

PAPER DETAILS

TITLE: Genetik Algoritma ile Mikrofon Dizilerinde Ses Kaynaginin Yerinin Bulunmasi

AUTHORS: Ekrem ÇONTAR,Hasan BILGE

PAGES: 0-0

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/75229>

Genetik Algoritma ile Mikrofon Dizilerinde Ses Kaynağının Yerinin Bulunması

Ekrem Çontar, Hasan Şakir Bilge

Bilgisayar Mühendisliği Bölümü, Gazi Üniversitesi, Ankara
ekremcontar@hotmail.com, bilge@gazi.edu.tr

Özet— Mikrofon dizileri birden fazla mikrofonun farklı uzaysal noktalara yerleştirilmesi ile elde edilen yapılardır. Farklı mikrofonlardan alınan ses sinyallerinin uygun zaman gecikmeleri ile ötelenip toplanması ile hüzme şekillendirme yapılır. Hüzme şekillendirmede yapılan yönlendirme ve odaklama işlemleri ile bir çeşit uzaysal süzgeçleme gerçekleştirilmiş olur. Böylece konuşmacının sesi, gürültü ve girişimlere göre daha fazla artırılmaktadır. Bu amaçla, konuşmacının yerinin (uzaysal koordinatlarının) bilinmesi gereklidir. Bu çalışmada, mikrofon dizilerine gelen sesler analiz edilerek genetik algoritma yöntemi ile konuşmacının konumu otomatik olarak bulunmuştur. Burada konuşmacıdan diğer ses kaynaklarına göre daha güçlü sinyal alındığı bilgisinden yararlanılmıştır. Yapılan deneysel çalışmalarda bu yöntem sayesinde çok az bir hatayla konuşmacının yerinin bulunduğu görülmüştür. Birçok uygulama alanı olan genetik algoritma, konuşmacı yerinin bulunması noktasında başarılı sonuçlar elde edilmesini sağlamıştır. Bu çalışmada, değişken genlikli mutasyon yöntemi kullanılması önerilerek performans artırılması sağlanmıştır. Geliştirilen bu sistem telekonferans gibi uygulamalarda etkin bir şekilde kullanılabilecektir.

Anahtar Kelimeler— Mikrofon dizileri, hüzme şekillendirme, genetik algoritma, ses kaynağının yerinin bulunması

Sound Source Localization in Microphone Arrays Using Genetic Algorithm

Abstract— Microphone arrays are sets of microphones placed geometrically at different locations. In delay-and-sum beam-forming, the signals received from different microphones are delayed by certain amounts and summed together. Beam-forming is a kind of spatial filtering that performed by steering and focusing operations. In this way, the speech signal is improved according to noise and interference signals. In order to perform beam-forming using microphone arrays, it is crucial to know the location of the sound source. In this study, the position of the sound source is automatically determined by analyzing the incoming sound signals, using genetic algorithm. In experimental studies, it is seen that it is possible to localize the sound source by using this method. Genetic algorithm, which has various applications, is successfully used for sound source localization problem. In this study, variable length mutation method is presented for performance improvement. The implemented system could be effectively used in many applications such as teleconferencing.

Keywords— Microphone array, beamforming, genetic algorithm, sound source localization

1. GİRİŞ

Ses algılama, ses kayıt ve ses tanıma sistemlerinde yaşanan en büyük problem, ortam şartlarından etkilenen ses sinyalindeki bozulmadır. Ses kaynağına çok yakın olan algılayıcılarda (mikrofonlarda) bu problem nispeten daha azdır. Ancak, yakın mikrofon kullanılmasının uygun olmadığı veya istenmediği birçok durum mevcuttur. Örneğin; bir toplantı salonunda her bir konuşmacı için farklı bir mikrofon bulundurulması pratik değildir. Bunun

yerine, toplantı masasının ortasına yerleştirilmiş veya tavana asılmış bir mikrofon kullanılmaktadır. Toplantı sırasında bazen aynı anda konuşan kişilerin veya karşılıklı konuşmaların olduğu durumlar da mevcuttur. Aracı ile hareket eden bir sürücünün, eller serbest cep telefonu ile görüşmesi diğer önemli bir uygulamadır. Duyma problemi olan bir kişinin, sesleri daha iyi duyabilmesi için konuşmacıdan yakın mikrofon kullanmasını istemesi ise pratik değildir. Bu ve benzeri durumlarda, ses kaynağından yayılan sesin mikrofona ulaşınca kadar

kalitesinde ortama göre değişen miktarlarda bozulma olmaktadır.

Ses kaynağından gelen sesin daha güçlü alınabilmesi amacıyla, birden fazla mikrofondan oluşan mikrofon dizileri kullanılmaktadır. Dizideki her bir mikrofondan alınan sinyallerin toplanması ile, tek bir mikrofondan alınan sinyalden daha güçlü bir sinyal elde edilmesi amaçlanmaktadır. Toplanarak elde edilen ses sinyalinin güçlü olabilmesi için, mikrofonlardan alınan seslere uygun gecikme miktarları uygulanmaktadır. Gecikme uygulanması fikri, sesin ortamda yayılması sırasında farklı mikrofonlara farklı zamanlarda ulaşması teorisine dayanmaktadır. Her bir mikrofon için uygulanacak gecikme miktarı, bu mikrofonun ses kaynağına olan uzaklığı ve mikrofonların birbirine göre konumlarına göre belirlenir. Uygun gecikmeler uygulandığında, elde edilen toplam ses sinyalinin, her bir mikrofondan alınan ses sinyalinden daha kaliteli olacağı beklenmektedir.

Mikrofonlara uygulanacak gecikme miktarının hesaplanabilmesi için öncelikle ses kaynağının yerinin bilinmesi gereklidir. Ancak, toplantı salonu gibi ortamlarda hangi zaman diliminde hangi konuşmacının konuşacağı önceden bilinmeyeceği için, bu bilginin önceden elde edilmesi mümkün değildir. Bu nedenle, gerçek zamanlı olarak ses kaynağının yerinin tespit edilmesi gereklidir.

Mikrofon dizilerinde ses kaynağının gerçek zamanlı olarak belirlenmesi amacıyla değişik yöntemler geliştirilmiştir. Çapraz iltintı yöntemi ve genetik algoritma bunlardan bazılarıdır [1-3].

Bu çalışmada, ses kaynağının yerinin tespit edilmesi amacıyla genetik algoritma kullanılmıştır. Ses kaynağının muhtemel yerleri için uygun gecikme miktarları uygulanarak elde edilen ses sinyalleri üst üste toplanmış ve bu ses sinyalinin en güçlü olduğu uzaysal koordinatın ses kaynağının gerçek yeri olduğu kabul edilmiştir.

Geliştirilen yöntem, daha önceden farklı mikrofon dizileri kullanarak kayıt edilmiş olan ses sinyallerine uygulanmış ve sonuçları değerlendirilmiştir. Optimizasyon yöntemi olarak, genetik algoritma kullanılmıştır. Değişken genlikli mutasyon yöntemi önerilmiş ve bu sayede hesaplama performansının artırılması sağlanmıştır.

2. MİKROFON DİZİLERİ

Mikrofon dizisi, birden fazla mikrofondan oluşan bir yapıdır. Bu yapıda mikrofonlar özel bir şekilde yerleştirilirler; örneğin doğrusal bir hat boyunca, parabolik bir şekil oluşturacak şekilde veya dairesel bir şekilde konumlandırılırlar. Mikrofonlar arası uzaklık aynı miktarda veya farklı miktarda olabilir. En yaygın

uygulama aralarında aynı uzaklık bulunan doğrusal bir hat boyunca yerleştirilen mikrofon dizileridir. Mikrofon dizilerindeki mikrofon sayısı ikiden başlar. Tipik uygulamalarda 4 veya 8 mikrofon kullanılmaktadır. Daha çok mikrofon kullanılan yapılar da mevcuttur. Bir mikrofon dizisinde mikrofon sayısı arttıkça sinyal gürültü oranı (SNR) daha fazla iyileştirilmiş olur.

