

**TOTAL LARENJEKTOMİ HASTALARI İÇİN  
KONUŞMA KODLAMA SİSTEMİ**

**Mesut Oytun OKTAY**

**Yüksek Lisans Tezi**

**Elektronik ve Haberleşme Mühendisliği Anabilim**

**Dalı**

**Danışman: Yrd. Doç. Dr. Rafet AKDENİZ**

**2014**

**T.C.**  
**NAMIK KEMAL ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS TEZİ**

**TOTAL LARENJEKTOMİ HASTALARI İÇİN KONUŞMA KODLAMA  
SİSTEMİ**

**Mesut Oytun OKTAY**

**ELEKTRONİK ve HABERLEŞME MÜHENDİSLİĞİ ANABİLİM DALI**

**DANIŞMAN: Yrd. Doç. Dr. Rafet AKDENİZ**

**TEKİRDAĞ-2014**

**Her hakkı saklıdır**

**Bu alıřma NKUBAP.00.17.YL.11.01 numaralı Namık Kemal Üniversitesi Bilimsel Arařtırma Projesi ile desteklenmiřtir.**

Yrd. Doç. Dr. Rafet AKDENİZ danışmanlığında, Mesut Oytun OKTAY tarafından hazırlanan “Total Larenjektomi Hastaları için Konuşma Kodlama Sistemi” isimli bu çalışma aşağıdaki jüri tarafından Elektronik ve Haberleşme Mühendisliği Anabilim Dalı’nda Yüksek Lisans tezi olarak oy birliği ile kabul edilmiştir.

Jüri Başkanı: Yrd. Doç. Dr. Rafet AKDENİZ

*İmza:*

Üye: Doç. Dr. Rıfat Koray ÇİFTÇİ

*İmza:*

Üye: Yrd. Doç. Dr. Erdiñ UZUN

*İmza:*

Fen Bilimleri Enstitüsü Yönetim Kurulu adına

Prof. Dr. Fatih KONUKCU  
Enstitü Müdürü

## ÖZET

Yüksek Lisans Tezi

### TOTAL LARENJEKTOMİ HASTALARI İÇİN KONUŞMA KODLAMA SİSTEMİ

Mesut Oytun OKTAY

Namık Kemal Üniversitesi

Fen Bilimleri Enstitüsü

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Danışman : Yrd. Doç. Dr. Rafet AKDENİZ

Bu çalışma, çeşitli nedenlerle gırtlak alınmış hastaların konuşmasına yardımcı olarak kullandıkları elektronik gırtlakın oluşturduğu mekanik sesli konuşmanın daha doğal bir konuşmaya dönüştürülmesi için bir işaret işleme yöntemi belirlemeyi amaçlar.

Bu çalışmada sesin oluşumu ve bu sırada görev yapan organlar açıklandıktan sonra bu organlardan gırtlakın yitimi sonucu ortaya çıkan sorun tanımlanmış ve bu soruna getirilen tıbbi çözümler hakkında bilgi verilmiştir. Konuşma üretiminin sağlanması için insan vücuduna en az müdahale ile kullanılabilen elektronik gırtlak desteğinin temel sorunu ve bunun ses işaretindeki yansıması hakkında bilgi verildikten sonra bu sorunun en aza indirgenmesi için uygulanabilecek yöntemlerden bahsedilmiştir ve başarımlarını karşılaştırmak için veri toplanmıştır.

Veri toplamada hastalığa sahip kişiler yerine sağlıklı bireylerin kullanımı tercih edilmiştir. Sağlıklı bireylerden elde edilen veriler işlenip başarımların değerlendirilmesi bu bireyler dışındaki bir grup insana yaptırılmıştır.

Çalışmanın sonunda en başarılı yöntem verilmiş ve gelecekte yöntemi iyileştirmek için yapılabilecekler değerlendirilmiştir.

**Anahtar kelimeler:** Sinyal İşleme, Ses İyileştirme, Larenjektomi

2014, 48 sayfa

## **ABSTRACT**

MSc. Thesis

### **SPEECH CODING SYSTEM FOR TOTAL LARYNGECTOMEES**

Mesut Oytun OKTAY

Namık Kemal University  
Graduate School of Natural and Applied Sciences  
Department of Electronics and Telecommunication Engineering

Supervisor : Assist. Prof. Dr. Rafet AKDENİZ

This work aims to obtain a signal processing method to neutralize the mechanical vibration sound from the electrolarynx which is used as a supporting device for speech restoration of people who have gone under resection operation over their larynxes due to various reasons.

In this work production of voice and the organs that take action are reviewed and then the problem arises after the resection of the larynx and medical approaches to voice regeneration after surgery are overviewed. After the problem with electrolarynx, noninvasive sound regeneration method, and its reflection on the sound signal are outlined, the approaches that can be taken are summarised and to asses their impact opinion score is collected.

In the data collection step, instead of laryngectomees, healthy individuals are preferred. The data obtained from healthy subjects are processed and evaluation of the methods are accomplished by the individuals that are not involved in data gathering step.

At the end of the work, the best method is given and probable approaches to achieve higher performance are mentioned.

**Keywords:** Signal Processing, Speech Enhancement, Laryngectomy

**2014, 48 pages**

## İÇİNDEKİLER

	Sayfa
<b>ÖZET</b> .....	<b>i</b>
<b>ABSTRACT</b> .....	<b>ii</b>
<b>İÇİNDEKİLER</b> .....	<b>iii</b>
<b>ÇİZELGE DİZİNİ</b> .....	<b>iv</b>
<b>ŞEKİL DİZİNİ</b> .....	<b>v</b>
<b>KISALTMALAR DİZİNİ</b> .....	<b>vi</b>
<b>TEŞEKKÜRLER</b> .....	<b>vii</b>
<b>1 GİRİŞ</b> .....	<b>1</b>
<b>2 SES ve KONUŞMANIN OLUŞUMU</b> .....	<b>5</b>
2.1 Ses Üretim Mekanizmasında Görev Alan Organlar .....	5
2.1.1 Akciğerler.....	5
2.1.2 Gırtlak.....	6
2.1.3 Ses Yolu.....	7
2.2 Sesin ve Konuşmanın Üretimi .....	8
2.3 Seslerin İşaretsel Ayrışması.....	9
<b>3 SORUNUN TANIMI ve ÇÖZÜM ÖNERİSİ</b> .....	<b>12</b>
3.1 Çözüm Önerisi.....	12
3.2 Veri Toplama .....	13
3.3 Önışleme.....	14
3.3.1 Gürültü Kestirimi .....	14
3.4 Spektral Gürültü Çıkarma Algoritmaları.....	14
3.4.1 Spektral Genlik Çıkarması .....	15
3.4.2 Spektral Güç Çıkarma .....	15
3.4.3 Aşırı Çıkarma Kullanılan Spektral Çıkarma .....	16
3.4.4 Doğrusal Olmayan Spektral Çıkarma .....	17
3.4.5 Çok Bantlı Spektral Çıkarma .....	18
3.4.6 Algısal Spektral Çıkarma .....	20
3.5 Frekans Uzayından Zaman Uzayına Geçiş .....	21
<b>4 SONUÇLAR ve VARGILAR</b> .....	<b>22</b>
4.1 Vargılar .....	23
4.2 İleri Çalışmalar .....	23
<b>5 KAYNAKLAR</b> .....	<b>25</b>
<b>EKLER</b> .....	<b>27</b>
<b>ÖZGEÇMİŞ</b> .....	<b>48</b>

## ÇİZELGE DİZİNİ

	<b>Sayfa</b>
Çizelge 2.1 Türkçe'nin Sesbilim Abecesi (Ergenç 2002) .....	10
Çizelge 4.1 Ortalama Fikir Değerlendirme Çizelgesi .....	22



