

**MİKROFON DİZİLERİ İLE
TELEKONFERANS SİSTEMLERİNİN İYİLEŞTİRİLMESİ**

Esmâ Sultan SEZEN

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**TEMMUZ 2010
ANKARA**

Esma Sultan SEZEN tarafından hazırlanan mikrofon dizileri ile telekonferans sistemlerinin iyileştirilmesi adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

Yrd. Doç. Dr., Hasan Şakir BİLGE
Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından oy birliği / oy çokluğu ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Prof. Dr. Mehmet R. TOLUN
Bilgisayar Mühendisliği, Çankaya Üniversitesi
Yrd. Doç. Dr., Hasan Şakir BİLGE
Bilgisayar Mühendisliği, Gazi Üniversitesi
Doç. Dr. M. Ali AKÇAYOL
Bilgisayar Mühendisliği, Gazi Üniversitesi

Tarih : 14/07/2010

Bu tez ile G.Ü. Fen Bilimleri Enstitüsü Yönetim Kurulu Yüksek Lisans / Doktora derecesini onamıştır.

Prof. Dr. Bilal TOKLU
Fen Bilimleri Enstitüsü Müdürü

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Esmâ Sultan SEZEN

**MİKROFON DİZİLERİ İLE
TELEKONFERANS SİSTEMLERİNİN İYİLEŞTİRİLMESİ
(Yüksek Lisans Tezi)**

Esmâ Sultan SEZEN

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
Temmuz 2010**

ÖZET

Mikrofon sistemlerinin kullanıldığı ses algıma, ses kayıt, ses tanıma gibi uygulamalarda yaşanan en büyük problemlerden biri ortam koşulları nedeni ile ses sinyalindeki bozulmalardır. Günümüzde ses kaynağından alınan ses sinyalinin daha güçlü olarak alınabilmesi için birden fazla mikrofondan oluşan mikrofon dizileri kullanılmaktadır. Mikrofon dizilerinin normal mikrofonlara göre iki temel üstünlüğü mevcuttur; kaynağa yakın olma zorunluluğunu ortadan kaldırmak ve her bir mikrofona sesin geliş zamanları arasındaki farkı kullanarak sesin kaynağını veya konuşmacının takibini gerçekleştirmek.

Mikrofon dizilerinde, sesin elde edilmesi için dizide yer alan mikrofonların yerlerinin bilindiği kabul edilerek geliş sürelerindeki gecikmelerin, ikili çiftler halinde incelendiği genelleştirilmiş çapraz korelasyon (GÇK) tekniği kullanılmaktadır. Bu teknik kullanılır iken ses sinyalinin daha iyi elde edilebilmesi için hüzme şekillendirme işlemi yapılması gerekir. Çalışmada ses sinyalinin elde edilebilmesi hüzme şekillendirme yöntemi olarak Yönlendirilmiş Tepki Gücü (YTG) yöntemi Faz Dönüşümü (FD) ağırlık fonksiyonu kullanan GÇK metodu ile birleştirilmiştir.

Bu çalışmada, beş farklı kişinin %75 oranla aynı anda konuştuğu bir ortamda konuşmacıların yerlerini 3 boyutlu uzayda bulunması hedeflenmiştir. Mikrofon dizilerine gelen seslerin kaynaklarının uzaysal konumlarının otomatik olarak belirlenmesi amaçlanmıştır. Mikrofon dizileri ile YTG-FD yöntemi kullanarak ses kaynaklarının yönlerinde hüzme şekillendirme yapılmış ve en güçlü çıkış sinyallerinin elde edildiği noktalar aranan ses kaynaklarının yeri olarak tahmin edilmiştir. İncelenecek noktaların belirlenmesi konusunda optimizasyon yöntemi olarak genetik algoritma kullanılmıştır.

Bilim Kodu : 902.1.067
Anahtar Kelimeler : Mikrofon Dizisi, Yönlendirilmiş Tepki Gücü – Faz Dönüşümü Metodu, Üst Üste Binen Konuşmalar
Sayfa Adedi : 88
Tez Yöneticisi : Yrd. Doç. Dr. Hasan Şakir BİLGE

IMPROVEMENT OF TELECONFERENCING SYSTEMS WITH MICROPHONE ARRAYS

(M.Sc. Thesis)

Esma Sultan SEZEN

**GAZİ UNIVERSITY
INSTITUTE OF SCIENCE AND TECHNOLOGY**

July 2010

ABSTRACT

The most important problem for microphone systems' applications like speech recognition, audio recording, voice recognition is distortion of sound signal caused by environmental conditions. Nowadays, to obtain more powerful sound signals, microphone arrays are used by combining more than one microphones. There are mainly two advantages of microphone arrays over normal microphones; the necessity of being close to the microphone has been cancelled and the estimation of sound source location or talker tracking by using the difference between time delay arrivals of each microphone.

For the microphone arrays, the generalized cross correlation (GCC) technique is used in that the time delays of arrivals are analyzed by accepting that the locations of microphones have been known. While using this technique, to obtain better quality of sound signals, beamforming process should be performed. In this study, Steered Response Power (SRP) has been used as beamforming technique by combining GCC technique with Phase Transform (PHAT) weighting function.

In this study, the locations of simultaneous (with 75%) five talkers are tried to be found in 3D space. It is aimed to find the sound source locations of source signals using microphone arrays. SRP-PHAT method is used to

beamforming application through the source locations directions with the help of microphone arrays. The most powerful output signals are accepted as the location of talkers. In investigation, genetic algorithm is used as optimization technique.

Science Code	:	902.1.067
Key Words	:	Microphone Array, Steered Response Power-Phase Transformation, Overlapping Speech
Page Number	:	88
Adviser	:	Yrd. Doç. Dr. Hasan Şakir BİLGE

TEŞEKKÜR

Bu çalışmada, değerli görüş ve katkılarıyla beni yönlendiren ve her konuda yardımcı olan hocam Sayın Yrd. Doç. Dr. Hasan Şakir Bilge'ye teşekkürü borç bilirim.

Ayrıca tez çalışması esnasında görüşleri ve motivasyonu ile yardımcı olan Sayın Dr. Kasım Öztoprak 'a ve Sayın Akif Yazıcı 'ya, çalışma boyunca sabırla yanımda olan eşime, anneme, babama ve biricik kardeşim Merve'ye, çalışmam boyunca düzenli aralıklarla neler yaptığımı sorarak beni motive etmeye çalışan dostlarım Burcu Özkal'a ve Sinem Budak'a, yüksek lisans eğitimimde beni destekleyen TÜBİTAK'a teşekkür ederim.

İÇİNDEKİLER

	Sayfa
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER	ix
ÇİZELGE LİSTESİ.....	xii
ŞEKİLLERİN LİSTESİ	xiii
RESİMLERİN LİSTESİ	xvi
SİMGELER VE KISALTMALAR.....	xvii
1. GİRİŞ.....	1
2. MİKROFON DİZİLERİ.....	5
2.1. Mikrofon	5
2.2. Mikrofon Dizileri.....	6
2.3. Mikrofon Dizilerinin Kullanım Alanları	6
3. DAHA ÖNCE YAPILMIŞ ÇALIŞMALAR.....	10
4. SES KAYNAĞININ YERİNİN BULUNMASI	17
4.1. Temel Akustik Koşulları	17
4.2. Doğrudan Yol Yayılımı (Direct Path Propagation)	18
4.3. Yayılım Yönü ve Erişim.....	20
5. GENELLEŞTİRİLMİŞ ÇAPRAZ KORELÂSYON (GÇK)	22
5.1. Genel Bilgiler.....	22
5.2. Genelleştirilmiş Çapraz Korelasyon Metodunun Uygulaması	23

Sayfa

5.3. En Çok Olabilirlik Ağırlık Fonksiyonu	26
5.4. Faz Değişim Ağırlık Fonksiyonu	26
5.5. Bant Geçirgen Ağırlık Fonksiyonu	27
6. YÖNLENDİRİLMİŞ TEPKİ GÜCÜ (YTG)	28
6.1. Yönlendirilmiş Tepki	28
6.2. Genelleştirilmiş Çapraz Korelasyonda Yönlendirilmiş Tepki Gücü Gösterimi	31
6.3. Faz Dönüşümü İle Yönlendirilmiş Tepki Gücünün Birleştirilmesi: YTG-FD Yöntemi	34
7. GENETİK ALGORİTMA.....	36
7.1. Genetik Algoritmanın Temel Yapısı ve İşlevi.....	36
7.2. Genetik Kodlama.....	40
7.3. Başlangıç Çözümlerinin Oluşturulması.....	41
7.4. Uygunluk Fonksiyonu.....	42
7.5. Yeniden Üretim Operatörü.....	42
7.6. Genetik Operatörler.....	47
7.6.1. Çaprazlama operatörü	47
7.6.2. Mutasyon operatörü	49
7.6.3. Seçkinlik (elitizm)	50
7.6.4. Onarım	50
7.7. Kontrol Parametreleri.....	50
8. GELİŞTİRİLEN YÖNTEM	53
9. DENEYSEL ÇALIŞMALAR	59
9.1. Genetik İşlemler.....	61

	Sayfa
9.2. Verinin Analizi	67
9.3. Genetik Algoritma ile Yapılan Uygulama	75
10. SONUÇ VE DEĞERLENDİRME	82
KAYNAKLAR	85
ÖZGEÇMİŞ	88

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.1. Çalışmanın sonuçları.....	11
Çizelge 3.2. Kelime hata oranı sonuçları (%)	12
Çizelge 3.3.Üç yöntemin karşılaştırma sonuçları	16
Çizelge 7.1. Bireylerin uygunluk değerleri ve seçilme olasılıkları.....	45
Çizelge 9.1. Mikrofonların 3 boyutlu uzayda yerleri.....	59
Çizelge 9.2. Konuşmacıların 3 boyutlu uzayda yerleri.....	60
Çizelge 9.3 Koordinat bit uzunluklarına göre elde edilen sonuçlar	62
Çizelge 9.4 Süre ve Sapma miktarlarının gösterimi.....	62
Çizelge 9.5. Kenar uzunluğu 5 cm olan kör bölge için elde edilen sonuçlar.....	65
Çizelge 9.6. Kenar uzunluğu 10 cm olan kör bölge için elde edilen sonuçlar.....	65
Çizelge 9.7. Kenar uzunluğu 20 cm olan kör bölge için elde edilen sonuçlar.....	66
Çizelge 9.8. Kenar uzunluğu 30 cm olan kör bölge için elde edilen sonuçlar.....	66
Çizelge 9.9. Kenar uzunluğu 40 cm olan kör bölge için elde edilen sonuçlar.....	66
Çizelge 9.10. Kenar uzunluğu 50 cm olan kör bölge için elde edilen sonuçlar.....	66
Çizelge 9.11. 20 cm netlik için gerçek değerler ile elde edilen değerlerin karşılaştırması	70
Çizelge 9.12. 5 cm netlik için gerçek değerler ile elde edilen değerlerin karşılaştırması	72
Çizelge 9.13. 3 cm netlik İÇİN gerçek değerler ile elde edilen değerlerin karşılaştırması	74
Çizelge 9.14. Bulunan konuşmacılar ile gerçek konuşmacı koordinatlarının mukayesesi	81

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 3.1. Çalışma sonucunda elde edilen sonuçlar	10
Şekil 3.2. Sistemin Yapısı	11
Şekil 3.3. Farklı konuşmacı sayıları için kelime hata oranları.....	12
Şekil 3.4. Yapılan çalışma sonucunda 4 konuşmacının yerlerinin tespiti.....	13
Şekil 3.5. Üç yöntemin karşılaştırma sonuçları	14
Şekil 3.6. Stokastik bölge konsantrasyonu yöntemi	15
Şekil 3.7. Gauss karışım metodu ile elde edilen sonuçlar (a) iki boyutlu enerji Haritası (b) uygunluk değerleri dağılımı.....	16
Şekil 3.8. Bölge sıfırlayarak kümeleme yöntemi ile elde edilen sonuçlar (a) elde edilen en iyi N nokta (b) bölge sıfırlayarak kümeleme yöntemi sonucunda elde edilen bölgeler[8]	16
Şekil 4.1. Kartezyen koordinatlarına göre ses kaynağının ve mikrofonun yerleri.....	19
Şekil 4.2. 3 elemanlı mikrofon dizisi için yayılım yönleri.....	21
Şekil 5.1. Gecikme zamanları arasındaki farklar sayesinde ses kaynağının yerini bulmak	22
Şekil 6.1. Geciktir-ve-topla hüzme şekillendiricisi.....	28
Şekil 7.1. Genetik algoritmanın genel yapısı	37
Şekil 7.2. Genetik algoritmanın temel akış diyagramı	39
Şekil 7.3. Kodlama uzayı ve çözüm uzayı	41
Şekil 7.4. Düzenli örnekleme uzayına örnek	43
Şekil 7.5. Genişletilmiş örnekleme uzayındaki seçim işlemi.....	44
Şekil 7.6. Rulet tekerleği ile olasılıkların gösterimi.....	46
Şekil 7.7. Tek noktalı çaprazlama işlemine örnek	48

Şekil	Sayfa
Şekil 7.8. İki noktalı çaprazlama işlemine örnek	48
Şekil 7.9. Rassal çaprazlama işlemine örnek	49
Şekil 7.10. Mutasyon işlemine örnek	49
Şekil 7.11. Onarım işlemine örnek	50
Şekil 8.1. Birinci konuşmacı ile oluşan kör bölge	54
Şekil 8.2. Birinci ve ikinci konuşmacılar ile oluşan kör bölgeler	55
Şekil 8.3. Birinci, ikinci ve üçüncü konuşmacılar ile oluşan kör bölgeler	56
Şekil 8.4. Birinci, ikinci, üçüncü ve dördüncü konuşmacılar ile oluşan kör bölgeler	57
Şekil 8.5. Birinci, ikinci, üçüncü, dördüncü ve beşinci konuşmacılar ile oluşan kör bölgeler	58
Şekil 9.1. Konuşmacı yerlerinin gösterimi	60
Şekil 9.2. Kromozom kodlamasının gösterimi	61
Şekil 9.3. Kromozomda yer alan bit sayısının işlem süresine etkisi	63
Şekil 9.4. Çaprazlama işlemine örnek	64
Şekil 9.5. 20 cm netlikte elde edilen uygunluk değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri ve (c) z değerindeki değişimleri	68
Şekil 9.6. 5 cm netlikte elde edilen uygunluk değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri, (c) z değerindeki değişimleri	70
Şekil 9.7. 3 cm Netlikte elde edilen uygunluk değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri ve (c) z değerindeki değişimleri	73
Şekil 9.8. Konuşmacılar ve mikrofon dizisine olan uzaklıkları	75
Şekil 9.9. Ses kaynağını bulurken farklı iterasyonlardan görünümüler	76

Şekil	Sayfa
Şekil 9.10. 5 Konuşmacının yer tespiti (a) ilk konuşmacının, (b) ikinci konuşmacının, (c) üçüncü konuşmacının, (d) dördüncü konuşmacının, (e)beşinci konuşmacının yer tespiti	77
Şekil 9.11. İterasyon sayısına göre uygunluk değeri değişimi.....	80
Şekil 10.1. Üçüncü konuşmacının x değerinin incelenmesi	84

RESİMLERİN LİSTESİ

Resim	Sayfa
Resim 2.1. Robotlarla entegre edilmiş mikrofon dizisi çalışması	7
Resim 2.2. Bir albüm çalışması esnasında kullanılan mikrofon dizilerinin görünümü	7
Resim 2.3. Canlı televizyon yayınında orkestra için kullanılan mikrofon dizisi	8
Resim 2.4. Araba ortamında eller serbest görüşme için mikrofon dizisi kullanımı.....	8

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Açıklama

cm	Santimetre
Hz	Hertz
KHz	Kilo Hertz
m	Metre
m/s	Metre/Saniye
ms	Milisaniye

Kısaltmalar

Açıklama

EÇO	En Çok Olabilirlik (Maximum Likelihood)
FD	Faz Dönüşümü (Phase Transform)
GA	Genetik Algoritma(Genetic Algorithm)
GÇK	Genelleştirilmiş Çapraz Korelasyon (Generalized Cross Correlation)
MEMS	Mikro Elektrik –Mekanik Sistemli (Micro-Electro-Mechanical Systems)
YTG	Yönlendirilmiş Tepki Gücü (Steered Response Power)

1. GİRİŞ

Sinyal işleme tekniklerinin gelişimi ile, mikrofon dizilerinin kullanılması birleştirildiğinde yüksek kaliteli ses sinyali elde edilmektedir. Dizilerde filtreleme işlemleri kullanarak, istenilen ses sinyalinin diğer ses sinyallerinden ayırt edilmesi ve istenmeyen ses sinyallerinin baskılanması sağlanmaktadır. Bu işlem hüzme şekillendirme olarak adlandırılmaktadır. Mikrofon dizileri, tek bir kaynağa veya farklı algoritmalar kullanılarak bir veya birden fazla konuşmacıya odaklanacak şekilde tasarlanmaktadır. Elektronik olarak yapılacak olan yönlendirme sayesinde, el ile kontrol edilen mikrofonlu kulaklıklara veya tüfek tipi mikrofonlara gerek kalmamaktadır. Hatta, mikrofon dizilerinin birçok uygulamada kullanılan el mikrofonlarının yerini alma potansiyeli mevcuttur.

Mikrofon dizileri birçok farklı uygulamada kullanılmaktadır. Bu uygulamalara örnek olarak telekonferans sistemleri [1], konuşmacı tanıma [2] konuşmacının yerinin belirlenmesi [3-5] ve gürültülü ortamlarda ses kaydı [6] verilebilir. Yapılan çalışmalarda farklı ve ilginç mikrofon dizi tasarımları yapılmıştır. Küre şeklindeki tasarımlar [7], küp şeklinde yapılan tasarımlar [4] veya yüksek sayıda mikrofon içeren tasarımlar [5] mevcuttur. Bu konuda yapılan diğer çalışmalar, gerçek zamanlı yazılımlar geliştirmek [8] ve donanım tasarlamak ve geliştirmek [9] alanlarını kapsamaktadır.

Yapılan çalışmalar sonucunda, bazı firmalar mikrofon dizilerini kullanan ticari ürünler geliştirdi. Acoustic Magic firmasının Voice Tracker ürünü ile Polycom firmasının HDX Tavon Mikrofon Dizisi [9, 10]. Her iki firmada, kapalı ortamlarda kaliteli ses elde edilmesi için mikrofon dizisini kullanmıştır.

Uygulamaların birçoğu, düşük hata payına sahip ve hızlı konuşmacı yeri tespiti tekniklerine ihtiyaç duymaktadır. Hareket halinde olan birden fazla konuşmacının takip edildiği durumlarda [11] ise, her bir saniyede birçok gerçekçi yer tahmininin yapılması gerekmektedir. Bu konuşmacılara odaklanacak olan hüzme şekillendiricinin, yer tahminlerini bulmak için kullandığı her bir veri parçasında

konuşmacıların hareketi görmezden gelinir. Buna karşın, güncelleme hızının, yanlış yönlenebilirliğe sebebiyet vermeyecek şekilde ayarlanması gerekir.

Ayrıca, konuşmanın elde edilebilmesi için hüzme şekillendirme işleminde kullanılan uzun veri parçasından kaynaklanan kabul edilemez gecikmeler oluşur. Gerçek zamanlı uygulamalar için, bu tip uzun gecikmeler rahatsız edici olur. Bu tip faktörler, mikrofon veri ihtiyaçları üzerinde kısıtlara sebebiyet verir. Bu nedenle yapılan çalışmalarda doğru uzunlukta veri parçasının seçimi önemlidir.

İşlem süresini ve işlemci gücü ihtiyaçlarını azaltmak için birçok kaynak yeri belirleme algoritması, mikrofon dizisini çiftlere ayırmaktadır [12]. İkili gecikme tahmini, ikişerli olarak gruplandırılan mikrofonlara ses dalgasının gelme süreleri arasındaki farktan elde edilir. Konuşmacının yerini tespit etmek için, her bir çift grubuna ses sinyalinin ulaşma zamanları arasındaki farkların tahminleri kullanılmaktadır. Buna karşın, çift olarak gruplandırmayı yöntem olarak kullanan teknikler, gürültülü ortamlarda istenilen sonuçları verememektedir. İkili tekniklerin performansı kullanılan veri miktarı artırılarak iyileştirilebilir.

Birden fazla mikrofonun kullanıldığı ve birden fazla ses kaynağının bulunduğu ortamlarda, gürültüye karşı dayanıklı algoritmaların geliştirilmesi gerekmektedir. Hata eğilimi olan geliş zamanları arasındaki farka dayalı tahmin uygulamaları, bu hususta yetersiz kalmaktadır. Bu duruma alternatif olarak, belirlenmiş bir bölgenin hüzme şekillendirici kullanarak taranması ile çıktı sinyalinin gücünün en yüksek değerini aldığı noktaların bulunması yöntemi kullanılabilir.

Bu yöntem ikili yöntemle göre hesaplama süreci daha yoğun olmasına karşın, veriyi tek bir çiftten elde edilen tek zaman gecikme parametresine çevirmek yerine çoklu mikrofonlardan elde edilen sinyalleri birleştirir. Bu yöntem, parametre tahmini için bütün mikrofonlardan elde edilen verileri entegre ederek konuşmacının yerini bulmak için her bir kanal verisinin kısa bir parçasını kullanır. Gecikme zamanına bağlı tahminlere göre hüzme şekillendirme yöntemlerinin diğer bir üstünlüğü ise aynı anda birden fazla konuşmacının yerini bulabilmektedir. Bu tip senaryolarda, bu tez

çalışmasında, yönlendirilmiş hüzme şekillendiricinin gücünün, birçok yerde en yüksek değerini alması gerekmektedir. Her bir en yüksek nokta, aktif bir konuşmacının yerini göstermektedir.

Tez çalışmasının amacı, aynı anda aktif olarak konuşan beş farklı konuşmacının üç boyutlu uzayda yerini tespit edilmeye çalışılmıştır. Bu amaçla 24 mikrofon ile veri toplamış olan bir veri kümesinden yararlanılmıştır [11]. Ses kaynaklarının yerini belirlemek için hüzme şekillendirme yöntemi ile ikili tekniğin birleştirildiği Yönlendirilmiş Tepki Gücü-Faz Dönüşümü tekniği kullanılmıştır. Çıktı olarak elde edilen Yönlendirilmiş Cevap Gücü'nün en yüksek değeri aldığı yerler konuşmacıların yeri olarak belirlenmiştir. Araştırma yapılacak noktaların belirlenmesinde optimizasyon tekniği olarak değişken genlikli mutasyona dayalı genetik algoritma kullanılmıştır.

Bu çalışmanın amaçları aşağıdaki gibi listelenebilir;

- Üç boyutlu uzayda %75 üst üste gelecek şekilde konuşan beş konuşmacının yerini mikrofon dizileri kullanarak bulmak.
- Hüzme şekillendirme yöntemi olarak YTG-FD yöntemini kullanmak.
- Değişken genlikli mutasyonuna dayalı genetik algoritma kullanarak arama uzayının optimizasyonunu sağlamak.

Tez çalışmasının 2. Bölümünde, çalışmaya konu olan mikrofon dizileri hakkında teorik bilgi verilmiştir. Ayrıca bu bölümde, mikrofon dizilerinin kullanım alanları detaylı olarak anlatılmıştır.

Daha önce yapılmış olan çalışmaların yer aldığı 3. Bölümde ise bugüne kadar tez konusu ile ilgili yapılmış olan bilimsel çalışmalar ve sonuçları yer almaktadır.

4. Bölümde ise çalışmanın temelini oluşturan ses kaynağının yerinin bulunması konusu incelenmiştir. Öncelikli olarak çalışmanın yapılabilmesi için kabul edilen temel akustik koşullar anlatılmıştır. Sonra, ses kaynağından ses dalgasının yayılımı

incelenmiştir ve son olarak ise ses dalgasının yayılım yönü ve mikrofonlara olan erişimi incelenmiştir.

5. Bölümde, ses kaynağının yerinin tespitinde kullanılacak olan hüzme şekillendirici ile entegre edilecek olan mikrofon çiftleri halinde gecikme zamanlarının hesaplanmasında kullanılan Genelleştirilmiş Çapraz Korelasyon (GÇK) yöntemi hakkında bilgi verilmiştir. Ayrıca bu bölümde GÇK yönteminin teorik olarak uygulamasının yanı sıra bu yöntemde kullanılan ağırlık fonksiyonları hakkında da bilgi verilmiştir.

6. Bölümde, hüzme şekillendirici olarak kullanılacak olan yönlendirilmiş tepki gücü hakkında detaylı bilgi verilmiştir. Ayrıca bu bölümde, genelleştirilmiş çapraz korelasyonda yönlendirilmiş cevabın gösterimi verilmiştir. Son olarak ise faz dönüşümü ile yönlendirilmiş tepki gücünün birleştirilmesi ile elde edilen YTG-FD yöntemi anlatılmıştır.

7. Bölümde, ses kaynaklarının yerinin tespitinde kullanılan genetik algoritma ile ilgili bilgiler sunulmuştur. 8. Bölümde, geliştirilen yöntem hakkında detaylı bilgi verilmiştir. 9. Bölümde deneysel çalışmalar ve çalışmalar sonucunda elde edilen veriler anlatılmıştır. Son bölüm olan 10. Bölümde ise elde edilen sonuçlar yorumlanarak değerlendirilmiştir.

2. MİKROFON DİZİLERİ

2.1. Mikrofon

Mikrofon, ses sinyallerini elektrik sinyaline çeviren sensörler olarak tanımlanabilir. İlk defa 1876 yılında Emile Berliner tarafından telefonlarda kullanılmak için ses sinyali vericisi olarak icat edilmiştir. İcat edildiği günden bugüne kadar birçok alanda ve cihazda kullanılmaktadır. Örneğin; telefon, ses kayıt sistemleri, işitme cihazları, canlı ve kayıtlı video mühendislikleri, radyo ve televizyon yayınları.

Bir ses dalgasındaki titreşimlerin, ses dalgasının benzeri olan elektriksel sinyali üretmeye yarayan birçok fiziksel prensip vardır. Bunlar, bağlantı direncinin değişimi, piezo elektrik, elektromanyetik ve manyetostriksiyon (mıknatıslandığı zaman bir cismin boyunda meydana gelen değişiklik) prensipleri olarak sıralanabilir. Bütün bu prensipler ve diğerleri yıllarca yapılan çalışmalar sonucunda piezo-elektrik, elektromanyetik, elektrostatik ve kapasitif prensipleri uygulamada kullanılmıştır.

Bütün mikrofonlar ses dalgalarına tepki gösteren çeşitli şekillerde yapılmış diyafram ya da benzeri bir elemana sahiptir. Mikrofona gelen ses dalgaları diyaframa çarpar ve ses basıncındaki değişikliklere göre diyafram içe veya dışa doğru hareket ederek mekanik titreşimi oluşturur. Bu titreşimler sonucunda mikrofonun çıkış uçlarında bir gerilim meydana gelir. Çıkış uçlarında meydana gelen gerilim, hareket eden parçanın ya hızı ya da titreşimlerinin genliği ile orantılıdır.

Üretiminde kullanılan prensiplere ve yapılarına göre birçok farklı tip mikrofon bulunmaktadır. Mikrofon çeşitleri olarak Kondansatörlü (Kapasitif) Mikrofonlar, Elektrodinamik Mikrofonlar, Şerit Mikrofonlar, Karbonlu Mikrofonlar, Manyetik Mikrofonlar, Kristalli Mikrofonlar, Piezo Elektrik Mikrofonlar, Fiber Optik Mikrofonlar, Lazer Mikrofonlar, MEMS (Mikro Elektrik –Mekanik Sistemli) Mikrofonlar sayılabilir [13].

2.2. Mikrofon Dizileri

Yapılan alıřmalar sonucunda tek bir mikrofonunun kullanılmasının yanı sıra aynı anda birden fazla mikrofonun kullanılmasının faydalı olabileceęi ortaya konulmuřtur. Bu kapsamda birden fazla mikrofondan oluřan yapılar yani mikrofon dizileri ortaya konulmuřtur. Mikrofon dizisi, uzaysal olarak farklı yerlere yerleřtirilmiř olan mikrofonlardan oluřan, bütn mikrofonların eř-zamanlı olarak alıřtıęı bir sistemdir.

Mikrofon dizilerinin normal mikrofonlara göre temelde iki stnlę bulunmaktadır. Bunlardan ilki kaynaęa yakın olma zorunluluęunu ortadan kaldırmalarıdır. Ses sinyali elde edilir iken, en ok dikkat edilen husus ses sinyali ile grltnn etkileřiminin en az seviyede tutulmasıdır. Normal mikrofonlarda bu soruna zm olarak mikrofon, ses kaynaęına olabildięince yakın bir yere yerleřtirilmektedir. Bu sayede mikrofon doęrudan istenilen ses sinyalini alırken, grlt sinyalini daha az almaktadır. Fakat birden fazla konuřmacının bulunduęu konferans salonları gibi ortamlarda bu tip uygulamalar yetersiz kalmaktadır. Buna karřılık mikrofon dizileri uzaysal olarak istenildięi gibi yerleřtirilebilmektedir. Bu mikrofonlardan elde edilen sinyaller ile sinyal iřleme sreci gerekleřtirildięinde sinyal kalitesini bozan grlt ortadan kaldırılabilmekte ve metrelerce uzaktan ses sinyali elde edilebilmektedir.

Dięer bir stnlę ise mikrofon dizileri sayesinde geliř aısı hesaplama iřlemleri yapılabilmektedir ve bu sayede sesin kaynaęı veya konuřmacının takibi gerekleřtirilebilmektedir. Saęladıęı stnlkler sayesinde birok farklı alanda kendilerine kullanım alanı bulmuřlardır.