Mikrofon dizilerinin geniş uygulama alanları bulunmaktadır. Bunlardan biri işitme cihazlarıdır; işitme cihazlarının başarımının artırılmasında mikrofon dizileri kullanılmaktadır [4]. İnsanlar tek kulakla sesin ne taraftan geldiğini anlayamazlar, ancak 2 kulak sayesinde her iki kulağa gelen sesin arasındaki zaman gecikmesinden yola çıkarak beyin ses kaynağının yönünü bulabilir. Bir mikrofon içeren işitme cihazı ile ses kaynağının yönünü belirlemek mümkün olmamaktadır. Oysa mikrofon dizileri ile işitme cihazı kullanan kişilerin ses kaynağının yönünü anlamaları mümkün olmaktadır.

Mikrofon dizileri, ses kaynağının hareketli olduğu uygulamalarda ses kaynağının yerinin gerçek zamanlı olarak izlenmesi için de kullanılmaktadır. Mikrofon dizisine göre yönü ve uzaklığı bulunan ses kaynağına hüzmeye şekillendirme (beamforming) yapılarak, alınan ses sinyalinin kalitesi (SNR) artırılmaktadır.

Diğer önemli bir uygulama, mikrofon dizileri ile ses tanıma sistemlerinin başarımının artırılmasıdır [1]. Birden fazla mikrofon kullanılarak kaydedilen ses sinyalleri ile harflerin/kelimelerin tanınması daha isabetli olmaktadır.

Cep telefonlarının çok yaygın bir şekilde kullanıldığı zamanımızda eller serbest olarak konuşma tercih edilen bir özelliktir. Bu teknolojiye ses kalitesinin artırılması konusunda çalışmalar yapılmaya devam edilmektedir. Özellikle hareketli taşıtların içinde trafik ve araç gürültüsünün yoğun olmasından dolayı eller serbest modunda konuşma, mükemmel bir şekilde yapılamamaktadır. Gürültüsüz ortamlarda bile bu şekilde konuşurken yankı, seslerin algılanamaması gibi çeşitli sorunlar ortaya çıkmaktadır. Cep telefonları ile eller serbest olarak konuşmada kabul edilebilir bir düzeyde ses iletişiminin sağlanması, günümüzün önemli çalışma konularından biridir.

Bir mikrofon dizisi kullanarak istenen yönden gelen sinyaller güçlendirilirken, diğer yönlerden gelen sinyaller zayıflatılarak bastırılmaktadır. Bu nedenle burada kullanılan hüzmeye şekillendirme işlemi, uzaysal süzgeçleme olarak da değerlendirilmektedir. Dar bantlı bir hüzmeye şekillendiricide N adet alıcıdan alınan sinyallerin k zamanındaki çıktısı, alıcılardan alınan sinyallerin toplanması ile hesaplanmaktadır [5].

$$y(k) = \sum_{i=1}^N w_i^* x_i(k) \quad (1)$$

Eşitlik-1’de w her bir alıcı için kullanılabilecek ağırlık fonksiyonunu, “*” işareti ise karmaşık eşleniği göstermektedir. Ağırlıklar, hüzenin (beam) şeklini belirlemek amacıyla kullanılmaktadır. Temel uygulamalarda her mikrofon kanalı için ağırlık fonksiyonu 1 olarak alınabilir.

Geniş bantlı bir hüzenin şekillendiricisi ise y çıktısı aşağıdaki şekilde hesaplanmaktadır [5].

$$y(k) = \sum_{i=1}^N \sum_{p=0}^{K-1} w_{i,p}^* x_i(k-p) \quad (2)$$

Eşitlik-2’de $K-1$, her bir alıcıdaki gecikme miktarını belirtmektedir.

3. MEVCUT ÇALIŞMALAR

Mikrofon dizileri ile ilgili literatürde birçok bilimsel çalışma yayınlanmıştır. Bunlardan biri Hamahara ve ark. tarafından yapılan çalışmadır. Bu çalışmada, “dalga boyu sabit matrisi” diye adlandırılan bir yöntemin ve genetik algoritmanın birlikte kullanılmasına dayalı üç boyutlu bir mikrofon dizisi önerilmektedir [2]. Burada “dalga boyu sabit matrisi” yönteminin uygulanabilirliği gösterilmiş ve 3 farklı ses kaynağının yerinin aynı anda tespit edilebildiği belirtilmiştir.

Mikrofon dizileri konusunda yapılan bir diğer çalışma ise, MIT (Massachusetts Institute of Technology) bilgisayar bilimleri laboratuvarlarında geliştirilen ve 1020 adet mikrofondan oluşan dizi ile yapılan çalışmadır [6]. Bu çalışmada, dizideki mikrofonların, uzaysal konumları (pozisyonları) bilinen ses kaynaklarına hüzenin şekillendirme yöntemi ile yönlendirilmesi sonucu ses kalitesinde artış sağlanması amaçlanmıştır. Hüzenin şekillendirme yöntemi olarak geciktir-ve-topla yöntemi kullanılmıştır. Mikrofon dizisinden kaydedilen ses sinyalinin ses tanıma işlemi gerçekleştirilmiştir. Çalışmada, ses kaynaklarının yerinin bilindiği kabul edilerek bu noktalara hüzenin şekillendirme yapılmıştır. Hüzenin şekillendirme sonucu elde edilen ses bilgisi, ses tanıma algoritması için giriş olarak kullanılmıştır.

Diğer bir çalışmada, Widrow mikrofon dizilerini işleme cihazlarında ses kaynağı yönünün tespit edilmesi ve ses tanıma amacıyla kullanmıştır [4]. Widrow, mikrofon dizilerinin işleme cihazlarında kullanılması sonucunda; sinyal gürültü oranında yaklaşık 10 dB değerinde bir artış sağlandığını, dizinin istenen tarafa yönlendirilmesi sonucu yankılanma etkisinin önemli ölçüde azaldığını, ve geri beslemeden dolayı oluşan gürültünün yaklaşık 15 dB değerinde azaldığını belirtmiştir.

Aguilar ve ark., 2 adet mikrofondan oluşan dizi yardımıyla ses kaynağının yönünü tespit etmeye çalışmışlardır. Mikrofonlar, aralarındaki uzaklık 1 metre olacak şekilde yerleştirilmiştir [7]. Mikrofonlardan alınan ses bilgileri sayısallaştırıldıktan sonra geciktir-ve-topla yöntemi ile üst üste toplanmış, en iyi çıkış sinyalini veren gecikme açısı ses kaynağının açısı olarak kabul edilmiştir. 2 farklı doğrudan arama yöntemi kullanılarak deneysel çalışmalar yapılmıştır. Çıkış sinyalinin maksimum genliği sağlanacak şekilde çözüm aranmıştır. Mikrofon sayısı 2 ve aralarındaki uzaklık 1 metre olduğundan dolayı, 10 KHz frekansında örneklenen sayısal vektörlerde uygulanabilecek gecikme -29 ile 29 arasında hesaplanmıştır. Bu aralıktaki olası tüm değerler başlangıç çözüm kümesine dahil edildiğinden, 2. iterasyonda hızlıca sonuca ulaşılmıştır.

Birinci ve Leblebicioğlu, genetik algoritma yardımı ile mikrofon dizilerindeki mikrofonların ideal yerleşim noktalarını belirlemeye çalışmışlardır [8]. Birden fazla gürültü kaynağının, 1 adet ses kaynağının bulunduğu 100 metre karelik bir odada 8 adet mikrofonun en uygun yerleşim yerini tespit etmişlerdir. Mikrofonlardan elde edilen sinyalleri toplamak amacıyla, geciktir ve topla yöntemine dayalı hüzenin şekillendirici kullanmışlardır.

