısaca özetlenirse uzaydaki bir nokta için yönlendirilmiş bir cevap bulma işlemidir [13]. Bu içerikte pek çok algoritma ortaya konmuştur. Bu algoritmalar arasında en basit ancak başarılı olan yöntem geciktir ve topla tekniğidir. Adından da anlaşıldığı üzere, bu teknikte her bir mikrofondan alınan işarete mesafe farklılıklarını telafi edecek gerekli gecikmeler uygulanır ve böylece tüm işaretler zamansal olarak üst üste getirilmiş olur. Üst üste gelen ses sinyalleri toplandığında ve/veya ortalaması alındığında ise her bir orijinal ses işaretinden daha zengin ve daha temiz bir işaret elde edilmiş olur. Geciktir ve topla algoritması temsili olarak Şekil 3'te gösterilmiştir. Burada $y(t,q)$ ile gösterilen yönlendirilmiş cevap işaretidir ve hem mesafeye hem de zamana bağlı olan bir fonksiyon olarak ifade edilmiştir.



Şekil 3 Geciktir ve topla tekniği

Ses kaynağı yer bulma algoritmalarında, bulunan ya da belirlenen konumun doğru konum olup olmadığı, algoritmanın performansını büyük ölçüde etkiler. Zira elde edilen işaretin enerjisi yalnızca konum doğru olarak bilindiğinde maksimum olacaktır. Demet oluşturulan konum için çıkış işaretinin enerji karakteristiğinden yararlanma da bu kapsamda geliştirilen bir yöntemdir. Çıkiştaki işaretin enerji karakteristiğinden yararlanmak için de belirlenen ya da bulunan konum için yönlendirilmiş cevap gücü hesaplanır. Belirli bir q konumu için yönlendirilmiş cevap gücü (1) denklemi ile ifade edilir.

$$P(\mathbf{q}) = \int_{-\infty}^{\infty} |y(t, \mathbf{q})|^2 dt \quad (1)$$

Yönlendirilmiş cevap gücü, frekans domeninde de tanımlıdır. Zira Parseval teoremine göre tüm t zamanı boyunca toplanan dalğanın içerdiği toplam enerji, tüm ω frekans bileşenleri üzerinden toplanan Fourier

dönüşümü dalgasının sahip olduğu toplam enerjiye eşittir. Bu durumda frekans domeninde tanımlanan yönlendirilmiş cevap gücü (2) denklemi ile hesaplanır.

$$P(\mathbf{q}) = \int_{-\infty}^{\infty} |Y(\omega, \mathbf{q})|^2 d\omega \quad (2)$$

Eğer $Y(\omega, \mathbf{q})$ ifadesinin yerine denklemin açılımı konursa, filtre eklenmiş geciktir ve topla algoritmasının yönlendirilmiş cevap gücünü elde edilmiş olunur, ve bu da (3) denkleminde gösterilmiştir.

$$P(\mathbf{q}) = \int_{-\infty}^{\infty} \left| \sum_{m=0}^{M-1} W_m(\omega) X_m(\omega) e^{-j\omega\tau_{am}^{\mathbf{q}}} \right|^2 d\omega \quad (3)$$