2.3. Mikrofon Dizilerinin Kullanım Alanları

Mikrofon dizilerinin en ok kullanıldıęı alanların bařında konuřmacı tanıma alıřmaları gelmektedir. Bu alanda yapılan alıřmalar iki gruba ayrılmaktadırlar. Konuřmacının yerinin tespiti ve konuřmacının tanımlanması. Konuřmacı yerinin

tespit etme çalışmalarında karşılaşılan en yaygın uygulama, konferans odalarında konuşmacının yerini tespit ettikten sonra kamera hareketi ile entegre ederek konuşmacının görüntüsünü kaydetme çalışmalarıdır. Ayrıca robotlara entegre edilmiş mikrofon dizileri ile konuşmacının yerini tespit ettikten sonra robotu, konuşmacının bulunduğu yere gönderilmesi de yapılan çalışmalar arasında yer almaktadır. Çalışma ile ilgili detay, Resim 2.1’de verilmiştir.

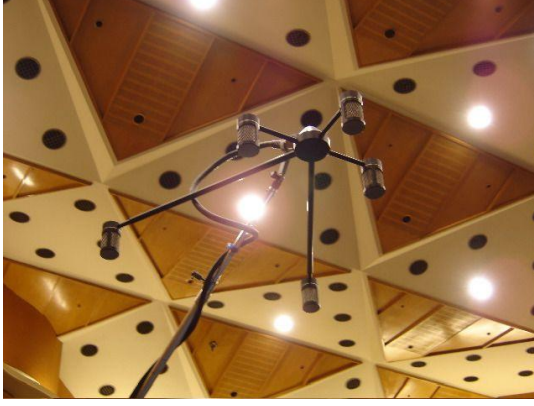


Resim 2.1. Robotlarla entegre edilmiş mikrofon dizisi çalışması [4]

Film, televizyon programları ve hatta canlı gösterilerin kayıt altına alınmasında gürültünün ortadan kaldırılarak en iyi ses kalitesine elde etmekte de kullanılmaktadır. Bu konuda yapılan çalışmalara örnekler Resim 2.2 ve Resim 2.3’te verilmiştir.



Resim 2.2. Bir albüm çalışması esnasında kullanılan mikrofon dizilerinin görünümü [14]



Resim 2.3. Canlı televizyon yayınında orkestra için kullanılan mikrofon dizisi [15]

Bazı işitme cihazları da minyatür mikrofon dizileri kullanarak gürültülü ortamlarda konuşmaların daha net elde edilmesini sağlamıştır [16]. Bu sayede işitme cihazlarında başarı ölçümü olarak nitelendirilen kelime anlaşılma oranında artış yaşanmıştır.

Ayrıca araba ortamlarında eller serbest (hands-free) telefon konuşmalarının yapılabilmesi için mikrofon dizileri kullanılmıştır [17]. Bu çalışmaya örnek Resim 2.4'te verilmiştir. Son günlerde oldukça popüler olan akıllı ev konseptinde her bir odaya ayrı ayrı yerleştirilen mikrofon dizileri sayesinde kullanıcının istediği yerden komut vermesi sağlanmıştır. Alınan komut evde yer alan ana bilgisayara iletilerek istenilen işlemin hayata geçmesi sağlanmıştır.



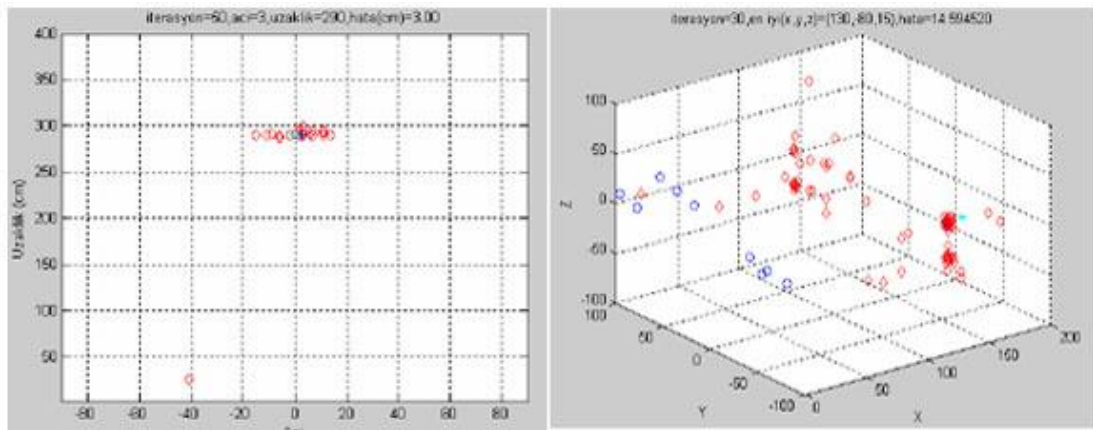
Resim 2.4. Araba ortamında eller serbest görüşme için mikrofon dizisi kullanımı [17]

Uzun zamandır kullanılabilir halde bulunan mikrofon dizileri, dijital sinyal işleme tekniklerinin gelişmesi ve her geçen gün gelişıyor olması nedeniyle popülerlik kazanmıştır. Her geçen gün, hayattaki kullanım alanları artmaktadır.

3. DAHA ÖNCE YAPILMIŞ ÇALIŞMALAR

Bu bölümde mikrofon dizileri ile ilgili daha önce yapılan ve tez çalışmasında yol gösterici olarak kullanılmış olan akademik çalışmalardan bahsedilecektir. Kullandıkları yöntemler ve elde ettikleri sonuçlar incelenecektir. İncelenen konular; mikrofon dizilerinde genetik algoritmanın çözüm kümesi bulunmasında kullanımı, mikrofon dizilerinde kullanılan uygunluk değeri fonksiyonlarını ve mikrofon dizileri ile çoklu konuşmacıların yer tespiti.

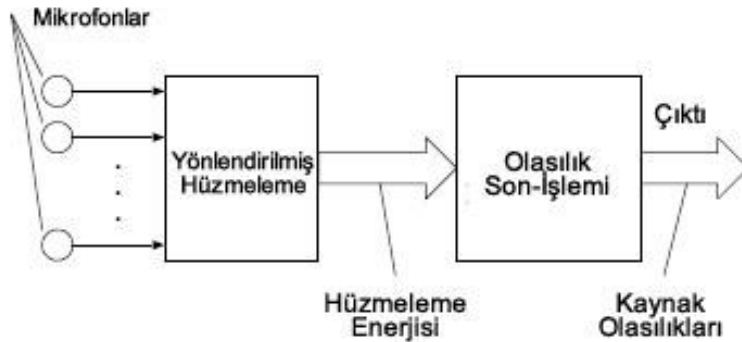
Mikrofon dizilerinde çözüm adaylarının belirlenmesi için stokastik yöntemlerden genetik algoritmayı kullanan bir çalışmada Çontar [18] tarafından yapılmıştır. Bu çalışmada, tek konuşmacının yerinin tespiti amaçlanmıştır. Uygunluk değeri olarak geciktir-ve-topla hüzme şekillendiricisi kullanılmıştır. Farklı mikrofon grupları ile öncelikle 2 boyutlu olarak sonrasında ise farklı bir dizi ile 3 boyutlu olarak ses kaynağının yerinin bulunması amaçlanmıştır. Şekil 3.1’de gösterildiği gibi, 2 boyutlu yapılan çalışmada kaynağın yeri 3 cm bir kayma ile başarılı olarak bulunurken, 3 boyutlu düzlemde hata oranı biraz daha yükselmiştir. Bu çalışmada geciktir-ve-topla hüzme şekillendirici ve yaka mikrofonları kullanılmıştır.



Şekil 3.1. Çalışma sonucunda elde edilen sonuçlar [18]

Valin ve ark. yaptıkları çalışmada mikrofon dizilerini kullanarak bir robotun hareket ettirilmesini hedeflemişlerdir. Bu çalışmada Şekil 3.2’de görüldüğü gibi önce elde

edilen sinyalleri geciktir-ve-topla hüzme şekillendiriciden geçirmişler ve sonrasında olasılık son-işlemine tabi tutarak konuşmacının yerini belirlemişlerdir. Olasılık son-işlemi esnasında, hüzme şekillendiriciden elde edilen değerlere göre konuşmacının yerleri olasılıklar kullanılarak bulunmaktadır. Yapılan çalışmada robotun üzerine küp şeklinde bir kasanın her köşesine yerleştirilmiş 8 mikrofondan yararlanılmışlardır. Farklı uzaklıklarda (3m, 5m ve 7m) ve farklı sesler için (el çırpma, konuşma, gürültü) denemeler yapmışlardır. Sonuçta, Çizelge 3.1’de verilen değerler elde edilmiştir [4].



Şekil 3.2. Sistemin Yapısı [4]

Çizelge 3.1. Çalışmanın sonuçları [4]

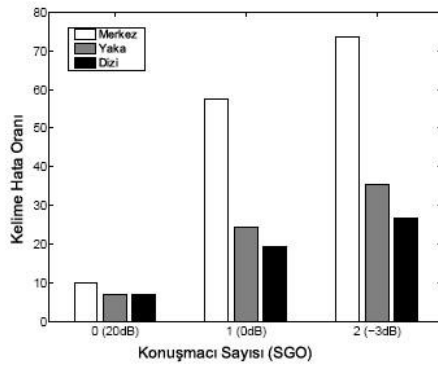
Ses Kaynağı	3 Metre	5 Metre	7 Metre
El Çırpma	92%	94%	84%
Konuşma	100%	90%	42%
Gürültü	100%	100%	100%

Moore ve ark. tarafından yapılan çalışmada, tek kişinin, iki kişinin aynı anda veya üç kişinin aynı anda konuştuğu bir çalışma geliştirilmiştir. Bu çalışmada uygunluk fonksiyonu olarak geciktir-ve-topla hüzmeleme şekillendiricisi kullanılmıştır. 20 cm çapa sahip bir dairenin etrafına eşit uzaklıklarda 8 mikrofon yerleştirilmiştir, her bir kullanıcıya yaka mikrofoni kullanıdırılmıştır ve masanın orta kısmına merkezi bir mikrofon yerleştirilmiştir. Sayıların seslendirilmiş olduğu bir veritabanı kullanılmıştır. Konuşmalar ayrıştırılarak, kelime hata oranlarına bakılmıştır. En iyi sonuç, Şekil 3.3’de ve Çizelge 3.2’de gösterildiği gibi mikrofon dizileri ile elde edilmiştir [19]. Ancak çalışmada kullanılan uygunluk değeri için her bir

konuşmacının yaka mikrofona sahip olması gerekmektedir. Her konuşmacı için yaka mikrofona kullanımı gerçekçi değildir.

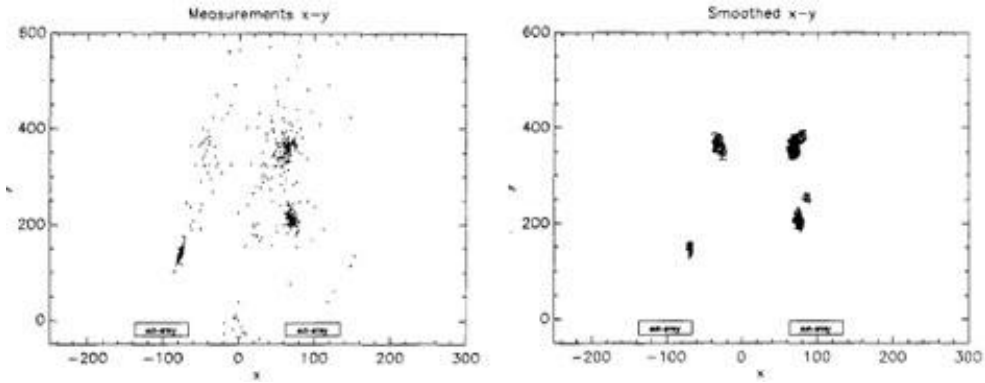
Çizelge 3.2. Kelime hata oranı sonuçları (%) [19]

Senaryo	Merkez	Yaka	Dizi
S ₁	10.06	7.01	7.00
S ₁₂	60.45	26.69	19.37
S ₁₃	54.67	22.17	19.26
S ₁₂₃	73.55	35.25	26.64



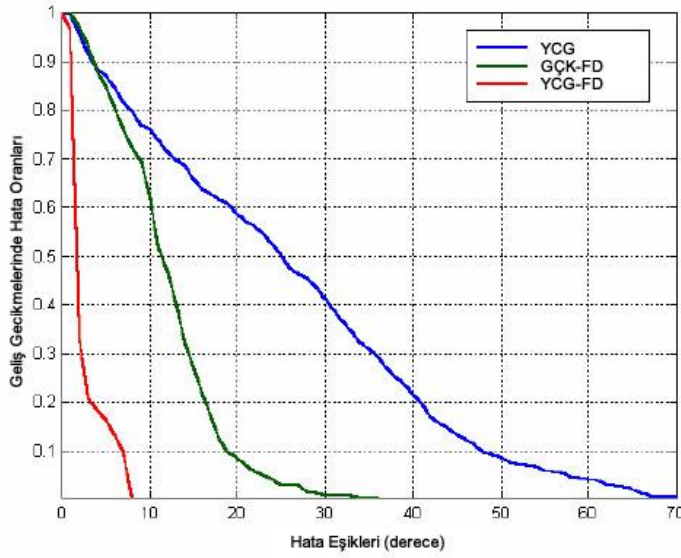
Şekil 3.3. Farklı konuşmacı sayıları için kelime hata oranları [19]

Benzer bir çalışmada Sturim ve ark. [20] tarafından hareket eden konuşmacılar için yapılmıştır. Bu çalışmada hareketlerin yönünü ve doğrultularını belirlemek için Kalman filtresinden faydalanılmıştır. Geliştirilen metotla aynı anda birden fazla Kalman filtresi paralel olarak çalışarak uygunluk değerlerini vermiştir. Şekil 3.4’de ilk şekilde genel olarak bulunan konuşmacı yerleri gösterilirken, ikinci şekilde Kalman filtreleri tarafından yumuşatılmış konuşmacı yerleri görülmektedir. Elde edilen sonuçlardan anlaşıldığına göre, konuşmacı yerlerinin tespitinde netliğin yeterli olmaması nedeniyle yöntem geliştirilmeye ihtiyaç duymaktadır.



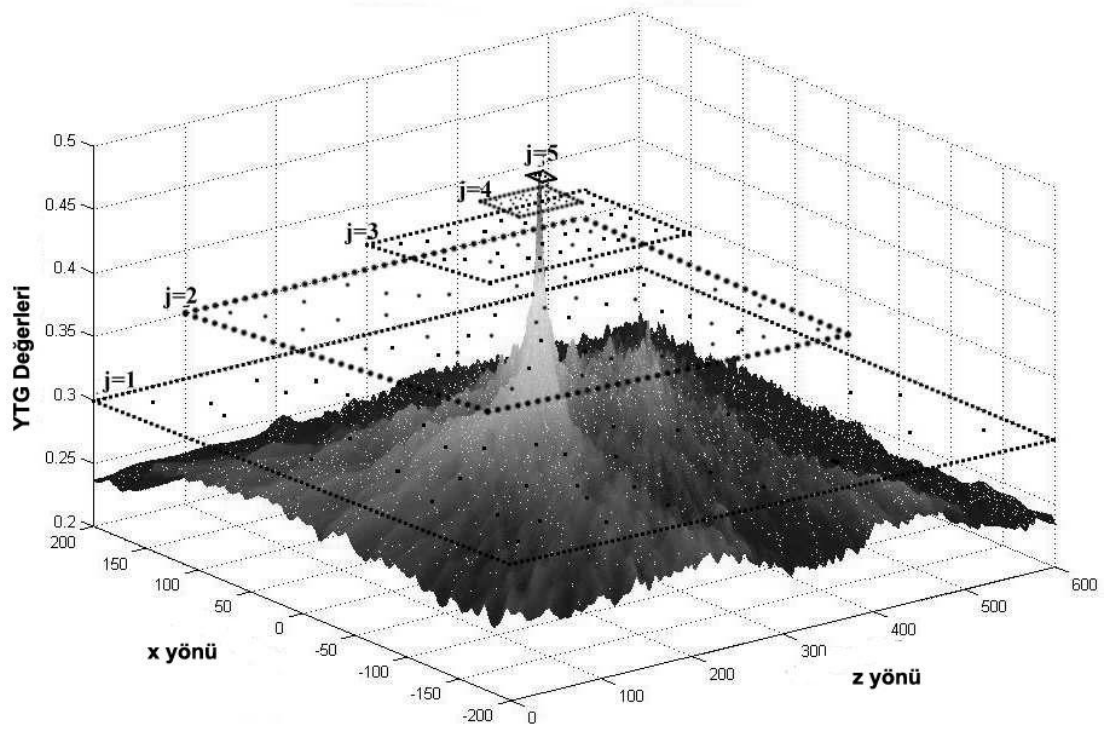
Şekil 3.4. Yapılan çalışma sonucunda 4 konuşmacının yerlerinin tespiti [20]

Dibiase yaptığı çalışmada konuşmacının yerlerini bulmak için kullanılan farklı uygunluk fonksiyonlarını, farklı sayılarda mikrofonlardan oluşan diziler ile deneyerek değerlendirmiştir [5]. Bu çalışmada, ses kaynağının yerini bulmak için yaka mikrofonu yardımı ile elde edilen referans sinyal yerine mikrofon dizilerinin uzaysal koordinatlarının bilindiği kabul edilmiştir. Bu doğrultuda uygunluk fonksiyonlarından yönlendirilmiş tepki gücü, yönlendirilmiş tepki gücü- faz dönüşümü ve genelleştirilmiş çapraz korelasyon- faz dönüşümü metotları mukayese edilmiştir. Çalışmada 8 mikrofondan oluşan, 15 mikrofondan oluşan ve 256 mikrofondan oluşan üç düzenek üzerinde deneyler yapılmıştır. Sonuçta, yönlendirilmiş tepki gücü- faz dönüşümü yöntemi, diğer iki yöntemden üstün çıkmıştır. Şekil 3.5'te 8 mikrofondan oluşan dizi ile yapılan çalışmanın sonuçları gösterilmiştir.



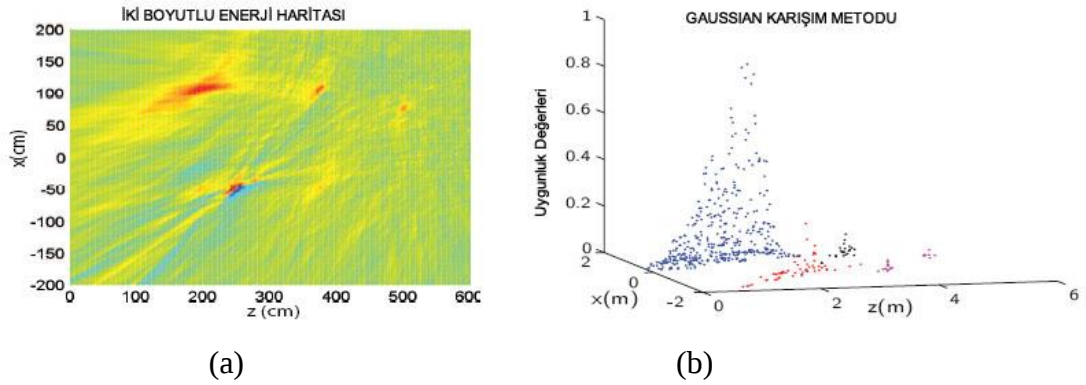
Şekil 3.5. Üç yöntemin karşılaştırma sonuçları [5]

Do ve ark. tarafından yapılan bir çalışmada YTG-FD yöntemi uygunluk değeri fonksiyonu olarak ve stokastik bölge konsantrasyonu yöntemi de incelenecek olan çözümleri bulmak için kullanılmıştır [11]. Şekilde 3.6'da yöntemin çalışma prensibi gösterilmiştir. Bu yöntemde yüksek uygunluk değerine sahip bölgeler belirlenerek her aşamada bölgenin boyutu küçültülmektedir. Bu çalışmada tek bir konuşmacının yerinin tespiti hedeflenmiştir. Çalışma, bu tez çalışmasına ışık tutmuştur. Stokastik bölge konsantrasyonu yönteminde üç farklı parametre için çalışma yapılmıştır. Sonuçta 40 veri çerçevesinde konuşmacının yeri tespit edilebilmiştir.

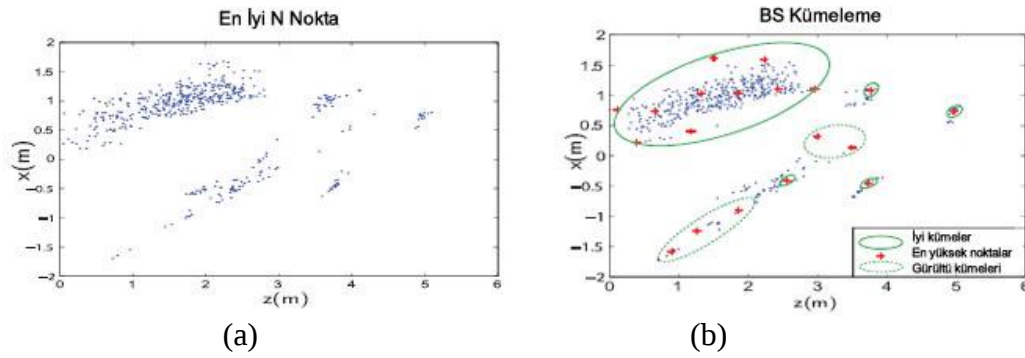


Şekil 3.6. Stokastik bölge konsantrasyonu yöntemi [11]

Bu bölümde son olarak yer alacak olan çalışma Do ve ark. tarafından yapılan ve aynı anda beş konuşmacının yerini bulmayı hedefleyen çalışmadır [8]. Bu çalışmada uygunluk fonksiyonu olarak YTG-FD tekniği kullanılmıştır. Çözümlerin bulunması için ise iki farklı yöntemden faydalanılmıştır. İlk metot Gauss Karışımı Metodu (Gaussian Mixture Method) olarak adlandırılan ve enerji haritalarında oluşan en yüksek değerlere göre konuşmacıların yerini belirlemektedir. İkinci metot ise bölge sıfırlayarak kümeleme (Region Zeroing) olarak adlandırılan ve elde edilen uygunluk fonksiyon değerlerinde benzer sonuçları gruplayarak konuşmacıların yerlerini tespit eden yöntemdir. Bu yöntemle göre uygunluk değerlerinin kümелendiği bölgeler belirlenerek, o bölgede tek bir konuşmacının yer aldığı kabul edilir. Çalışmada aynı anda konuşan 5 konuşmacı yer almıştır. Bu tez çalışmasında kullanılan veri tabanı bu çalışmadan alınmıştır. Şekil 3.7’de Gauss Karışım Metodu ile elde edilen sonuçlar verilmiştir. Şekil 3.8’te bölge sıfırlayarak kümeleme yöntemi ile elde edilen sonuçlar verilmiştir.



Şekil 3.7. Gauss karışım metodu ile elde edilen sonuçlar (a) iki boyutlu enerji haritası (b) uygunluk değerleri dağılımı [8]



Şekil 3.8. Bölge sıfırlayarak kümeleme yöntemi ile elde edilen sonuçlar (a) elde edilen en iyi N nokta (b) bölge sıfırlayarak kümeleme yöntemi sonucunda elde edilen bölgeler[8]

Çizelge 3.3’de ise her iki metotla bir başka çalışmada geliştirilen toplaklama kümeleme (agglomerative clustering) olarak adlandırılan yöntemle elde edilen sonuçların karşılaştırılması verilmiştir. Bu çizelgede “doğru pencere” tanımı ile asıl yerler bulunduğu, “extra” tanımı asıl yerler tam olarak bulunamadığında ve “kaçırılan” tanımı ise yerler bulunamadığında kullanılmıştır. Çizelgedan da anlaşıldığı gibi %20 oranında konuşmacının yeri tespit konusunda kayıp mevcuttur.

Çizelge 3.3.Üç yöntemin karşılaştırma sonuçları [8]

Algoritma	Doğru Pencere %	Ekstra %	Kaçırılan%
GKM	73.5	18.1	17.5
BSK	60.1	11.1	23.6
TK	51.3	8.2	30.1

4. SES KAYNAĞININ YERİNİN BULUNMASI

Ses dalgaları titreşim kaynağından enerji taşırlar. Bu enerji, çeşitli ortamlar tarafından iletilir. Bu esnada ortam yer değiştirmez, hareket eden madde değil, hareket enerjisidir. Ses kaynağından çıkan ses maddenin taneciklerini titreştirir. Bu nedenle ses yayılır. Ses, tanecikler halinde yayılır, tanecikler ne kadar sık ise o kadar hızlıdır. Bu nedenle ses boşlukta yayılamaz.

Yapılan sinyal işleme işlemlerinde ses dalgalarının doğrusal dalgalar olarak yayıldığı kabul edilir [5]. Bu kabul sayesinde ses kaynağı ile mikrofonlar arasındaki akustik yol doğrusal sistem olarak kabul edilebilir. Bu tip doğrusal modellerin geçerli kabul edilebilmesi için, yayılma ortamının temel akustik koşullarını sağlaması gerekmektedir. Bu koşullar küçük boyutlu konuşma ortamları için gerçekçi olup, mikrofon dizileri sinyal işleme tekniklerinde kullanılmaktadır [21].

4.1. Temel Akustik Koşulları

Temel akustik koşulları, hem hoparlör gibi akustik kaynakların hem de hava gibi akustik ortamların özelliklerini tanımlamak için kullanılır. Temel akustik koşulları beş temel elemandan oluşmaktadır [5].

1. Kaynaklar küresel ses dalgaları yaymaktadır. Kaynağın şeklinin ve büyüklüğünün ses dalgalarının yayılımı üzerinde etkisinin olmadığı kabul edilmektedir.
2. Doppler etkisi ihmal edilebilir. Kaynak hareket halinde olsa bile sesin hızı ile karşılaştırıldığında kaynağın hareketi önemsiz kalmaktadır. Bu nedenle, frekansta Doppler ötelemesi ihmal edilebilir.
3. Ortam homojendir. Sesin yayılma hızının her yerde ve her zaman aynı olduğu kabul edilmektedir. Sesin hızı olarak bilinen 331.4 m/s olarak alınmaktadır. Diğer bir deyişle ortam yansız olduğu kabul edilmektedir.

4. Ortam saçınımsızdır (nondispersive). Saçımı frekansa bağlı olarak yayılım hızını etkilemektedir. Bu durum doğrusal sistemler için kabul edilebilir bir durum değildir. Bu nedenle doğrusal sistemlerde saçınım önemsenmemektedir.
5. Ortam kayıpsızdır. Kayıpsız ortam yayılma dalgalarının enerjisini emerek kayba sebebiyet vermez.

4.2. Doğrudan Yol Yayılımı (Direct Path Propagation)

Ses dalgalarının zor ve karışık geometriye sahip ortamlardaki yayılımını sadeleştirmek istendiğinde, fiziksel modellerden yararlanılmaktadır. Doğrudan yol yayılım modeli de bu modellerden bir tanesidir.

Temel akustik koşullarda, uzaysal olarak sabit kabul edilen bir ses kaynağının oluşturduğu ses dalgası, ses sinyali, $s(t)$, ile doğrudan ilişkilidir. Bu kabul boş alanlar için geçerli olduğu gibi, oda gibi kapalı alanlar içinde geçerlidir. Açık alanlarda, sesin yayılımı engelleyecek insan, ağaç, duvar gibi objeler bulunmaz. Dolayısıyla ses hiçbir engel ile karşılaşmadan yayılır. Ancak mikrofon dizilerinin kullanıldığı ortamlar için bu durum geçerli değildir. Bu tip ortamlarda ses, farklı objelere çarparak hem zayıflar, hem de gürültü oluşturur. Buna karşın, interferansların olduğu ortamlarda da doğrudan yol yayılım modeli, kaynaktan mikrofona olan yayılımı tanımlamakta kullanılmaktadır.

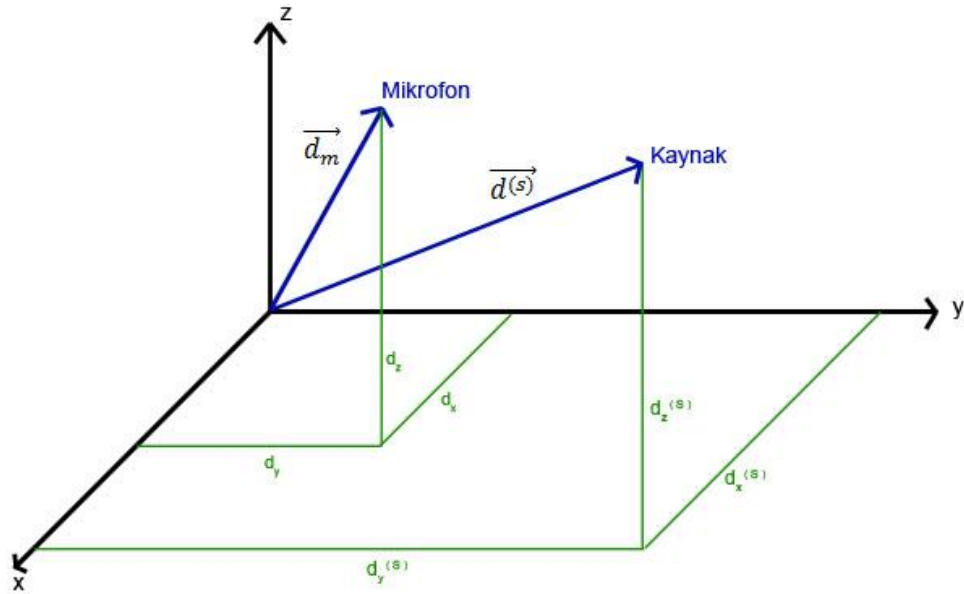
Ortamin doğrusallığı sayesinde, mikrofonda elde edilen sinyal, doğrudan yol bileşenlerinin toplamı ile odanın yüzeylerinde yansıyan ses dalgalarının toplamı olarak modellenir. Sinyal işleme algoritmaları da doğrudan yol bileşenleri ile yansımaları ve gürültüyü ayırt ederek konuşmacının yerini bulmayı hedeflerler [5].