Hacıvelioğlu ve Dinçer, akıllı anten sistemlerinde işaret geliş açısı kestirim yöntemlerini incelemişlerdir. Hüzenin şekillendirme teknikleriyle beraber kullanılan, işaret geliş açısı (DOA – Direction of Arrival) hesaplama yöntemlerinden; geleneksel hüzenin şekillendirici, MUSIC (Multiple Signal Classification) ve ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) yöntemleri üzerinde performans karşılaştırması yapılmıştır. Geleneksel hüzenin şekillendirici olarak geciktir ve topla yöntemini kullanmışlardır [9]. Sonuçlar incelendiğinde, geleneksel hüzenin şekillendirici algoritmasının düşük çözünürlüklü algoritmalar sınıfında yer aldığı görülmektedir. Çözünürlüğü artırmak için, dizideki anten eleman sayısını arttırmak gerekmektedir. MUSIC ve ESPRIT algoritmaları ise yüksek çözünürlüklü alt uzay tabanlı algoritmalar [9].

Tamer ve ark., temel tahmin algoritmalarından biri olan en büyük olabilirlik yöntemini uygulayarak anten dizisine ulaşan geliş açısını belirlemeye çalışmışlardır [10]. Anten sayısı arttıkça işaretin geliş açısı daha belirgin bir hale gelerek komşu değerlerden ayrılmaktadır. Ancak, anten sayısının artırılması işlenen vektörlerin boyutunu büyütmede ve dolayısı ile işlem zamanını artırmaktadır.

Direk ve ark., sualtı hidrofon dizi alıcılarında geleneksel ve uyarlamalı hüzenin şekillendiriciler kullanarak kaynak sinyallerin geliş açılarını kestirmeye çalışmışlardır. Kaynakların koordinatlarının (derinlik ve mesafe

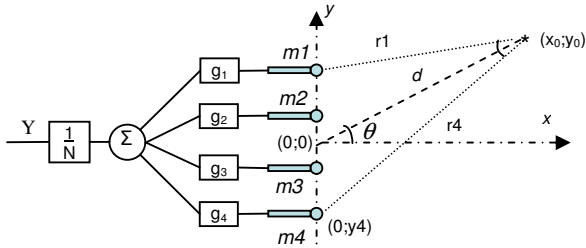
bilgilerinin) kestirilmesi benzetimlerini yapmışlardır. Benzetimler dalga kılavuzu özelliği taşıyan İstanbul Boğazı sualtı ortam şartları dikkate alınarak gerçekleştirilmiştir. Kullanılan yapay verilerin dikey doğrusal bir dizi alıcı ile elde edildiği varsayılmıştır. Benzetimler sonucunda uyarlamalı hüzme şekillendiricilerin geleneksel olanlara göre daha yüksek çözünürlük sağladıkları görülmüştür. Uyarlamalı hüzme şekillendirici amacıyla en küçük varyans tekniği kullanılmıştır [11].

4. GELİŞTİRİLEN YÖNTEM

Mikrofon dizileri ile ses kaynağı yerinin tespit edilebilmesi için, öncelikle ses kaynağının olabileceği muhtemel noktalar içinde arama yapılmıştır. Arama amacıyla Genetik Algoritma kullanılmıştır. Denenen her bir nokta için geciktir-ve-topla yöntemi ile hüzme şekillendirme yapılmış, en güçlü sinyali veren nokta odak noktası, yani ses kaynağının yeri olarak kabul edilmiştir.

4.1. Geciktir-ve-Topla Hüzme Şekillendiricisi

Problemde her bir çözüm kümesi olarak açı ve uzaklıktan (θ, d) oluşan bir değer kullanılmıştır. “d” değeri orijin noktasından $(0;0)$ ses kaynağının bulunduğu noktaya olan uzaklığı göstermektedir. “ θ ” değeri ise, orijin noktasından $(0;0)$ ses kaynağının bulunduğu noktaya çizilen doğrunun “x” eksenine yaptığı açığı göstermekte olup, -90 ile $+90$ arasında bir değerdir (Şekil 1). Bu durumda ses kaynağının koordinatları $(x_0; y_0)$, Eşitlik-3 ile hesaplanmıştır.



Şekil 1. Geciktir ve topla yöntemi ile hüzme şekillendirme yapılması.

$$x_0 = d * \cos(\theta), \quad y_0 = d * \sin(\theta) \quad (3)$$

Odak noktasının, diğer bir ifade ile ses kaynağının bulunduğu noktanın, her bir mikrofona uzaklığını (r_i) hesaplamak için eşitlik (4) kullanılmıştır.

$$r_i = \sqrt{(y_0 - y_i)^2 + (x_0)^2} \quad (4)$$

Eşitlik (4)'de y_i , her bir mikrofona “y” eksenindeki konumunu göstermektedir. Gecikme değerlerinin

sayısallaştırılmış ses bilgisi dizisi içerisinde indislere dönüştürülmesi işlemi ise, eşitlik (5) yardımı ile yapılmıştır.

$$g_i = \frac{r_i}{c} \cdot f \quad (5)$$

Eşitlik (5)'te r_i , i nci mikrofona odak noktasına uzaklığını, c ses hızını, f ise örnekleme frekansını göstermektedir. g_i ise i nci mikrofona oluşan gecikme miktarını indis olarak göstermektedir. Burada sesin havadaki yayılma hızı (c) olarak 340 ms^{-1} alınmıştır.

Her bir mikrofona için eşitlik (5) ile hesaplanan gecikme miktarları mikrofondan alınan ses dizilerine uygulanarak diziler üst üste toplanmış ve toplam ses dizisi elde edilmiştir:

$$Y_k = \frac{1}{N} \sum_{j=1}^N x_{k+g_j}^j \quad (6)$$

Eşitlik (6)'da Y_k , elde edilen çıkış dizisinin k ' nci elemanını, N mikrofona sayısını, x ise giriş dizilerini göstermektedir.

4.2. Genetik İşlemler

Genetik algoritma, yönlendirilmiş rasgele arama algoritmalarının bir türüdür [12]. Yöntem; kalıtım, mutasyon, doğal seçim ve çaprazlama gibi biyolojik evrim kurallarından esinlenerek geliştirilmiştir. Algoritma, arama uzayında mevcut olan çözümlerin oluşturduğu bir başlangıç yoğunluğunu (popülasyon) kullanır. Bu başlangıç yoğunluğu, her bir kuşakta (jenerasyon) tabii seçme ve tekrar üreme işlemleri vasıtası ile ard arda geliştirilir. En son kuşağın en uygun yani en kaliteli bireyi, problem için en iyi bir çözümdür. Bu çözüm, her zaman optimum olmayabilir, ancak kesinlikle optimuma en yakın olan çözümdür. Genetik Algoritmanın farklı problemlere uyarlanmasında en önemli adım, probleme özgü genetik kodlama ve uygunluk fonksiyonunun belirlenmesidir.

Standart bir genetik algoritmanın temel adımları aşağıdaki gibidir [13]:

Adım 1. Muhtemel çözümlerden başlangıç çözümleri (popülasyonu) oluştur.

Adım 2. Popülasyondaki her çözümün uygunluk değerini hesapla.

Adım 3. Durdurma şartları sağlanıyorsa durdur. Yoksa aşağıdaki adımları gerçekleştir;

3.1. Uygunluk ölçekleme işlemini gerçekleştir.

3.2. Seçme işlemini uygula (uygunluk değeri daha yüksek olan bireyler yeni popülasyonda daha fazla temsilciye, daha düşük olanlar ise daha az temsilciye

sahip olacaklar).

3.3. Rasgele seçme işlemini uygula (popülasyonun çoğunluğu tabii seçme, az bir kısmı da rasgele seçme işlemine tabi tutulur).