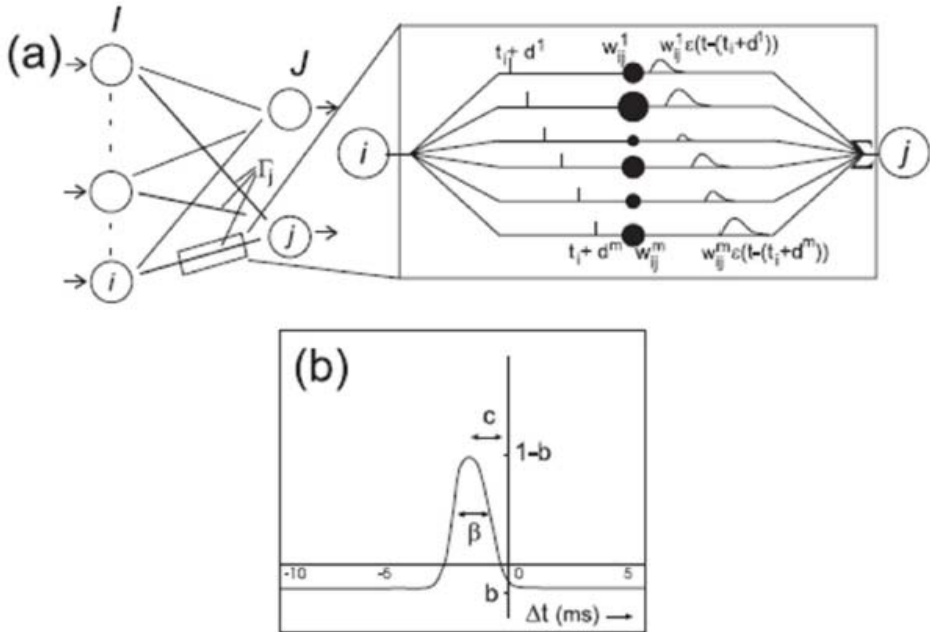
Eğer (3) denklemindeki $W_m(\omega)$ filtresi, faz dönüşümü işlevini yapan bir filtre olarak seçilir ise, elde edilen sonuç faz dönüşümlü yönlendirilmiş cevap gücü olacaktır. Bu yaklaşım [16] numaralı referansta öne sürülmüş ve gürültülü ve/veya yankılı ortamlarda ses kaynağının yerini bulmak için halen bilinen en iyi yöntemlerden biri olma özelliğini elinde bulundurmaktadır. Faz dönüşüm filtresi (4) denklemi ile hesaplanabilir.

$$W_m(\omega) = \frac{1}{|X_m(\omega)|} \quad (4)$$

Bu dönüşüm (3) denkleminde yerine konursa, frekans domeninde faz dönüşümlü yönlendirilmiş cevap gücü ifadesi elde edilmiş olur.

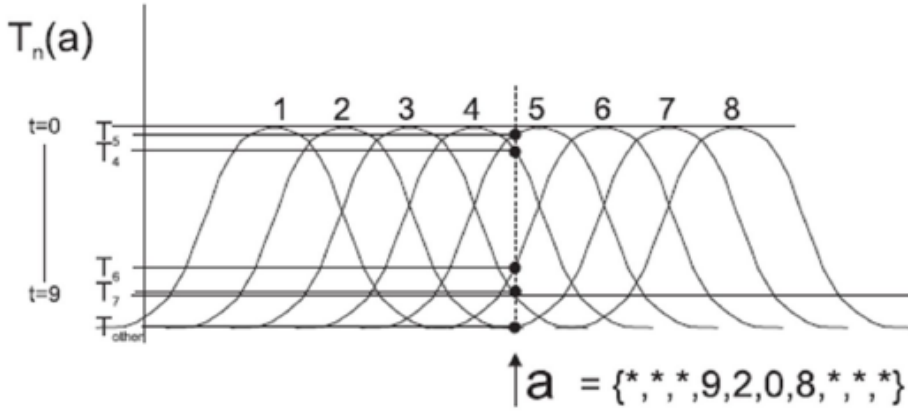
B. İğnecikli Sinir Ağı

Literatürde verilen iğnecikli sinir ağı yapısı ve ilgili yöntemler [22], [23], [24] kullanılmıştır. Buna göre ağ yapısında geleneksel ileri beslemeli yapay sinir ağlarında olduğu gibi, katmanlar arası tam bağlantı kullanılmıştır. İlk katmanda sinir hücreleri girişe göre alıcı alan sinir hücrelerinden oluşmakta ve ikinci katmanda ise çıkış sinir hücreleri radyal taban işlevli sinir hücrelerinden oluşmaktadır.