Doğrudan yol yayılımı, dalga denkleminde elde edilmektedir [21]. Ses kaynağından, $s(t)$, r uzaklıkta oluşan dalga alanı Eş. 4.1’de verildiği gibidir

$$f_{direk}(r, t) = \frac{a}{r} s(t - r/c) \quad (4.1)$$

Bu eşitlik, dalga alanı, orijinal ses sinyalinin ölçeklendirilebilir ve zamana bağlı halini elde etmek için kullanılmaktadır. Zayıflama faktörü, kaynağa olan uzaklık ile ters orantılıdır. Gecikme ise uzaklığın ses hızına, c , olan oranına eşittir. Sabit a değişkeni ise ortama bağlıdır.

Mikrofon ile ses kaynağının birbirine olan uzaklığının, üç boyutlu Kartezyen Koordinat sisteminde gösterimi Şekil 4.1’de verilmiştir.



Şekil 4.1. Kartezyen koordinatlarına göre ses kaynağının ve mikrofonun yerleri

Mikrofonun ses kaynağına olan uzaklığı Eş. 4.2 ile hesaplanmaktadır.

$$r_m \equiv \left| \vec{d}_m - \vec{d}^{(s)} \right| \quad (4.2)$$

Bu eşitlikte yer alan $\vec{d}_m$, mikrofonun ve $\vec{d}^{(s)}$ ses kaynağının Kartezyen Koordinat sistemindeki yerlerinin üç elemanlı vektörleridir ve Eş. 4.3 verildiği gibi gösterilir

$$\vec{d}_m \equiv \begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} \quad \vec{d}^{(s)} \equiv \begin{bmatrix} d_x^{(s)} \\ d_y^{(s)} \\ d_z^{(s)} \end{bmatrix} \quad (4.3)$$

Kaynaktan mikrofonta olan yayılım süresi ise Eş. 4.4 ile hesaplanmaktadır.

$$\tau_m \equiv \frac{r_m}{c} \quad (4.4)$$

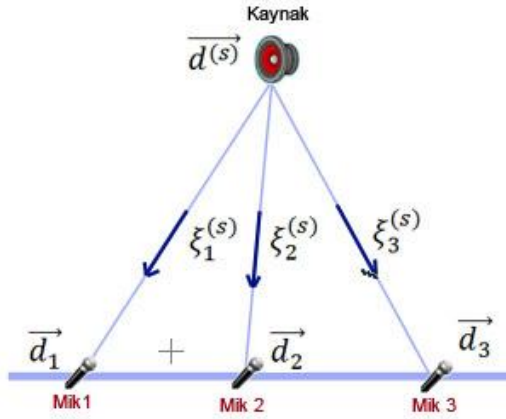
Bu nedenle, $\vec{d}^{(s)}$ ses kaynağında üretilmiş $\vec{d}_m$ noktasındaki dalga alanı, Eş. 4.5'te verildiği gibidir.

$$f_{direk}(\vec{d}_m, \vec{d}^{(s)}, t) = \frac{a}{r_m} s(t - \tau_m) \quad (4.5)$$

4.3. Yayılım Yönü ve Erişim

Eşitlik 4.5 verilen doğrudan yol bileşenleri, sesin kaynağından mikrofon m' e giderken doğrusal bir hat üzerinden yayıldığını göstermektedir. Mikrofonta doğru, ses dalgasının izlediği yolun yönü, yayılım yönü olarak adlandırılır. Bu yönün tam ters yönü ise geliş yönü olarak adlandırılır ve mikrofonta gelen ses dalgalarının toplamına eşittir. Her iki terimde ses kaynağı tarafından oluşturulan ses dalgası ile mikrofon arasındaki ilişkiyi tanımlamak için kullanılır.

Uzaysal alanda farklı koordinatlara yerleştirilmiş M adet mikrofondan oluşan mikrofon dizilerinde, her bir mikrofonta giden M adet üst üste binmiş bileşenlerden oluşan ses dalgaları mevcuttur. Bu dalgaların koordinat sisteminde doğrusal olarak tanımlanmasına yayılım yönü denilmektedir.



Şekil 4.2. 3 elemanlı mikrofon dizisi için yayılım yönleri

Şekil 4.2 'de üç elemandan oluşan bir mikrofon dizisi için yayılım yönleri gösterilmiştir. Yayılım yönleri, ses kaynağının yeri ve mikrofonların yeri ile Eş. 4.6'da verildiği gibi ifade edilmektedir.

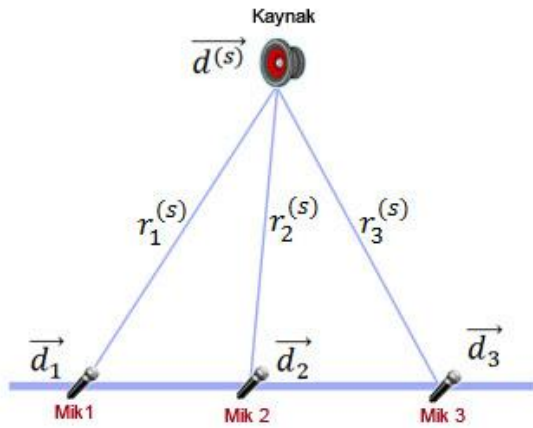
$$\vec{\xi}_m^{(s)} \equiv \frac{\vec{d}_m - \vec{d}^{(s)}}{|\vec{d}_m - \vec{d}^{(s)}|}; m = 1, \dots, M \text{ kadar} \quad (4.6)$$

Geliş yönü, yayılım yönünün tam ters yönünü gösteren vektördür ve ses kaynağını göstermektedir. Her bir mikrofon için ayrı bir geliş yönü vardır. Yayılım yönü, $\vec{\xi}_m^{(s)}$ olarak tanımlanırsa, geliş yönü $-\vec{\xi}_m^{(s)}$ vektörü ile tanımlanır.

5. GENELLEŞTİRİLMİŞ ÇAPRAZ KORELÂSYON (GÇK)

5.1. Genel Bilgiler

Genelleştirilmiş Çapraz Korelasyon (GÇK) metodu, iki mikrofona gelen ses dalgalarının arasındaki zaman farkını bulmak için kullanılmaktadır. Birden fazla mikrofon çiftinin geliş zamanları arasındaki farklar grubu, ses kaynağının yerinin noktasal olarak bulunması için kullanılmaktadır. Üç mikrofondan oluşan ve gecikme zamanları arasındaki farklar sayesinde ses kaynağının yerini bulmak için kullanılan yapının örneği Şekil 5.1’de verilmiştir. Şekilde görüldüğü gibi her bir mikrofonun kaynağa olan uzaklığı $r_m^{(s)}$ olarak $m=1,2,3$ şeklinde ifade edilmektedir.



Şekil 5.1. Gecikme zamanları arasındaki farklar sayesinde ses kaynağının yerini bulmak

Eşitlik 4.4’te verildiği gibi, yayılım gecikmesi uzaklığın ses hızına olan oranı ile ölçülmektedir. A ve b olmak üzere iki mikrofon arasındaki zaman farkı ise τ_{ab} şeklinde gösterilir ve Eş. 5.1’de verildiği gibi yayılım gecikmelerinin farkı olarak hesaplanır.

$$\tau_{ab} = \tau_a - \tau_b \quad (5.1)$$

Eşitlikte τ_a ve τ_b tanımları yerlerine yerleştirilirse, Eş. 5.2 elde edilir.

$$\tau_{ab} = \frac{r_a^{(s)} - r_b^{(s)}}{c} \quad (5.2)$$

Eşitlik, uzaklıklarla gecikmelerin ilişkisi şeklinde yazılmak istenildiğinde Eş. 5.3 elde edilir.

$$r_a^{(s)} - r_b^{(s)} = c \tau_{ab} \quad (5.3)$$

Buradan da anlaşıldığı gibi, kaynağın yerini birden fazla mikrofon çiftinin gecikmelerinin yardımı ile noktasal olarak hesaplanabilmektedir.

GÇK metodu mikrofon dizileri ile yapılan çalışmalarda sıklıkla kullanılmaktadır. Tercih edilmesindeki en önemli sebep, referans sinyale ihtiyaç duyulmadan sadece dizideki mikrofonların koordinatlarının bilinmesi ile konuşmacının yeri gerçekçi olarak bulunabilmesidir.

5.2. Genelleştirilmiş Çapraz Korelasyon Metodunun Uygulaması

Mikrofon sinyal modeli Eş 5.4 ile ifade edilmektedir.

$$x_m(t) = \frac{1}{r_m} s(t - \tau_m) * \gamma_m(\vec{d^s}, t) + \widehat{v}_m(t) \quad (5.4)$$

Bu eşitlikte $\gamma_m(\vec{d^s}, t)$, m. mikrofon kanalının frekans ve faz tepkilerini ortaya koyan doğrusal filtredir. $\widehat{v}_m(t)$ ise sinyaldeki yansıma gürültüsünü ve gerçek gürültüyü içeren toplam gürültüdür. r_m , kaynağın mikrofona olan uzaklığını ifade ederken, τ_m gecikme süresini ifade etmektedir.

Mikrofon çifti olan mikrofon-1 ve mikrofon-2 için geliş zaman farkı Eş 5.5'te gösterilmiştir.

$$\tau_{12} \equiv \tau_2 - \tau_1 \quad (5.5)$$

Bu eşitlikten faydalanarak τ_2 Eş. 5.6'teki gibi ifade edilebilir.

$$\tau_2 = \tau_1 + \tau_{12} \quad (5.6)$$

Elde edilen denklemler mikrofon sinyal modelinde yerine yerleştirilir ise Eş. 5.7 ve 5.8 elde edilir.

$$x_1(t) = \frac{1}{r_1} s(t - \tau_1) * \gamma_1(\vec{d^s}, t) + \widehat{v}_1(t) \quad (5.7)$$

$$x_2(t) = \frac{1}{r_2} s(t - \tau_1 - \tau_{12}) * \gamma_2(\vec{d^s}, t) + \widehat{v}_2(t) \quad (5.8)$$

Mikrofonların aynı sistemin parçaları olduğu göz önünde bulundurulursa $s(t - \tau_1)$ 'ın mikrofon-1 sinyalinin ölçeklendirilmiş hali olduğu ve mikrofon-2 sinyali için ise ölçeklendirilmiş ve aynı zamanda ötelenmiş hali olduğu görülür. Bu iki sinyalin çapraz korelasyonu, gecikme süresi olan τ_{12} kadar $s(t)$ sinyalinin ötelenmiş halindeki faz kaymasını gösterir. $x_1(t)$ ve $x_2(t)$ sinyalleri arasındaki çapraz korelasyon Eş. 5.9'da verilmiştir.

$$c_{12}(\tau) = \int_{-\infty}^{+\infty} x_1(t) x_2(t + \tau) dt \quad (5.9)$$

Çapraz korelasyon fonksiyonun Fourier dönüşümü, çapraz spektrum yoğunluğu veya çapraz spektrum olarak adlandırılır ve Eş. 5.10'da verildiği şekilde hesaplanır.

$$C_{12}(\omega) = \int_{-\infty}^{+\infty} c_{12}(\tau) e^{-j\omega\tau} d\tau \quad (5.10)$$

Eş. 5.9'dan Eş. 5.10 çıkarıldığında ve Fourier dönüşümünün konvolusyon özelliğinden faydalananarak Fourier dönüşüm $x_1(t)$ ve $x_2(t)$ değişkenleri Eş. 5.11'de verilmiştir.

$$C_{12}(\omega) = X_1(\omega)X_2^*(\omega) \quad (5.11)$$

$X_1(\omega)$, $x_1(t)$ 'nin Fourier dönüşümü ve $X_2^*(\omega)$ ise $x_2(t)$ 'nin karmaşık eşleniğidir. Eş. 5.11 de verilen çapraz korelasyon fonksiyonunun $X_1(\omega)$ ve $X_2'(\omega)$ cinsinden ters Fourier dönüşümü Eş. 5.12'de verilmiştir.

$$c_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X_1(\omega)X_2'(\omega)e^{j\omega\tau} d\omega \quad (5.12)$$

Genelleştirilmiş Çapraz Korelasyon Fonksiyonu, $R_{12}(\tau)$, $x_1(t)$ ve $x_2(t)$ sinyallerinin filtrelenmiş hallerinin çapraz korelasyonudur. Filtrelerin Fourier dönüşümünden elde edilmiş halleri olan $G_1(\omega)$ ve $G_2(\omega)$ halindeki GÇK fonksiyonu Eş. 5.13'te verilmiştir.

$$R_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} ((G_1(\omega)X_1(\omega))(G_2(\omega)X_2(\omega))'e^{j\omega\tau} d\omega \quad (5.13)$$

Sinyaller ve filtreler bir araya gelecek şekilde yeniden yazıldığında Eş. 5.14 elde edilir.

$$R_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G_1(\omega)G_2'(\omega)X_1(\omega)X_2'(\omega)e^{j\omega\tau} d\omega \quad (5.14)$$

Frekansa bağımlı ağırlık fonksiyonunu $\Psi_{12}(\omega) \equiv G_1(\omega)G_2'(\omega)$ olarak tanımlanırsa, GÇK fonksiyonu Eş. 5.15'teki gibi yazılabilir.

$$R_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi_{12}(\omega)X_1(\omega)X_2'(\omega)e^{j\omega\tau} d\omega \quad (5.15)$$

Doğru ağırlık fonksiyonun seçilmesi ile, $R_{12}(\tau)$ mikrofon 1 ile 2 arasındaki gecikme zamanına denk gelen bölgede pik yapması gerekmektedir. Mikrofonlar arasındaki geliş zamanı farkı faz kaymasında $R_{12}(\tau)$ en yüksek olduğu değere eşittir.

$$\hat{\tau}_{12} = \underset{\tau \in D}{\operatorname{argmax}} R_{12}(\tau) \quad (5.16)$$

Eşitlik 5.15 birçok yerel en yüksek değere sahiptir. Bu yerel yüksek değerlerin faz kaymaları ve genlikleri; mikrofonlar arasındaki uzaklık, ses kaynağının yapısı, gürültü sinyalinin yapısı, ağırlık fonksiyonu gibi birçok faktöre bağlıdır.

5.3. En Çok Olabilirlik Ağırlık Fonksiyonu

Yansımanın olmadığı, kaynak ve gürültü seslerinin Gauss sinyallerinin ilişkisiz olduğu durumlarda, en çok olabilirlik (EÇO) ağırlık fonksiyonu asimtotik, tarafsız ve verimli bir değer üretir. Kaynak sinyalin güç spektral yoğunluğu $S(\omega)$, ve gürültü sinyallerinin güç spektral yoğunluğu, $V_1(\omega)$ ve $V_2(\omega)$, olarak EÇO ağırlık fonksiyonu Eş. 5.17’de verilmiştir.

$$\Psi_{12}^{E\text{ÇO}}(\omega) = \frac{|S(\omega)|}{|V_1(\omega)||V_2(\omega)|} - \left\{ 1 + \frac{|S(\omega)|}{|V_1(\omega)|} + \frac{|S(\omega)|}{|V_2(\omega)|} \right\}^{-1} \quad (5.17)$$

EÇO ağırlık fonksiyonu, tek kaynak ve kısa segmentli konuşmalar için verimli olduğu görülmüştür [22]. Gürültünün ve yansımanın olduğu, uzun konuşmalar için verimli değildir.

5.4. Faz Değişim Ağırlık Fonksiyonu

Faz dönüşümü (Phase Transform - PHAT) ağırlık fonksiyonu, gürültünün ve yansımanın olduğu ortamlarda kullanılmaktadır ve EÇO ağırlık fonksiyonuna göre gerçek ortamlarda daha iyi performans sağlamaktadır. Gürültülü ortamlardaki veri üretimi sağlamlığı nedeniyle sıklıkla tercih edilmektedir. Faz dönüşümü tekniği,

mikrofon sinyallerini baskılayarak beyazlaştırır, bu durumda çapraz spektrumda Eş. 5.18’de verildiği gibi baskılayarak beyazlaşır.

$$\Psi_{12}(\omega) = \frac{1}{|X_1(\omega)X_2'(\omega)|} \quad (5.18)$$

5.5. Bant Geçirgen Ağırlık Fonksiyonu

Diğer bir ağırlık fonksiyon yöntemi olan bant geçirgen ağırlık fonksiyonu belirlenen bant aralığında dışındaki zayıflama ile ilgilenir. Konuşma sinyalleri için genel olarak 300 Hz – 6kHz olarak kabul edilir ve Eş. 5.19’da verildiği gibi tanımlanır.

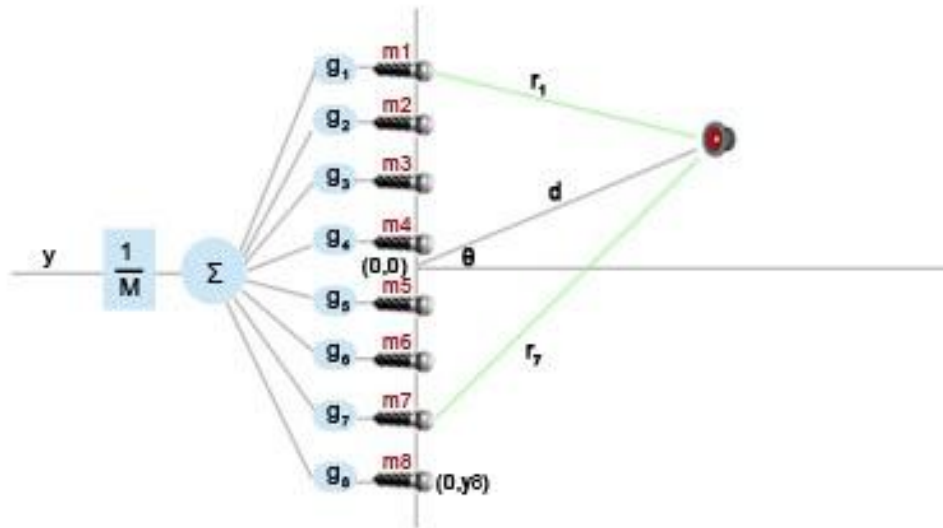
$$\Psi_{12}(\omega) = \begin{cases} 1 & 2\pi \cdot 300 \leq \omega \leq 2\pi \cdot 6000 \\ 0 & \text{Diğer} \end{cases} \quad (5.19)$$

300 Hz altında bulunan sinyalin alınmamasının faydası, bu frekansın altındaki sinyallerin genellikle arka fonda çalışan bilgisayar veya klima gibi cihazlara ait olan devamlı gürültüler olmasıdır. Bant geçirgenliği ağırlık fonksiyonu genellikle EÇO ve FD ağırlık fonksiyonlarının eşleniği ile, konuşma enerjilerinin en çok bulunduğu bant frekanslarına yoğunlaşmak için kullanılırlar.

6. YÖNLENDİRİLMİŞ TEPKİ GÜCÜ (YTG)

6.1. Yönlendirilmiş Tepki

Mikrofon dizilerinde elde edilen sinyalleri işlerken, sinyalin yerine ve uzaydaki yönüne odaklanılmaktadır. Mikrofon dizilerinin ses yakalama kabiliyetini kullanarak birçok çalışmada hüzme şekillendirme yöntemi kullanılmıştır [18]. Hüzme şekillendirme, filtrele-ve-topla işlemi ile yapılabilir. Bu işlemde her bir mikrofondan elde edilen sinyaller geçici bir filtreden geçirilerek, toplanır. Bu sayede güçlendirilmiş tek bir sinyal elde edilir. Hüzme şekillendirme işleminde kullanılan filtreler, istenilen sinyali güçlendirirken diğer sinyallerin zayıflamasını hedefler. Bu amaçla, kullanılan en yaygın ve basit yöntem, kaynak sinyalin yayılım gecikmeleri eşit olacak şekilde bütün sinyallerin ötelenmesidir. Bu metot, geciktir-ve-topla hüzme şekillendiricisi olarak adlandırılır. Şekil 6.1’de görüldüğü gibi her bir mikrofondan alınan sinyal, zaman ekseninde kaydırılarak toplanır.



Şekil 6.1. Geciktir-ve-topla hüzme şekillendiricisi

Hüzme şekillendirici teknikleri hem ses sinyalini yakalamak hem de ses kaynağının yerini tespit etmek için kullanılır. Ses kaynağının yeri biliniyor ise, hüzme şekillendirici kaynağa odaklanır ve mikrofon girdi sinyallerini kullanarak

güçlendirilmiş sinyal elde eder. Kaynağın yeri bilinmiyor ise, hüzme şekillendirici belirlenmiş bir bölgedeki noktaların gecikme değerlerini yönlendirerek tarama işlemi yapar. Bu yöntemle elde edilen sonuç, yönlendirilmiş sonuç olarak adlandırılır. Yönlendirilmiş Tepki Gücü (YTG), yayılım gecikmeleri ile yönlendirilmiş gecikmeler aynı olduğu durumlarda en yüksek değerini alır. Bu durumda ses kaynağının yeri tespit edilmiş olur. Yönlendirilmiş gecikmeler ile yayılım dalgasının özellikleri tahmin edilerek kaynağın yeri belirlenebilir.

Yönlendirilmiş tepki, genellikle M adet yönlendirilmiş gecikmeden, $\Delta_1 \dots \Delta_m$, oluşan bir fonksiyondur. Yönlendirilmiş gecikmeler, hüzme şekillendirme işlemi için kullanılır. Yönlendirilmiş tepki, hüzme şekillendiricinin odaklandığı bölgeyi tarayarak elde edilir. Ses kaynağının yeri odak bölgesinde ise yönlendirilmiş tepki gücü en yüksek değere ulaşır.

M adet elemandan oluşan filtre-ve-topla hüzme şekillendirici frekans bölgesinde Eş. 6.1'de verildiği gibi gösterilir.

$$Y(\omega, \Delta_1 \dots \Delta_M) = \sum_{m=1}^M G_m(\omega) X_m(\omega) e^{-j\omega \Delta_m} \quad (6.1)$$

Filtre-ve-topla hüzme şekillendiricinin sonucu olan yönlendirilmiş tepki gücü ise Eş. 6.2 verilmiştir.

$$P(\Delta_1 \dots \Delta_M) \equiv \int_{-\infty}^{+\infty} Y(\omega, \Delta_1 \dots \Delta_M) Y'(\omega, \Delta_1 \dots \Delta_M) d\omega \quad (6.2)$$

Eşitlikte $\Delta_1 \dots \Delta_M$ değerinin en yüksek olmasını sağlayan değerler, mikrofonlarda oluşan gecikme farklarını tahmin etmek için kullanılır. Bu durum, iki mikrofon için genelleştirilmiş çapraz korelasyon özelliği ile benzerlik gösterir. İki mikrofon arasındaki ses dalgasının geliş zamanları arasındaki fark, τ değerini aldığı anda en yüksek değer elde edilmektedir. Mikrofon a ve b için geliş zamanlarındaki fark, a ve

b mikrofonlarının yönlendirilmiş gecikmeleri arasındaki farkı en yüksek değerine çeker. Bu değer yönlendirilmiş tepki gücüne eşittir ve Eş. 6.3'te gösterilmiştir.

$$\hat{t}_{ab} \equiv \hat{\Delta}_a - \hat{\Delta}_b \quad (6.3)$$

Önceden belirlenmiş bir bölgedeki noktaların, geliş zamanları arasındaki farkın (GZAF) ortalama karekök (root mean square) hatasını en aza indirgeyerek konuşmacının yeri tespit edilebilir. Bu işlemi yapabilmek için, aday GZAF vektörleri, gecikme farkları kullanılarak hesaplanır. Gecikme farkları, filtre-ve-topla hüzme şekillendiriciden elde edilen yönlendirilmiş gecikmelere eşittir. Yönlendirilmiş tepki gücü, hüzme şekillendirici kullanılarak belirlenmiş bir bölgedeki yönlendirilmiş gecikmelerin en yüksek olduğu değer olarak kabul edilir. Elde edilen çıktıda, en yüksek güç değerine sahip nokta konuşmacının yeri olarak belirlenir. Bu doğrultuda, aday konum, $\vec{d}$, için yönlendirilmiş tepki gücü Eş. 6.4 verilmiştir.

$$P(\vec{d}) = P(\Delta_1 \dots \Delta_M) \text{ ve } \Delta_m = \tau_0 - \frac{|\vec{d}_m - \vec{d}|}{c} \quad (6.4)$$

Yayılm gecikmesi ise yayılım yönü, $\vec{\xi}_0$, ile tanımlanırsa Eş. 6.5 elde edilir.

$$\Delta_m = \frac{-\vec{\xi}_0 \cdot \vec{d}_m}{c} \quad (6.5)$$

Yayılm yönü ise azimuth ve elevation açıları, θ ve ϕ , ile Eş. 6.6'da olduğu gibi tanımlanır.

$$-\vec{\xi}_0 = \begin{bmatrix} \cos \phi & \sin \theta \\ \cos \phi & \cos \theta \\ \sin \phi & \end{bmatrix} \quad (6.6)$$

Yönlendirilmiş tepki gücünün açılarla ifade edilmiş hali Eş. 6.7'de verilmiştir.

$$P(\theta, \emptyset) = P(\Delta_1 \dots \Delta_M) \text{ ve } \Delta_m = \frac{-\vec{\xi}_0 \cdot \vec{d}_m}{c} \quad (6.7)$$

6.2. Genelleştirilmiş Çapraz Korelasyonda Yönlendirilmiş Tepki Gücü Gösterimi

Sinyal ve yönlendirilmiş tepki gücü eşitliklerinin birleştirilmesi ile filtre-ve-topla hüzme şekillendiricisinin YTG 'si Eş. 6.8 olduğu gibi yazılabilir.

$$P(\Delta_1 \dots \Delta_M) = \int_{-\infty}^{+\infty} \left(\sum_{a=1}^M G_a(\omega) X_a(\omega) e^{-j\omega \Delta_a} \right) \left(\sum_{b=1}^M G'_b(\omega) X'_b(\omega) e^{-j\omega \Delta_b} \right) d\omega \quad (6.8)$$

Toplam işaretlerinin çarpımından faydalanarak Eş. 6.9 elde edilir.

$$P(\Delta_1 \dots \Delta_M) = \int_{-\infty}^{+\infty} \sum_{a=1}^M \sum_{b=1}^M (G_a(\omega) X_a(\omega) e^{-j\omega \Delta_a}) (G'_b(\omega) X'_b(\omega) e^{-j\omega \Delta_b}) d\omega \quad (6.9)$$

Çift toplam içerisindeki terimler düzenlendiğinde Eş. 6.10 elde edilir.

$$P(\Delta_1 \dots \Delta_M) = \int_{-\infty}^{+\infty} \sum_{a=1}^M \sum_{b=1}^M G_a(\omega) G'_b(\omega) X_a(\omega) X'_b(\omega) e^{j\omega(\Delta_b - \Delta_a)} d\omega \quad (6.10)$$

Mikrofon sinyallerinin ve filtrelerinin enerjisinin sonlu olması nedeniyle integralleri yakınsardır. Bu nedenle toplamlar integral dışına alınabilir ve Eş. 6.11'deki gibi yazılabilir.

$$P(\Delta_1 \dots \Delta_M) = \sum_{a=1}^M \sum_{b=1}^M \int_{-\infty}^{+\infty} G_a(\omega) G'_b(\omega) X_a(\omega) X'_b(\omega) e^{j\omega(\Delta_b - \Delta_a)} d\omega \quad (6.11)$$

Ağırlık fonksiyonu Eş. 6.12'de olduğu gibi tanımlanır.

$$\Psi_{ab}(\omega) \equiv G_a(\omega) G'_b(\omega) \quad (6.12)$$

Filtreler ağırlık fonksiyonu ile değiştirildiğinde YTG, Eş. 6.13'te verildiği gibi tanımlanır.

$$P(\Delta_1 \dots \Delta_M) = \sum_{a=1}^M \sum_{b=1}^M \int_{-\infty}^{+\infty} \Psi_{ab}(\omega) X_a(\omega) X_b'(\omega) e^{j\omega(\Delta_b - \Delta_a)} d\omega \quad (6.13)$$

Eşitlik 5.15'te verildiği gibi iki mikrofonun genelleştirilmiş çapraz korelasyonu Eş. 6.14'te verildiği gibi tanımlanır.

$$R_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi_{12}(\omega) X_1(\omega) X_2'(\omega) e^{j\omega\tau} d\omega \quad (6.14)$$

Bu eşitlik a ve b mikrofonuna, zamansal kayma, Δ_{ab} olarak tanımlayarak uygulandığında, a ve b mikrofonları için genelleştirilmiş çapraz korelasyon Eş. 6.15'te verildiği gibi tanımlanır.

$$R_{ab}(\Delta_{ab}) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi_{ab}(\omega) X_a(\omega) X_b'(\omega) e^{j\omega\tau} d\omega \quad (6.15)$$

$\Delta_{ab} \equiv \Delta_a - \Delta_b$ denklemini kullanarak, yönlendirilmiş tepki gücü, genelleştirilmiş çapraz korelasyon fonksiyonu olarak Eş. 6.16 verildiği gibi tanımlanır.

$$P(\Delta_1 \dots \Delta_M) = 2\pi \sum_{a=1}^M \sum_{b=1}^M R_{ab}(\Delta_a - \Delta_b) \quad (6.16)$$

Bu denklemde, yönlendirilmiş gecikmelerin farkı kadar zamanda kaydırılmış bütün olası GÇK çiftlerinin toplamı elde edilir. Toplam $\Delta_{ab} = 0$ olduğu andaki GÇK içeren M adet oto korelasyonun toplamıdır. Bu durum, bağımsız yönlendirilmiş gecikmelerden dolayı, yönlendirilmiş tepki gücünde DC kaymaya neden olur ve Eş. 6.17' de verildiği gibi ifade edilir.

$$\sum_{m=1}^M R_m(0) = \frac{1}{2\pi} \sum_{m=1}^M \int_{-\infty}^{+\infty} |G_m(\omega)|^2 |X_m(\omega)|^2 d\omega \quad (6.17)$$

Eşitlik 6.16, her bir çaprazlamanın permütasyonunu içerir. Buna karşın, her bir permütasyonun yönlendirme gecikmeleri ters olduğundan, GÇK kombinasyonlarının toplamı, permütasyonun 2 ile ölçeklendirilmesi ile Eş. 6.18’de olduğu gibi elde edilir.

$$R_{ab}(\Delta_a - \Delta_b) = R_{ba}(\Delta_b - \Delta_a) \Leftrightarrow \quad (6.18)$$

$$R_{ab}(\Delta_a - \Delta_b) + R_{ba}(\Delta_b - \Delta_a) = 2R_{ab}(\Delta_a - \Delta_b)$$

Bu nedenle, Eş. 6.16, ölçeklendirme faktörünü ve sabit kaymayı içerecek şekilde bütün olası GÇK kombinasyonlarının toplamıdır.