## ŞEKİL DİZİNİ

	<b>Sayfa</b>
Şekil 2.1	Konuşma üretiminde aktif rol oynayan organlar ..... 5
Şekil 2.2	Gırtlakın transvers kesitten görünüşü ..... 6
Şekil 2.3	Glottal Hava Basıncı Değişiminin (a) zamana göre değişimi ve (b) frekans uzayındaki genlik bileşenleri. Ses yolunun rezonans karakteristiği (c) ve ses işaretinin tüm değişimlerden sonraki halinin genlik spektrumu(d). ..... 7
Şekil 2.4	Ses üretimi sistem modeli ..... 8
Şekil 3.1	Elektronik Gırtlak ve Kullanım Şekilleri (Griffin Labs 2008) ..... 12
Şekil 3.2	Kullanılan ses iyileştirme sistemi özeti ..... 13
Şekil 3.3	Önişlemede kullanılan alçak geçiren ve yüksek geçiren filtrelerin frekans yanıtı ..... 14
Şekil 3.4	Spektral Genlik Çıkarım Algoritması Blok Diyagramı ..... 15
Şekil 3.5	Spektral Güç Çıkarım Algoritması Blok Diyagramı ..... 16
Şekil 3.6	Aşırı Çıkarım Kullanılan Spektral Güç Çıkarım Algoritması Blok Diyagramı 17
Şekil 4.1	Fikir puanlarının dağılımları ..... 22
Şekil 4.2	İşaret Kalite Değerlendirme Uygulaması ..... 23

## KISALTMALAR DİZİNİ

İGO : İşaretin Gürültüye Oranı

## TEŞEKKÜRLER

Bu çalışmanın gerçekleşmesinde emeđi geçen başta sabırlı sađlıklı denekler olmak üzere herkese ve özellikle

danışmanım Yrd. Doç. Dr. Rafet Akdeniz'e,

NKUBAP.00.17.YL.11.01 numaralı Bilimsel Araştırma Projesi'yle çalışmada kullandığımız ekipmanı sađlayan Namık Kemal Üniversitesi'ne

teşekkürü bir borç bilirim.

Bu çalışmayı yaparken kaybettiğimiz, irdedeğimiz sorunun çözümü için seçtiğimiz yolu aydınlatan eserin yazarı Philipos C. Loizou'nun (10 Kasım 1965-22 Temmuz 2012) anısına...



# 1 GİRİŞ

Konuşma, iletişim araçlarının çeşitlendiği ve yaygınlaştığı günümüzde bile insanlar arası iletişimde en çok tercih edilen araçtır. Konuşmada, dilin belirlediği sınırlı sayıda ünlü ve ünsüzün, sınırlı sayıda ses birimiyle, fonemle, ifade edilmesiyle oluşturulan kelimeler, dil söz dizimine uygun olarak dizilerek aktarılmak istenen bilgi karşı tarafa iletilir. Bu yolla iletilen bilgi miktarı ses birimi başına seslerin bir örnek dağılıma sahip oldukları varsayımıyla yaklaşık 5 bittir. Normal hızda konuşulduğunda ise bu 50 bit/s gibi bir hıza karşılık gelmektedir (Flanagan ve ark. 2008) (Ek 4.2). Konuşmanın hızlı veya verimli bir iletişim aracı olmamasına rağmen insanlar arası iletişimde en çok tercih edilen yol olması konuşmanın doğal bir iletişim aracı olmasından kaynaklanmaktadır.

İnsanlar arası iletişimde temel kipi oluşturan konuşma kaybedildiğinde sestən öte en sık ve rahat kullanılan bilgi aktarım kipinin kaybı söz konusu olduğundan yeniden kazanımı önem ve aciliyet gerektirir.

Konuşmanın bozulması veya yitimi, konuşma üretiminde görev alan organ veya organların işlevlerini yerine getirememesi sonucu farklı biçim ve farklı seviyelerde olur. Bozulmanın sebebi de basit bir patojenden beyin fonksiyonlarının yitimine kadar değişen bir yelpazededir. Dolayısıyla bozulmanın çeşitleri ve bunları düzeltmenin yolları da farklılaşır.

Bu çalışmada gırtlak kanseri sonrasında gırtlak tamamen alınmış (total larenjektomi geçirmiş) hastalarda meydana gelen konuşma bozulması incelenecek ve total larenjektomi sonrası ses rehabilitasyon yöntemlerinden elektronik gırtlak konuşmasının iyileştirilmesi üzerinde durulacaktır.

Gırtlak kanseri baş ve boyunda görülen en sık görülen kötü huylu tümördür (Genden ve ark. 2007). Bu hastalıkla mücadelede radyoterapi ve koruyucu gırtlak cerrahisi (konservatif larengeal cerrahi) yöntemlerinin yetersiz kaldığı durumlarda parsiyal, yani kısmi, veya total larenjektomi, gırtlakın tamamen alınması yoluna başvurulur (Aksoy ve ark. 2010).

Total larenjektomi, tümörün alınması için gırtlakın alınması, reseksiyonu, fıkri 19. yüzyılın başından beri vardı fakat anestezi eksikliği, hava yolunun kontrolünü sağlayacak tıbbi cihaz eksikliği, kan nakli ve antibiyotiklerin gelişmemesi nedeniyle gerçekleştirilemiyordu. Bu olanaksızlıklara rağmen Theodor Billroth, kısmi gırtlak alınması (parsiyal larenjektomi (hemilaryngectomy)) sonrası tümörü tekrar eden hastaya 31 Aralık 1873 yılında total larenjektomi uygulamış, hasta 3 Mart 1874 yılında taburcu olmuş fakat tümörün yayılması (metastaz, metastatik hastalık) nedeniyle ameliyattan 7 ay sonra kaybedilmiştir. İlk başarılı total larenjektomi ile tedavi, İtalya Turin'den Enrico Bottini tarafından 1875 yılında gerçekleştirilmiştir ve hasta ameliyattan sonra 15 yıl yaşayabilmiştir (Genden ve ark. 2007).

Total larenjektomiden sonra sesi oluşturan aktivatör, vibratör ve artikulatör organlardan vibratör organın kaybı söz konusudur (Erişir ve İnci 2001). Ameliyattan sonra hastanın sesine kavuşması için bir yol aranır. Günümüzde bu yollardan en çok tercih edilenleri, özefagus (yemek borusu) konuşması, elektrolarenks (elektronik gırtlak) yardımıyla konuşma ve cerrahi ses restorasyonudur.

Yemek borusu konuşması, ses telleri olmayan hasta için (alaryngeal hasta) en eski terapi yöntemidir. Bu yöntemde hava farengoözefegal segmente alınır ve birden serbest bırakılarak farengoözefegal kısmın titreşimi sağlanır (Barkmeier ve ark. 2003). Doğal bir yöntem olduğundan ilk tercih edilenidir fakat eğitim süreci ve yeteneğe dayanması nedeniyle başarısızlık oranı yüksektir.

Elektronik gırtlak ise avuç içi büyüklüğünde, boyuna veya ağzın içine yerleştirilerek kullanılan bir cihazdır. Konuşma için gereken temel titreşimi sağlamak amacıyla kullanılır. Kullanım sırasında bir eli meşgul etmesi, mekanik bir ses üretmesi, pile bağımlılık oluşturması ve pahalı olması nedeniyle tercih edilmemektedir (Aksoy ve ark. 2010). Fakat uzun cümlelerde çok daha az çabaya ihtiyaç duyması ve çoğu durumda daha efektif olması da avantajları arasındadır (Liu ve ark. 2006).

Cerrahi ses rehabilitasyonunda en çok kullanılan yöntem trakeözefagal zımbadır (tracheoesophageal puncture). 1979 yılında Blom ve Singer'in bir valf kullanarak yeni bir trakeözefegal ameliyat tekniği geliştirdiği teknikte protez, trakeden gelen akciğer havasını özefagusa şantlar ve konuşma için gerekli sesin oluşmasını sağlar (Barkmeier ve ark. 2003). Prosedür, soluk borusu üzerinde bir delik açmak ve yemek borusuyla trake arasında fistül oluşturmak şeklindedir. Konuşmak istendiğinde boğazdaki stoma kapatılarak havanın trakeyle özefagus arasına yerleştirilen ses protezinden yemek borusuna geçmesi buradan da ağza ilerlemesi sağlanır. Özefagus konuşmasındaki gibi özefagusun yukarısındaki boğazın aşağısındaki dokuların (pharyngoesophageal segment, neoglottis veya pseudoglottis) titreşimiyle konuşma için gerekli titreşim sağlanır (Sharifzadeh ve ark. 2009).