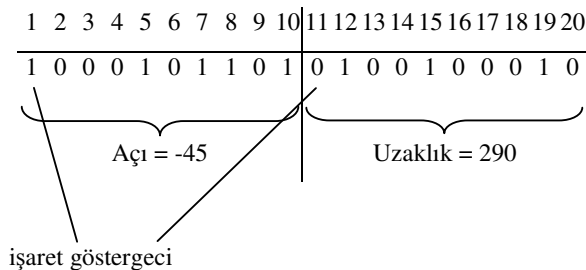
3.4. Önceki popülasyonda var olan en iyi bireyi (çözümü) muhafaza et (doğal olmayan bir işlem).

3.5. Çaprazlama işlemini uygula (Mevcut iki bireyden yeni iki birey üretilir).

3.6. Mutasyon işlemini uygula (Her bit belli bir olasılıkla değişime tabi tutulur).

4.3. Genetik Kodlama

Genetik algoritmanın bir probleme uygulanması sırasında ilk adım, problemin çözüm kümesinin genetik kodlamasının yapılması, yani kromozomların her birinin nasıl temsil edileceğinin belirlenmesidir. Bir problem için bir çözüm kümesi oluşturan karakter veya sayı dizisi kromozom veya birey olarak adlandırılır. Kromozomlar genellikle "0" ve "1" lerden oluşan ikili sayı katarı ile temsil edilir. Burada problemin çözüm kümesi olan (θ, d) değerlerini temsil etmek için ikili (binary) kodlama yapılmıştır. Açık (θ) değeri -90 ile +90 arasında, uzaklık (d) değeri ise 0 ile 400 (santimetre) arasındadır. Bu sebeple, genişliği 20 bit olan bir kromozom yapısı kullanılmıştır. Onluk sistemdeki sayılar ikili sayıya dönüştürülerek kromozomlar (bireyler) oluşturulmuştur. Kromozom içindeki her bir genin birinci biti işaret göstergesi olarak kullanılmıştır. ($\theta = -45$, $d = 290$) olan bir çözümün ikili düzende kromozom olarak kodlanması Şekil-2'deki gibi yapılmıştır.



Şekil 2. İkili düzende kromozomun kodlanması.

4.4. Başlangıç Çözümlerinin Oluşturulması

Başlangıç çözümleri, problemin çözüm uzayından seçilen ve mümkün olduğunca problem uzayını tam temsil edebilen çözümlerdir. Başka bir ifadeyle, seçilen başlangıç çözümleri problemin çözüm uzayına homojen olarak dağılmış olması tercih edilir. Böylece çözüme daha çabuk ulaşma olasılığı artacaktır. Başlangıç çözümlerinin oluşturulması, genellikle rasgele seçimlerle gerçekleştirilir. Ancak, problemle ilgili bazı çözümler kabaca biliniyorsa, başlangıç popülasyonu bu çözümler kullanılarak da oluşturulabilir. Böylece en iyi çözüm

bulmada zaman açısından tasarruf sağlanmaktadır [13].

Başlangıç çözüm sayısı (popülasyon sayısı) probleme göre değişiklik göstereceğinden, genellikle deneme yanılma yöntemi ile belirlenmektedir. İkili kodlanmış kromozomların kullanıldığı durumlarda ideal popülasyon büyüklüğünün ne olması gerektiğinin teorik araştırması Goldberg (1985) tarafından yapılmıştır. Araştırma sonucunda, popülasyon büyüklüğü ile kromozom uzunluğu arasındaki bağıntı eşitlik (7) ile tanımlanmıştır [14]:

$$\text{popülasyonbüyüklüğü} = 1,65 \cdot 2^{0,2 \cdot \text{uzunluk}} \quad (7)$$

Popülasyon büyüklüğü için seçilen değer, algoritmanın performansını iki şekilde etkilemektedir. Birincisi, popülasyon büyüklüğünün aşırı küçük seçilmesi, arama uzayının yetersiz örneklenmesine sebep olacağından, kontrollü iraksamayı sağlamak zorlaşacak ve arama belli bir alt en iyi noktaya doğru sürüklenecektir. İkincisi, popülasyon büyüklüğü için aşırı yüksek bir değer seçildiğinde bir nesillik gelişim oldukça uzun süreye ihtiyaç duyacaktır. Bu ise, özellikle gerçek zamanlı problem uygulamalarında arzu edilmeyen bir durumdur. Bu yüzden popülasyon büyüklüğü için uygun bir değer belirlenmelidir [13].

4.5. Çaprazlama

Çaprazlama, bir popülasyon içindeki bireylerden (kromozomlardan) iki tanesinin seçilerek bunlardan iki adet yeni birey oluşturma işlemidir. Mevcut popülasyondan çaprazlanacak bireylerin seçimi amacıyla, rulet tekerleği yöntemi uygulanmıştır. Bu yöntemde seçilecek bireyler rasgele seçilir, ancak uygunluk değeri daha iyi olan bireylerin seçilme olasılıkları daha fazladır. Tek noktalı ve iki noktalı çaprazlama yöntemleri kullanılmıştır. Tek noktalı çaprazlamada, 1 ile 19 arasında rasgele bir sayı seçilir. Seçilen sayıdan önceki bölüm 1 nci bireyden, sonrası ise 2 nci bireyden alınarak yeni birey oluşturulur. Örnek: çaprazlama noktası 5 ise, k1 ve k2 kromozomlarının çaprazlanması sonucu y1 ve y2 kromozomları aşağıdaki şekilde elde edilir.

k1 : 1 1 0 0 1 0 1 0 1 1
k2 : 0 1 1 0 0 1 1 1 0 0

y1 : 1 1 0 0 1 1 1 1 0 0
y2 : 0 1 1 0 0 0 1 0 1 1

İki noktalı çaprazlama yönteminde ise, iki farklı çaprazlama noktası belirlenir ve yeni bireyin iki bölümü bir ebeveyninden, bir bölümü ise diğer ebeveyninden alınır. Yukarıdaki örnekte ebeveynler 3 ve 8 noktalarından çaprazlanacak olursa aşağıdaki yeni bireyler elde edilir.

y1 : 1 1 0 0 0 1 1 0 1 1

y2 : 0 1 1 0 1 0 1 1 0 0

Çaprazlamanın hangi koşullarda yapılacağını tespit etmek amacıyla çaprazlama oranı parametresi kullanılmıştır. Seçilen iki kromozomun çaprazlanabilmesi için, 0 ile 1 arasında rasgele seçilen bir sayının, çaprazlama oranından daha düşük olması gerekmektedir. Aksi takdirde seçilen bu iki birey çaprazlamaya tabi tutulmadan mutasyona uğrattılır ve sonraki popülasyona aktarılır. Bu çalışmada çaprazlama oranı 0.80 olarak alınmıştır.

4.6. Mutasyon

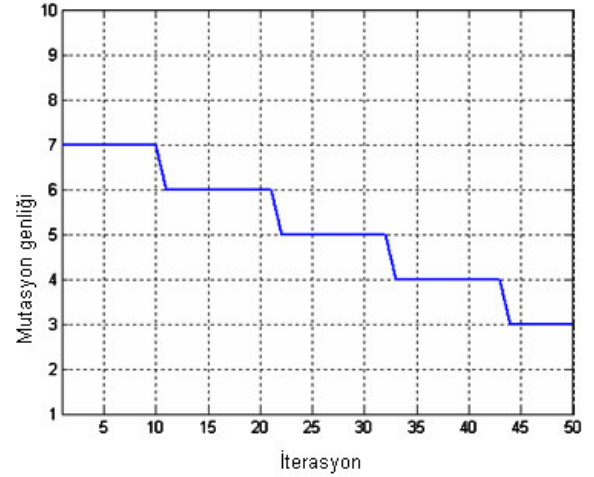
Çaprazlama sonucu elde edilen yeni bireyler yeni popülasyona aktarılmadan önce mutasyon işlemine tabi tutulmuştur. Mutasyon işlemi, yeni nesildeki bireylerin ebeveynlerinden aldıkları özelliklerin çok küçük oranda değişime uğratılması şeklinde gerçekleştirilir. Mutasyon oranı olarak çok küçük bir sayı (0.05) belirlenmiştir. Kromozomdaki her bir değer (bit) için 0 ile 1 arasında rasgele bir sayı seçilir. Seçilen sayı mutasyon oranından daha küçük ise, ilgili bitin değeri 0 ise 1'e, 1 ise 0'a değiştirilir.