Sonuç olarak M adet elemana sahip bir mikrofon dizisinin yönlendirilmiş tepki gücü, ikili şekilde gruplandırılmış bütün olası mikrofon çiftlerinin genelleştirilmiş çapraz korelasyonlarının toplamına eşittir. Bu nedenle, mikrofon sayısı arttıkça, yönlendirilmiş tepki gücü, genelleştirilmiş çapraz korelasyon tekniğini çiftli gruplardan çoklu mikrofon gruplarına taşır.

6.3. Faz Dönüşümü İle Yönlendirilmiş Tepki Gücünün Birleştirilmesi: YTG-FD Yöntemi

Dibiase tarafından yapılan çalışmada, faz dönüşüm ağırlık fonksiyonunun, konuşma sinyallerinde GÇK yöntemi uygulamak için verimli bir metot olduğu görülmüştür. Buna karşın, orta dereceli gürültülü ortamlarda, çiftler için GÇK metodunda kısa veri bloklarının kullanılmasında performansın düştüğü ortaya konulmuştur. Ayrıca bölüm 6.2’ de GÇK yöntemi ile YTG yöntemi arasındaki ilişki açıklanmıştır. Bu ilişki ve uygun filtreler kullanılarak, bütün olası YTG-FD kombinasyon çiftlerinin toplamı, filtrele-ve-topla hüzmeye şekillendirici kullanılarak hesaplanabilir. Bu filtreler, faz dönüşümü ve yönlendirilmiş tepki gücüne yönlendirme yaparlar.

a ve b mikrofonlarının sinyalleri için genelleştirilmiş çapraz korelasyon fonksiyonu yeniden incelenirse;

$$R_{ab}(\Delta_{ab}) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi_{ab}(\omega) X_a(\omega) X_b'(\omega) e^{j\omega\tau} d\omega \quad (6.19)$$

Ayrıca faz değişim ağırlık fonksiyonu da Eş. 6.20’deki gibi tanımlanmıştır.

$$\Psi_{ab}(\omega) = \frac{1}{|X_a(\omega) X_b'(\omega)|} \quad (6.20)$$

Bölüm 6.2’de, elemanlı mikrofon dizisinde, yönlendirilmiş tepki gücünün, bütün olası GÇK çiftlerinin toplamına eşit olduğu görülmüştür. Eğer filtre-ve-topla hüzmeye şekillendiricinin filtreleri, yönlendirilmiş gecikmeleri bulabilecek şekilde seçilir ise, yönlendirilmiş tepki toplamı bütün olası YTG-FD çiftlerinin toplamına eşit olur. GÇK ağırlık fonksiyonu ile bu filtreler arasındaki ilişki Eş. 6.21’de verilmiştir.

$$\Psi_{ab}(\omega) \equiv G_a(\omega) G_b'(\omega) \quad (6.21)$$

Her iki eşitliğin birleştirilmesi ile Eş. 6.22’de tanımlanan faz değişim aralığı fonksiyonu ortaya konulur.

$$G_a(\omega)G'_b(\omega) = \frac{1}{|X_a(\omega)X'_b(\omega)|} \quad (6.22)$$

$\frac{1}{|X_a(\omega)X'_b(\omega)|} = \frac{1}{|X_a(\omega)|} \frac{1}{|X'_b(\omega)|}$ ve $\frac{1}{|X'_b(\omega)|} = \frac{1}{|x_b(\omega)|}$, Eş. 6.22 yeniden Eş. 6.23'de verildiği gibi yazılabilir.

$$G_a(\omega)G'_b(\omega) = \frac{1}{|x_a(\omega)|} \frac{1}{|x_b(\omega)|} \quad (6.23)$$

Bu eşitlik sayesinde filtreler Eş. 6.24'de verildiği gibi tanımlanır.

$$G_a(\omega) = \frac{1}{|x_a(\omega)|} \quad G_b(\omega) = \frac{1}{|x_b(\omega)|} \quad (6.24)$$

Bu filtreler, YTG-FD filtreleridir. M mikrofonlu dizi için Eş. 6.25'deki gibi tanımlanır.

$$G_m(\omega) = \frac{1}{|x_m(\omega)|} \quad b = 1 \dots M \quad (6.25)$$

Faz dönüşümünde olduğu gibi, bu filtreler mikrofon sinyallerini beyazlaştırır. Beyazlaştırma yöntemi, faz dönüşümündeki en yüksek değerlerin verimli olarak güçlendirir ve aynı etkiyi yönlendirilmiş tepki gücü içinde yapar.

7. GENETİK ALGORİTMA

Genetik Algoritma (GA), 1975 yılında John Holland tarafından ortaya konulmuştur [24]. Canlılar arasında yiyecek, su, barınak gibi kısıtlı kaynaklar için veya eşler için yapılan mücadelelerde, yüksek uyumlu veya güçlü bireyler, zayıf bireylere göre daha başarılı sonuçlar elde ederler. Güçlü bireylerin yaşayarak yeni nesiller oluşturması, zayıf bireylerin yaşayarak yeni nesiller oluşturmasından daha yüksek olasılığa sahiptir. Dolayısıyla güçlü bireylerin genleri bir sonraki nesillere daha fazla yayılacaktır. Farklı bireylerden gelen, farklı özellikler için güçlü olarak tanımlanan genlerin, tek bir bireyde toplanması söz konusudur ve bu sayede daha güçlü nesiller elde edilebilmektedir [25].

Sezgisel bir metot olan GA, bir problem için genel optimum çözümü bulamayabilir. Ancak bilinen metotlarla çözüm bulunamayan veya çözülmesi oldukça uzun zaman alan problemlerde, en iyi çözüme yakın çözümler bulur.

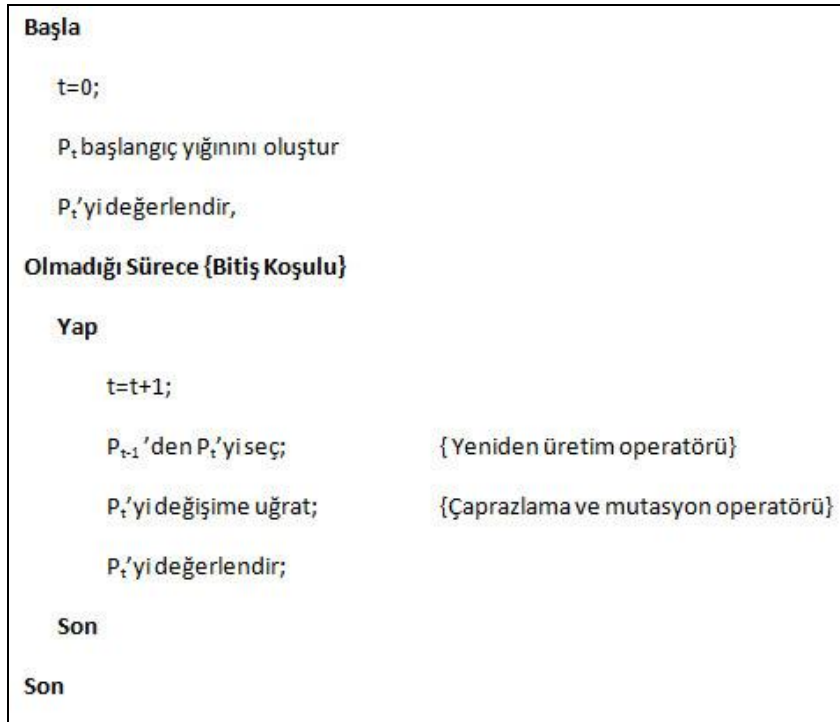
GA, bir noktadan diğer noktaya arama yapmak yerine noktalar yığını araştırması, parametrenin kendisi yerine parametreleri temsil eden dizileri kullanması ve stokastik olması nedeniyle bilinen optimizasyon metotlarından ayrılmaktadır. Bu nedenle GA, deterministik optimizasyon metotlarının başarısız olduğu ve arama uzayının büyük ve doğrusal olmadığı problemlerin çözümünde kullanılmaktadır.

Literatür incelendiğinde gezgin satıcı probleminden atölye çizelgelemeye, su dağıtım şebekelerinin tasarımından haberleşme şebekelerinin planlanmasına, hava hareketi angajmanlarına kadar birçok farkı probleminin çözümünde kullanılmaktadır [24,26-29].

7.1. Genetik Algoritmanın Temel Yapısı ve İşlevi

Bir GA, iki farklı setten oluşur; yığın olarak adlandırılan mümkün çözümlerin kodlandığı dizi seti ve biyolojik özelliği taklit eden operatörler seti. Her hangi bir

problemin çözümünde kullanılan basit bir GA' nın genel yapısı Şekil 7.1 'de verilmiştir.



Şekil 7.1. Genetik algoritmanın genel yapısı

GA genel yapısına bakıldığında ilk aşama olarak, tüm olası çözümlerin alt kümesinden oluşan belirli bir sayıda çözümü içeren başlangıç yığını veya başlangıç popülasyonu oluşturulur. Yığında bulunan her bir eleman bir dizi olarak kodlanır. Biyolojik olarak her bir dizi, bir kromozoma eşdeğerdir.

GA' nın her bir adımındaki popülasyon, nesil (generation) olarak adlandırılır. Popülasyondaki her bir birey uygunluk değerine (fitness value) sahiptir.

Uygunluk değeri, popülasyonda yer alan bireylerden hangilerinin bir sonraki nesle aktarılacağını belirlemek için kullanılır. Bir dizinin uygunluk değeri, çözülmeye çalışılan problemin amaç fonksiyonunun değerine eşittir. Bir dizinin uygunluk değeri için problem maksimizasyon problemi ise en yüksek değere sahip olması, minimizasyon problemi ise en düşük değere sahip olması hedeflenir.

GA' ın her bir iterasyonda yeniden üretim, genetik operatörler ve uygunluk değerinin hesaplanması olmak üzere üç ana adım mevcuttur. İterasyon işlemi sonlandırma kriteri ise, belirli bir sayıda iterasyon, uygunluk değerinin beklenen bir değere ulaşması veya belli sayıda iyileşme olmadan sürdürülen iterasyon sayısı olarak belirlenebilir.

İterasyon adımlarının detayları aşağıda verildiği gibidir:

1) Yeniden Üretim İşlemi: Bu işlem ile mevcut popülasyondan, bir sonraki popülasyona taşınacak olan diziler seçilir. Taşınacak olan diziler, mevcut popülasyonda en iyi uygunluk değerine sahip olan dizilerdir. Bu sayede, güçlü genetik özelliklerin bir sonraki popülasyona taşınması sağlanır.

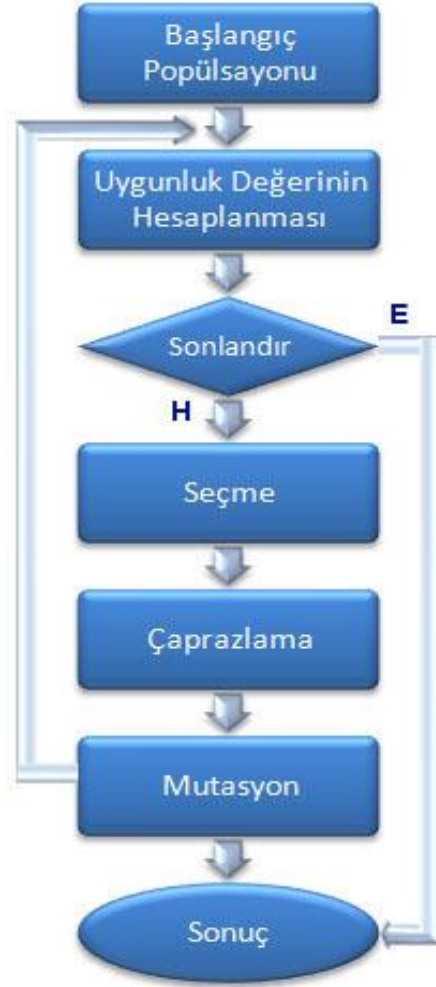
2) Genetik Operatörler: Seçim işlemi sonucunda oluşturulan popülasyonda yer alan dizilerden bir kısmına çaprazlama ve mutasyon gibi genetik operatörler uygulanır. Bu operatörler sayesinde, var olan genetik bilgiler kullanılarak, yeni popülasyon için yeni dizilerin (olası yeni çözümlerin) oluşturulması sağlanır. Çaprazlama operatörü, iki farklı dizi arasında bilgi değişimi yaparak yeni çözümler elde eder. Mutasyon operatörü ise mevcut dizinin bir kısmında değişimler yaparak, çözüm uzayında yeni noktalar elde edilmesini sağlar.

3) Uygunluk Değeri: Yeni popülasyona taşınacak olan dizilerin belirlenmesi için kullanılan araçtır. Bu nedenle, algoritmanın her bir iterasyon da yer alan popülasyonunun her bir dizisi için uygunluk değeri hesaplanır.

Herhangi bir problemin çözümünde kullanılan GA, aşağıdaki bileşenlerden oluşmaktadır.

- Popülasyonu oluşturacak olan bireylerin dizi olarak gösterimi
- Başlangıç yığının oluşturulması
- Her bir dizinin uygunluğunun belirlenmesi için değerlendirme fonksiyonu

- Yeniden üretim için seçim mekanizması
- Yeni bireylerin elde edilmesi için genetik operatörler
- Kontrol parametreleri



Şekil 7.2. Genetik algoritmanın temel akış diyagramı

GA' nın temel akış diyagramı Şekil 7.2'de verilmiştir. Standart bir genetik algoritmanın temel adımları aşağıda verilmiştir [30]:

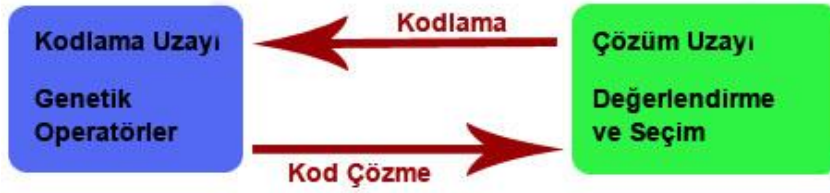
1. Muhtemel çözümlerden başlangıç çözümleri (popülasyonu) oluşturun.
2. Popülasyondaki her çözümün uygunluk değerini hesaplayın.
3. Durdurma şartları sağlanıyorsa durdur. Yoksa aşağıdaki adımları gerçekleştir;
 - 3.1. Uygunluk değerlerine göre popülasyon sıralama işlemini gerçekleştir.

- 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. Rastgele seçme işlemini uygula (popülasyonun çoğunluğu tabii seçme, az bir kısmı da rastgele 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).

7.2. Genetik Kodlama

GA' ı diğer arama algoritmalarından ayıran en temel özelliği, parametrelerin kendisi yerine, parametreleri temsil eden dizilerin kullanılmasıdır. Bu nedenle, GA' nın bir probleme uygulanması sırasında ilk adım, problemin çözüm kümesinin genetik kodlamasının yapılması, yani kromozomların her birinin nasıl temsil edileceğinin belirlenmesidir. GA' da bir problem için çözüm kümesi oluşturan karakter veya sayı dizisi, kromozom veya birey olarak adlandırılır. Literatürde en yaygın kullanılan kodlama, ikili düzende kodlamadır. Bu kodlamada, kromozomlar "0" ve "1" lerden oluşur. Dizinin uzunluğu, parametre ya da parametrelerin alt ve üst sınırları arasındaki tüm noktaları temsil edecek şekilde belirlenmektedir.

GA 'nın temel özelliklerinden birisi, değişimli olarak kodlama uzayı ve çözüm uzayı üzerinde çalışmasıdır. Yani genetik operatörler kodlama uzayında (kromozomlar) gerçekleşirken, değerlendirme ve seçim işlemi çözüm uzayında gerçekleşmektedir. Doğal seçim, diziler ile uygunlukları arasındaki ilişkiyi kurmaktadır. Kodlama uzayı ile çözüm uzayı arasındaki ilişki Şekil 7.3.'te verilmektedir.



Şekil 7.3. kodlama uzağı ve çözüm uzağı

Alt uzunluk sınırını belirten U_{min} ve üst uzunluk sınırını belirten U_{max} olarak verilen

ve L ikili sistemde kullanılacak olan bit sayısını belirten parametreler için uzunluk Eş. 7.1’de verildiği gibi hesaplanır.

$$\Pi = \frac{U_{max} - U_{min}}{2^L - 1} \quad (7.1)$$

Birden fazla parametreye sahip bir fonksiyon için dizi uzunluğu ise, her bir parametre için dizi uzunluklarının toplamına eşittir [25].

7.3. Başlangıç Çözümlerinin Oluşturulması

GA’ ın diğer metotlardan ayıran önemli bir özelliği de, noktadan noktaya değil, bir popülasyon içerisinde arama yapmasıdır [25]. Bu nedenle, GA’ ın ilk adımı olarak dizileri içeren başlangıç yığını oluşturulur. Başlangıç çözümleri, problemin çözüm uzayından seçilir ve mümkün olduğunca problem uzayını tam temsil etmesine özen gösterilir. Böylece çözüme ulaşım süresi kısalmaktadır. Başlangıç çözümlerinin oluşturulması, genellikle rasgele seçimler yapılır. Buna karşın, problemle ilgili bazı çözümler kabaca biliniyorsa, başlangıç popülasyonu bu çözümler kullanılarak ta oluşturulabilir. Böylece optimum çözüm bulmak daha verimli hale gelir.

Başlangıç çözüm sayısı (popülasyon sayısı) probleme göre değişiklik göstereceğinden, genellikle deneme yanılma yöntemi ile popülasyon sayısı belirlenir.

İkili kodlanmış kromozomların kullanıldığı durumlarda ideal popülasyon büyüklüğü Goldberg (1985) tarafından teorik olarak araştırılmıştır. Sonuç olarak, popülasyon büyüklüğü ile kromozom uzunluğu arasındaki bağıntı Eş. 7.2 ile tanımlanmıştır [25].

$$\text{PopülasyonBüyükülüğü} = 1,65 \cdot 2^{0,2 \cdot \text{kromozom uzunluğu}} \quad (7.2)$$

Popülasyon büyüklüğü, algoritmanın performansını iki açıdan etkilemektedir. Birincisi, popülasyon büyüklüğünün aşırı küçük seçilmesi durumunda, araştırma yapılacak olan uzayın yetersiz örneklenmesine sebep olacağından, kontrollü uzaklaşmayı sağlamak zorlaşacak ve araştırma belli bir alt optimal noktaya doğru sürüklenecektir. İkincisi, popülasyon büyüklüğü için aşırı yüksek bir değer seçilmesi durumunda ise bir nesillik gelişim için oldukça uzun süreye ihtiyaç duyacaktır. Bu ise, özellikle gerçek zamanlı problem uygulamalarında arzu edilmeyen bir durumdur. Bu yüzden popülasyon büyüklüğü için uygun bir değerin belirlenmesi algoritmanın verimliliği için önemlidir [30].

7.4. Uygunluk Fonksiyonu

Her iterasyonda, popülasyondaki dizilerin bir değerlendirme fonksiyonu yardımıyla uygunluk değerleri hesaplanır. Uygunluk değeri, bir sonraki popülasyonu oluşturacak yeni aday çözümlerin elde edilmesi için mevcut yığından hangi aday çözümlerin kullanılacağına belirlenmesinde rol oynamaktadır. GA da kullanılan değerlendirme fonksiyonu problemin amaç fonksiyonudur.

7.5. Yeniden Üretim Operatörü

Başlangıç popülasyonu oluşturulduktan sonra algoritmanın her iterasyonunda, yeni popülasyonun dizileri bir seçim süreci ile mevcut popülasyonun dizileri arasından seçilir. Yüksek uygunluk değerine sahip diziler, yeni dizilerin (yeni çözümlerin) elde edilmesinde yüksek olasılığa sahiptir. Bu operatör, doğal seçimi yapay olarak gerçekleştirmeyi hedefler. Amaç fonksiyonu, dizilerin bir sonraki adıma taşınmasına veya elenmesine karar vererek, "Doğal Seçim" yapay olarak gerçekleştirilir.

Dolayısıyla GA seçim işlemi ile beraber en uygun dizilerin bulunabileceği bölgelere yönlendirilmektedir [24].

Seçim aşamasında üç önemli konu söz konusudur. Bunlar;

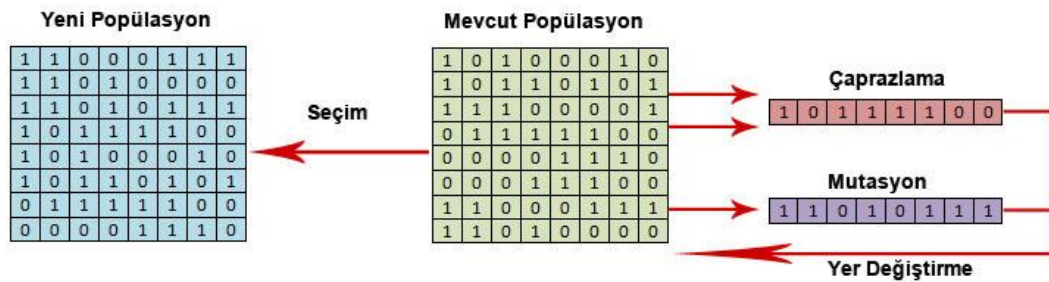
- Örnekleme uzayı
- Örnekleme mekanizması
- Seçim olasılığı

Örnekleme Uzayı

Seçim işlemi esnasında yeni popülasyonun, mevcut popülasyondaki tüm bireyler, üretilen yeni bireyler veya bunların bir kısmı alınarak elde edilir. Örnekleme uzayı, iki temel özelliğe sahiptir; genişlik ve içerik.

Düzenli Örnekleme Uzayı

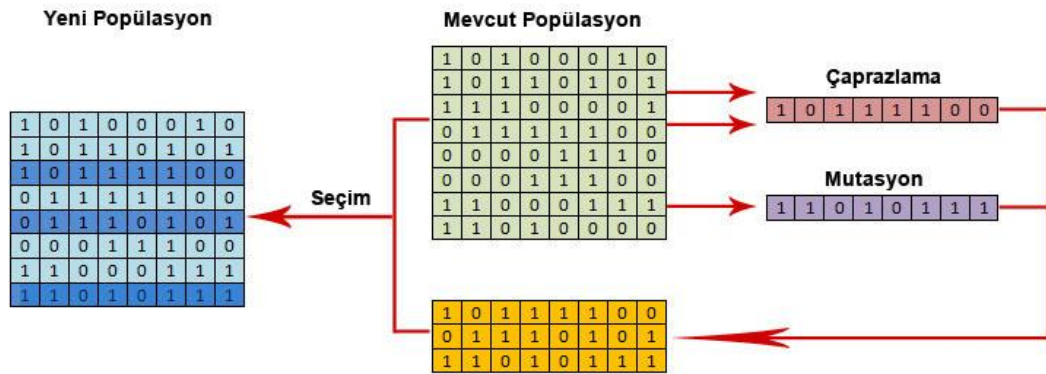
Holland tarafından önerilen basit GA'da, genetik operatörler ile elde edilen bireyler aileleri (mevcut bireyler) ile yer değiştirmektedir. Bu işleme, genel yer değiştirme denilmektedir. Genetik operatörler, arama işlemi sırasında uygunlukla ilgili hiçbir bilgiyi kullanmadığından dolayı elde edilen yeni bireyler ailelerinden kötü olabilir. Bu durumda, arama sırasında elde edilen iyi bireyler kaybedilebilir. Bu problemin üstesinden gelmek için, literatürde çeşitli yer değiştirme mekanizmaları önerilmiştir. Düzenli örnekleme uzayına dayalı seçim işlemi örneği Şekil 7.4'te verilmektedir.



Şekil 7.4. Düzenli örnekleme uzayına örnek

Genişletilmiş Örneklem Uzakı

Seçim işlemleri genişletilmiş örneklem uzakında yapıldığında, mevcut yığındaki bireyler ve genetik operatörler ile elde edilen yeni bireyler bir sonraki nesilde bulunmak için aynı şansa sahiptirler. $(\mu + \lambda)$ seçim işlemi genişletilmiş örneklem uzakında yapılan seçime tipik bir örnektir. Bu stratejide, μ adet aile ve λ adet çocuk bir sonraki nesilde yaşamak için yarışmakta olup, bunlar içerisinde en iyi μ adedi bir sonraki nesil için seçilmektedir. Bu yaklaşımın en önemli avantajı, arama sırasında elde edilen iyi bireylerin kaybedilmemesidir [31]. Şekil 7.5'te genişletilmiş örneklem uzakındaki seçim işlemini göstermektedir.



Şekil 7.5. Genişletilmiş örneklem uzakındaki seçim işlemi

Örneklem Mekanizması

Örneklem mekanizmaları, örneklem uzakından bireylerin nasıl seçileceği problemi ile ilgilenmektedir. Bireylerin örneklemeğinde üç temel yaklaşım vardır. Bunlar;

- Stokastik örnekleme.
- Deterministik örnekleme.
- Karma örnekleme.

Stokastik Örneklem. Bu isim, mevcut yığından dizilerin uygunluk değerlerine göre seçildiği seçim mekanizmalarının bir grubunu tanımlamaktadır. Rulet çemberi (roulette wheel) mekanizması bu örneklem sınıfında yer alan en basit örneklem mekanizmasıdır [18]. Bu yöntemde seçilecek bireyler yine rastgele seçilir ancak,

uygunluk değeri yüksek olan bireylerin seçilme olasılıkları yüksektir. Seçilecek bireyi belirlemek için, ruletin bütününün hesaplanması için Eş. 7.3 ile uygunluk değerlerinin toplamı bulunur.

$$UT = \sum_{i=1}^{PS} U_i \quad (7.3)$$

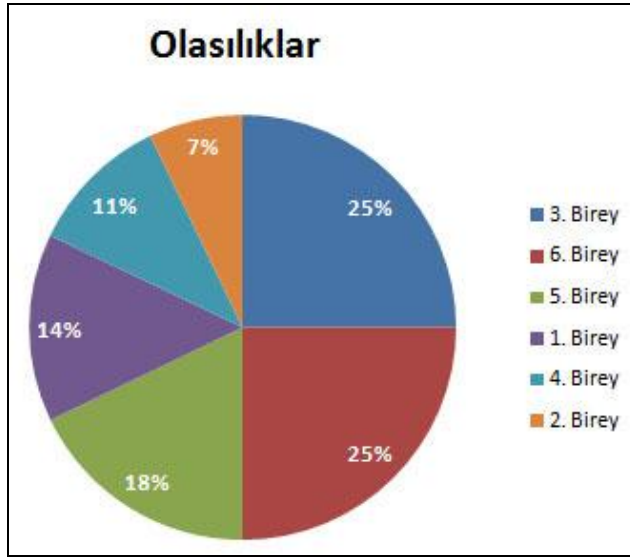
Eşitlikteki PS popülasyondaki birey sayısı (popülasyon büyüklüğü), U_i ise her bir bireyin uygunluk değerini göstermektedir. Elde edilen toplam uygunluk değeri kullanılarak her bir bireyin seçilme olasılığı (O_i) yüzde olarak Eş. 7.4 ile hesaplanır.

$$O_i = \frac{U_i}{UT} * 100 \quad (7.4)$$

Örnek olarak, 6 adet bireyden oluşan bir popülasyon için bireylerin uygunluk değerleri ve seçilme olasılıkları Çizelge 7.1’de görüldüğü gibi hesaplanmıştır. Hesaplanan olasılık değerlerinin sonucu rulet tekerleğinde Şekil 7.6’da görüldüğü gibi gösterilebilir.