Şekil 4 Ağ yapısı ve adaptasyon

Sinirler arası çoklu bağlantı (sinaps) kullanılan bu yapı Şekil.4 (a)'da kesikli olarak ayrıntıda gösterilmiştir ve buradaki her alt bağlantıdaki gecikme farklıdır ve bu da toplam cevabı etkilemektedir. Sonuç olarak, bağlantı sonrası potansiyel, alt bağlantı potansiyellerinin zamanda toplanması ile elde edilmektedir [25]. Gözetimsiz sınıflandırma aşamasında temel olarak Bohte'nin çalışmalarında verilen [22], [23] yöntem ve yapı kullanılmıştır. Buna göre ağ yapısında geleneksel ileri beslemeli yapay sinir ağlarında olduğu gibi, katmanlar arası tam bağlantı kullanılmıştır. İlk katmanda sinir hücreleri girişe göre alıcı alan sinir hücrelerinden oluşmakta ve ikinci katmanda ise çıkış sinir hücreleri radyal taban işlevli sinir hücrelerinden oluşmaktadır.



Şekil 5 Alıcı alanlar

Sinirler arası çoklu bağlantı (sinaps) kullanılan bu yapı Şekil.4 (a)'da kesikli olarak gösterilmiştir ve buradaki her alt bağlantıdaki gecikme farklıdır ve bu da toplam cevabı etkilemektedir. Sonuç olarak, bağlantı

(sinaps) sonrası potansiyel, alt bağlantı potansiyellerinin zamanda toplanması ile elde edilmektedir.

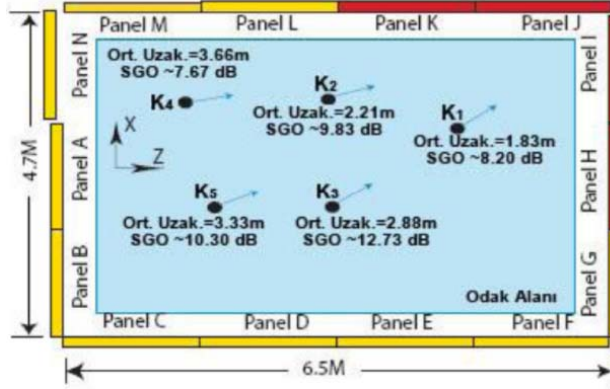
Giriş katmanındaki sinir hücreleri, çakışan Gaussçu alıcı alanlar ile zamanda kodlama yapmaktadır. 8 sinir hücresi için bir kodlama örneği Şekil.5'te gösterilmiştir. Burada "a" ile gösterilen giriş değişkeni örnek bir veri noktasına karşılık gelmektedir ve her bir sinir hücresi için ateşleme zaman gecikmesine çevrilmiştir. Buna göre en büyük değeri alan 5. Sinir 0'a en yakın zamanda ateşlenirken en küçük değeri alan 7. sinir en son ateşlenmiş olacaktır.

III. UYGULAMA

Bu bölümde, telekonferans tipi bir uygulamadan elde edilen veriler kullanılarak, önerilen faz dönüşümlü yönlendirilmiş cevap gücü yöntemi ile yer bulma ve iğnecikli sinir ağı yöntemi ile de ses kaynaklarının sınıflandırılması test edilmiştir.

Kullanılan veri Brown Üniversitesi, LEMS Laboratuvarı (Laboratory for Engineering Man/Machine Systems), tarafından sağlanmıştır. Bu veride 5 ayrı konuşmacı bir oda içerisinde konuşmalar yapmakta ve çevresel olarak yerleştirilmiş 24 adet mikrofondan oluşan mikrofona dizisi ile eş zamanlı olarak kayıt yapılmaktadır. Bu kayıtlarda konuşmalar doğal olarak birbirlerine karışmakta ve anlaşılabilir değildir. Telekonferans uygulamaları için örnek olabilecek bu kayıt ortamı Şekil 6'da görülmektedir. Burada konuşmacıların ortalama uzaklıkları ve desibel cinsinden çıkan seslerin güçleri de yaklaşık olarak verilmektedir. Alınan

ses kayıtları 20 kHz ile örneklenmiş ve kayıt yaklaşık 20 saniye kadar sürdürülmüştür.