Nefes alma sırasında stoma kapatıldığında hastanın konuşma için iyi bir ses elde ettiği rapor edilmiştir. Bunun yanında trake içine secretion sızıntısı sorunu, istenmeyen fistula problemleri ve trake içine solunumun devamı yan etkiler olarak bildirilmiştir (Barkmeier ve ark. 2003).

Kullanım açısından bakıldığında gelişen protezler nedeniyle en çok kullanılan yöntem ses protezleridir. Ameliyat sonrası değerlendirmesinde özefegal konuşma %14-76 arasında değişen bir başarıya sahipken (Anadolu ve ark. 1998), ses protezlerinin başarımı daha yüksektir. Aksoy ve arkadaşları SB Haseki Eğitim ve Araştırma Hastanesi'nde 2006-2010 yılları arasındaki ses protezi kullandıkları 58 hastada 51 akıcı anlaşılabilir konuşma elde etmişlerdir (%87.9) (Aksoy ve ark. 2010). Erişir ve arkadaşları Cerrahpaşa Tıp Fakültesi'nde 1993-2000 yılları

arasında ses protezlerinde takibi mümkün 29 hastada 21 başarı (%72) sağlamışlardır (Erişir ve İnci 2001). Anadolu ve arkadaşlarının Ankara Üniversitesi'nde 1994-1997 yılları arasında gerçekleştirdikleri ses protezi kullanımında başarı, 24 hastadan 21'inde (%87.5) sağlanmıştır.

2000 yılı itibariyle dünya üzerinde 600000 total larenjektomi geçirmiş hasta bulunduğu tahmin edilmiştir ve bunların 3'te 1'inin trakeözefagal veya özefagus konuşması yerine elektronik gırtlak konuşmasını anatomik veya kişisel nedenlerden tercih ettiği düşünülmektedir (Niu ve ark. 2003). Literatürde elektrolarenks konuşmasının daha kolay ve efektif olduğu durumların da olduğu belirtilmiştir ve ayrıca trakeostami gibi geçici durumlarda da elektrolarenks, konuşmayı sağlamak için faydalı bir araç olduğundan bahsedilmiştir (Niu ve ark. 2003).

Total larenjektomi sonrası hangi ses restorasyonu kullanılırsa kullanılsın, bu ses hastanın eski sesinden ve normal insan sesinden uzaktır. Sayısal işaret işlemenin yaygınlaşmasıyla bu konuşmaların iyileştirilmesi çalışılan bir problem haline gelmiştir. Bu konuşmalardan trakeözefagal ve özefagal konuşmanın normal konuşmadan farkı, daha düşük ve sürekli değişiklik gösteren temel frekans, düşük şiddet, formant frekanslarında ses yolunun kısılmasından dolayı yüksek frekansa doğru öteleme şeklinde sıralanabilir (Cole ve ark. 1997).

Genelde çalışmalar bu tipte konuşmalara odaklanmıştır. Bunları iyileştirmek için kullanılan yaklaşım analiz-sentez yöntemlerinden doğrusal öngörüm kodlamasıyla elde edilen formantın simüle edilmiş glottal dalga yapısıyla uyarılması şeklindedir. Örnek çalışmalar için (Türkmen 2008) ve (Tarakçıoğlu 2010) incelenebilir. Bu yolla iyileştirmenin formant frekans uyarımından daha iyi sonuçlar verdiği gözlenmiştir (Cole ve ark. 1997).

Elektronik gırtlak konuşmasının normal konuşmaya göre temel farklılığı uyarımın mekanik olarak üretilmesinden dolayı değişmeyen temel frekansa sahip olmasıdır. Larenjektomi sonrası sesi yeniden kazanma metotlarıyla elde edilen konuşmalardan da bu temel özelliğiyle ayrılır. Bu özelliği yüzünden ötümsüz fonemleri oluştururken ihtiyaç duyulan durakları oluşturmak, çok kısa aralıklar için uyarımın kesilmesi anlamına geldiğinden bu cihazla bunu gerçekleştirmek şimdilik mümkün değildir. Ayrıca normal konuşmada duygu aktarımı ve vurgu için temel frekans değişimini de bu cihazla gerçekleştiremeyiz (Niu ve ark. 2003). Araştırmalar cihazın mekanik sesinin bastırılarak normal konuşmaya yaklaştırma üzerine odaklanmıştır.

İlk araştırmalar yükseltmenin etkisini incelemişlerdir. Ağıza yakın yerleştirilen mikrofonla denemeler yapılmış, sesteki mekanikliğe cihazdan yayılan sesin neden olup olmadığı araştırılmıştır (Cole ve ark. 1997). Bunun belirlenmesi üzerine Norton ve Bernstein'in mekanik sesi bastırmak için 1993 yılında elektronik gırtlığın köpükle yalıtımı üzerine bir iyileştirme denemiştir fakat cihazı tutmanın zorlaşması, köpüğün, elle tumanın sağladığı sönüm mekanizmasını ortadan kaldırması ve cihaz kullanımını zorlaştırması gibi nedenlerle başarısız olmuştur.

Cihazdan yayılan girişimin işaret seviyesinde bastırılmasında ise temel olarak iki yol

izlenmiştir. Bunlar çok algılayıcı ve tek algılayıcının kullanıldığı yöntemlerdir. İki durumda da girişimin doğru tespiti önem kazanmaktadır. Çok algılayıcı yöntemlerde girişimin ana bileşen olduğu algılayıcı sayesinde gürültü bileşeni elde edilerek, gürültünün pasif olduğu algılayıcı işaretinden çıkarılması şeklindedir. Tek algılayıcı yöntemler ise sessiz kalınan durumlarda gürültü kestirimi yapılmaya çalışılır.

Girişimin miktarı cihazdan cihaza farklılık gösterdiği gibi hastadan hastaya da değişiklik göstermektedir. Bu farklılığın aynı cihazla, ağız kapalı olarak 30 hastayla yapılan ölçümlerde 7-26dB arasında değiştiği gözlenmiştir(Niu ve ark. 2003). Bu farklılıklar cihazın boyna nasıl kuplajlandığı, titreşim karakteristiği, boynun yayılım karakteristiği, kullanıcının cihazı kullandaki yetkinliği ile ilişkilendirilmiştir(Niu ve ark. 2003).

Tek algılayıcının bulunduğu durumlarda tespit edilen gürültü, işareten çeşitli yaklaşımlarla çıkarılmaya çalışılır. Bu çalışmada ise farklı gürültü çıkarım yöntemlerinin doğal ses elde etmedeki yetkinlikleri değerlendirilmiştir.

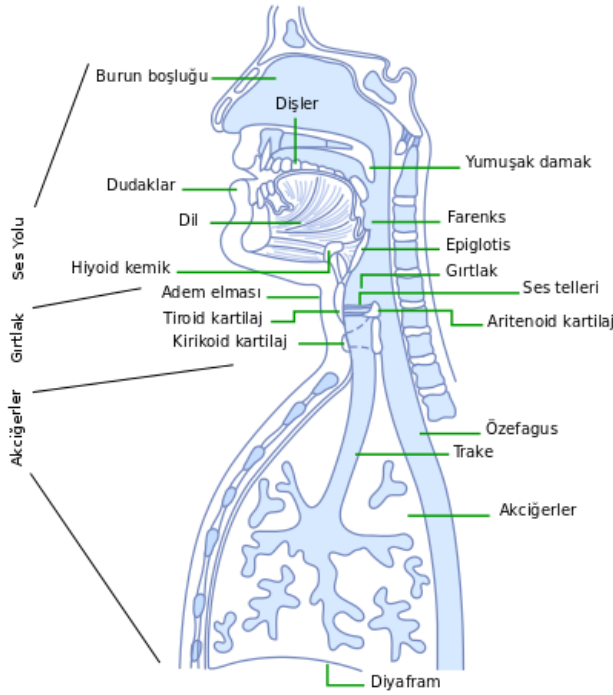
## 2 SES ve KONUŞMANIN OLUŞUMU

Konuşma ve duyma işlevsel olarak farklı gibi görünse de konuşma, duyma geribeslemesi olmaksızın doğru bir şekilde oluşturulamaz. Konuşmanın öğrenilmesi aşamasında duyduğumuz seslere konuşma üretiminde görev alan organlarımızın kullanımını şekillendirerek benzer sesler çıkarmaya çalışıyoruz. Çevremizdekilerden aldığımız geribesleme ile sözcükleri öğrenmeye ve bilgi aktarımına başlıyoruz (Roy 2011).