Çizelge 7.1. Bireylerin uygunluk değerleri ve seçilme olasılıkları

Birey No	Birey Dizisi	Uygunluk Değeri	Seçilme Olasılığı (%)
1	00110100	100	14.25
2	01010010	50	7.15
3	10110011	175	25
4	10010000	75	10.75
5	00100010	125	17.85
6	11100010	175	25
Toplam		700	100



Şekil 7.6. Rulet tekerleği ile olasılıkların gösterimi

Popülasyondan seçilecek bireyi belirlemek için, 1 ile 100 arasında rastgele bir rakam seçilir. Daha sonra, seçilen rakama eşit veya daha büyük oluncaya kadar Eş. 7.3 ile elde edilen değerler üst üste toplanır. Bu işlem, Şekil 7.6’da rulet tekerleği üzerinde “0” dan başlamak üzere saat yönünde hareket etmek anlamına gelmektedir. Örnek olarak; rastgele seçilen rakam 30 ise, 6.inci birey seçilecektir. Şekil 7.6’da açık olarak görüldüğü gibi, uygunluk değeri yüksek olan bireylerin seçilme olasılıkları aynı oranda daha yüksektir.

Deterministik Örnekleme: Bu yaklaşımda örnekleme uzayından en iyi popülasyon genişliği seçilir. Genişletilmiş örnekleme uzayında anlatılan $(\mu+\lambda)$ ve (μ,λ) seçim stratejileri bu sınıfta yer almaktadır. Bu stratejilerde seçim aşamasında aynı dizinin birden fazla seçilmesi engellenmektedir. Kesintili seçim ve blok seçim de deterministik örnekleme içerisinde yer alan diğer iki mekanizmadır.

Karma Örnekleme: Bu yaklaşım stokastik ve deterministik örnekleme yaklaşımlarının özelliklerini kapsamaktadır. Turnuva seçim mekanizması, bu sınıfta yer alan en temel mekanizmadır. Bu mekanizmada popülasyondan rassal olarak bir grup dizi seçilir. Bu grup içindeki en iyi uygunluk değerine sahip dizi yeni popülasyona kopyalanır. Popülasyon genişliğine ulaşıncaya kadar bu işleme devam

edilir. Genellikle, grup genişliği ikidir. Ancak, bu sayının problemin yapısına göre artırılması da mümkündür [24].

Seçim Olasılığı: Bu konu, yığındaki her dizinin seçilme olasılığının nasıl belirleneceği ile ilgilidir. Orantılı seçim işleminde, bir dizinin seçim olasılığı bu dizinin uygunluğu ile orantılıdır.

7.6. Genetik Operatörler

GA' da yeniden üretim operatörü ile arama, yüksek uygunluğa sahip bölgelere doğru yönlendirilir. Bu bölgelerdeki yeni çözüm noktalarına ulaşmada ise genetik operatörler kullanılır. Genetik operatörler, yığının genetik bilgilerini kullanarak yeni çözümleri elde eder. GA' da kullanılan en genel iki genetik operatör çaprazlama ve mutasyondur [24].

7.6.1. Çaprazlama operatörü

Çaprazlama, farklı çözümler arasında bilgi değişimini sağlayarak arama uzayının benzer ancak araştırılmamış bölgelerine ulaşmayı sağlayan bir arama operatörüdür. Çaprazlama operatörü, yığından rastgele olarak seçilen iki dizinin belirli bölümlerini karşılıklı değiştirerek, arama uzayında yeni noktaları verecek yeni iki farklı diziyi elde etmek işlemidir. Bir yığında çaprazlama operatörü, çaprazlama olasılığı ile uygulanır. İkili sistemde kodlama yapılmış diziler için literatürde bilinen ve çok sık kullanılan 3 çaprazlama operatörü vardır; 1 noktalı, 2 noktalı ve rassal çaprazlama operatörleridir.

Tek (Bir) noktalı çaprazlama operatörü: Bu operatörde, çaprazlama noktası 1 ile L-1 arasından rastgele olarak seçilir. Eşleşen iki dizide, bu çaprazlama noktasından sonraki bölümler yer değiştirerek yeni iki dizi elde edilir [31]. Örneğin, bir yığından anne ve baba olarak iki dizi seçilsin ve dizi uzunlukları 8 ($L = 8$) olsun. 1 ile 7 arasından rastgele olarak seçilen çaprazlama noktasının değeri 4 olsun. Bu

çaprazlama noktasına göre Çocuk1 ve Çocuk2 yeni iki dizi Şekil 7.7’ deki gibi elde edilir.

	Çaprazlama Noktası = 4							
	1	2	3	4	5	6	7	8
Anne:	1	0	0	1	0	0	1	1
Baba:	0	0	1	1	1	1	0	0
Çocuk 1:	1	0	0	1	1	1	0	0
Çocuk 2:	0	0	1	1	0	0	1	1

Şekil 7.7. Tek noktalı çaprazlama işlemine örnek

İki noktalı çaprazlama operatörü: Bu operatörü tek noktalı çaprazlama operatöründen ayıran tek fark 1 ile L-1 arasında rastgele olarak 2 farklı kesme noktasının seçilmesidir. Yeni diziler, eşleşen dizilerin bu noktalar arasındaki bölgelerinin yer değiştirmesi ile elde edilir [32]. İki noktalı çaprazlama işlemine örnek Şekil 7.8’de verilmiştir. Bu örnek için seçilen kesme noktaları 3 ve 6’dır.

	Çaprazlama Noktaları = 3 ve 6							
	1	2	3	4	5	6	7	8
Anne:	1	0	0	1	0	0	1	1
Baba:	0	0	1	1	1	1	0	0
Çocuk 1:	1	0	0	1	1	1	1	1
Çocuk 2:	0	0	1	1	0	0	0	0

Şekil 7.8. İki noktalı çaprazlama işlemine örnek

Rassal (Uniform) çaprazlama operatörü. İlk defa Syswerda (1989) tarafından önerilen operatörde, öncelikle geçici olarak ikili sistemde bir dizi oluşturulur. Uzunluğu yığındaki dizilerin uzunluğuna eşit olan bu dizide, 0 ve 1 değerleri önceden belirlenen bir olasılık değerine göre elde edilir [32].

Bu dizideki 1 değerleri, çocuk 1 dizisindeki anne dizinin elemanlarının, çocuk 2 dizisinde ise baba dizisinin elemanlarının kullanılacağını gösterir. Aynı şekilde 0 değerleri, çocuk 1 dizisinde seçilen baba dizisinin elemanlarının, çocuk 2 dizisinde

ise seçilen anne dizisinin elemanlarının kullanılacağını gösterir. Şekil 7.9'da bu işleme örnek verilmiştir.

	1	2	3	4	5	6	7	8
Anne:	1	0	0	1	0	0	1	1
Baba:	0	0	1	1	1	1	0	0
Geçici Dizi:	1	1	0	0	0	1	0	1
Çocuk 1:	1	0	1	1	1	0	0	1
Çocuk 2:	0	0	0	1	0	1	1	0

Şekil 7.9. Rassal çaprazlama işlemine örnek

7.6.2. Mutasyon operatörü

GA 'da mutasyon operatörü, küçük bir olasılıkla bir dizi içindeki bir veya birkaç değeri rassal olarak değiştirerek yığında yeni dizilerin elde edilmesini sağlar [32]. İkili sistem kodlamasının kullanıldığı bir dizide mutasyon işlemi, rastgele seçilen elemanı 1 ise 0, 0 ise 1 yaparak yeni diziyi elde etme işlemidir. Şekil 7.10'da verilen örnek, bir diziyi mutasyon operatörünün uygulanışını göstermektedir. Anne dizisinin 4. ve 8. elemanları mutasyona uğratılarak Çocuk1 dizisi elde edilir.

Özellikle, GA 'nın son iterasyonlarında mutasyonun etkinliği artmaktadır. Çünkü son iterasyonlarda yığın iyi çözümlere yakınsadığından, diziler birbirlerine çok benzemektedir. Bu durum ise çaprazlama operatörünün aramasını kısıtlar. Nitekim, çaprazlama sonucu elde edilen diziler de birbirlerine benzer olacaklardır. Bu aşamada mutasyon, yığındaki değişkenliği gerçekleştirerek arama uzayında yeni çözüm noktalarının elde edilmesini sağlamaktadır. Her yığına mutasyon operatörü, mutasyon olasılığı ile uygulanır.

	1	2	3	4	5	6	7	8
Anne:	1	0	0	1	0	0	1	1
Çocuk 1:	1	0	0	0	0	0	1	0

Şekil 7.10. Mutasyon işlemine örnek

7.6.3. Seçkinlik (elitizm)

Bir popülasyondan çaprazlama yöntemi ile yeni popülasyonun oluşturulması sırasında, bulunan en iyi çözümün kaybedilmesi olasılığı vardır. Bu riski ortadan kaldırmak amacıyla, her popülasyondaki en iyi bir veya birkaç birey doğrudan yeni popülasyona atılır. Doğal ortamda karşılığı olmayan bu işleme seçkinlik (elitizm) adı verilir. Bir algorithmada elit eleman sayısının yüksek olması, iyi bireylerin popülasyonda baskınlık sağlayarak genetik çeşitliliğin azalmasına ve dolayısı ile algoritmanın erken yakınsamasına yani yerel optimuma takılmasına sebep olabilmektedir [18].

7.6.4. Onarım

Çaprazlama ve mutasyon sonucu elde edilen yeni bireyler bazı durumlarda probleme ait gerçek durumları ifade etmemektedir. Örneğin gezgin satıcı probleminde bazı noktalara 2 defa gidilmesi, bazı noktalara hiç gidilmemesi çözümün geçersiz olduğunu gösterir. Bu gibi durumların önüne geçmek için, çaprazlama ve mutasyon işleminden sonra her bir bireye (kromozoma) geçerlilik kontrolü yapılır. Belirlenen sınırın dışında olan veya geçerli olmayan kromozomlar üzerinde gerekli düzeltme işlemi yapılarak geçerli hale getirilir [23]. Şekil 7.11’de 8 noktaya gitmesi gereken bir satıcı için elde edilen geçersiz dizi ve bu diziye uygulanan onarım işlemi gösterilmektedir.

Geçersiz Kromozom	2	5	4	7	8	1	3	7
Geçerli Kromozom (Onarılmış)	2	5	4	6	8	1	3	7

Şekil 7.11. Onarım işlemine örnek

7.7. Kontrol Parametreleri

Bir GA ’da popülasyon genişliği, çaprazlama oranı, mutasyon oranı, nesil ayrılığı (generation gap), seçim stratejisi ve ölçeklendirme fonksiyonu olmak üzere 6 farklı kontrol parametresi vardır.

Bir GA'nın performansı bu parametrelerin değerlerine çok bağlıdır. Bu nedenle, bir probleme GA uygulanmadan önce bu parametrelerin en iyi kombinasyonunun belirlenmesi gerekir. Bu kısımda, kontrol parametreleri ve parametre değerlerinin algoritmanın performansındaki etkileri incelenecektir [32].

- *Popülasyon genişliği:* Yığın genişliği algoritmanın yakınsaması ile ilgilidir. Popülasyon genişliği küçük olduğunda GA 'nın performansı azalır. Çünkü küçük popülasyon arama uzayını örneklemede yetersiz kalmaktadır ve zamansız yakınsamaya sebep olmaktadır. Popülasyon genişliğinin büyük olması ise, çözüm uzayının çok iyi örneklenmesini sağladığı için aramanın etkinliğini artırır ve zamansız yakınsamayı önler. Ancak popülasyon genişliği büyük olduğunda, her iterasyonda dizilerin değerlendirilmesi çalışma zamanını artırmakta ve yakınsama çok yavaş gerçekleşmektedir.
- *Çaprazlama oranı:* Çaprazlama oranı, çaprazlama operatörünün kullanım sıklığını kontrol eder. Her yeni yığında, *çaprazlama oranı* x N adet diziye çaprazlama uygulanır. Yüksek çaprazlama oranı, yığın değişkenliğini hızlı bir şekilde gerçekleştirir. Düşük çaprazlama oranı, aramanın çok yavaş gerçekleşmesine sebep olur.
- *Mutasyon oranı:* Seçim işlemi ile elde edilen yeni yığındaki her dizinin her elemanı, mutasyon oranına eşit bir olasılıkla rastgele değişime uğrar. Sonuçta yaklaşık olarak, her iterasyonda *mutasyon oranı* x *popülasyon genişliği* x *dizi uzunluğu* adetinde mutasyon gerçekleşir. Yüksek mutasyon oranı GA 'daki aramanın bir rassal aramaya eşdeğer olmasına neden olduğundan, seçilen mutasyon oranının düşük olması tercih edilir.
- *Nesil ayrılığı:* Her iterasyonda yığında meydana gelecek değişimin yüzdesini kontrol eder.
- *Seçim stratejileri:* GA'da kullanılan seçim stratejileri 7.5'te incelenmiştir. Ancak GA literatüründe bilinen en genel ve basit iki strateji "rulet çemberi" ve "elitist seçim" stratejileridir.

- *Ölçeklendirme faktörü:* GA 'nın çalışma süresince yığındaki değişkenliğin korunması önemlidir. Bu nedenle yığındaki uygunluk değerlerinin kullanılacak bir metot yardımı ile ölçeklendirilmesi gerekir.

8. GELİŞTİRİLEN YÖNTEM

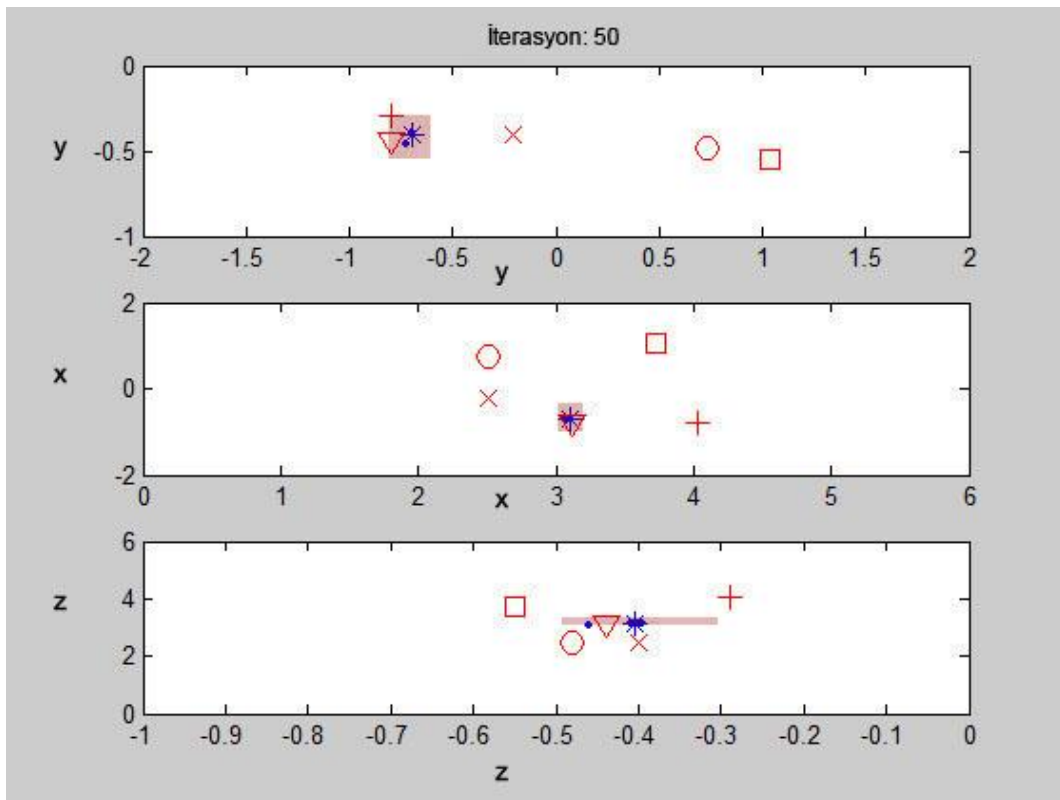
Birden fazla konuşmacının yerini tespit etmek için mikrofon dizileri kullanılır iken, öncelikle ses kaynaklarının bulunabileceği muhtemel noktalar için arama yapılmıştır. Arama işleminin optimizasyonu için genetik algoritma kullanılmıştır. Denenen her bir nokta için yönlendirilmiş tepki gücü- faz dönüşümü (YTG-FD) tekniği ile hüzme şekillendirme işlemi yapılmıştır ve çıktı sinyalinin en güçlü olduğu noktalar ses kaynağının yeri olarak kabul edilmiştir.

Geliştirilmiş olan yöntem üç temel aşamadan oluşmaktadır, GÇK-FD yöntemine göre gecikme değerlerinin hesaplanması, YTG değerini elde etmek için filtrele-ve-topla hüzme şekillendiricisi ve incelenecek noktaların belirlenmesi için genetik algoritmanın kullanılmasıdır. GÇK-FD yöntemi ile hesaplamalar Bölüm 5’te detaylı olarak verilmiştir. YTG değerinin hesaplanması için yapılan hesaplamalar ise Bölüm 6.3’de detaylı olarak verilmiştir.

Çalışmada değerlendirme kriteri olarak yönlendirilmiş tepki gücü kullanılmıştır. Bu değer en yüksek değeri aldığı nokta veya noktalar, konuşmacı veya konuşmacıların yeri olarak belirlenmiştir.

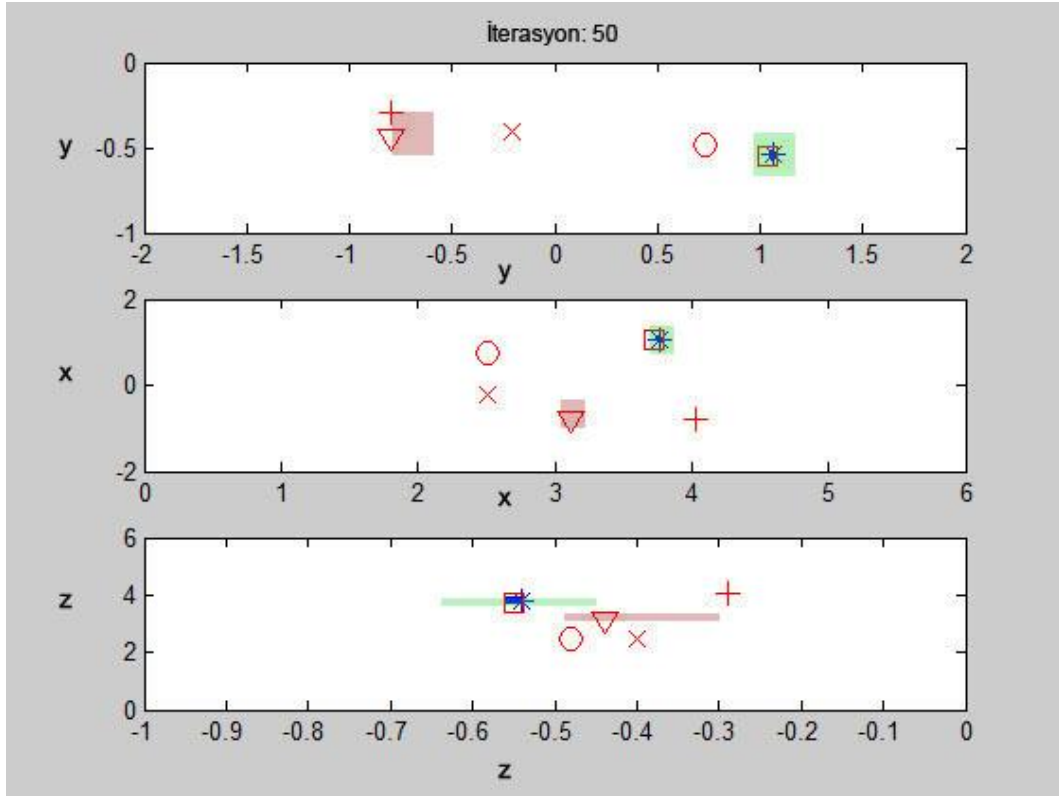
Birden fazla ses kaynağının yer aldığı ortamlarda, her bir ses kaynağının yeri tespit edilmesi hedeflenmektedir. Bu amaç kapsamında genetik algoritma kullanılmıştır. Fakat genetik algoritma yapısı gereği sadece bir kaynağa yakınsamakta ve onun yerini en optimum şekilde bulmayı hedeflemektedir. Bu nedenle yapılan çalışmada yeni bir metot geliştirilmiştir. Bu metotta genetik algoritmanın bir önceki aşamada bulunduğu yerler kör alan olarak ilan edilmektedir. Çalışmada kör alan 40 cm kenar uzunluğuna sahip bir küp olarak tasarlanmıştır. Örneğin birinci turda bulunan ses kaynağının noktasal yeri merkeze gelecek şekilde küp bir alan kör bölge olarak seçilmektedir. Bir sonraki kaynak aramalarında bu bölgede çıkan kromozomlar değerlendirilmemektedir. Bu sayede yeri bulunmuş bir kaynaktan başka bir kaynağa doğru genetik algoritmanın yönlendirilmesi sağlanmıştır.

Yapılan işlemi anlatmak için Şekil 8.1'den 8.5'e kadar olan çizimlerden faydalanılacaktır. Şekilde kırmızı renkle gösterilen “+,” ∇ , X, O, $\square$ ” simgeler beş farklı konuşmacının yerlerini göstermektedir. Ayrıca mavi renkle gösterilen “*” ise en iyi çözümü ve “.” simgeleri ise sonraki en iyi dokuz sonucu göstermektedir. Şekil 8.1'de programın çalıştırıldıktan sonra ilk konuşmacının yeri tespit edildikten sonra ilan edilen kör bölge kırmızı alanlar ile gösterilmektedir. Şekildeki alanların her üç resimde farklı olmasının sebebi eksenlerin uzunluklarındaki farklılıklardır.



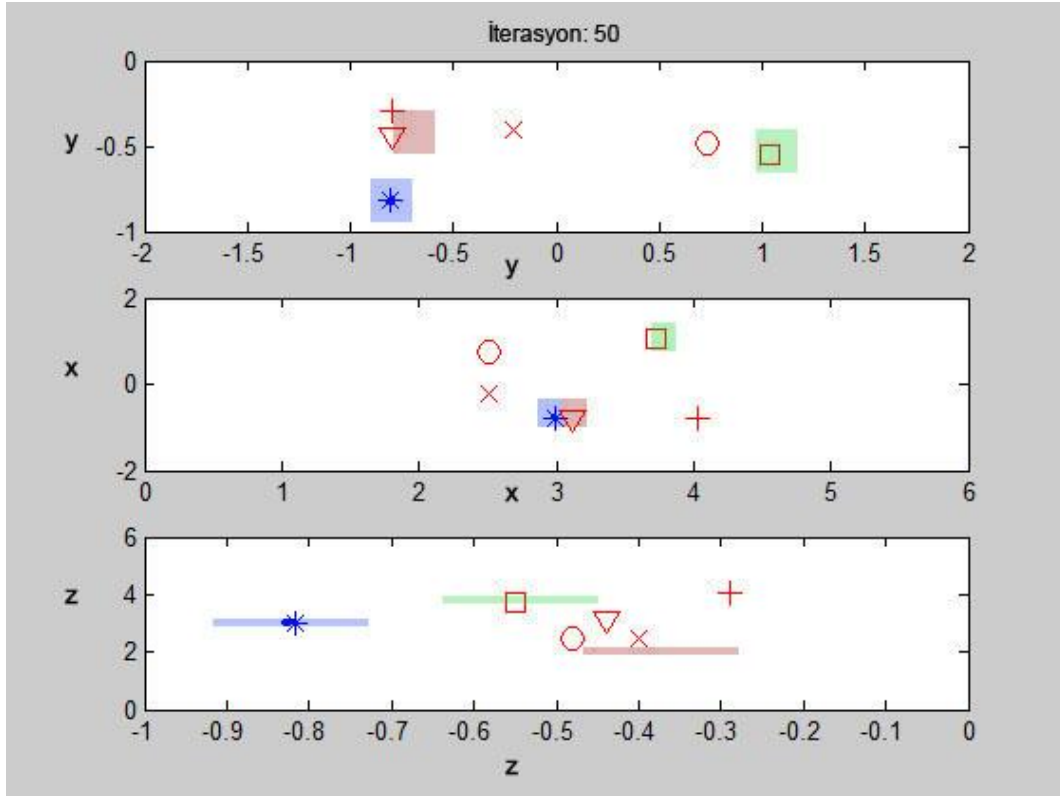
Şekil 8.1. Birinci konuşmacı ile oluşan kör bölge

Birinci konuşmacının belirlenmesinden sonra ikinci konuşmacının yerini tespit etmek için 50 iterasyondan oluşan arama işlemi tamamlanır. Bu esnada birinci konuşmacının belirlenmesi sonucunda oluşan kör bölgede yer alan sonuçlar aramaya tabi tutulmaz. Bu çalışmanın sonucunda Şekil 8.2'de verilen şekil elde edilir. İkinci konuşmacının bulunması sonucunda oluşturulan kör bölgeler yeşil ile gösterilmektedir.



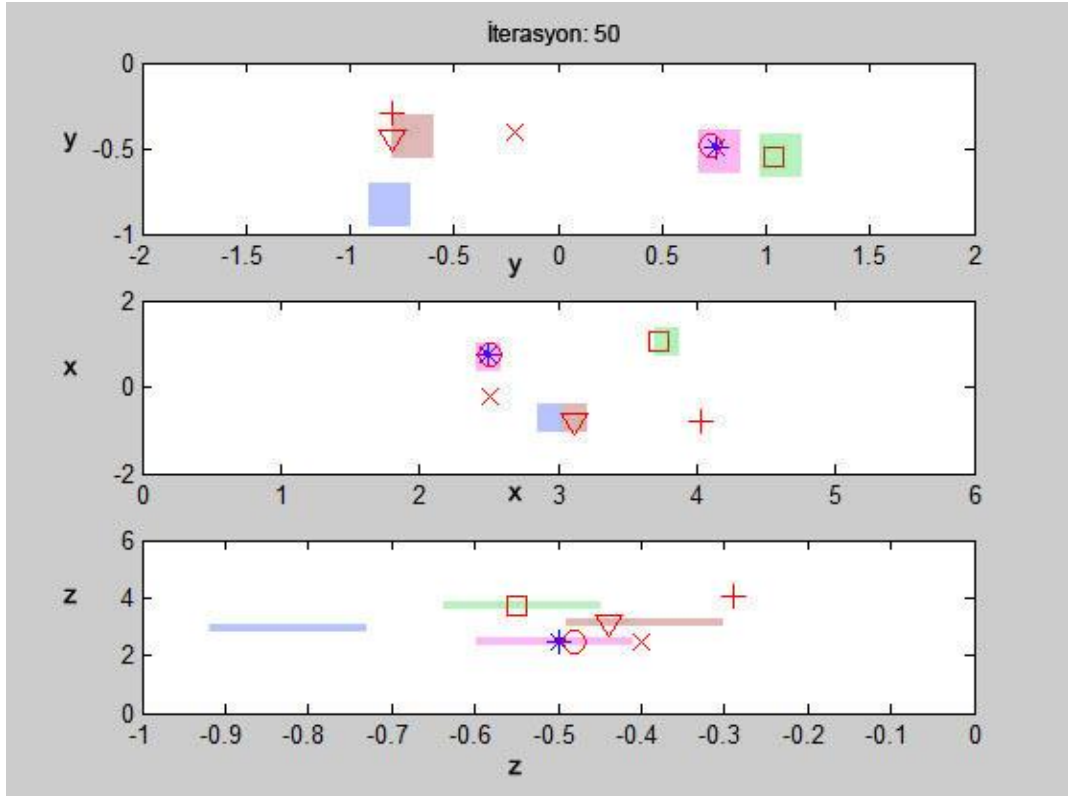
Şekil 8.2. Birinci ve ikinci konuşmacılar ile oluşan kör bölgeler

İkinci konuşmacının yeri tespit edildikten ve kör bölgesi ilan edildikten sonra, üçüncü konuşmacının bulunması için arama yapılır. Şekil 8.3'te üçüncü konuşmacının yeri ve kör bölgesi gösterilmektedir. Üçüncü konuşmacı için kör bölge mavi ile gösterilmiştir.



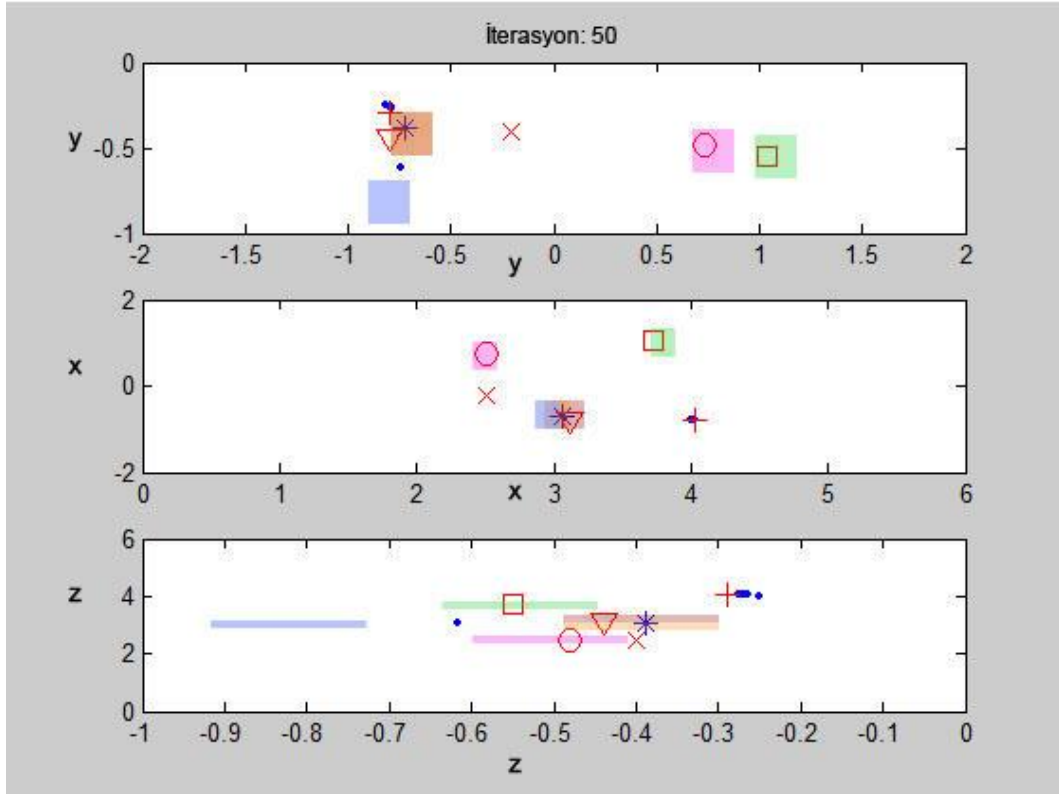
Şekil 8.3. Birinci, ikinci ve üçüncü konuşmacılar ile oluşan kör bölgeler

Üçüncü konuşmacının yerinin tespitinden sonra dördüncü konuşmacının yeri tespit edilmiştir. Konuşmacının yeri tespit edilirken, önceden belirlenmiş olan üç kör bölgede olmayan bir nokta belirlenmesi hedeflenmektedir. Şekil 8.4'te dördüncü konuşmacının yeri ve kör bölge gösterilmiştir. Kör bölge pembe renkle gösterilmektedir.