Mutasyonun her adımda her yeni bireye rastsal olarak uygulanması yerine farklı uygulamalar da mevcuttur. Hacıoğlu ve ark., iterasyonun sadece belli adımlarında (örneğin her 5 adımda bir) tüm bireylerin tüm genlerine mutasyon uygulamışlardır. Aradaki adımlarda mutasyon uygulamayarak çözümün kendisini toparlamasına izin vermişlerdir. Uygulanacak titreşimin genliğini, ortalama uygunluk değerinin değişimine bağlı olarak uyarlamışlardır. Titreşimli genetik algoritma (TGA) olarak adlandırdıkları bu yöntemin, geleneksel mutasyon uygulanan yöntemle göre daha başarılı olduğu belirtilmiştir [15].

Mutasyon işleminin tanımına uygun olması için, mutasyon genliği kavramı önerilmiştir. Bu kavram sayesinde, programın başlangıcında mutasyon işleminin etkisi daha büyük iken, programın sonlarına doğru mutasyon genliği azaltılarak, mevcut çözümün yakınlarında arama yapma zorunluluğu getirilmiştir. Bunun sonucu olarak, genetik çeşitlilik sağlanmakla birlikte, algoritmanın daha düzgün bir şekilde yakınsaması sağlanmıştır. Değişken mutasyon genliği işlemi Şekil 3'te detaylı olarak gösterilmiştir. Mutasyon genliği başlangıçta "7" iken, sonunda "3" olmaktadır. 50 iterasyonluk bir çalışma için mutasyon genliğinin değişim grafiği Şekil 4'de görüldüğü gibidir.



Şekil 3. Değişken mutasyon genliği uygulaması.



Şekil 4. Mutasyon genliği değişim grafiği.

4.7. Uygunluk fonksiyonu

Popülasyondaki her bir çözüm için açı ve uzaklık değerine göre mikrofon sinyallerine gecikme uygulanarak üst üste toplandıktan (Eşitlik 6) sonra uygunluk değerlendirmesi yapılmıştır. Uygunluk fonksiyonu için, elde edilen çıkış sinyali değişik kriterlere göre değerlendirilerek, en güçlü sinyali veren noktanın ses kaynağının yeri olduğu kabul edilmektedir. Değerlendirme kriteri olarak aşağıdakiler kullanılabilir;

- Ses sinyalinin gücünün gürültü sinyalinin gücüne oranı (SGO),

$$P_s = \frac{\sum_{i=1}^N s_i^2}{N}, \quad P_n = \frac{\sum_{j=1}^M n_j^2}{M} \quad (8)$$

$$SGO = 10 \cdot \log_{10} \left(\frac{P_s}{P_n} \right) \quad (9)$$

Burada, P_s konuşmanın gücü ve P_n gürültünün gücüdür.

- Çıkış sinyalinin ortalama değeri (OD),

$$OD = \frac{1}{N} \sum_{k=1}^N Y_k \quad (10)$$

Burada, Y ortalaması hesaplanacak sinyal dizisi, N dizinin eleman sayısıdır.

Açı										Uzaklık									
1	0	0	0	1	0	1	1	0	1	0	1	0	0	1	0	0	0	1	0

- Çıkış sinyalinin (Y) referans sinyali (X) ile arasındaki çapraz ilinti katsayısı (korelasyonu),

$$X_k = X_k - OD_X$$

$$Y_k = Y_k - OD_Y$$

$$kor(X,Y) = \frac{\sum_{k=1}^N X_k \cdot Y_k}{\sqrt{\left(\sum_{k=1}^N X_k \cdot X_k\right) \cdot \left(\sum_{k=1}^N Y_k \cdot Y_k\right)}} \quad (11)$$

- Sinyalin maksimum genliği.

Test edilen uygunluk fonksiyonlarının hepsi odak noktasına yakın bir sonuç vermelerine rağmen, en iyi sonucu çapraz ilinti katsayısı yönteminin verdiği tespit edilmiştir.

5. DENEYLER VE BULGULAR

Geliştirilen yöntem, Carnegie Mellon Üniversitesi bilgisayar laboratuvarında 15 adet mikrofon kullanılarak kayıt edilmiş olan ses sinyalleri üzerinde denenmiştir. Mikrofonlara gelen ses bilgileri 16 KHz frekansında örneklenmiştir. Mikrofonların dizilişleri Şekil 5'te görüldüğü gibidir. Şeklin solundaki 1 ile 15 arasındaki numaralar mikrofon numarasını göstermektedir. 8 numaralı mikrofonun bulunduğu koordinat (0;0) olarak kabul edilmiştir.

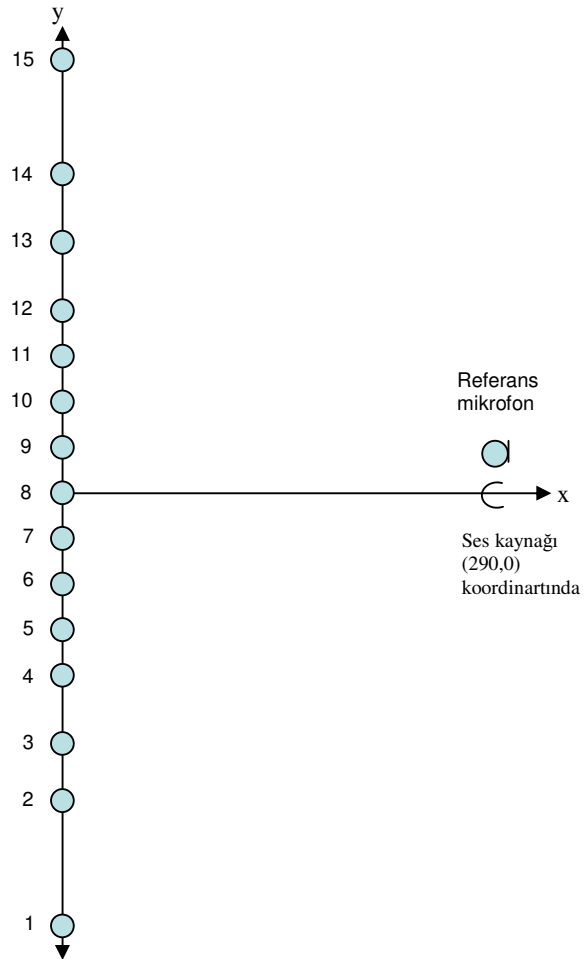
Mikrofonlar “y” eksenini üzerinde sıralı olarak dizilmiş olmalarına rağmen, aralarındaki uzaklık sabit değildir. Bu nedenle, mikrofon pozisyonları tek tek hesaplanmıştır. Mikrofonların tamamının kullanılması yerine, bir takım alt kümelerin kullanılması da mümkündür. Bu çalışmada, mikrofonlardan alt mikrofon grupları oluşturularak, geliştirilen yöntem her bir alt grup için denenmiş ve sonuçları değerlendirilmiştir.