Konuşma üretiminde aktif rol oynayan organlar buldukları konumlara göre aşağıdan yukarıya akciğerler, gırtlak ve ses üretim yoludur.

### 2.1 Ses Üretim Mekanizmasında Görev Alan Organlar

Vücudun diyafram ile burun arasındaki kısmı ses üretmede görev alan organları barındırır.

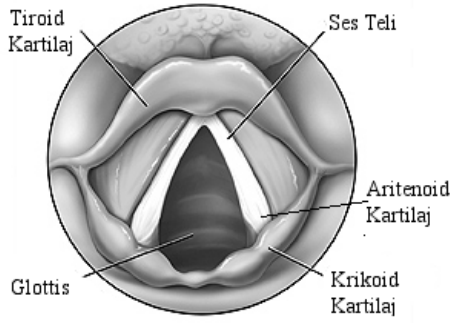


Şekil 2.1: Konuşma üretiminde aktif rol oynayan organlar

#### 2.1.1 Akciğerler

Ses üretiminde görev alan organlarının en büyüğü akciğerlerdir ve Şekil 2.1'da üst vücut bölgesinin sagittal kesitinin en altındaki hacmi oluşturan kısımdır. Ses üretim mekanizmasında hava depolanmasını sağlayarak titreşecek havayı depolama görevi görürler ve bir bakıma kaynak olarak düşünülebilirler. Hava depolama ve salınım hızı *diyafram* yardımıyla kontrol edilir. Trake yoluyla akciğerlere giren ve depolanan hava ses üretmek istendiğinde gene trake yoluyla





Şekil 2.2: Gırtlığın transvers kesitten görünüşü

kontrollü olarak gırtlığa ulaştırılır.

### 2.1.2 Gırtlak

Gırtlak, kas, bağ doku ve kıkırdaktan oluşan bir yapıdır. Ses telleri olarak adlandırılan yapıyı da barındıran organdır. Ses telleri, gırtlığın arkasından önüne doğru ulaşan iki bağ dokudan ibarettir. Ses telleri arasındaki boşluğa *glottis* adı verilir. Ses tellerinin arka tarafı aritenoide (aritenoid kırkıdağa) bağlıdır, aritenoid kırkıdaklar krikoid kırkıdak üzerinde yanıl hareket yaparak kontrollü titreşimi sağlarlar.

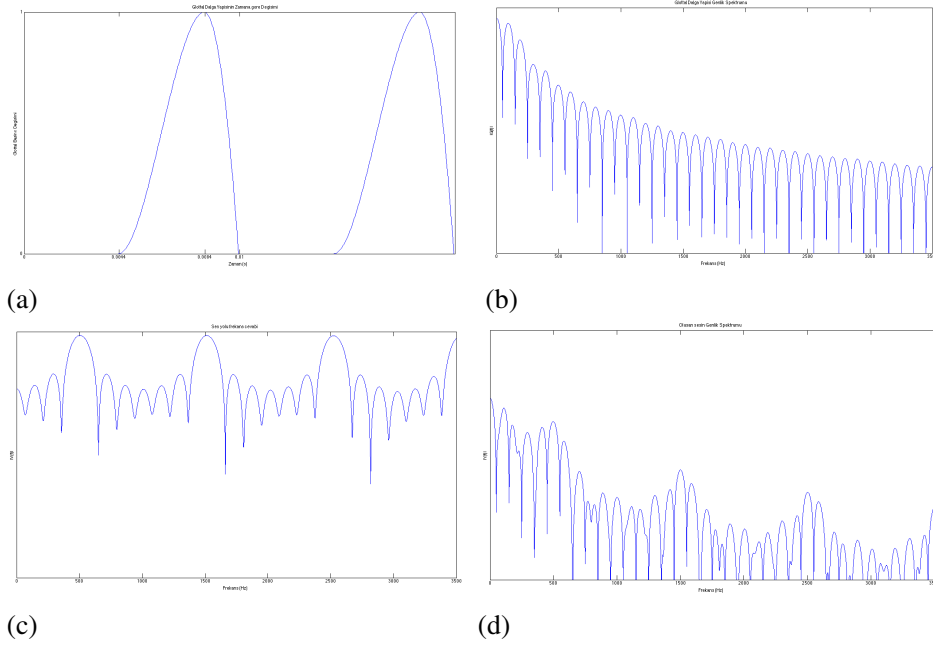
Ses tellerinin üç farklı temel durumda olduğu düşünülebilir. Bunlar nefes almadaki, sesli (ötümlü) ve sessiz (ötümsüz) sesleri çıkartmadaki durumlar olarak düşünülebilir.

Nefes alma durumunda ses telleri ayırık, dolayısıyla glottis, havanın rahatça girip çıkacağı şekilde açık durumdadır.

Sesli (ötümlü) durumda aritenoid kırkıdakları birbirine yaklaşarak havanın geçmesine engel olacak şekilde ses tellerinin kapanması sağlar. Ses tellerinin altındaki hava basıncındaki artış bir noktaya ulaşınca açılarak basıncın bir anda düşmesine neden olurlar. Basıncı düştüğünde yeniden kapanabilir duruma gelirler ve bu bir döngü halinde devam eder. Bu açılıp kapanma hareketinin oluşturduğu basınç değişimi duyulabilir sesin oluşmasını sağlar ve açılıp kapanma süresi sesin perde periyodu (pitch periyodu) olarak adlandırılır ve bu periyodun bir bölümü temel frekans olarak adlandırılır. Erkeklerin temel frekansı, daha uzun ve daha ağır ses tellerine sahip oldukları için kadın ve çocuklarınkinden düşüktür.

Ötümlü ses üretiminde glottal dalga yapısı benzetimlerinin doğala yakınlığını ortalama fikir puanlamasıyla değerlendiren bir çalışmada en yüksek puanı alan uyarım ifadesi (2.1) denkleminde gösterilmiştir (Rosenberg 1971).

$$g(n) = \begin{cases} 3 \left(\frac{n}{T_0}\right)^2 - 2 \left(\frac{n}{T_0}\right)^3 & 0 \leq n \leq T_0 \\ 1 - \left(\frac{n-T_0}{T_N}\right)^2 & T_0 < n \leq T_0 + T_N \end{cases} \quad (2.1)$$



Şekil 2.3: Glottal Hava Basıncı Değişiminin (a) zamana göre değişimi ve (b) frekans uzayındaki genlik bileşenleri. Ses yolunun rezonans karakteristiği (c) ve ses işaretinin tüm değişimlerden sonraki halinin genlik spektrumu(d).

Bu ifadeye  $T_0$  basıncın artış süresi,  $T_N$  basıncın azalış süresini belirtmektedir. Bunların optimum değeri  $T_0 = 0.4T$  ve  $T_N = 0.16T$  olarak bulunmuştur (Loizou 2007). Bu ifadelerdeki  $T$ , perde periyoduna karşılık gelmektedir.

100Hz'lik temel frekansa sahip glottal dalga benzetiminin zamana göre aldığı değerler Şekil 2.3a'da gösterilmiştir.