Şekil 8.4. Birinci, ikinci, üçüncü ve dördüncü konuşmacılar ile oluşan kör bölgeler

Dördüncü konuşmacının yerinin tespitinden sonra, beşinci konuşmacının yeri tespit edilirken diğer dört konuşmacının yerinin tespiti sonucunda belirlenen kör bölgelerde yer almaması gerekmektedir. Şekil 8.5'te beşinci konuşmacının yeri gösterilmektedir. Beşinci konuşmacının yerinde z ekseninde oluşan farklılık sayesinde farklı bir noktadan elde edilmiştir. Beşinci konuşmacının kör bölgesi turuncu ile gösterilmektedir.



Şekil 8.5. Birinci, ikinci, üçüncü, dördüncü ve beşinci konuşmacılar ile oluşan kör bölgeler

Şekilde kırmızı renkle gösterilen “+,” ∇ , X, O, $\square$ ” beş farklı konuşmacıdan, “ ∇ , O, $\square$ ” ile gösterilen üç konuşmacının yeri net olarak tespit edilmiştir. “+” ile gösterilen konuşmacı hata payı yüksek olarak bulunmuştur. “X” ile gösterilen konuşmacı ise bulunamamıştır.

9. DENEYSEL ÇALIŞMALAR

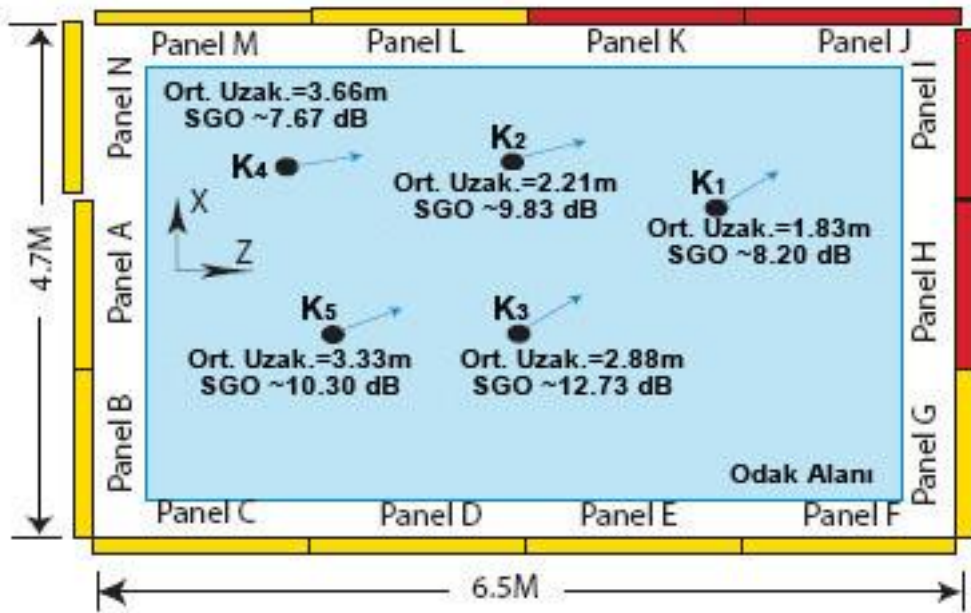
Bu çalışma kapsamında geliştirilen yöntem, Brown Üniversitesinin Lems Laboratuvarında 24 adet mikrofondan oluşan bir dizi ile beş kişinin aynı anda konuşmaların kayıt edildiği ses sinyalleri üzerinde denenmiştir [11]. Çalışma boyutlu Kartezyen uzayında yapılmıştır. Mikrofonlardan elde edilen ses bilgileri 20 KHz frekansta örneklenmiştir. Mikrofonların dizilişleri Çizelge 9.1’de verilmiştir. Mikrofonların her biri üç boyutlu Kartezyen uzayında farklı noktalarda yer almaktadır.

Çizelge 9.1. Mikrofonların 3 boyutlu uzayda yerleri

Mikrofon Numarası	X Koordinatı (cm)	Y Koordinatı (cm)	Z Koordinatı (cm)
1	-35,379	53,222	601,390
2	-32,274	44,279	603,930
3	2,724	58,469	600,940
4	7,034	67,629	598,770
5	13,208	81,217	595,660
6	45,404	74,006	596,390
7	96,876	43,144	604,510
8	103,290	55,335	601,230
9	136,740	52,258	602,630
10	157,240	54,452	601,830
11	161,510	67,120	598,690
12	193,990	69,129	597,960
13	202,310	70,324	558,990
14	202,780	62,200	558,960
15	203,320	62,135	510,690
16	202,210	74,557	491,090
17	203,510	50,771	491,160
18	203,350	65,919	461,040
19	202,160	68,559	422,890

20	203,460	41,663	414,080
21	202,480	62,695	383,950
22	202,680	56,795	369,330
23	202,750	47,737	369,180
24	201,660	59,599	321,430

Beş farklı konuşmacının yer aldığı sistemin görünümü Şekil 9.1’ de verilmiştir. Çizelge 9.2’ de ise konuşmacıların yerleri 3-boyutlu düzlemde verilmiştir.



Şekil 9.1. Konuşmacı yerlerinin gösterimi [11]

Çizelge 9.2. Konuşmacıların 3 boyutlu uzayda yerleri

	X (cm)	Y (cm)	Z (cm)
Kaynak 1:	72,96	-48,00	250,52
Kaynak 2:	-79,44	-29,00	402,92
Kaynak 3:	-20,48	-40,00	250,52
Kaynak 4:	-79,44	-44,00	311,48
Kaynak 5:	103,44	-55,00	372,44

9.1. Genetik İşlemler

Kodlama

Problemin çözüm kümesi olan ve bir noktanın koordinatlarını belirten x, y, z değerlerini temsil etmek için ikili sistem kullanılmıştır. Her bir koordinatın gösterimi için 8 veya 16 bit kullanılmıştır. Her bir koordinatın onluk sistemindeki değeri, ikili sayı sistemine çevrilerek, kromozomdaki yerine yerleştirilmiştir. Koordinat değerleri arama uzayının büyüklüğüne olan oranına, var olduğu yerin eklenmesi ile elde edilmiştir. Değerleri hesaplamak için kullanılan formül Eş. 9.1’de verilmiştir.

Kromozom değeri

$$= \text{Onluk Sistemden İkili Sisteme Geçiş (Yuvarla } (2^n \times (\text{değer} - \text{en alt değer})) / (\text{en üst değer} - \text{en alt değer}))) \quad (9.1)$$

Her bir koordinatı 8 bit ile temsil edilen $x = 0.7944$ m, $y = -0.29$ m ve $z = 4.0292$ m olan bir çözümün ikili düzlemde kromozom olarak kodlaması Şekil 9.2’te verilmiştir.

x								y								z							
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	1	0	0	1	1	0	1	1	0	1	1	0	1	1	0	1	0	1	0	1	1	0	0

Şekil 9.2. Kromozom kodlamasının gösterimi

Uzayı doğru temsil edecek olan bit sayısının belirlenmesi için analiz yapılmıştır. Analize göre bit sayıları 8 ila 16 olarak ayarlanmış ve ilk çözüme ulaşma süresi ve bulunan sonucun çözüme olan uzaklığı değerlendirilmiştir. Süreler, aynı bilgisayar ortamında Matlab programından başka yazılım çalıştırılmadan elde edilmiştir. Çizelge 9.3’de elde edilmiş sonuçlar verilmiştir. Bu çizelgede x, y ve z elemanları için veri bulunmayan satır kaynağın yerini göstermektedir. Buna karşın farklı bit uzunluklarında elde edilen farklı değerler ve elde edilme süreleri çizelgede verilmiştir.

Çizelge 9.3 Koordinat bit uzunluklarına göre elde edilen sonuçlar

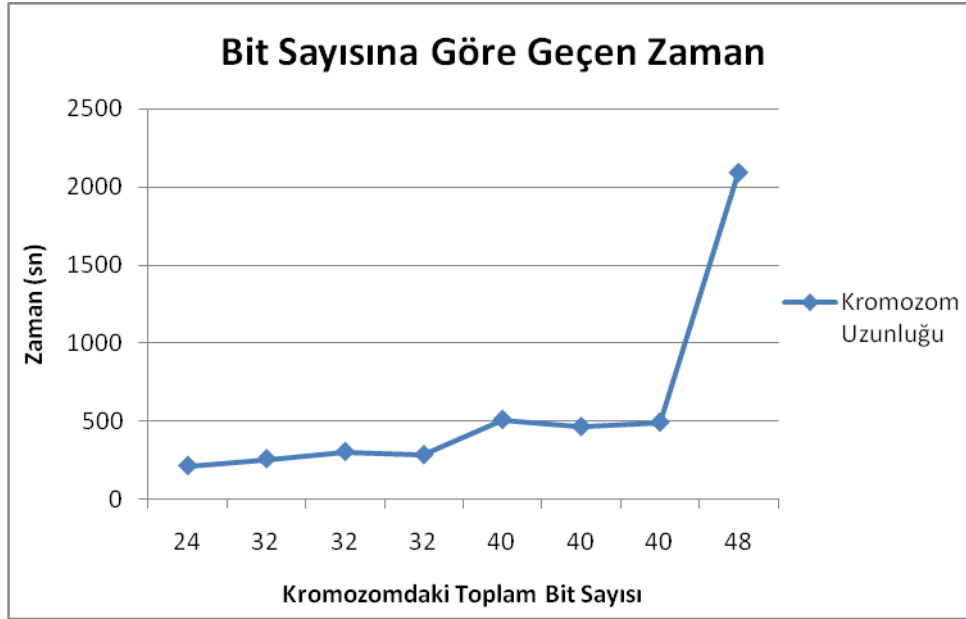
X Uzunluk (bit)	Y Uzunluk (bit)	Z Uzunluk (bit)	Süre(sn)	X (cm)	Y(cm)	Z(cm)
8	8	8	215,00	-69,80	-40,00	310,59
8	8	16	305,00	-71,37	-42,35	309,00
8	16	8	285,00	-66,67	-27,94	314,29
8	16	16	510,00	-71,37	-43,71	308,91
16	8	8	259,00	-71,93	-44,31	308,24
16	8	16	465,00	-71,31	-42,35	309,07
16	16	8	493,00	-69,85	-40,90	310,59
16	16	16	2091,00	-70,73	-40,95	309,77
Hedef:				-79,44	-44,00	311,48

Çizelge 9.4’de ise elde edilen sonuçlar, kaynağın yeri ile mukayese edilmiştir. Elde edilen sonuçlar ile gerçek kaynağın yeri arasındaki farklar öncelikli olarak her bir koordinat için ayrı ayrı ve ortalama olarak belirtilmiştir.

Çizelge 9.4 Süre ve sapma miktarlarının gösterimi

X Uz. (bit)	Y Uz. (bit)	Z Uz. (bit)	Süre (sn)	Uygunluk Değeri	X Fark (cm)	Y Fark (cm)	Z Fark (cm)	Ortalama Sapma (cm)
16	8	8	259,00	22,8928	9,45	0,70	1,04	3,73
8	16	16	510,00	23,0406	10,16	0,66	0,83	3,88
8	8	16	305,00	23,0502	10,16	3,75	0,80	4,90
16	8	16	465,00	23,0499	10,23	3,75	0,77	4,92
16	16	16	2091,00	22,9676	10,96	6,93	0,55	6,15
16	16	8	493,00	22,7609	12,07	7,05	0,29	6,47
8	8	8	215,00	22,7475	12,13	9,09	0,29	7,17
8	16	8	285,00	18,8935	16,08	36,50	0,90	17,83

Yapılan analizin sonucunda elde edilen kromozom grupları ortalama kayma miktarlarına ve sonrasında sürelerine göre küçükten büyüğe doğru sıralanmıştır. Bu sıralama Çizelge 9.4’de gösterilmektedir. Sonuçta iki kromozom grubunun aynı kabul edilebilecek benzerlikte ortalama kaymaya sahip olduğu gözlenmiştir. Bu kromozomlarda işlem sürelerine bakılmıştır. Kromozomdaki bit sayısının işlem süresine olan etkisi Şekil 9.3’te verilmiştir. Bit sayısının artışı ile işlem süresi üstel bir artış gösterdiği görülmektedir.



Şekil 9.3. Kromozomda yer alan bit sayısının işlem süresine etkisi

Toplam işlem süresine göre kromozom grupları sıralandıktan sonra, ilk iki grup için uygunluk değerlerinin benzer olduğu görülmüştür. İlk grubun işlem süresinin diğerinin yarısına eşit olduğu gözlenmiştir. Bu nedenle, süre anlamında kısa olan kromozom seçilmiştir. Sonuç olarak X bitinin uzunluğu 16 bit, Y ve Z bitlerinin uzunluğu 8 bit olarak belirlenmiş ve her bir kromozom 32 bit ile gösterilmiştir.

Çaprazlama

Çaprazlama yapılacak olan bireylerin seçimi için Bölüm 6.5'te detaylı olarak anlatılan rulet tekerleği yöntemi kullanılmıştır.

Çaprazlama işlemi her bir ekse için tek noktali çaprazlama olarak yapılmıştır. Çaprazlama işlemi için 1 ila 8 arasında bir sayı, 9 ila 16 arasında bir sayı, 17 ila 24 arasında bir sayı olmak üzere üç sayı seçilmiştir. Seçilen her bir sayı için öncesi ilk kromozomdan sonrası ise ikinci kromozomdan alınarak yeni bireyler oluşturulmuştur. Şekil 9.4 'te anne ve baba kromozomlarından Çocuk 1 ve Çocuk 2 kromozomları çaprazlama yöntemi ile elde edilmesi gösterilmiştir.

	1. Nokta = 3								2. Nokta = 14								3. Nokta = 17							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Anne:	1	0	0	1	0	0	1	1	1	0	0	1	1	1	0	1	0	0	1	0	1	1	0	1
Baba:	0	0	1	1	1	1	0	0	0	1	1	0	0	1	1	1	1	0	0	1	1	1	0	0
Çocuk 1:	1	0	0	1	1	1	0	0	1	0	0	1	1	1	1	1	0	0	0	1	1	1	0	0
Çocuk 2:	0	0	1	1	0	0	1	1	0	1	1	0	0	1	0	1	1	0	1	0	1	1	0	1

Şekil 9.4. Çaprazlama işlemine örnek

Çaprazlama işleminin ne zaman yapılacağını belirlemek amacı ile çaprazlama oranı parametresi kullanılmıştır. Çaprazlama işleminin yapılabilmesi için 0 ila 1 arasında rastgele seçilen bir sayının, çaprazlama oranından küçük olması gerekmektedir. Çaprazlama yapılmayan bireyler mutasyon işlemine tabi tutulur. Çalışmada kullanılan çaprazlama oranı 0,70 olarak alınmıştır.

Mutasyon

Çaprazlama işleminden sonra bireyler mutasyon işlemine tabi tutulur. Mutasyon işleminde iki temel kavram vardır; mutasyon oranı ve mutasyon genliği. Mutasyon oranı, mutasyon işleminin yapılıp yapılmamasına karar vermeye yarayan parametredir. Kromozomun her bir değeri için 0 ila 1 arasında rastgele bir sayı seçilir. Seçilen sayı mutasyon oranından küçük ise mutasyon işlemi yapılır. Aksi durumda değer olduğu gibi bırakılır. Örneğin mutasyon oranı 0,2 olan bir çalışmada, rastgele sayı olarak 0,15 elde edilirse, kromozomun seçili değeri 1 ise 0, 0 ise 1 olarak değiştirilir. Çalışmada mutasyon oranı olarak 0,1 kabul edilmiştir.

Mutasyon işlemi, doğada her zaman daha güçlüyü elde etmeyi hedeflemiştir. Bu amacın, algoritmaya yansıtılması için mutasyon genliği kavramı ortaya konulmuştur. Bu kavram sayesinde, ilk iterasyonlarda mutasyonun etkisi daha büyük iken, iterasyon sayısının sonlarına doğru mutasyon genliğinin azalması sonucunda mutasyon etkisi azalır. Bu sayede, sonlara doğru mevcut çözümlerin yakınlarında arama yapılması sağlanır. Mutasyon genliği sayesinde başlangıçta genetik çeşitlilik sağlanırken, sonlara doğru gereksiz atlamaların yapılarak zaman kaybına neden olacak adımlar engellenmiştir. Yapılan çalışmada mutasyon genliği her bir koordinat için, her 10 iterasyonda 1 bit azaltılmıştır. Yani başlangıcı iterasyonunda her bir

koordinat için mutasyon genliği 7 bit iken, 30. İterasyonda ise mutasyon genliği 5 bit olur.

Elitizm

Çalışmada elitizm sayısı bütün popülasyonun %10 u kadar alınmıştır. Aynı zamanda bu sayı, aranan konuşmacı sayısının 3 misline karşı gelmektedir.

Kör Bölge

Her bir ses kaynağının bulunması sürecinde bulunan ses kaynağının etrafında ne kadar büyüklükte bir alanın kör bölge ilan edilmesini belirlenmek için bir çalışma yapılmıştır. Çalışmada sıra ile 5 cm, 10 cm, 20 cm, 30 cm, 40 cm ve 50 cm’lik kenar uzunlukları olan küpler incelenmiştir. Sonuçları Çizelge 9.5-9.10 da verilmiştir.

Çizelge 9.5. Kenar uzunluğu 5 cm olan kör bölge için elde edilen sonuçlar

	X Gerçek	Y Gerçek	Z Gerçek	X	Y	Z	ΔX	ΔY	ΔZ	$\Delta \Sigma$
Kaynak 1:	72,96	-48,00	250,52	-72,36	-50,20	308,24	0,6	2,2	57,72	60,52
Kaynak 2:	-79,44	-29,00	402,92	-71,92	-43,14	308,24	7,52	14,1	94,68	116,34
Kaynak 3:	-20,48	-40,00	250,52	-67,77	-35,69	308,24	47,29	4,31	57,72	109,32
Kaynak 4:	-79,44	-44,00	311,48	-80,09	-60,78	308,24	0,65	16,8	3,24	20,67
Kaynak 5:	103,44	-55,00	372,44	-81,25	-65,88	308,24	22,19	10,9	64,2	97,27
							Toplam:			404,1

Çizelge 9.6. Kenar uzunluğu 10 cm olan kör bölge için elde edilen sonuçlar

	X Gerçek	Y Gerçek	Z Gerçek	X	Y	Z	ΔX	ΔY	ΔZ	$\Delta \Sigma$
Kaynak 1:	72,96	-48,00	250,52	-67,91	-29,80	312,94	5,05	18,2	62,42	85,67
Kaynak 2:	-79,44	-29,00	402,92	-70,71	-40,00	310,59	8,73	11	92,33	112,06
Kaynak 3:	-20,48	-40,00	250,52	-65,55	-18,82	315,29	45,07	21,2	64,77	131,02
Kaynak 4:	-79,44	-44,00	311,48	-80,32	-62,75	298,82	0,88	18,8	12,66	32,29
Kaynak 5:	103,44	-55,00	372,44	-72,08	-50,20	308,24	31,36	4,8	64,2	100,36
							Toplam:			461,40

Çizelge 9.7. Kenar uzunluğu 20 cm olan kör bölge için elde edilen sonuçlar

	X Gerçek	Y Gerçek	Z Gerçek	X	Y	Z	ΔX	ΔY	ΔZ	$\Delta \Sigma$
Kaynak 1:	72,96	-48,00	250,52	-66,29	-9,80	315,29	6,67	38,2	64,77	109,64
Kaynak 2:	-79,44	-29,00	402,92	-70,05	-29,80	310,59	9,39	0,8	92,33	102,52
Kaynak 3:	-20,48	-40,00	250,52	-62,45	-43,92	315,29	41,97	3,92	64,77	110,66
Kaynak 4:	-79,44	-44,00	311,48	-80,55	-70,59	298,82	1,11	26,6	12,66	40,36
Kaynak 5:	103,44	-55,00	372,44	-71,92	-50,20	308,24	31,52	4,8	64,2	100,52
							Toplam:			463,7

Çizelge 9.8. Kenar uzunluğu 30 cm olan kör bölge için elde edilen sonuçlar

	X Gerçek	Y Gerçek	Z Gerçek	X	Y	Z	ΔX	ΔY	ΔZ	$\Delta \Sigma$
Kaynak 1:	72,96	-48,00	250,52	76,27	-53,3	247,1	3,31	5,33	3,46	12,1
Kaynak 2:	-79,44	-29,00	402,92	-80,3	-75,7	298,8	0,9	46,7	104,1	151,69
Kaynak 3:	-20,48	-40,00	250,52	-68,4	-14,5	312,9	47,88	25,5	62,42	135,79
Kaynak 4:	-79,44	-44,00	311,48	-71,9	-45,5	308,2	7,58	1,49	3,24	12,31
Kaynak 5:	103,44	-55,00	372,44	104,6	-49,4	376,5	1,2	5,59	4,03	10,82
							Toplam:			322,7

Çizelge 9.9. Kenar uzunluğu 40 cm olan kör bölge için elde edilen sonuçlar

	X Gerçek	Y Gerçek	Z Gerçek	X	Y	Z	ΔX	ΔY	ΔZ	$\Delta \Sigma$
Kaynak 1:	72,96	-48,00	250,52	76,29	-53,33	247,06	3,33	5,33	3,46	12,12
Kaynak 2:	-79,44	-29,00	402,92	-82,17	-29,02	400,00	2,73	0,02	2,92	5,67
Kaynak 3:	-20,48	-40,00	250,52	-68,05	-20,78	312,94	47,57	19,2	62,42	129,21
Kaynak 4:	-79,44	-44,00	311,48	-80,09	-60,78	298,82	0,65	16,8	12,66	30,09
Kaynak 5:	103,44	-55,00	372,44	106,36	-54,90	376,47	2,92	0,1	4,03	7,05
							Toplam:			184,1

Çizelge 9.10. Kenar uzunluğu 50 cm olan kör bölge için elde edilen sonuçlar

	X Gerçek	Y Gerçek	Z Gerçek	X	Y	Z	ΔX	ΔY	ΔZ	$\Delta \Sigma$
Kaynak 1:	72,96	-48,00	250,52	76,33	-53,3	247,1	3,37	5,33	3,46	12,16
Kaynak 2:	-79,44	-29,00	402,92	-83,2	-30,2	296,5	3,76	1,2	106,5	111,41
Kaynak 3:	-20,48	-40,00	250,52	-102	-94,9	272,9	81,76	54,9	22,42	159,08
Kaynak 4:	-79,44	-44,00	311,48	-72,2	-44,7	308,2	7,29	0,71	3,24	11,24
Kaynak 5:	103,44	-55,00	372,44	106	-53,7	376,5	2,55	1,27	4,03	7,85
							Toplam:			301,7

Çalışmanın sonucunda gözlenmiştir ki en iyi sonuç 40 cm kenar uzunluğu olan kör bölge için elde edilmiştir. 40 cm'in altında olan uzunluklar için baskın konuşmacıya yönelimin tekrarlandığı için istenilen sonuçlar elde edilememiştir.

Her bir turda belirlenen ses kaynağı yerinin etrafına her kenarının uzunluğu 40 cm olan küp şeklinde bir kör bölge belirlenmiştir. 40 cm seçilmesindeki temel mantık insan başının ve omuzlarının 40 cm'lik bir küp olarak kabul edilebilmesidir. İki konuşmacının aynı anda aynı 40 cm'lik küpün içerisinde yer almasının oldukça güç olmasıdır.

Ayrıca çalışma esnasında kör bölgede bulunan kromozomlar onarılarak yeniden incelemeye tabi tutulmuştur. Bu sayede popülasyonda kayıp verilmemektedir.

Uygunluk Fonksiyonu

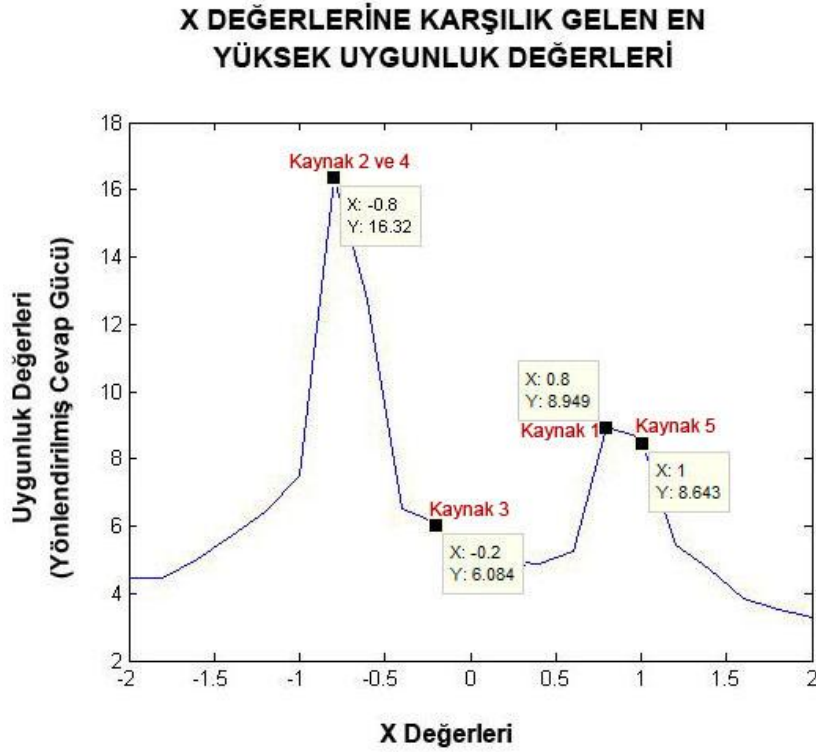
Popülasyondaki her bir çözüm için, mikrofon çiftlerine gelen seslerin geliş süreleri arasındaki farkları esas alarak filtreleyen ve sonra sinyalleri toplayarak elde edilen yönlendirilmiş tepki gücü uygunluk fonksiyonu olarak kullanılmıştır.

9.2. Verinin Analizi

Elde edilen sonuçların doğru bir şekilde yorumlanabilmesi ve yapılan çalışmanın doğruluğunun kontrol edilebilmesi için, çalışmada kullanılan veri analiz edilmiştir. Analiz işlemi sırasında 20 cm, 5 cm ve 3 cm'lik kaymalar ile araştırma uzayında belirlenen bütün noktalar için uygunluk değeri hesaplama işlemi yapılmıştır. Elde edilen sonuçlara göre, geliştirilen yönetimin sonucunda elde edilen sonuçlar değerlendirilmiş ve karşılaşılan durumlar yorumlanabilmiştir.

Öncelikli olarak 20 cm netlikte elde edilen 3906 kromozom değerleri incelenmiştir ve elde edilen grafikler Şekil 9.5'te verilmiştir. Bu inceleme sonucunda x ekseninde uygunluk değerlerinin -0,8, 0,8 ve 1 değerlerinde belirgin bir şekilde en yüksek değerlerini aldığı gözlenmiştir. Ayrıca -0,2 noktasında da belirgin olmayan bir

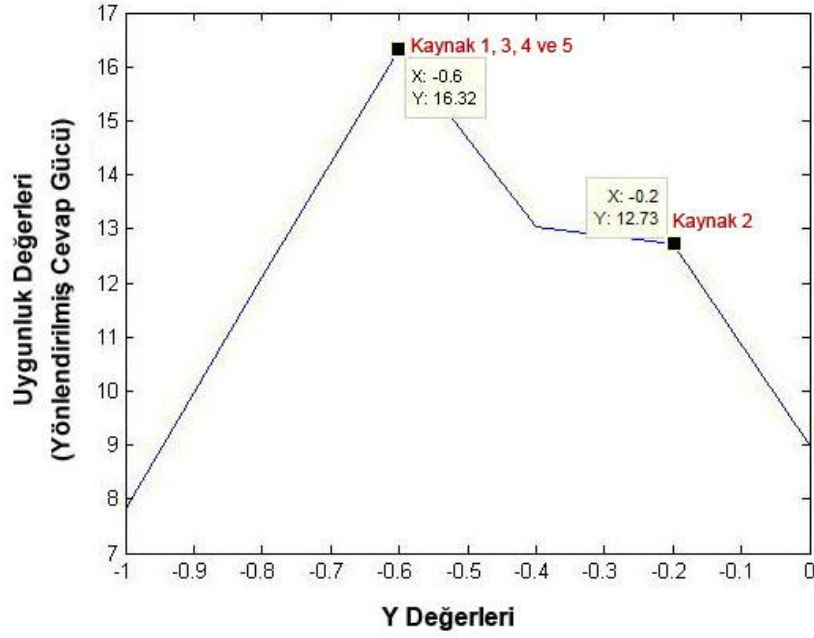
yükselmenin olduğu gözlenmiştir. Y ekseninde ise -0,6 ve -0,2 noktalarında uygunluk değerleri için en yüksek değerlerini aldığı gözlenmiştir. Z ekseninde ise 2,6, 3 ve 4 noktalarında uygunluk değerlerinin en yüksek değerlerini aldığı görülmektedir ve bu yüksek değerler birbirlerine çok yakındır.