Şekil 6 Verinin toplandığı kayıt ortamı

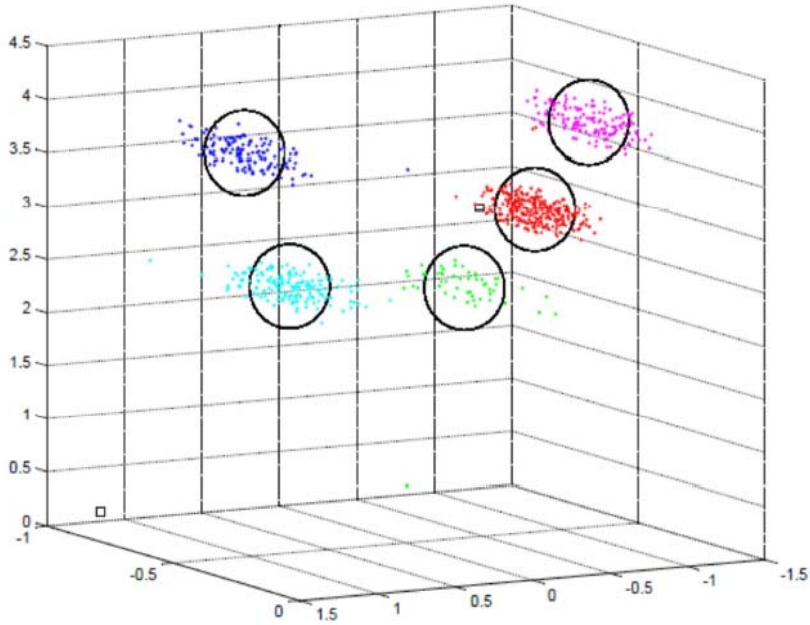
Yöntem kısmında, burada kullandığımız yer bulma algoritması olan faz dönüşümlü yönlendirilmiş cevap gücü tanımlanmış ve formülize edilmişti. Bu kısımda yukarıda açıklanan veri kümesine ilk aşamada faz dönüşümlü yönlendirilmiş cevap gücü - yer bulma algoritması kullanılarak sınıflandırmada kullanılacak ikincil veri kümesi elde edilmektedir.

Veri kümesinden her 1024 örnekleme bir olmak üzere 512 örneklemlerle ilerleyerek kısa dönemli ses verisi yer bulma algoritmasına uygulanmıştır. Bu durumda 20 kHz ile örneklenen veri kümesinden toplamda 780 adet kısa dönem ses kayıt verisi elde edilmiştir.

Ortamda 5 adet konuşmacı bulunmakta ve bu konuşmacılar, konuşmalarında farklı içerikleri farklı zamanlama ve farklı vurgulamalar

ile gerçekleştirmektedirler. Bundan dolayı her bir kısa dönem için çalıştırılan yer bulma algoritması, her bir çıktısında öne çıkan, dolayısı ile farklı, konumları tespit etmektedir. Bu konumlar o an için sesi, vurgusu veya zamanlaması baskın olan konuşmacıya ait olan konumlar olarak değerlendirilmelidir.

Sonuç olarak mevcut örnekleme frekansı ve kayıt süresi ile seçilen kısa dönem periyotları, 780 adet kısa dönemli ses örnekleme cümlesine dönüştürülmüş, tüm bu kısa dönemli örneklemlerin her biri için faz dönüşümlü yönlendirilmiş cevap gücü - yer bulma algoritmasının çalıştırılması ile belirlenen konumlardan da 780 adet 3'er eksenli kartezyen koordinat verisi elde edilmiştir.



Şekil 7 Konumların sınıflandırılması

Bu ikincil veride konuşmacıların kartezyen koordinatlarının gürültü, anlık yönelimler, karşılıklı etkileşim ve girişimler gibi nedenlerden ötürü dağınık ve düzensiz bir haldedir. Oysa ki, demet oluşturarak en iyi sesi alabilmek için konum bilgisinin gerçeğe en yakın konumu göstermesi gerekmektedir. Bu nedenle bu ikincil veri üzerinde bu aşamada bir sınıflandırma yapılması ve konuşmacıların kesin yerlerinin belirlenmesi gerekmektedir.

Sınıflandırma aşamasında dört, beş ve altıncı bölümlerde anlatılan iğnecikli sinir ağı kullanılmıştır. Yöntem bölümünde gösterilen örnek uygulamadaki yapı ve algoritma kullanılmış, ek olarak çıkış katmanında 5 konuşmacıyı temsilen 5 adet radyal tabanlı hücreye yer verilmiştir.