### 2.1.3 Ses Yolu

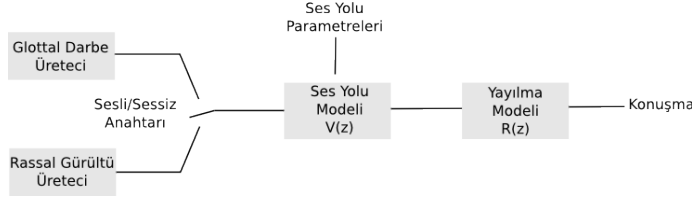
Ses yolu, gırtlakla dudak arasındaki ağız boşluğunu ve *velum*, yumuşak damak, ile başlar, burun boşluğunu da kapsyarak sonlanır. Ağız boşluğu dil, diş, dudak ve çene gibi *artikülatör*leri kapsar. Ağız boşluğunun erkeklerdeki ortalama uzunluğu 17cm'dir.

Ünlü harfleri seslendirirken gırtlak ötümlü konumda ses yolu ise en engelsiz durumdadır. [a] sesi üretilirken ses yolu bir ucu kapalı bir boru gibi davranır.

Bir ucu kapalı boru ses için bir rezonatör görevi görür. Bu yapı için rezonans frekansları denklem 2.2 ile verilmiştir.

$$f_n = \frac{nv}{4L} \quad n = 1, 3, 5, \dots \quad (2.2)$$

Burada  $n$ , mod,  $v$  dalga yayılım hızı,  $L$  boru uzunluğunu göstermektedir.



Şekil 2.4: Ses üretimi sistem modeli

Bu rezonatöre genişbant bir ses işareti uygulandığında, mod frekanslarındaki dışındaki frekansları bastırma, mod frekanslarında ise birlikte salınma haline geçtiğinden bir filtre gibi davranır. Örnek olarak ortalama erkek ses yolunu ele aldığımızda temel rezonans frekansı  $f_1 = 34320/(4 * 17) = 504.706 Hz$ 'e eşit olur.

## 2.2 Sesin ve Konuşmanın Üretimi

Dildeki sesler, ses oluşturma birimlerinin (artikülatörler) farklı farklı yapılandırılmalarıyla oluştururlar. Bir sesi oluştururken akciğerlerde depolanmış hava trakeden gırtlığa kontrollü olarak salınır. Oluşturulacak sesin ötümlü veya ötümsüz olmasına göre ses telleri konumlandırılır. Ötümlü ses üretileceği zaman havanın geçişine izin vermeyen ses telleri, biriken hava basıncıyla açılır ve ardından hemen kapanır. Şekil 2.3a'de ötümlü ses üretimindeki glottis basınç değişiminin bir benzetimi gösterilmiştir.

Oluşan bu basınç değişimi, ses yolunda şekillendirilir. Velum ile bu basınç değişimi ağıza ve buruna değişen miktarlarda yönlendirilir. Ağız boşluğuna gelen ses burada dil, damak, dişler ve dudaklar ile son şeklini alır ve burun bileşeni de varsa ses yolu dışında birleşerek son sesi oluşturur. Yani ses şekillendiricileri sese son halini verir.

Bu işlem kabaca Şekil 2.4'deki sistem ile modellenilebilir. Ses üretiminde uyarıcı kaynak, gırtlak, temelde iki durumda bulunabilir; ötümlü ve ötümsüz ses üretme durumları. Daha sonra kaynaktan yayılıma kadar ses yolu devreye girer. Buradaki organların farklı şekillenişleri sesin oluştuğu noktayı belirler. Son olarak da yayılma devreye girer.

Türkçe konuşurken kullanılan sesleri sesbilimciler, ötüme, ağızın durumuna ve seslerin oluştukları noktalara göre Çizelge 2.1'de özetlemişlerdir (Ergenç 2002).

Örnek olarak [ɑ] sesini ele alalım. Bu sesi oluşturmak için Çizelge 2.1'de bu ötümlü (ünlü) ses geniş, düz, artdamaksıl (postdorsal) olarak sınıflandırılmıştır. Bu ses yolunun en engelsiz durumuna karşılık geldiğinden ses yolunun bir ucu kapalı bir boruya en benzediği hale denk düşer. Ünlü seslerden  $a$  üretilirken Şekil 2.4'deki sesli/sessiz anahtarı ötümlü kaynak dalga formunu iletecek konumdadır. Ses yolu ve yayılma blokları bir ucu kapalı bir borununkine benzer bir dönüştürme sağlarlar. Ses işaretinin şekillenışı Şekil 2.3'de a, b, c, d ile gösterilmiştir.

## 2.3 Seslerin İşaretsel Ayrışması

Daha önceden de bahsedildiği gibi konuşma üretme-konuşma algılama eğitimi doğrudan sonra başlıyor. Üretilen sesin aileden alınan geri beslemeye uyumu zaman içinde artikülasyonların yapılandırılmalarının doğru şekilde değişiminin öğrenilmesiyle artırılıyor. Burada referans işaretine uyumun belli ölçütleri sesleri ayırt ederken de kullanıyoruz. Bu ipuçları artikülasyon konumlarının ses işaretine yansımalarıdır.

Çizelge 2.1’de verilen Türkçe’nin ses abecesi çizelgesinin orta sütunu gırtlak ve ses yolundaki organların durumunu ve sesin oluştuğu noktayı belirtmektedir. Bu sesleri oluşturmak üzere gırtlak zamanla değişen farklı işaretler üreten bir kaynak, ses yolu ise frekans cevabı zamanla değişen bir kavite olarak düşünülebilir. Bu noktada problem ana ses sınıflarının birbirinden ayrılması değil, aynı sınıftaki seslerin ayrıştırılması olduğundan daha ince ayrıntılar kullanılmalıdır.

Geniş, düz, artdamaksıl ünlü ses [a]’nın kaynak-filtre modeliyle üretilmesi, Şekil 2.3’de gösterilmiştir. Dönüşümde görülebileceği gibi çıkış işaretinin genlik spektrumunun zarfı ses yolunun durumuyla büyük ölçüde etkilenmiştir. Frekans spektrumunun zarfını, Doğrusal Öngörüm Kodlamasıyla elde edebileceğimizden ses yolunun durumu hakkında bu yolla fikir elde edilebilir.

Ses yolu ile belirlenen rezonans tepeleri DC’den yüksek frekansa doğru numaralandırılır.  $F_1$ ,  $F_2$ ,  $F_3$ ’e sırasıyla birinci, ikinci, üçüncü formant frekansı adı verilir. İlk formant frekansı  $F_1$ , ağzın açıklığıyla ilgilidir. Daha yüksek frekansta oluşan  $F_1$ , ağız açıklığının daha fazla olduğunu gösterir.  $F_2$  ise ağız boşluğuyla ilgilidir, dil konumu, dudak aktivitesi ile ilgili de fikir verir.  $F_3$ , ağız boşluğundaki engellenmenin konumu hakkında fikir verir (Loizou 2007).

Seslilerin ayrıştırılmasında ses birimlerinin  $F_1$ - $F_2$  değerlerinin ve sürelerinin yeterli ayrıştırma sağladığı 1950’li yıllarda Peterson ve Barney tarafından yapılan çalışmadan 40 yıl sonra yapılan çalışma bunun yanında spektral değişimin de incelenmesinin gerektiğini ortaya koymuştur (Loizou 2007). Aradaki farkın denek grubunun ağız-şive farklılıklarından kaynaklanmasının muhtemel olduğu belirtilmiştir.

Geniz sesleri (nazallar), genel olarak yüksek formantların bastırılması ve düşük frekans rezonansının varlığıyla ayrışır. Önce ve sonrasındaki formant geçişleri de ek bilgi sağlar.

Patlamalı sesler, hava geçişinin kapatılmasıyla oluşturulduklarından ses işaretinde kısmen veya tamamen bir sessizlik anı vardır. Bu ipucu gürbüzdür ve diğer sınıflarda bulunmayan bir özellik taşır. Hava geçişinin engellenmesi sonunda ortaya patlamalı bir ses çıkar. Bu da önemli bir ipucudur.