(a)

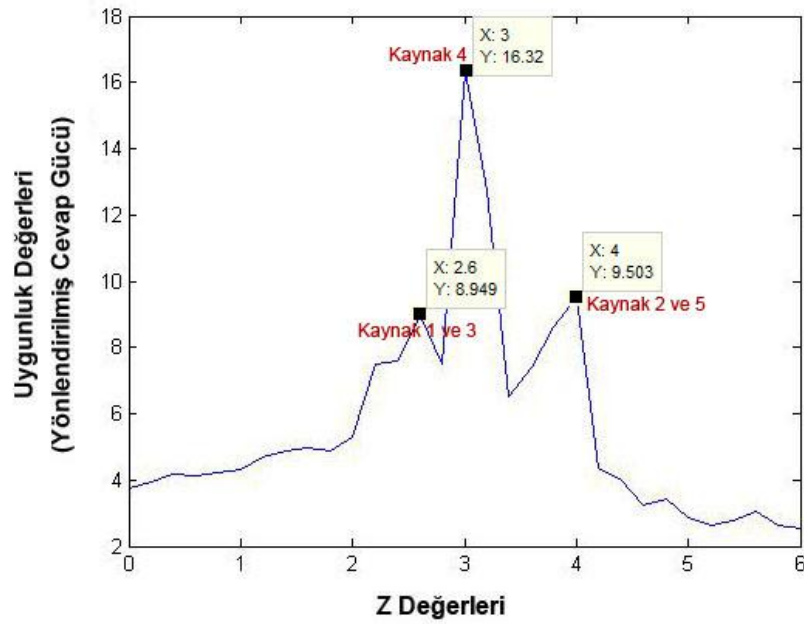
Şekil 9.5. 20 cm netlikte elde edilen uygunluk değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri ve (c) z değerindeki değişimleri

Y DEĞERLERİNE KARŞILIK GELEN EN YÜKSEK UYGUNLUK DEĞERLERİ



(b)

Z DEĞERLERİNE KARŞILIK GELEN EN YÜKSEK UYGUNLUK DEĞERLERİ



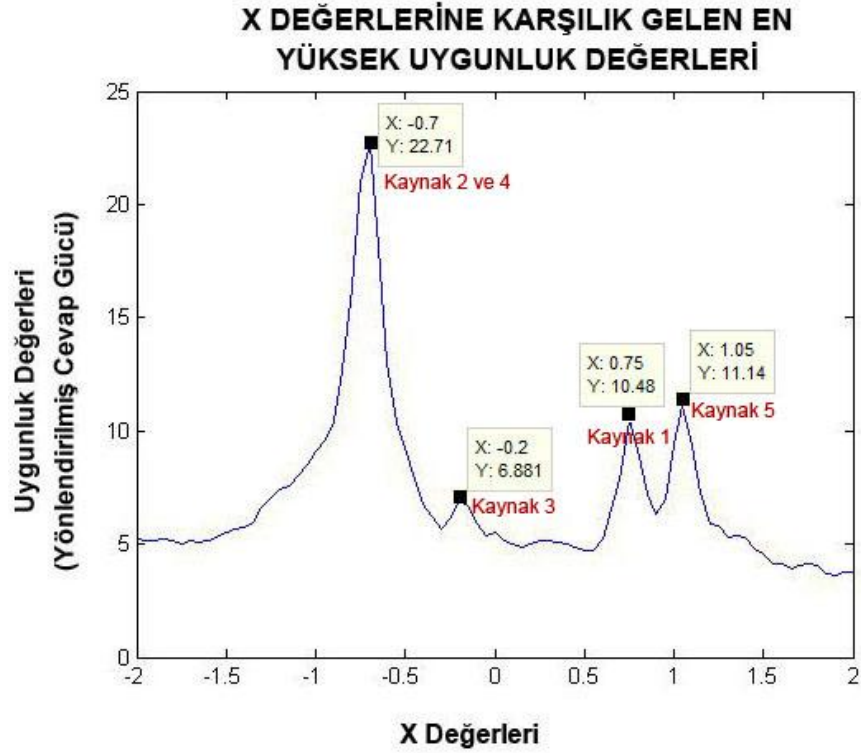
(c)

Şekil 9.5. (Devamı) 20 cm Netlikte Elde Edilen Uygunluk Değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri ve (c) z değerindeki değişimleri

Çizelge 9.11. 20 cm netlik için gerçek değerler ile elde edilen değerlerin karşılaştırması

X Eksen			Y Eksen			Z Eksen		
Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)	Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)	Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)
-79,44	-80,00	0,56	-55,00	-60,00	5	250,52	260,00	9,48
-20,48	-20,00	0,48	-48,00	-60,00	12	311,48	300,00	11,48
72,96	80,00	7,04	-44,00	-60,00	16	372,44	300,00	72,44
103,44	100,00	3,44	-40,00	-60,00	20	402,92	400,00	2,92
			-29,00	-20,00	9			

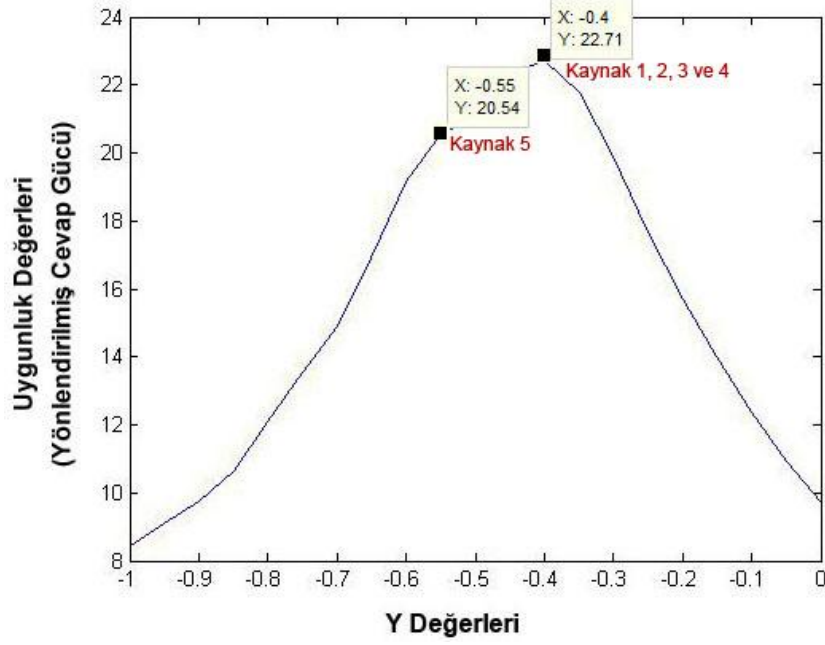
Bu incelemeler sonucunda, 20 cm’lik bir netliğin yetersiz olduğuna karar verilmiştir. Bunun sebebi elde edilen şekillerdeki gösterim doğrularının keskin hatlara sahip olmasıdır. Ayrıca Çizelge 9.11’de verilmiş olan gerçek değerler ile bulunan değerler arasındaki farklılıklardır. Bir sonraki aşama olarak 5 cm netlikte veriler analiz edilmiştir ve Şekil 9.6’te verilmiştir.



(a)

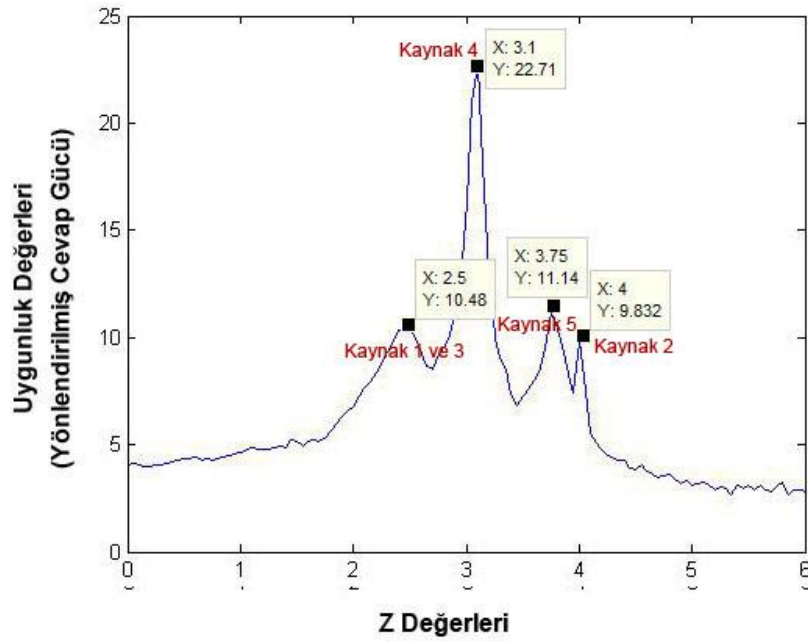
Şekil 9.6. 5 cm netlikte elde edilen uygunluk değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri, (c) z değerindeki değişimleri

Y DEĞERLERİNE KARŞILIK GELEN EN YÜKSEK UYGUNLUK DEĞERLERİ



(b)

Z DEĞERLERİNE KARŞILIK GELEN EN YÜKSEK UYGUNLUK DEĞERLERİ



(c)

Şekil 9.7 (Devamı) 5 cm Netlikte Elde Edilen Uygunluk Değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri, (c) z değerindeki değişimleri

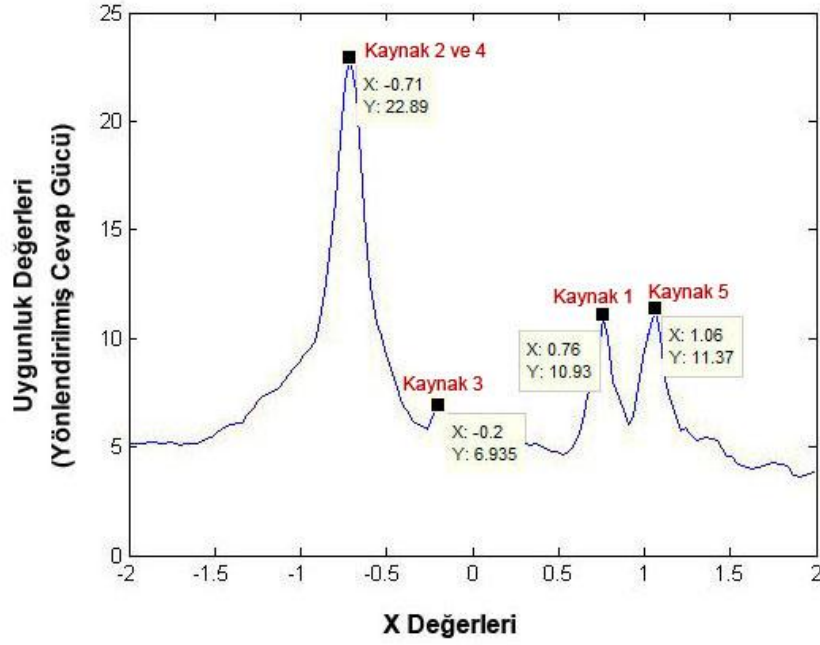
Çizelge 9.12. 5 cm netlik için gerçek değerler ile elde edilen değerlerin karşılaştırması

X Eksen			Y Eksen			Z Eksen		
Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)	Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)	Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)
-79,44	-70,00	9,44	-55,00	-55,00	0	250,52	250,00	0,52
-20,48	-20,00	0,48	-48,00	-40,00	8	311,48	310,00	1,48
72,96	75,00	2,04	-44,00	-40,00	4	372,44	375,00	2,56
103,44	105,00	1,56	-40,00	-40,00	0	402,92	400,00	2,92
			-29,00	-40,00	11			

5 cm netlikte elde edilen 205821 kromozom incelenmiş ve elde edilen grafikler Şekil 9.6'da verilmiştir. Bu inceleme sonucunda x ekseninde uygunluk değerlerinin -0,7,-0,2, 0,75 ve 1,05 değerlerinde belirgin bir şekilde pik yaptığı gözlenmiştir. Y ekseninde ise -0,55 ve -0,4 noktalarında uygunluk değerleri için piklerin olduğu gözlenmiştir. Z ekseninde ise 2,5, 3,1, 3,75 ve 4 noktalarında uygunluk değerlerinin en yüksek değerleri aldığı görülmektedir ve bu değerlerin birbirlerine çok yakındır.

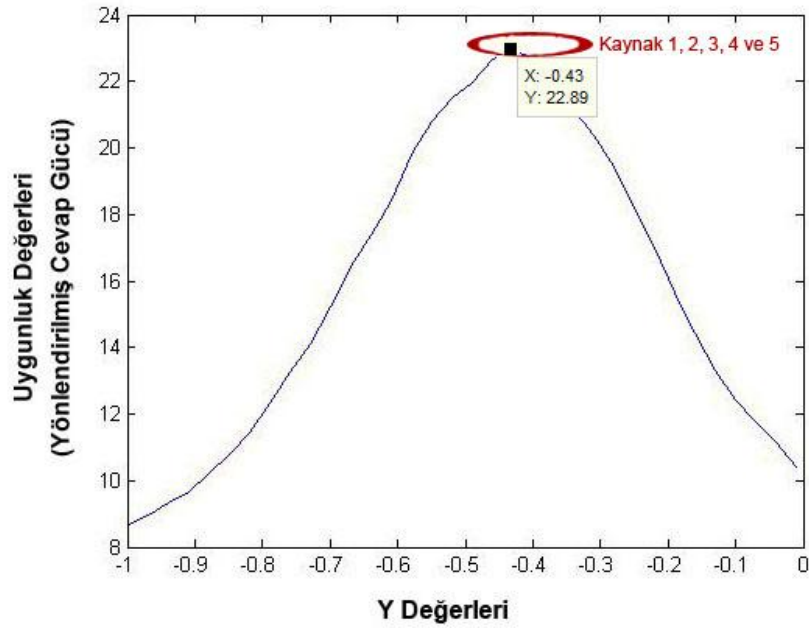
5 cm netlikte, 20 cm netliğe göre farkların Çizelge 9.12'den görüldüğü gibi azaldığı gözlenmiştir. Ancak 3 cm'lik netlikte gözlenmesinde fayda olacağı düşünülmüştür. Şekil 9.7'de 3 cm netlikte veri için yapılan analizler sonucunda elde edilen grafikler verilmiştir.

X DEĞERLERİNE KARŞILIK GELEN EN YÜKSEK UYGUNLUK DEĞERLERİ



(a)

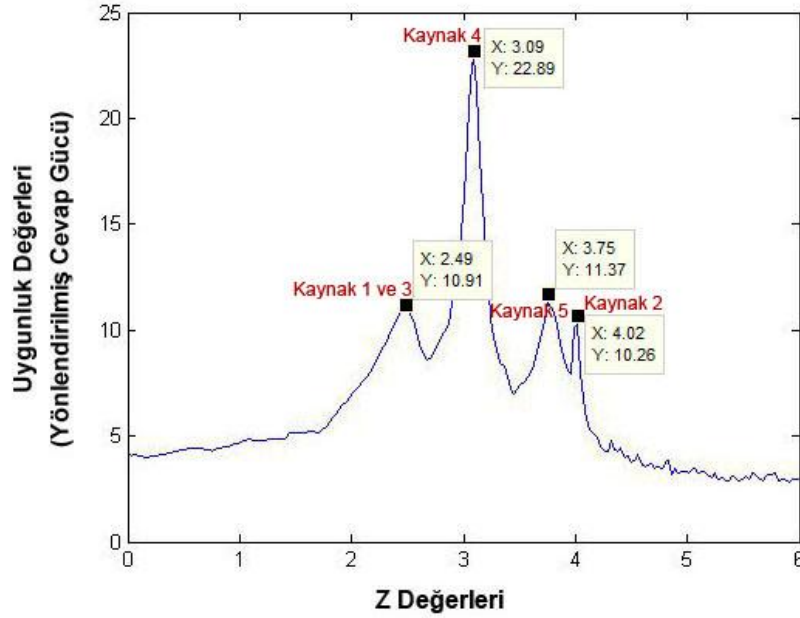
Y DEĞERLERİNE KARŞILIK GELEN EN YÜKSEK UYGUNLUK DEĞERLERİ



(b)

Şekil 9.8. 3 cm netlikte elde edilen uygunluk değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri ve (c) z değerindeki değişimleri

**Z DEĞERLERİNE KARŞILIK GELEN EN
YÜKSEK UYGUNLUK DEĞERLERİ**



(c)

Şekil 9.7 (Devamı) 3 cm Netlikte Elde Edilen Uygunluk Değerleri (a) x değerindeki değişimleri (b) y değerindeki değişimleri ve (c) z değerindeki değişimleri

Çizelge 9.13. 3 cm netlik için gerçek değerler ile elde edilen değerlerin karşılaştırması

X Eksenini			Y Eksenini			Z Eksenini		
Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)	Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)	Gerçek Değer (cm)	Bulunan Değer (cm)	Fark (cm)
-79,44	-71,00	8,44	-55,00	-43,00	12	250,52	249,00	1,52
-20,48	-20,00	0,48	-48,00	-43,00	5	311,48	309,00	2,48
72,96	76,00	3,04	-44,00	-43,00	1	372,44	375,00	2,56
103,44	106,00	2,56	-40,00	-43,00	3	402,92	402,00	0,92
			-29,00	-43,00	14			

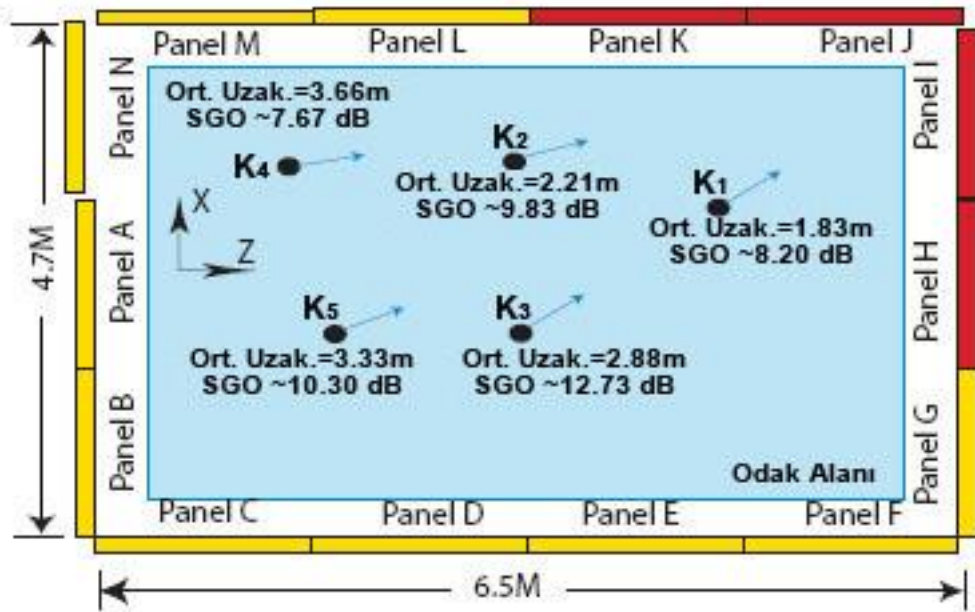
3 cm netlikte elde edilen 927034 değerler incelenmiş ve elde edilen grafikler Şekil 9.7’da verilmiştir. Bu inceleme sonucunda x ekseninde uygunluk değerlerinin -0,71,-0,2, 0,76 ve 1,06 değerlerinde belirgin bir şekilde en yüksek değerleri aldığı gözlenmiştir. Y ekseninde ise -0,43 noktasında uygunluk değerinde en yüksek

olduğu gözlenmiştir. Z ekseninde ise 2,49, 3,09, 3,75 ve 4,02 noktalarında uygunluk değerlerinin en yüksek değerleri aldığı görülmektedir.

3 cm netliğin geliştirilen yöntem ile elde edilecek sonuçların incelenmesi için Çizelge 9.13'te verilen fark değerlerinin yeterli olduğu kabul edilmiştir. Ayrıca incelenecek olan nokta sayısının çok artması nedeniyle Matlab programında çok zaman almaktadır.

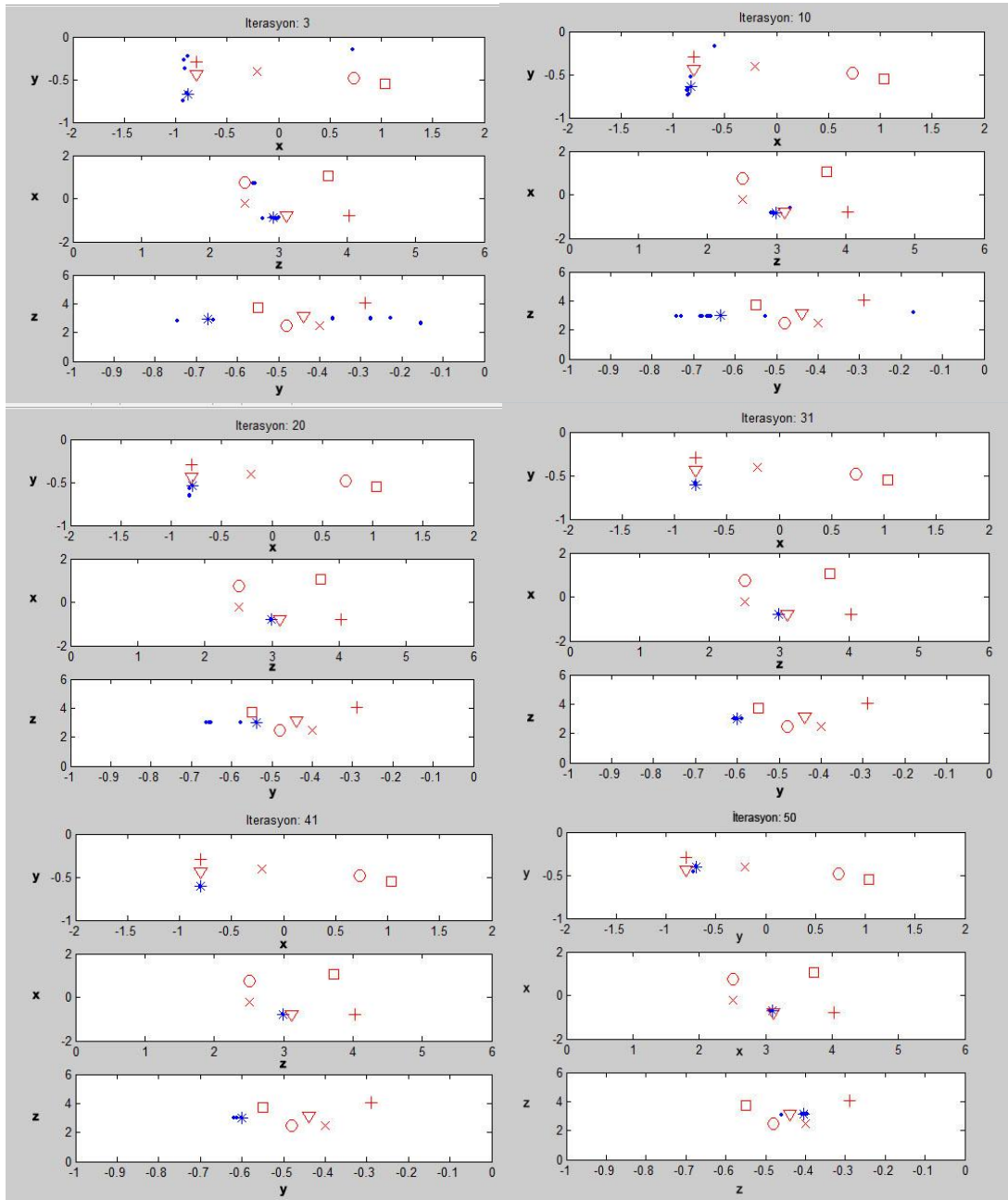
9.3. Genetik Algoritma ile Yapılan Uygulama

Uygulamada kullanılan veriler Büyük Mikrofon Dizisi olarak adlandırılan bir sistem [33] tarafından, 24 mikrofondan oluşan dizi yardımı ile kaydedilmiştir. Odaklanılan odanın ebatları 4m x 1m x 6m'dir. Beş adet sabit ve canlı konuşmacı mikrofon dizisine bakarak aynı anda konuşmuşlardır. Şekil 9.8'de ortalama uzaklıklar ve sinyal gürültü oranları verilmiştir [33]. Sinyal Gürültü Oranları laboratuvar ortamında yapılan ölçümler sonucunda elde edilmiştir. Şekildeki oklar konuşmacıların konuşma yönlerini göstermektedir.



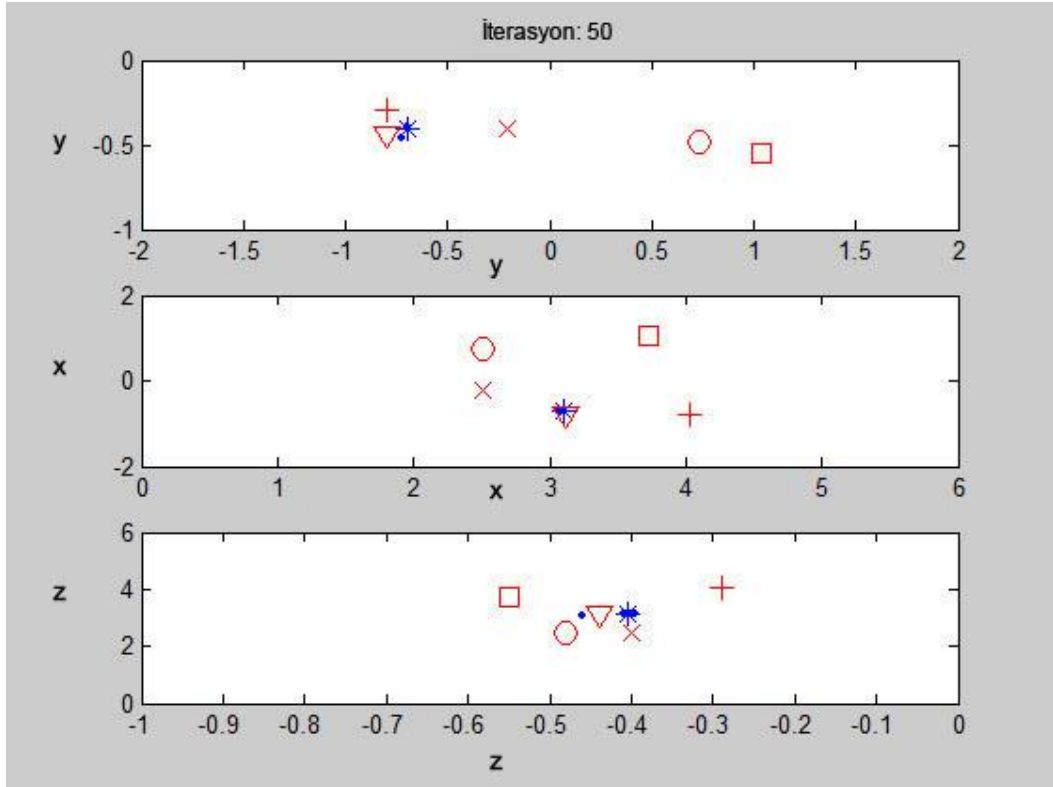
Şekil 9.9. Konuşmacılar ve mikrofon dizisine olan uzaklıkları

Örnekleme oranı 20 KHz'dir. Çalışmada popülasyon büyüklüğü 127 olarak kabul edilmiştir. Şekil 9.9'da algoritmanın bir konuşmacıyı bulurken iterasyonlardan bazılarında alınmış görüntüler yer almaktadır. Burada “.” ile ifade edilen en iyi 9 çözüm yer alırken, “*” ile en iyi çözüm gösterilmektedir. Ayrıca kırmızı ile gösterilmiş her bir şekil bir konuşmacıyı göstermektedir.