	X (metre)	Y (metre)	Z (metre)
1. Konuşmacı	0.0368	0.0101	0.0046
2. Konuşmacı	0.0058	0.0224	0.0208
3. Konuşmacı	0.0222	0.0191	0.0563
4. Konuşmacı	0.0797	0.0124	0.0245
5. Konuşmacı	0.0037	0.0436	0.0610

Tablo 1 Belirlenen konumlar ile gerçek konumlar arasındaki farklar

Veri kümesi, iğnecikli sinir ağı girişine uygulandıktan sonra çıkışta beş merkeze kümelenen konum bilgileri Şekil 7’de görüldüğü gibi elde edilmiştir. Bu kümelenen konum bilgilerinin ortalamaları alınarak nihai konum bilgilerine ulaşılmıştır. Tespit edilen konum bilgileri gerçek

konum değerlerine oldukça yakındır, zira Tablo 1’de hatalar gösterilmiştir.

IV. SONUÇ

Bu çalışmada, birçok ayrı alanda uygulanabilirliği gösterilmiş olan, dizisel işaret işleme alanının akustik veri işleme üzerine çalışılan kolu olan mikrofon dizileri ile ilgilidir. Mikrofon dizileri ile, teorik olarak mikrofon sayısı kaynak sayısından büyük eşit olmak üzere, ses kaynaklarının kartezyen koordinatları oldukça yüksek bir hassasiyet ile bulunabilmektedir.

Ses kaynağı konumu belirlemede başarıyı en yüksek yöntemlerden biri olan faz dönüşümlü yönlendirilmiş cevap gücü algoritması üzerine çalışılmış ve uygulama ve benzetim bölümünde de bu algoritmaya yer verilmiştir. Gerçek zamanlı uygulamalar göz önünde bulundurularak algoritmaya 512 örneklem adımları ile 1024 örnekle bölümler halinde girdiler uygulanmıştır. 20 kHz’lik örnekleme frekansı düşünüldüğünde 1024 örnek yaklaşık olarak 0.05 saniyeye karşılık gelmektedir. Bu da, algoritmanın 25 milisaniyelik adımlar ile çalıştırıldığı anlamına gelmektedir. Bir telekonferans örneği düşünüldüğünde bu süre gerçek zamanlı kontrol için oldukça uygundur.

Ses kaynağı yer bulma algoritması çıktı olarak 3 boyutlu uzayda Kartezyen koordinat bilgisi vermektedir. 25 milisaniyelik adımlarla alınan 50 milisaniyelik veri parçaları her periyotta farklı konum bilgilerini çıktı olarak vermektedir çünkü konuşmacıların konuşma

zamanları, tonlamaları, ses yükseklikleri birbirinden farklıdır. Bunun yanı sıra ortamda gürültü ve yankı gibi bozucu etkenler de bulunmaktadır.

Mikrofon dizileri ile yapılan kayıtlarda kaynakların ayrı ayrı dinlenebilmesi için odaklama yapılması – dolayısı ile demet oluşturulması gerekmektedir ve bunun için de ses kaynağı konumunun kesin bir şekilde bulunabilmesi gerekmektedir. Bu amaçla faz dönüşümlü yönlendirilmiş cevap gücü algoritmasından yukarıda belirtilen sıklıklarda alınan konum bilgilerinin tekrar değerlendirilerek ses kaynağı konumlarının tayin edilmesi gerekmektedir. Bu yeniden değerlendirme işlemi için iğnecikli sinir ağı ile sınıflandırma yapılması uygun görülmüştür. Bu şekilde gerçek zamanlı uygulamalara uygunluk gözetilerek sıralı öğrenme tekniğine yer verilen iğnecikli sinir ağı da aynı zamanda paralel olarak işleyerek yer bulma algoritmasının çıktıları olan 3 boyutlu konum bilgilerini girdi olarak almaktadır. İğnecikli sinir ağının çıktısı olan 5 öbek ve harici veri noktaları da belli periyotlarla değerlendirilerek bir sınıflandırma gerçekleştirilmektedir. Sınıflandırma sonucunda belirlenen öbeklerin ortalaması 3 boyutlu uzaysak koordinatlarda konuşmacının konumu olarak tayin edilmektedir.