Patlamalı sınıf, ötümlü ve ötümsüz olarak alt sınıflara ayrılır. Bunların işarete yansımaları,

Çizelge 2.1: Türkçe'nin Sesbilim Abecesi (Ergenç 2002)

Ünlüler		
[a]	geniş, düz, artdamaksıl (postdorsal): <i>anı</i>	[a'nī]
[a]	geniş, düz, öndamaksıl (predorsal): <i>laf</i>	[ˈlaf]
[e]	geniş, düz, öndil (kapalı): <i>elma</i>	[el'ma]
[ɛ]	geniş, düz, öndil (açık): <i>dere</i>	[de'rɛ]
[i]	dar, düz, öndil (açık): <i>iğde</i>	[i:'dɛ]
[ɪ]	dar, düz, öndil (kapalı): <i>simit</i>	[sɪ'mɪt]
[i]	dar, düz, arkadil (ortadil): <i>ısı</i>	[i'si]
[ɔ]	geniş, yuvarlak, arkadil (açık): <i>soru</i>	[sɔ'ru]
[o]	geniş, yuvarlak, arkadil (kapalı): <i>oğlak</i>	[o:'ɫak]
[ʊ]	dar, yuvarlak, arkadil (açık): <i>kulak</i>	[kulak]
[u]	dar, yuvarlak, arkadil (kapalı): <i>uğur</i>	[u:ɣ]
[œ]	geniş, yuvarlak, öndil (açık): <i>örtü</i>	[œr'ty]
[ø]	geniş, yuvarlak, öndil (kapalı): <i>öğren</i>	[ø:'rɛn]
[y]	dar, yuvarlak, öndil (açık): <i>ümit</i>	[y'mɪt]
[y]	dar, yuvarlak, öndil (kapalı): <i>düğme</i>	[dy:'mɛ]
Ünsüzler		
[b]	ötümlü, patlamalı, çift dudak: <i>balık</i>	[balɪk]
[d]	ötümlü, patlamalı, dil ucu-diş ardı: <i>dede</i>	[de'dɛ]
[g]	ötümlü, patlamalı, dil-artdamak: <i>karga</i>	[kar'ga]
[ʃ]	ötümlü, patlamalı, dil-artdamak (ön): <i>genç</i>	[ʃɛntʃ]
[h]	ötümsüz, sızıcı, gırtlak: <i>hasta</i>	[hasta]
[ʒ]	ötümlü, sızıcı, dil-öndamak: <i>müjde</i>	[myʒ'dɛ]
[k]	ötümsüz, patlamalı artdamak: <i>akıl</i>	[a'kiɫ]
[c]	ötümsüz, patlamalı, artdamak (ön): <i>keci</i>	[ce'dɪ]
[l]	ötümlü, yan dar., dilucu-ön damak: <i>lale</i>	[la:'lɛ]
[ɫ]	ötümlü, yan dar., dilucu-ön damak: <i>kul</i>	[ˈkuɫ]
[m]	ötümlü, genizsi, çift dudak: <i>dam</i>	[ˈdam]
[n]	ötümlü, genizsi, dilucu-dişeti: <i>anı</i>	[a'nī]
[ŋ]	ötümlü, genizsi, dil-artdamak: <i>süngü</i>	[syŋ'ɣy]
[p]	ötümsüz, patlamalı, çift dudak: <i>ip</i>	[ˈɪp]
[r]	ötümlü, çok vuruşlu, dilucu-dişeti: <i>raf</i>	[ˈraf]
[r]	ötümlü, tek vuruşlu, dilucu-dişeti: <i>ırmak</i>	[ir'mak]
[β]	ötümsüz, sızıcı: <i>bir</i>	[ˈbɪɣ]
[s]	ötümsüz, sızıcı, dilucu-dişeti: <i>ses</i>	[ˈsɛs]
[ʃ]	ötümsüz, sızıcı, dil-öndamak: <i>aşı</i>	[a'ʃi]
[t]	ötümsüz, patlamalı, dilucu-diş ardı: <i>ütü</i>	[y'ty]
[v]	ötümlü, sızıcı, dudak-diş: <i>var</i>	[ˈvaɫ]
[v]	(ötümlü, sızıcı), yarı ünlü: <i>tavuk</i>	[ta'vʊk]
[j]	(ötümlü, dil-öndamak), yarı ünlü: <i>yat</i>	[ˈjat]
[z]	ötümlü, sızıcı, dilucu-dişeti: <i>azık</i>	[a'zik]
[z]	ötümsüz, sızıcı, dilucu-dişeti: <i>yoza</i>	[ˈjɔz]
[ʒ]	ötümlü, sızıcı, dil-öndamak: <i>cam</i>	[ˈʒam]
[ʃ]	ötümsüz, sızıcı, dil-öndamak: <i>seçim</i>	[se'tʃɪm]

kapalılık periyodunda temel frekansta,  $F_0$ , cılız bir işaretin varlığı veya yokluğu şeklindedir. Bu da kapalılık periyodu sonrasındaki patlamanın şiddetini etkiler. Ses Başlama Zamanı (Voice Onset Time) da bu sınıflara göre farklılık gösterir. Sınıf içi farklılaşma patlamalı seslerin son şekillendiği noktalarda belirginleşir. Bunun dudak, damak ve diş olmasına göre sinyaldeki izdüşümü  $F_2$  frekansının farklılaşmasıdır.

Sızıcı sesler, diğer sınıflardan daha uzun süren periyodik olmayan işaretlerle kendilerini gösterirler. Sınıf içi farklılıklar gürültü şeklindeki işaretin yoğunluk farkı ve spektrumun şeklidir. Bir kısmında yüksek frekanslı spektral tepeler oluşurken diğer kısım daha düşük frekanslı düz tepeli spektrum oluşturur. Son bir upucu olarak öncesinde ve sonrasında yaşanan  $F_2$  ve  $F_3$  geçişleri kullanılabilir.

### 3 SORUNUN TANIMI ve ÇÖZÜM ÖNERİSİ

Elektronik gırtlak ve kullanımı Şekil 3.1’de gösterilmiştir (Griffin Labs 2008). Avuç içine sığan büyüklüğe sahip cihaz içerisindeki servomotorun cihazın tepesindeki zarı titretilmesiyle işlevini yerine getirir. Cihazın boyuna veya ağza kuplajlanarak, yanındaki düğmeyle uyarım kontrol edilir.

Bu yolla üretilen konuşma, uyarımın mekanik olarak üretilmesinden dolayı değişmeyen temel frekansa sahip olmasıdır. Gerçi son yıllarda bu soruna ayarlanabilir düğmeyle konuşma sırasında değişen perde periyoduna sahip uyarım yapabilen modellerle kısmen çözüm getirilmiştir. Fakat mekanik tını sorunu devam etmektedir. Diğer bir problem de ötümsüz ses birimlerini oluştururken ihtiyaç duyulan durak ve yüksek frekanslı gürültü benzeri uyarımın oluşturulamamasıdır.



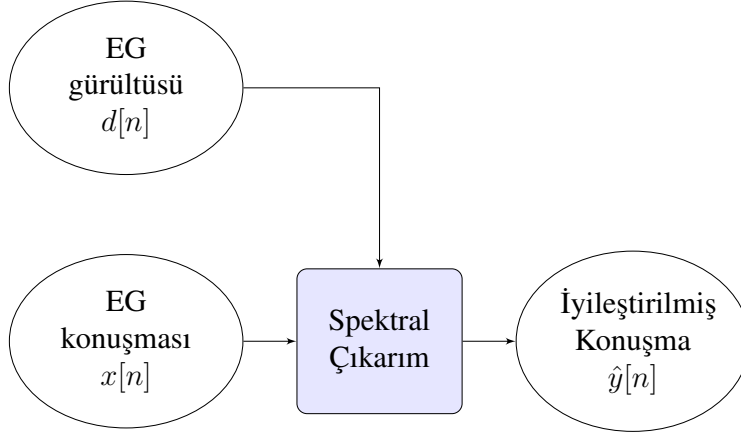
Şekil 3.1: Elektronik Gırtlak ve Kullanım Şekilleri (Griffin Labs 2008)

Konuşmaya bu mekanikliği kazandıranın büyük bir bileşenin cihazdan çevreye yayılan mekanik ses olduğu düşünülmektedir (Meltzner 2003).