5.1. Mikrofon Dizileri

Alt mikrofon gruplarının oluşturulması aşağıdaki şekilde yapılmıştır:

- 5, 6, 7, 8, 9, 10 ve 11 numaralı mikrofonlardan 7 mikrofonluk bir grup oluşturulmuştur ve aralarındaki uzaklık 0,04 metredir.
- 3, 4, 6, 8, 10, 12 ve 13 numaralı mikrofonlardan 7 mikrofonluk bir grup oluşturulmuştur ve aralarındaki uzaklık 0,08 metredir.
- 1, 2, 4, 8, 12, 14 ve 15 numaralı mikrofonlardan 7 mikrofonluk bir grup oluşturulmuştur ve aralarındaki uzaklık 0,16 metredir.
- 1, 8 ve 15 numaralı mikrofonlardan 3 mikrofonluk bir grup oluşturulmuştur ve aralarındaki uzaklık 0,48 metredir.

- 1 ve 15 numaralı mikrofonlardan 2 mikrofonluk bir grup oluşturulmuştur ve aralarındaki uzaklık 0,96 metredir.



Şekil 5. 15 adet mikrofonun ses kaynağına göre yerleşim düzeni.

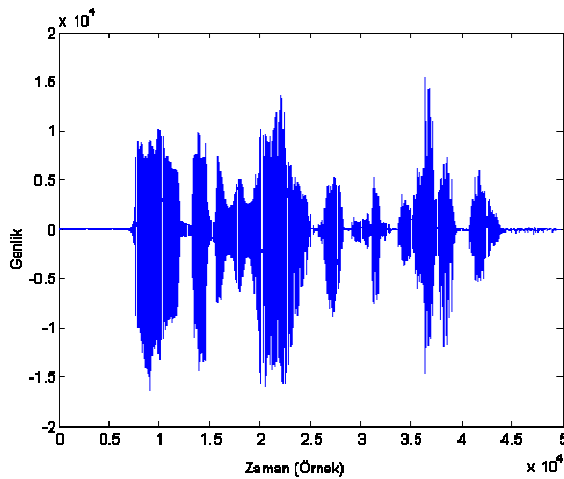
Ses kaynağı olarak kullanılan konuşmayı yapan kişi mikrofonların orta noktasının karşısında 290 cm uzaklıkta bulunmaktadır. Dolayısı ile bu noktanın (x,y) koordinatları (290 ; 0) olarak alınmıştır. Ses kaynağına çok yakın bir şekilde ağız mikrofonu yerleştirilerek referans sinyali de kaydedilmiştir.

5.2. Çeşitli Mikrofon Kombinasyonları İle Yapılan Uygulamalar

Mikrofon dizilerinden alınan ses sinyallerine her hangi bir işlem yapmadan önce ön filtreleme işlemi uygulanmıştır. Ön işlem olarak, 400 Hz altındaki frekansları bastırarak olan bir yüksek geçiren süzgeç kullanılmıştır. Şekil 6'da,

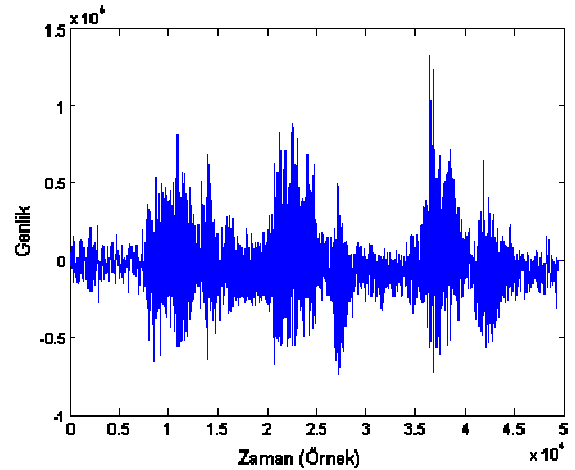
ses kaynağındaki konuşmacıya çok yakın bir mikrofondan alınmış ses sinyali görülmektedir. Burada yaka mikrofonu gibi yakın bir mikrofona kullanılmıştır. Dolayısıyla çevredeki gürültülerden çok az etkilenme söz konusudur. Ayrıca sesin kalitesi ideal düzeyde kabul edilebilir. Çalışmanın bundan sonraki kısımlarında bu referans sinyali ile hüzme şekillendiricinin çıkışındaki sinyal arasındaki ilinti katsayısı hesaplanmıştır.

Mikrofon dizisindeki 1 numaralı mikrofondan alınan sinyalin süzgeç uygulanmadan önceki grafiği Şekil 7.a'da, süzgeç uygulandıktan sonraki grafiği ise Şekil 7.b'de görülmektedir.

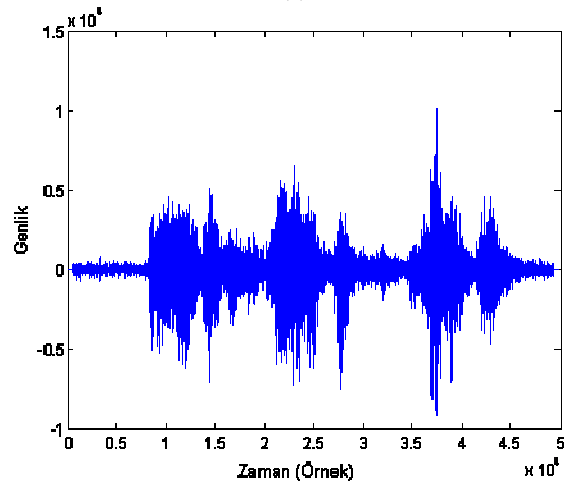


Şekil 6. Referans mikrofondan alınan ses sinyali.

Şekil 5'teki 15 mikrofondan alınan ses bilgilerinin tamamı kullanılarak elde edilen sonuç Şekil 8'de görülmektedir. Ses bilgisindeki 1,5 saniyelik gürültü ve konuşmanın bulunduğu bölüm alınmıştır. Örnekleme frekansı 16000 Hz olduğu için, 1,5 saniyelik ses bilgisi 24000 örneğe karşılık gelmektedir. Uygunluk fonksiyonu olarak ilinti katsayısı, popülasyon büyüklüğü olarak 20 seçilmiştir. Şekil 8(a)'da başlangıç çözümleri, Şekil 8(b)'de ise 50 iterasyon sonunda elde edilen sonuç verilmiştir. 50 iterasyon sonunda 0.03 metre hata ile çözüme ulaşılmıştır. Ancak, Şekil 8(c)'de görüldüğü gibi, yakınsama 15. iterasyon civarında gerçekleşmiştir. Bu şekillerde uygunluk değerinin azalması, ilinti katsayısının arttığını göstermektedir.



(a)

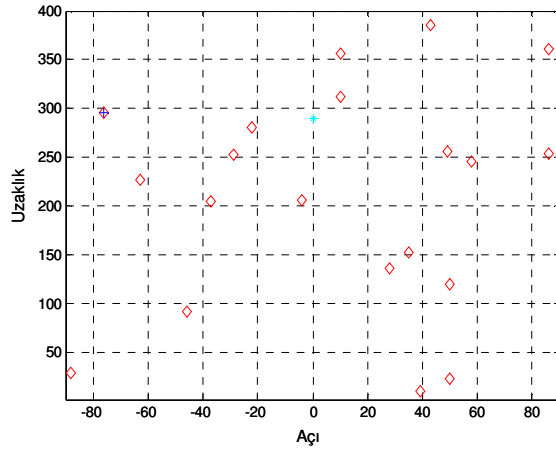


(b)

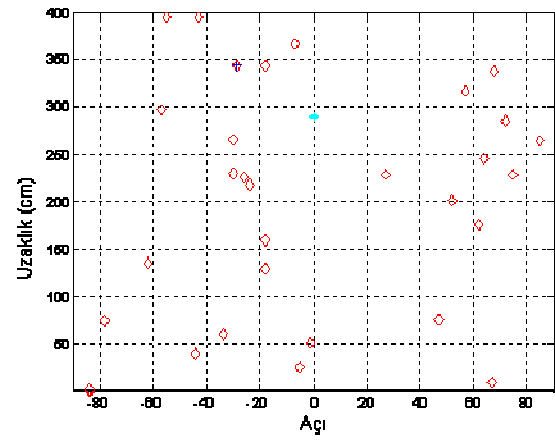
Şekil 7. (a) Giriş sinyaline süzgeç uygulanmadan önceki hali (b) Giriş sinyalinin süzgeçleme işleminden sonraki görünümü.