Elde edilen sonuçlar, oldukça yüksek bir başarıya işaret etmektedir. Belirlenen konumlar, bilinen gerçek konuşmacı konumları ile oldukça uyumludur ve hata oranları birkaç santimetreyi aşmamaktadır. Bu çalışmanın, kurulması amaçlanan bir deney düzeneğinden elde edilecek gerçek veriler ile doğrulanması ve geliştirilmesi planlanmaktadır.

KAYNAKÇA

- [1] P Stoica, R Moses. Spectral Analysis of Signals, NJ: Prentice Hall, 2005
- [2] Jha, RakeshMohan, Trends in Adaptive Array Processing, 2012
- [3] Viberg, Mats, Sensor Array Signal Processing: Two Decades Later, 1995
- [4] Visser H., Array and Phased Array Antenna Basics. New Jersey: Wiley, 2005
- [5] McGowan I., Microphone Arrays: A tutorial, April 2001.
- [6] Kushleyev A., Sound Localization. University of Maryland, Institute for Systems Research, Summer 2004.
- [7] Gregory E., Cole J., Audio Localization. Connexions Project, 2004.
- [8] Kazuhiro N., Hiroshi O., Hiroaki K., Auditory Fovea Based Speech Enhancement and its Application to Human-Robot Dialog System. Kyoto University, Graduate School of Informatics.
- [9] [9] Basu S., Schwartz S., Pentland A., Wearable Phased Arrays for Sound Localization and Enhancement. MIT Media Laboratory, IEEE 2000. EECE 503, Audio Engineering. Course Notes.
- [10] Zeren A., Müzik Fiziği, Pan Yayıncılık, 2003
- [11] Beranek L., Noise Reduction Precared for a Special Summer Program at the Massachusetts Institute of Technology. New York: McGraw-Hill, 1960
- [12] Benesty J., Chen J., Huang Y., Microphone Array Signal Processing, Springer,2008

- [13] Johnson D. H., Dudgeon D. E., Array Signal Processing: Concepts and Techniques, Prentice Hall, 1993
- [14] Brandstein W., Ward D., Microphone Arrays: Signal Processing Techniques and Applications, Springer, 2001
- [15] DiBiase J. H., A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays, Brown Üniversitesi, 2000
- [16] W.S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5:115–133, 1943
- [17] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by back-propagating errors. Nature, 323:533–536, 1986.
- [18] D.O. Hebb. The Organization of Behaviour. Wiley, New York, 1949.
- [19] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci., 79(8):2554–2558, 1982.
- [20] R. Van Rullen and S. Thorpe. Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. Neural Computation, 13:1255–1283, 2001.
- [21] A.L. Hodgkin and A.F. Huxley. A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. J. of Physiology, 117:500–544, 1952
- [22] L. Lapicque. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. J. Physiol. Pathol. Gen.,

- 9:620–635, 1907. cited by Abbott, L.F., in *Brain Res. Bull.* 50(5/6):303–304.
- [23] D.I. Standage and T.P. Trappenberg. Differences in the subthreshold Dynamics of leaky integrate-and-fire and Hodgkin-Huxley neuron models. In *IJCNN'2005, Int.Joint Conf. On Neural Networks*, pages 396–399. IEEE-INNS, 2005.
- [24] W.. Gerstner. Time structure of the activity in neural network models. *Physical Review E*, 51:738–758, 1995.
- [25] W.M. Kistler, W. Gerstner, and J.L. van Hemmen. Reduction of Hodgkin-Huxley equations to a single-variable threshold model. *Neural Computation*, 9:1015–1045, 1997.
- [26] W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531–2560, 2002.
- [27] H. Jaeger. Tutorial on training recurrent neural networks, covering BPTT, RTRL, EKF and the “echo state network” approach. Technical Report TR-GMD-159, German National Research Center for Information Technology, 2002.
- [28] W. Maass. Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, 10:1659–1671, 1997
- [29] Haykin S., (1994) *Neural Networks*, Macmillan, New York.
- [30] W. Gerstner and W. Kistler ‘Spiking Neuron Models – Single Neurons, Populations, Plasticity’ Cambridge Univ. Press (2002)