#### 3.1 Çözüm Önerisi

Bu çalışmada cihazdan yayılan sesin bastırılmasına odaklanılacağından ve tek mikrofonlu bir sistem kullanılacağından tasarlanan sistem kabaca Şekil 3.2’de gösterilmiştir.

Saf gürültüyü barındıran anlar işaret işlenmeden işaretlendi. Spektral çıkarım bloğu gürültü olarak işaretlenmiş anlara geldiğinde gürültü kestirimi yaparken gürültü olarak işaretlenmeyen anlarda kestirilmiş gürültünün işaretten belirli spektral çıkarım algoritmalarına uygun çıkarımı yapıldı. Bu çalışmada karşılaştırılmak üzere altı farklı yöntem kullanılmış ve bunların



Şekil 3.2: Kullanılan ses iyileştirme sistemi özeti

başarımı değerlendirilmiştir. Bu yöntemler;

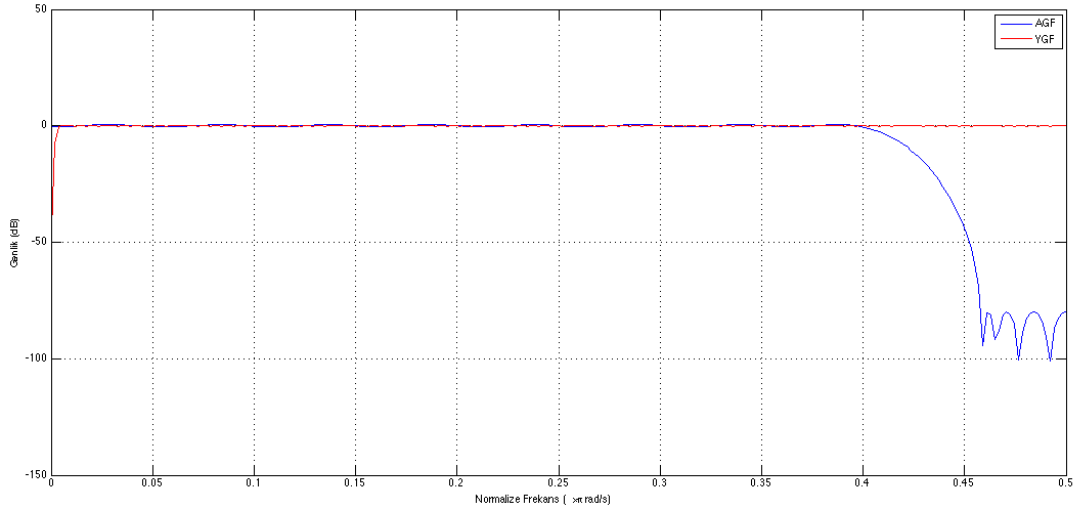
1. Genlik spektral çıkarma
2. Güç spektral çıkarma
3. Aşırı çıkarım kullanarak spektral çıkarma
4. Doğrusal olmayan spektral çıkarma
5. Çoklu bant spektral çıkarma
6. Algılama özelliklerini kullanarak spektral çıkarma

### 3.2 Veri Toplama

Yöntemlerin başarımı hakkında fikir edinmek için denekler sağlıklı bireyler arasından seçildi ve kısıtlı zamanda elektronik gırtlığa alışmaları beklendi. 20-30 yaş aralığındaki bir bayan dört erkek denekten bayan denek elektronik gırtlığa uyum sağlayamadı. Geriye kalan deneklerden ikisi, Çizelge 2.1'deki sesler için örneklerin kaydedilmesinde kullanıldı. Bu çalışmada 27 ve 29 yaşlarında iki sağlıklı erkek denekten, "bize reality show izlettireceğiz diye ortalığı salhaneye çevirmeye mezbahaya döndürmeye hiç niyetimiz yok" (Tarakçıoğlu 2010) söz dizimi elektronik gırtlakla tekrarlatıldı. Cümle yerine bu söz diziminin kullanılması anlaşılabilirliğin değerlendirilmesinde objektif kalınma ihtimalini artırma amacıyla.

Kayıt ortamı olarak boş ve sessiz büyük bir oda kullanıldı. Kayıt, M-Audio Fast Track harici ses kartına bağlı Samson C01 Studio Condenser mikrofonla denekten yaklaşık 20 santimetre uzakta sabitlenerek alındı. Audacity ses kayıt ve düzenleme programı kullanılarak yapılan kayıtlarda örnekleme hızı olarak 48000 Hz seçildi.





Şekil 3.3: Önışlemede kullanılan alçak geçiren ve yüksek geçiren filtrelerin frekans yanıtı

Ses işleme i5 işlemcili bir bilgisayar üzerinde çalıştırılan Matlab 2013a programında gerçekleştirildi.

### 3.3 Önışleme

Kayıt edilen işaret 50Hz'lik yüksek geçiren filtreden geçirildikten sonra 9.6 KHz'lik alçak geçiren filtreden geçirilmiştir. Kullanılan filtrelerin frekans cevabı Şekil 3.3'de gösterilmiştir. Filtrelenmiş işaret 30ms'lik %75 örtüşen pencerelere bölündü. Pencereleme fonksiyonu olarak Hamming kullanıldı.

#### 3.3.1 Gürültü Kestirimi

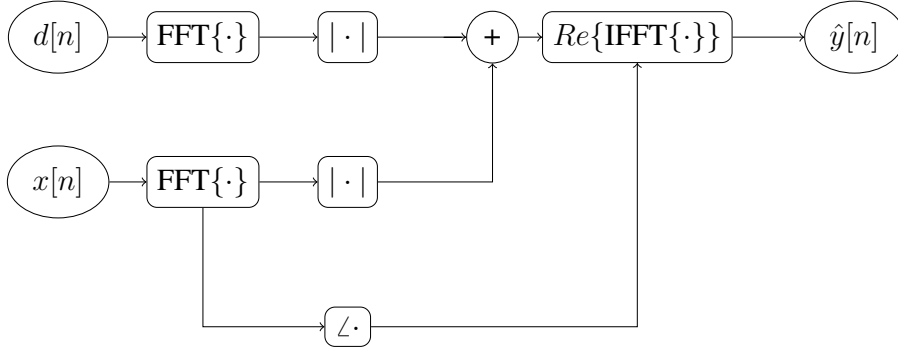
Gürültü kestirimi eğer  $i$ . çerçeve gürültüyse

$$\hat{D}[\omega] = \frac{\hat{D}[\omega] + D_i[\omega]}{2} \quad (3.1)$$

şeklinde güncellenerek gerçekleştirilmiştir.

### 3.4 Spektral Gürültü Çıkarma Algoritmaları

Gürültünün frekans bileşenlerini gürültülü işaretin frekans bileşenlerinden çıkarıp zaman uzayına geri dönüşümü amaçlayan sistem Şekil 3.2'de gösterilmiştir. Gürültü, cümleler arasındaki boşlukların zamanlarının belirlenmesi ve algoritmaya verilmesiyle işaretlenmiştir. İşlem sırasında ilgili anlarda gürültü kestirimi bu anlarda tekrarlanmıştır.