Mikrofonlardan sadece 7 tanesi (1, 2, 4, 8, 12, 14 ve 15 no'lu mikrofonlar) kullanılarak elde edilen sonuç Şekil 9'da verilmiştir. Bu kombinasyonda mikrofonlar arası sabit 0,16 metredir. Popülasyon büyüklüğü için, eşitlik (7) ile önerilen değer kullanılmıştır. Bu durumda popülasyon büyüklüğü 30, çaprazlama oranı 0.8, mutasyon oranı 0.05, uygunluk fonksiyonu olarak ilinti katsayısı alınmıştır. 50 iterasyon sonunda yaklaşık 0.04 metre hata ile hedefe ulaşılmıştır.

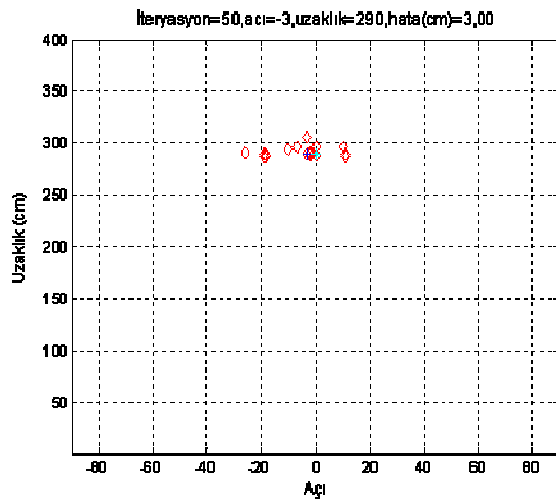
Daha az mikrofon kullanarak bir önceki adımda ulaşılan hedefe çok yakın bir sonuç elde edilmiştir.



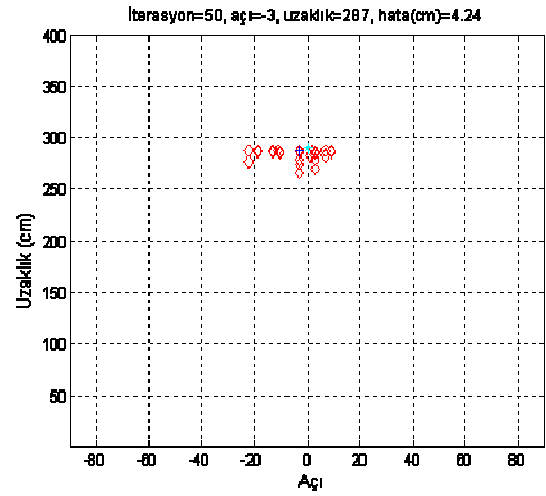
(a)



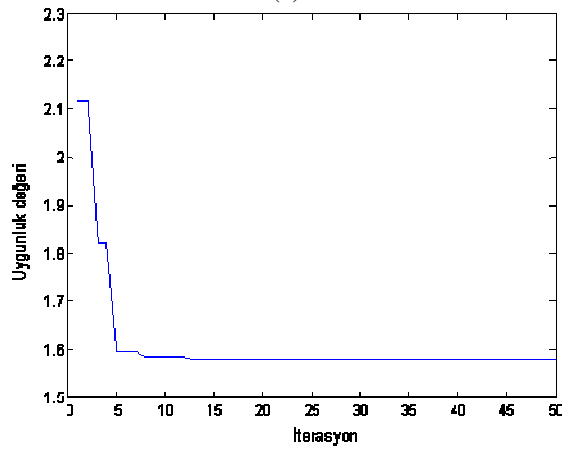
(a)



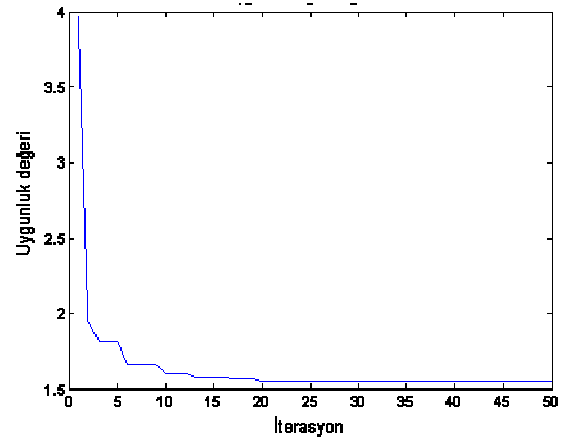
(b)



(b)



(c)



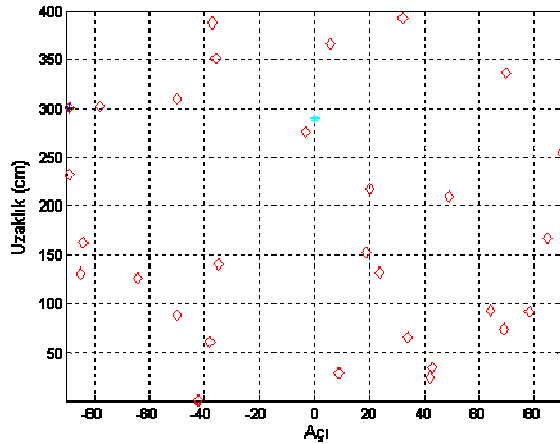
(c)

Şekil 8. 15 mikrofon kullanılması ile elde edilen sonuç (a) başlangıç çözümlerinin dağılımı, (b) 50 iterasyon sonunda elde edilen sonuç (c) iterasyon boyunca uygunluk değerinin değişimi. $\diamond$ bireyler, + en uygun birey, * hedef konum

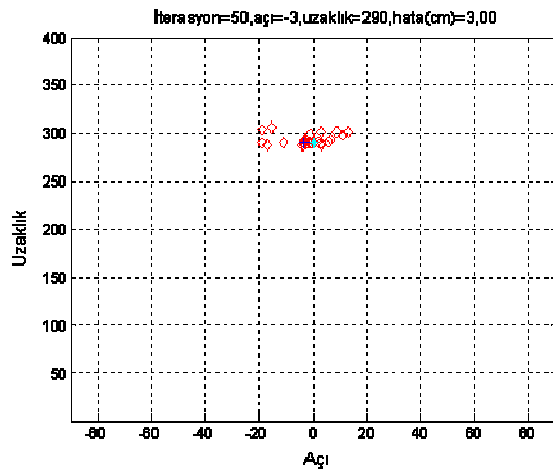
Şekil 9. 7 mikrofon kullanılarak elde edilen sonuç (a) başlangıç çözümlerinin dağılımı, (b) 50 iterasyon sonunda elde edilen sonuç, (c) iterasyon boyunca uygunluk değerinin değişimi. $\diamond$ bireyler, + en uygun birey, * hedef konum

Şekil 10'da ise, sadece 3 mikrofon (1, 8 ve 15 numaralı mikrofonlar) kullanılarak elde edilen sonuç görülmektedir. Popülasyon büyüklüğü 30, çaprazlama

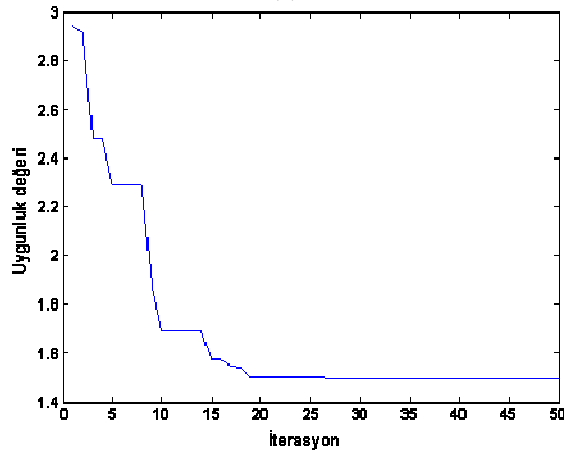
oranı 0.8, mutasyon oranı 0.05, uygunluk fonksiyonu olarak ilinti katsayısı alınmıştır.