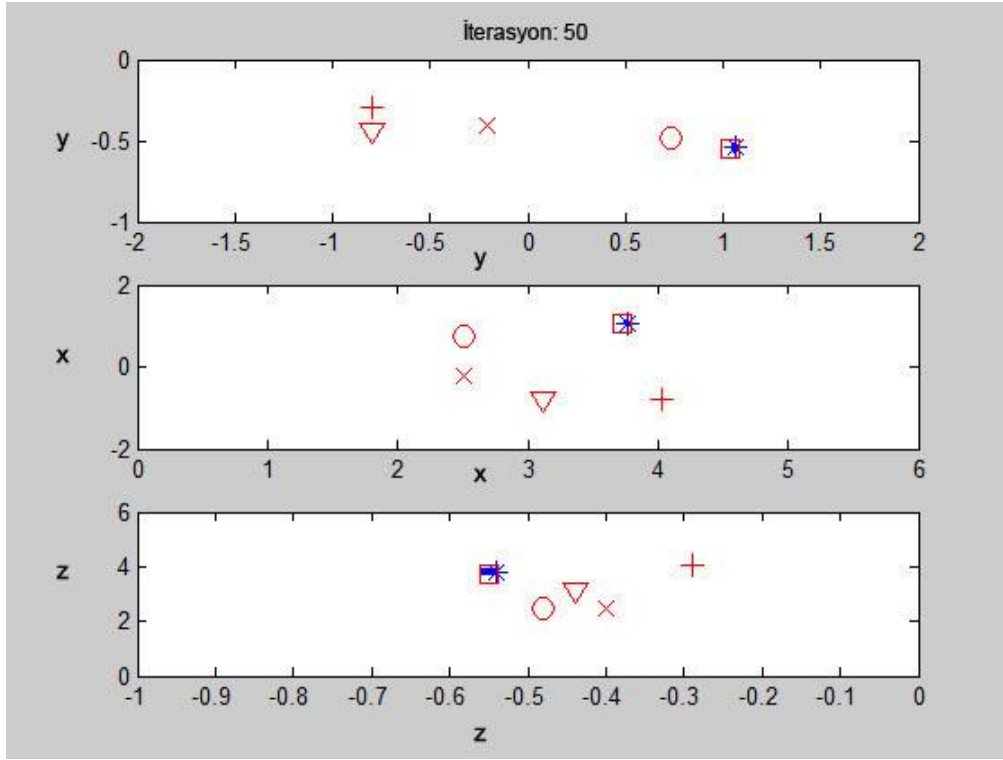
Şekil 9.10. Ses kaynağını bulurken farklı iterasyonlardan görünüm

Şekil 9.10'da ise beş konuşmacının yerlerinin tespiti sırasında elde edilen değerler gösterilmektedir. Aynı şekilde “*” ile en iyi çözüm, “.” ile ise en iyi 9 çözüm gösterilmektedir. Ayrıca kırmızı ile gösterilmiş her bir şekil farklı bir konuşmacıyı göstermektedir.

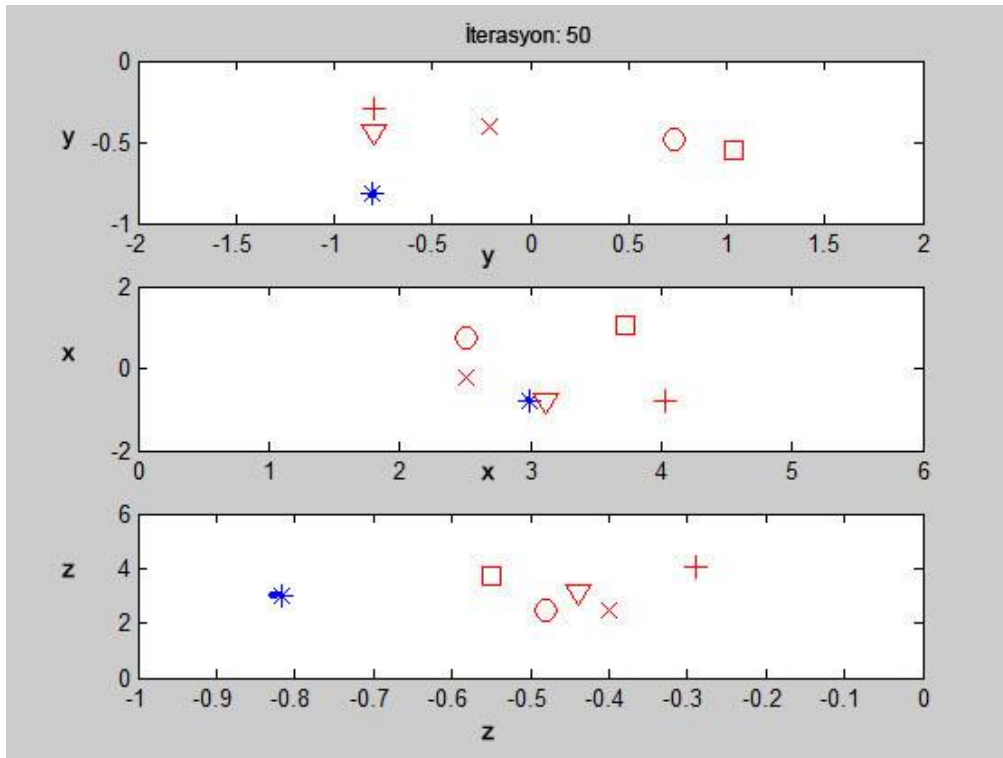


(a)

Şekil 9.11. 5 Konuşmacının yer tespiti (a) ilk konuşmacının, (b) ikinci konuşmacının, (c) üçüncü konuşmacının, (d) dördüncü konuşmacının, (e) beşinci konuşmacının yer tespiti

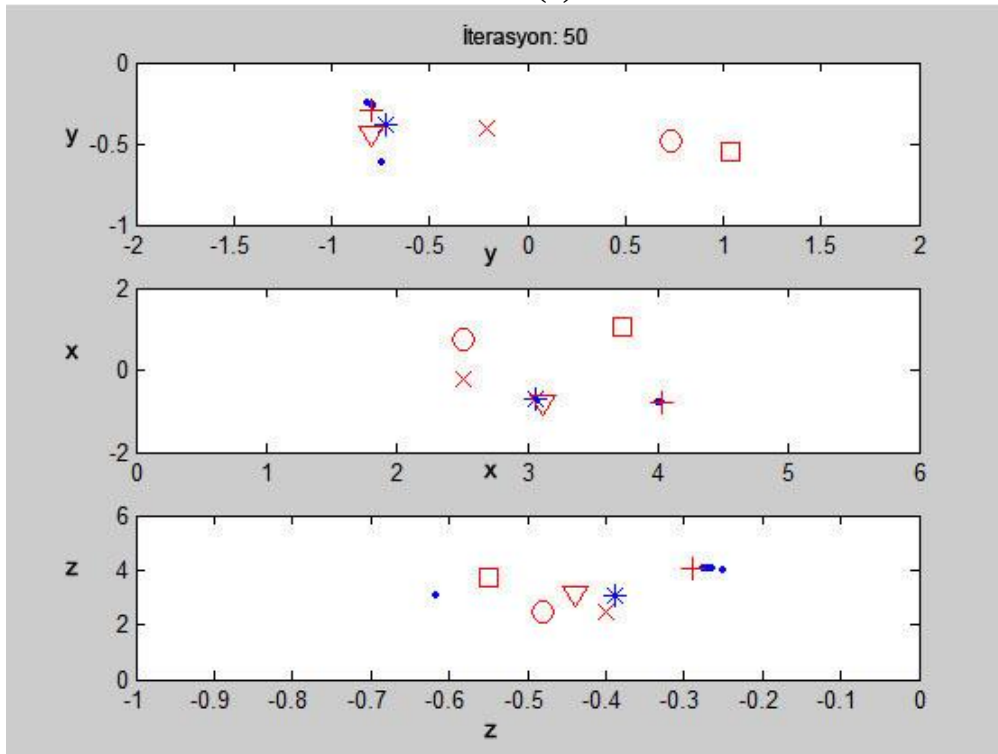
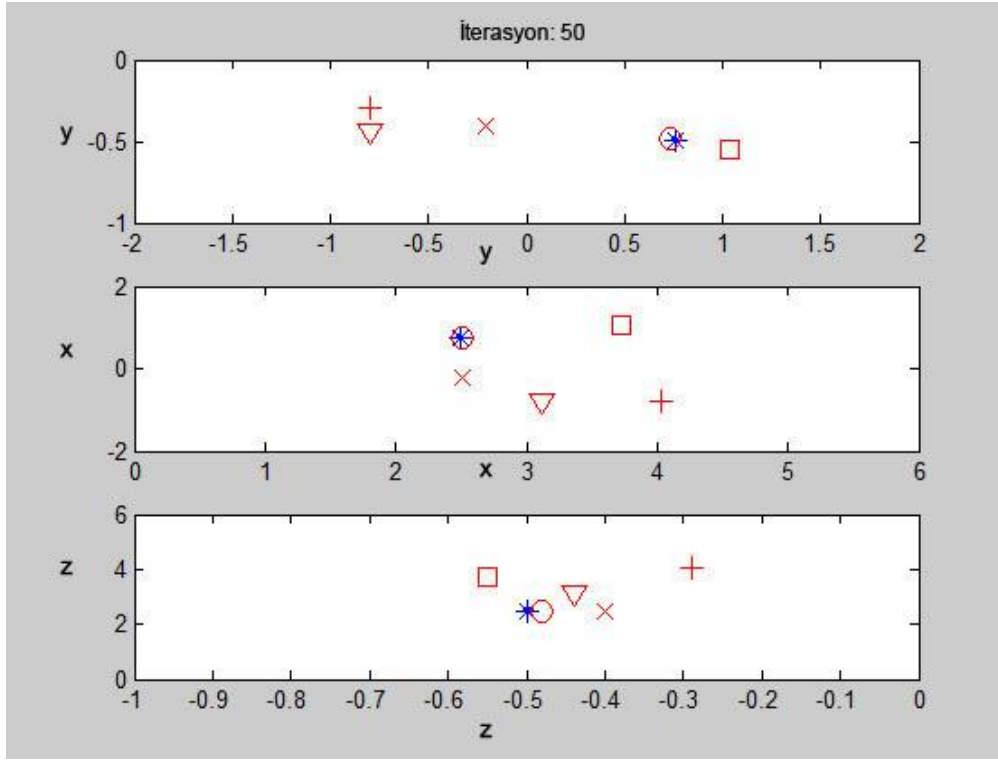


(b)

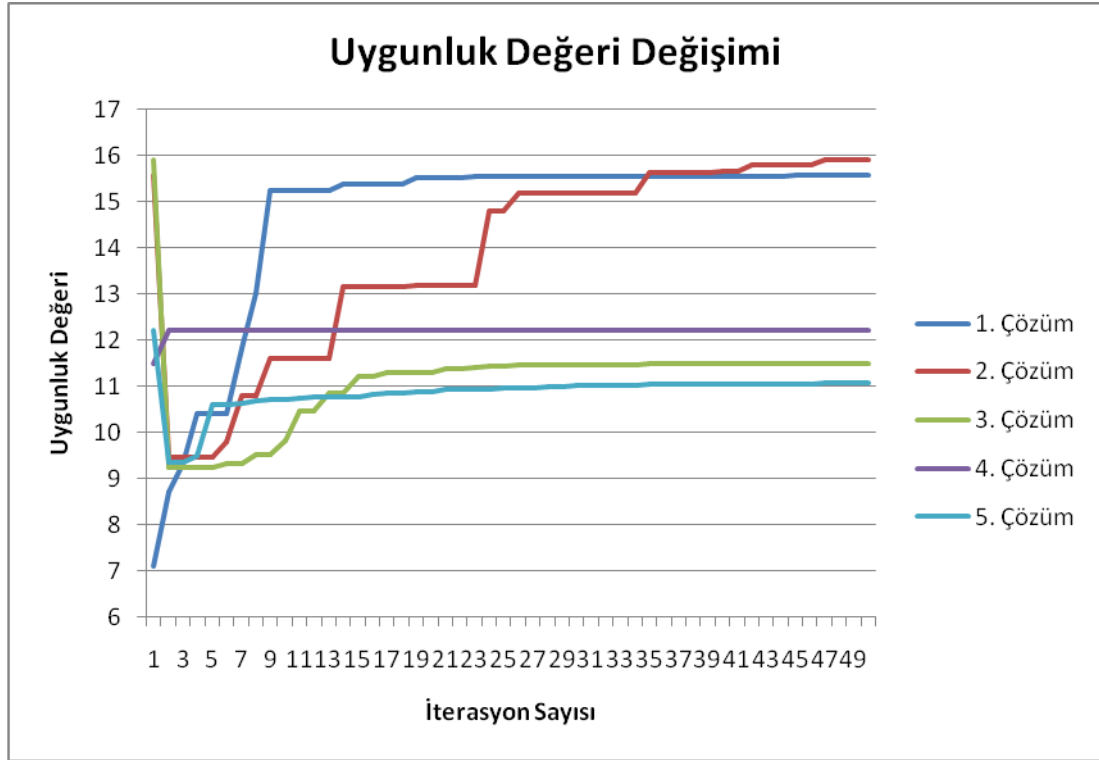


(c)

Şekil 9.10 (Devamı) 5 Konuşmacının yer tespiti (a) ilk konuşmacının, (b) ikinci konuşmacının, (c) üçüncü konuşmacının, (d) dördüncü konuşmacının, (e) beşinci konuşmacının yer tespiti



Şekil 9.10 (Devamı) 5 Konuşmacının yer tespiti (a) ilk konuşmacının, (b) ikinci konuşmacının, (c) üçüncü konuşmacının, (d) dördüncü konuşmacının, (e) beşinci konuşmacının yer tespiti



Şekil 9.12. İterasyon sayısına göre uygunluk değeri değişimi

5 konuşmacının yerinin tespiti için uygunluk değeri değişiminin, iterasyon sayısına göre değişimi Şekil 9.11’de verilmiştir. Genel olarak 25 ’inci iterasyondan sonra uygunluk değerinin belirlendiği görülmektedir. İlk çözümlerde değişimlerin daha yüksek olduğu görülmektedir. Son üç çözümde ise değişimlerin daha az olduğu gözlenmiştir.

Her bir konuşmacı için 50 iterasyon sonucunda yer tespiti gerçekleşmiştir. Çizelge 9.14’te verilmiştir. Bu çizelgeden de anlaşıldığı üzere dört konuşmacının yeri net olarak tespit edilirken, bir konuşmacının yeri doğru olarak tespit edilememiştir. Bu durumun sebeplerine sonuç ve değerlendirme kısımda detayla verilmesine karşın temel sebep, konuşmacının x değerinden daha yüksek uygunluk değerine sahip olan ama konuşmacının bulunmadığı noktaların bulunmasıdır.

Çizelge 9.14. Bulunan konuşmacılar ile gerçek konuşmacı koordinatlarının mukayesesi

	X Gerçek	Y Gerçek	Z Gerçek	X Bulunan	Y Bulunan	Z Bulunan	ΔX	ΔY	ΔZ
Kaynak 1:	72,96	-48,00	250,52	76,36	-49,80	249,41	3,4	1,8	1,11
Kaynak 2:	-79,44	-29,00	402,92	-72,48	-38,82	305,88	6,96	9,82	97,04
Kaynak 3:	-20,48	-40,00	250,52	-80,30	-81,57	298,82	59,82	41,57	48,30
Kaynak 4:	-79,44	-44,00	311,48	-69,69	-40,39	310,59	9,75	3,61	0,89
Kaynak 5:	103,44	-55,00	372,44	106,28	-54,12	376,47	2,84	0,88	4,03

10. SONUÇ VE DEĞERLENDİRME

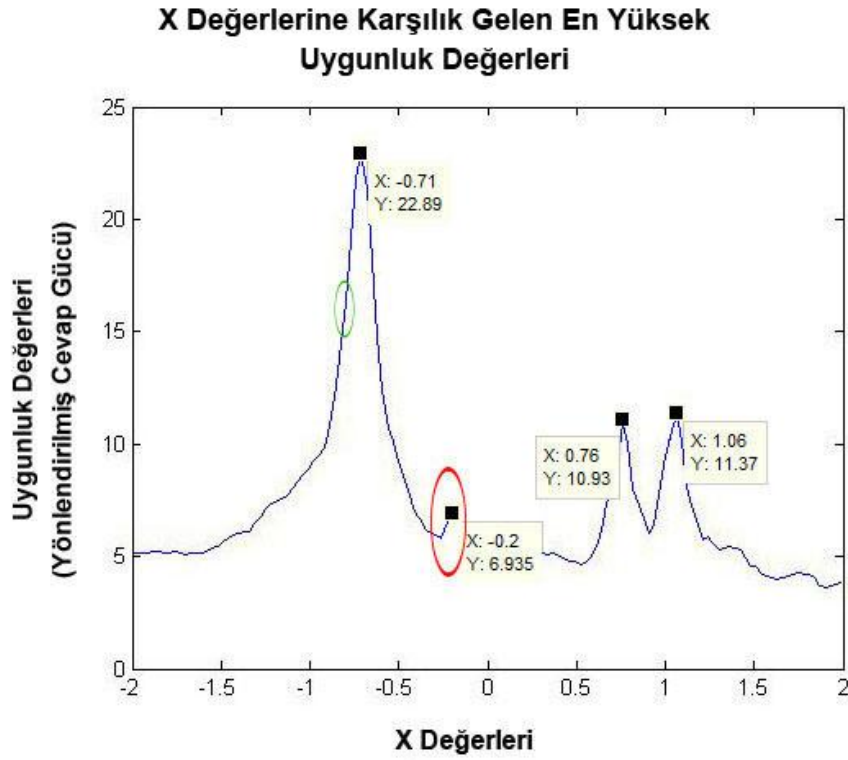
Bu tez çalışmasında, gürültülü ortamlarda aynı anda konuşan, sabit konuşmacıların yerlerinin mikrofon dizileri kullanılarak 3 boyutlu uzayda bulunması hedeflenmiştir. Mikrofon dizilerinin geniş olarak kullanıldığı konuşmacı yeri tespiti problemi, konuşmacı sayısının artması ve konuşmalarının aynı anda olması sonucunda karmaşık hale gelen bir problemdir. Problemin çözümünde konuşmacı sayısının ve mikrofon dizisinde yer alan mikrofonların yerlerinin bilindiği kabul edilmiştir. Aynı anda 5 kişinin %75 üst üste gelecek şekilde konuştuğu ve veri toplama işleminin 24 mikrofondan oluşan bir dizi sayesinde topladığı bir sistemde geliştirilen yöntem denenmiştir.

Problemin çözümü için iki temel metottan oluşan bir sistem kullanılmıştır. İncelenecek olan noktaların belirlenmesinde stokastik bir yöntem olan genetik algoritmadan faydalanılmıştır. Genetik algoritma, koşullar karşısında en güçlü bireylerin sonraki nesillere aktarılmasını sağlamaktadır. Bu kapsamda, her zaman en iyi bireylerin elde edilmesini sağlamaktadır. İkinci temel yöntem ise bireyler arasında güçlülüğü ölçmek için gerekli olan uygunluk değeridir. Uygunluk değeri yüksek olan bireyler konuşmacıların olası yerlerini göstermektedir. Uygunluk değeri hesaplamada, mikrofon dizilerinde sıklıkla kullanılan filtrele-ve-topla hüzme şekillendiricisi kullanılmıştır. Filtreleme işleminde, dizideki mikrofonların yerlerinin bilinmesi ile ses sinyalinin kaynaktan mikrofonlara geliş süreleri arasındaki farkı ikili olarak inceleyen Genelleştirilmiş Çapraz Korelasyon metodu kullanılmıştır. Mikrofonlardan elde edilen ve işleme tabi tutulan sinyallerin daha iyi elde edilmesi için Genelleştirilmiş Çapraz Korelasyon metoduna Faz Dönüşümü ağırlık fonksiyonu eklenmiştir. Filtrelerden elde edilen sinyallerin toplanması ile çıktı sinyali, Yönlendirilmiş Tepki Gücü sinyaline eşittir. Elde edilen sinyalin en yüksek değerleri aldığı noktalar, konuşmacıların yerleri olarak kabul edilmiştir. Sonuç olarak bu çalışmada hüzme şekillendirici olarak Yönlendirilmiş Tepki Gücü - Faz Dönüşümü metodu kullanılmıştır.

Ayrıca çalışmada kör bölge uygulaması olarak adlandırılan ve geliştirilmesi proje kapsamında yapılan bir metottan faydalanılmıştır. Genetik algoritma doğası gereği tek bir optimum sonuca yönelmektedir. Dolayısıyla ilk çalışımından sonra bir kaynağın yerini tespit ettiğinde, ilk seferinde en güçlü kaynağı bulacağından her zaman aynı kaynağa yönelmektedir. Bu yönelmeyi aşmak amacı ile kör bölge uygulaması kullanılmıştır. Bu uygulamada bir kaynağın yeri tespit edildiğinde, kaynak merkezde kalacak şekilde 40 cm kenar uzunluğuna sahip küp şeklinde bir alan kör bölge olarak ilan edilmektedir. Bir sonraki arama işlemlerinde kör bölge ilan edilen yerlerde bulunan çözüm adayları incelenmemektedir. Böylece diğer ses kaynaklarına erişim sağlanmıştır.

Çalışma MATLAB ile kodlanmış ve çalıştırılmıştır. Çalışmada kullanılan ses parçalarının her biri 102.4 ms'dir ve her seferinde 25.6 ms kaydırılmıştır. Örnekleme oranı 20 KHz'dir. Çalışmada popülasyon büyüklüğü 127 olarak kabul edilmiştir ve 50 iterasyona göre çalıştırılmıştır.

Çalışmanın sonucu 4 konuşmacının yeri oldukça net bulunurken bir konuşmacının yeri bulunamamıştır. Bunun sebebi, verinin analizi bölümünde yer alan, x değerlerine karşılık gelen en yüksek uygunluk değeri grafiğinde görülmektedir. 3 numaralı konuşmacının yeri $[-0,2048;-0,4000;-0,5052]$ dir. Şekil 10.1'de verilmiş olan grafikte kırmızı daire ile gösterilen ve konuşmacıya ait olan x değeri diğer x değerlerinin oldukça altındadır. Her iki yanında 20'şer cm'lik kör bölge ilan edilen diğer değerlerde bile kör bölgenin dışında kalan ama uygunluk değeri halen 3. konuşmacının x değerine ait uygunluk değerinden yüksek olan birçok nokta bulunmaktadır. Genetik algorithmada bu değerlerden birine yakınsamaktadır. Çizelge 8.6'da verilen örnek çalışmada, x değeri -81,57 noktasına yakınsamıştır. Şekil 10.1'de bu nokta yeşil daire ile gösterilmiştir. Bu nedenle konuşmacının yerinin bulunamaması kabul edilebilir bir durumdur.



Şekil 10.1. Üçüncü konuşmacının x değerinin incelenmesi

Bu çalışmada elde edilen sonuçlara göre ilerleyen çalışmalarda yapılabilecekler aşağıdaki gibi listelenebilir;

- Farklı stokastik yöntemler ile çalışmanın yeniden yapılması. Yapay Arı Koloni yöntemi amaç fonksiyonunda birden fazla sonucun olduğu problemler için uygundur. Birden fazla konuşmacının yer aldığı sistemde bu yöntemin faydalarından yararlanılabilir.
- Çalışma için farklı veri tabanları kullanılarak, geliştirilen yöntem farklı ortamlarda denenmiş olacaktır.

KAYNAKLAR

1. Fujimoto, M.; Ariki, Y.” Noise robust hands-free speech recognition using microphone array and Kalman filter as front-end system of conversational TV” **Multimedia Signal Processing, 2002 IEEE Workshop**, 9-11 , 268 - 271, (2002).
2. Yu, Y., “Distant Talker Labelling for a Conferencing Situation Using Data from a Large Aperture Microphone Array in a Noisy Reverberant Environment”, Doktora Tezi, **Brown Üniversitesi Mühendislik Ana Bilim Dalı**, Amerika Birleşik Devletleri (2007).
3. Sturim, DE, Brandstein, MS., Silverman, HF., “Tracking multiple talkers using microphone-array measurements”, **In Proceedings of ICASSP97**, IEEE, (1997).
4. Valin, J.M., Michaud, F., Rouat, J., Létourneau, D., “Robust sound source localization using a microphone array on a mobile robot,” **in Proceedings International Conference on Intelligent Robots and Systems**, (2003).
5. DiBiase, J. H., “High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays” Doktora Tezi, **Brown Üniversitesi Mühendislik Ana Bilim Dalı**, Amerika Birleşik Devletleri (2000).
6. Silverman, H.F., Patterson, W.R., Sachar, J.M., “First Measurements of a Large-aperture Microphone Array System for Remote Audio Acquisition,” **in Proceedings of IEEE International Conference on Multimedia and Expo**, (2000).
7. Meyer, J., Elko, G., “A Highly Scalable Spherical Microphone Array Based On Anorthonormal Decomposition Of The Soundfield”, **IEEE ICASSP-02**, USA, (2002).
8. Do, H., Silverman, H. F., Yu, Y., “A real-time SRP-PHAT Source Location Implementation Using Stochastic Region Contraction (SRC) on a Large-Aperture Microphone Array,” **IEEE International Conference on Acoustics, Speech, and Signal Processing 2007 (ICASSP 2007)**, 1, 121 – 124, (2007).
9. İnternet: Acustic Magic Firması Voice Talker ürünü, <http://www.acousticmagic.com/>
10. İnternet: Polycom Firmasının HDX Tavan Mikrofon Dizisi Ürünü: http://www.polycom.com/support/video/hdx_series/hdx_ceiling_microphone_array.html
11. Do, H., Silverman, H.F., “A method for locating multiple sources using a frame of a large-aperture microphone array data without tracking,” **in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing 2008 (ICASSP 2008)**, 301 – 304, (2008).

12. Yu, Y., Silverman, H.F., “An improved TDOA-based location estimation algorithm for large-aperture microphone arrays,” *in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing 2004 (ICASSP 2004)*, 4: 77 – 80, (2004).
13. İnternet: Wikipedia Ansiklopedisi: Microphone
http://en.wikipedia.org/wiki/Microphone#MEMS_microphone
14. İnternet:
http://www.guyfletcher.co.uk/index.php/getluckyrecording2009/March_2009_Update
15. İnternet: http://www.hitproductions.net/blog_index.php
16. Widrow B. “A Microphone Array for Hearing Aids”, *IEEE Circuits and Systems Magazine*, 1(2):26-32, (2001).
17. İnternet: Watri Araştırma Grubu Mikrofon Dizisi Araştırma Bölümü
<http://www.watri.org.au/research-labs-watri/signal-proc.html>
18. Çontar, E., “Mikrofon Dizilerinde Ses Kaynağının Yerinin Genetik Algoritma Kullanılarak Bulunması”, Yüksek Lisans Tezi, *Gazi Üniversitesi Fen Bilimleri Enstitüsü*, (2007)
19. Moore, D. C., McCowan , I. A., “Microphone Array Speech Recognition : Experiments on Overlapping Speech in Meetings” *IEEE 2003 International Conference on Acoustics*, 5: 497-500, (2003)
20. Sturim, D.E., Branstein, M.S., Silverman, H.S., “Tracking Multiple Talkers Using Microphone-Array Measurements” *IEEE 1997 International Conference on Acoustics*, 5: 371-374, (1997)
21. Johnson, D.H., Dudgeon, D.E., “Array Signal Processing: Concepts and Techniques”. *P T R Prentice Hall, Englewood Cliffs*, New Jersey 07632, (1993).
22. Brandstein, M.S., Adcock, J.E., Silverman, H.F., “A practical time-delay estimator for localizing speech sources with a microphone array”, *Computer, Speech and Language*, 9: 153-169, (1995).
23. Altıparmak, F., “Genetik Algoritma ile Haberleşme Şebekelerinin Topolojik Tasarımı”, Doktora Tezi, *Gazi Üniversitesi Fen Bilimleri Enstitüsü*, (1996).
24. Goldberg, D.E., “Genetic Algorithms in Search, Optimization and Machine Learning”, *Addison Wesley*, 1 – 85, (1989).
25. Kaya, M., “Genetik Algoritma Ve Gezgin Satıcı Probleminin Çözümü”, Yüksek Lisans Tezi, *Fırat Üniversitesi Fen Bilimleri Enstitüsü* (1999).

26. Güçlü, İ., "A Genetic Algorithm Approach To The Tool Allocation Problem İn Flexible Manufacturing Systems", Yüksek Lisans Tezi, **Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü** , (2006).
27. Güç, G., "Optimization Of Water Distribution Networks Using Genetic Algorithm", Yüksek Lisans Tezi, **Orta Doğu Teknik Üniversitesi Fen Bilimleri Enstitüsü**, (2006).
28. Tokgöz, H., "Hava Harekatında Görüş Ötesi Angajman İçin Genetik Algoritma Uygulaması" Yüksek Lisans Tezi, **Gazi Üniversitesi Fen Bilimleri Enstitüsü**, (2007).
29. Karaboğa, D., "Yapay Zeka Optimizasyon Algoritmaları", **Atlas Yayın Dağıtım**, İstanbul, 75-112 (2004).
30. Gen, M., Cheng, R., "Genetic Algorithms & Engineering Design", **Wiley: New York**, 1 – 14 (1997).
31. Mahmood, H., "Derece Kısıtlı Minimum Yayılan Ağaç Problemi İçin Genetik Algoritmalar", Yüksek Lisans Tezi, **Gazi Üniversitesi Fen Bilimleri Enstitüsü**, (2005).
32. Do, H., Silverman, H. F., Yu, Y., "A real-time SRP-PHAT source location implementation using stochastic region contraction (SRC) on a large-aperture microphone array," **IEEE International Conference on Acoustics, Speech, and Signal Processing 2007 (ICASSP 2007)**, 1: 121-124, (2007).

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : SEZEN, Esmâ Sultan
 Uyruğu : T.C.
 Doğum tarihi ve yeri : 31.07.1983 Ankara
 Medeni hali : Evli
 Telefon : 0 (312) 222 96 35
 Faks : 0 (312) 222 96 35
 e-mail : esmasultan@gmail.com.

Eğitim

Derece	Eğitim Birimi	Mezuniyet tarihi
Lisans	Atılım Üniversitesi / Bilgisayar Mühendisliği Bölümü	2006
Lisans	Atılım Üniversitesi / Endüstri Mühendisliği Bölümü	2006
Lise	Açık Öğretim Lisesi	2001

İş Deneyimi

Yıl	Yer	Görev
2006-2007	Türksat AŞ	Uzman Yardımcısı

Yabancı Dil

İngilizce

Yayınlar

1. Kaplan,H., Sezen, E.S., Akdoğan, M.R, Kuzu, S., “Türksat VSAT IP Projeleri: Uygulama Alanları”, I. NATİonal Aviation And Space Conference, September 2006, METU, ANKARA