(a)



(b)



(c)

Şekil 10. 3 mikrofon kullanılarak elde edilen sonuç (a) başlangıç çözümlerinin dağılımı, (b) 50 iterasyon sonunda elde edilen sonuç, (c) iterasyon boyunca uygunluk

değerinin değişimi. $\diamond$ bireyler, + en uygun birey, * hedef konum

Sadece 3 adet mikrofon kullanılarak 0.03 metre hata ile hedefe ulaşılmıştır. Ses kaynağı yerinin belirlenmesinde az sayıda mikrofonun yeterli olması hesaplama maliyetini düşürmektedir.

Ses kaynağının yerinin (geliş açısı ve uzaklık) belirlenmesinden sonra, mikrofon dizisini bu noktaya yönlendirmek ve odaklamak için mikrofonlara uygulanacak gecikme miktarları geometrik yöntemlerle hesaplanmaktadır.

6. SONUÇLAR VE DEĞERLENDİRME

Bu çalışmada, gürültülü ortamlarda ses kaynağının yerinin genetik algoritmaya dayalı mikrofon dizileri kullanılarak bulunması amaçlanmıştır. Çalışmanın temeli, mikrofon dizileri ile hüzmeye şekillendirme yapmaya dayanmaktadır. Hüzmeye şekillendirme, mikrofon dizisinin hedef bölgede belirli bir noktaya odaklanması şeklinde tanımlanabilir. Bu çalışmada, geciktir-ve-topla yöntemi kullanılmıştır. Yöntemde, dizi içerisindeki her bir mikrofondan alınan ses sinyali bilgisine belirli gecikmeler uygulanarak elde edilen sinyal bilgilerinin üst üste toplanması ile çıkış sinyali elde edilmektedir.

Ses kaynağının yerinin aranması için, makul bir zamanda en iyi veya en iyiye yakın bir çözüm bulan bir yöntem olan genetik algoritma kullanılmıştır. Genetik algoritma yardımı ile hesaplanan noktalara hüzmeye şekillendirme yapılarak elde edilen çıkış sinyali değerlendirmeye tabi tutulmuştur. Değerlendirme sonucunda en uygun sinyalin elde edildiği nokta, diğer bir ifadeyle uzaysal koordinat ses kaynağının yeri olarak kabul edilmiştir.

Genetik algoritmanın uygunluk değerlendirme fonksiyonu olarak Sinyal-Gürültü-Oranı (SGO), ilinti katsayısı, ortalama değer, maksimum genlik gibi parametreler kullanılabilir. SGO ve ilinti katsayısının hesaplanabilmesi amacıyla, referans mikrofondan alınan ses sinyali ile hüzmeye şekillendirme sonucu elde edilen ses sinyali kullanılmıştır. Mutasyon işleci olarak ise, değişken genlikli mutasyon işlemi önerilmiştir. Bu yöntemde, ikili kodlanmış kromozomların genlerine, iterasyon sayısı ile azalan bir genlikle mutasyon uygulanmış, böylece başlangıç adımlarında daha geniş bir çözüm uzayı aranırken, iterasyonun sonuna yaklaştıkça, her bir çözümün daha dar çevresi aranmıştır. Bu şekilde, algoritmanın yerel minimuma düşme olasılığı azaltılmıştır. Yapılan deneyler sonucunda, çok küçük bir hata ile ses kaynağının yerinin hesaplanabildiği görülmüştür.

TEŞEKKÜR

Bu çalışma, Gazi Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından 06/2005-37 numaralı “Mobil Cihazlar İçin Akıllı Mikrofon” adlı proje kapsamında desteklenmiştir.

KAYNAKLAR

- [1] Sullivan, M., T., “Multi-Microphone Correlation-Based Processing for Robust Automatic Speech Recognition”, Ph.D. Thesis, Carnegie Mellon University Department of Electrical and Computer Engineering, Pennsylvania, 8-100 (1996).
- [2] Hamahara, T., Kuroiwa, K., Hoshino, O., “Field Measurement Of Multi-Sound Directions By Using Microphone Array System Based On Genetic Algorithm”, **Proceedings of the 9th International Conference on Neural information Processing (ICONIP’02)** , 1216-1220, Singapore (2002).
- [3] Kawano, J., Ogawa, T., Kuroiwa, K., Hoshino, O., “Estimation of Sound Direction for Multi-Sound Field Using Genetic Algorithm”, **Proceedings of the 8th International Conference on Neural information Processing (ICONIP’01)**, Shanghai-China (2001).
- [4] Widrow, B., “A Microphone Array for Hearing Aids”, **IEEE Circuits and Systems Magazine**, 1 (2): 26-32 (2001).
- [5] Veen, V., Buckley, K., “Beamforming: A Versatile Approach to Spatial Filtering”, **IEEE Acoustics, Speech and Signal Processing (ASSP) Magazine**, 4-24 (1988).
- [6] Weinstein, E., Steele, K., Agarwal, A., Glass, J., “LOUD: A 1020-Node Modular Microphone Array and Beamformer for Intelligent Computing Spaces”, **MIT/LCS Technical Memo MIT-LCS-TM-642**, Cambridge, 1-17 (2004).
- [7] Aguilar, J., R., Arias, M., Salinas, R., Abidi, M., A., “Direct search of time delay in beamforming applications”, **Proc. of the III Encuentro Chileno de Acústica INGEACUS 2004**, Valdivia-Chile, 186-192 (2004).
- [8] Birinci, I., Y., Leblebicioğlu, K., “Mikrofon Dizisi Yerleşiminin Genetik Algoritma ile Optimizasyonu”, **Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU’2006)**, İstanbul, 1-4 (2006).
- [9] Hacivelioglu, İ., Dinçer, H., “Akıllı Anten Sistemlerinde İşaret Geliş Açısı Kestirim Yöntemleri”, **II. İletişim Teknolojileri Ulusal Sempozyumu**, Adana (2005).
- [10] Tamer, Ö., Köktürk, G., Özmehmet, K., “Düzgün Doğrusal Anten Dizilerinde En Büyük Olabilirlik Yöntemi Kullanarak Geliş Açısı Belirlenmesi”, **DEÜ Mühendislik Fakültesi Fen Ve Mühendislik Dergisi**, 4(3): 37-43 (2002).
- [11] Direk, Ş., Baykut, S., Akgül, T., “Sualtı Doğrusal Dizi Alıcılarda Hüzme Şekillendiriciler ile Geliş Yönlerinin Kestirimi”, **Sinyal İşleme ve İletişim Uygulamaları Kurultayı**, Eskişehir, 280-283 (2007).
- [12] Goldberg, D., E., Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, USA , 1-25 (1989).
- [13] Karaboğa, D., **Yapay Zeka Optimizasyon Algoritmaları**, Atlas Yayın Dağıtım, İstanbul, 75-112 (2004).
- [14] Goldberg, D., E., Optimal Initial Population Size for Binary-Coded Genetic Algorithms, TCGA Report Number 851001, University of Alabama, Alabama (1985).
- [15] Hacıoğlu, A., “Kanat Profili Dizaynında Genetik Algoritma Kullanımı”, **İTÜ Dergisi**, 2(6) : 73-81 (2003).