Şekil 3.4: Spektral Genlik Çıkarım Algoritması Blok Diyagramı

### 3.4.1 Spektral Genlik Çıkarması

Bu yöntemde yapılan, frekans uzayında işaretin genliğinden öngörülen gürültünün genliğinin bütün pencerelerde çıkarılmasıdır. Bu çıkarma işlemi sonucunda hatalı gürültü kestirimi nedeniyle negatife düşen spektral genlik bileşenleri sıfıra ötelenir. Yöntemin matematiksel ifadesi (3.2)'de verilmiştir. Yöntemin blok şeması Şekil 3.4'de verilmiştir.

$$|\hat{Y}[\omega]| = \begin{cases} |X[\omega]| - |\hat{D}[\omega]| & \text{eğer } |X[\omega]| > |\hat{D}[\omega]| \\ 0 & \text{diğer} \end{cases} \quad (3.2)$$

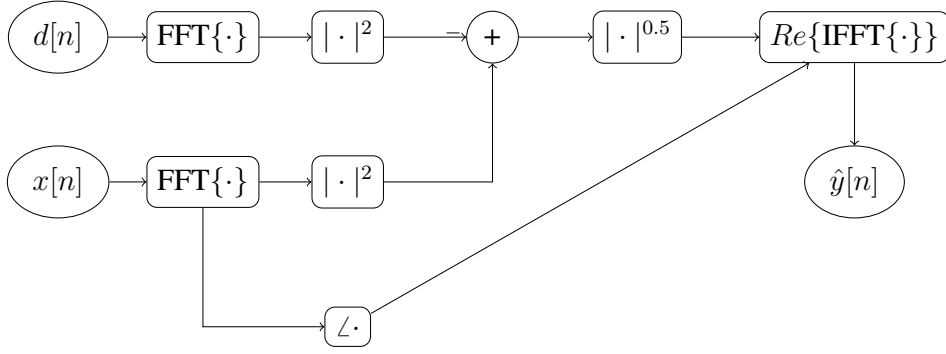
Elde edilen iyileştirilmiş genlik spektrumundan zaman işaretine geçmek için dönüşüm yaparken, gürültülü işaretin faz bilgisi eklenerek dönüşüm yapılarak iyileştirilmiş ses işareti elde edilir.

### 3.4.2 Spektral Güç Çıkarma

Bu yöntemde kestirilen gürültünün spektral gücü, işaretin spektral gücünden çıkarılarak ses iyileştirme yapılır. Elde edilen spektral gücün kareköküne gürültülü işaretin fazı eklenerek zaman uzayı bileşenleri elde edilir. Yöntemin matematiksel ifadesi (3.3)'de verilmiştir. Yöntemin blok şeması Şekil 3.5'de verilmiştir.

$$|\hat{Y}[\omega]|^2 = \begin{cases} |X[\omega]|^2 - |\hat{D}[\omega]|^2 & \text{eğer } |X[\omega]|^2 > |\hat{D}[\omega]|^2 \\ 0 & \text{diğer} \end{cases} \quad (3.3)$$

Bu yöntemde de spektral güç çıkarma adımında doğrultucu uygulanmıştır.



Şekil 3.5: Spektral Güç Çıkarım Algoritması Blok Diyagramı

### 3.4.3 Aşırı Çıkarma Kullanılan Spektral Çıkarma

Genlik ve güç spektral çıkarma yöntemleri, elde edilen spektrumla ilgilenmez. Kestirilmiş gürültünün spektral genliğinin/gücünün konuşma işaretinin genliğinden/gücünden çıkarılması söz konusudur. Sadece negatif genlik değerinin oluşmaması için doğrultucu uygulanmıştır.

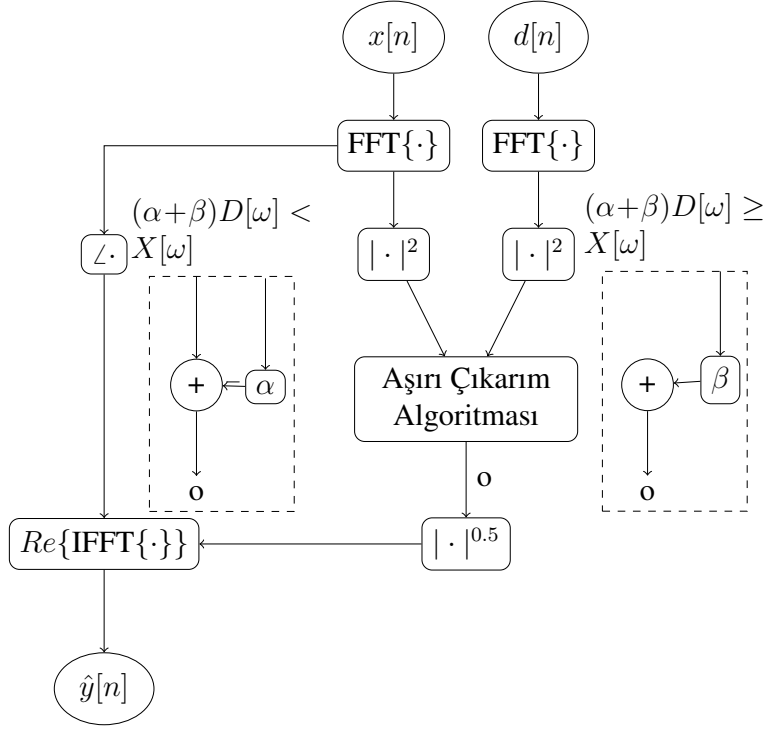
Gürültü kestirimindeki hatalar nedeniyle doğrultma işlemi genlik spektrumunda yalıtılmış tepeler oluşmasına neden olur. Belirli frekanslardaki yalıtılmış tepelerin zaman uzayında karşılığı sinüzoidal işarettir ve algıda ciyaklama sesini andıran ton sesine karşılık gelir. Bu oluşan gürültüye müzikal gürültü denir. Geniş bant yalıtılmış tepeler de kalabilir.

Yalıtılmış tepeleri ortadan kaldırmak için aşırı çıkarma yöntemi ortaya atılmıştır (Loizou 2007). Yöntem, kestirilen gürültünün işlenen sestten fazlaca çıkarılmasıyla dar ve geniş bant yalıtılmış tepe bırakmamayı amaçlar. Aşırı çıkarma rağmen kalan tepelerin etkisini azaltmak için de çıkarılan gürültünün bir kısmı tekrar eklenir. Yöntemin matematiksel ifadesi (3.4)'de verilmiştir.

$$|\hat{Y}[\omega]|^2 = \begin{cases} |X[\omega]|^2 - \alpha |\hat{D}[\omega]|^2 & \text{eğer } |X[\omega]|^2 > (\alpha + \beta) |\hat{D}[\omega]|^2 \\ \beta |D[\omega]| & \text{diğer} \end{cases} \quad (3.4)$$

Bu ifadelerdeki  $\alpha$  aşırı çıkarma çarpanı,  $\beta$  spektral taban parametresi olarak adlandırılır.  $\alpha > 1$  ve 1'e yakın seçilir.  $\beta$ ,  $0 < \beta \ll 1$  şeklinde seçilir (Loizou 2007). Yöntemin blok şeması Şekil 3.6'te verilmiştir.

Bu yöntemde kullanılan  $\alpha$  ve  $\beta$  çarpanları Berouti ve arkadaşlarının önerdiği şekilde, işaretin gürültüye oranına (İGO) bağlı seçilmiştir (Berouti ve ark. 1979). Bu yöntemde kullanılan ilişkiler aşağıda verilmiştir.



Şekil 3.6: Aşırı Çıkarım Kullanılan Spektral Güç Çıkarım Algoritması Blok Diyagramı

$$\alpha = \begin{cases} \alpha_0 - \frac{-5}{s} & \text{eğer } \dot{\text{IGO}} \leq -5dB \\ \alpha_0 - \frac{\dot{\text{IGO}}}{s} & \text{eğer } -5dB < \dot{\text{IGO}} < 20dB \\ 1 & \text{eğer } 20dB \geq \dot{\text{IGO}} \end{cases} \quad (3.5)$$

$$\beta = \begin{cases} 0.02 & \text{eğer } 0dB < \dot{\text{IGO}} \\ 0.06 & \text{eğer } 0dB \geq \dot{\text{IGO}} \end{cases} \quad (3.6)$$

Aşırı çıkartma çarpanı hesaplama ifadesindeki  $s$ ,  $\dot{\text{IGO}}$ 'nun üst sınırında  $\alpha = 1$  yapan değerdir. Uygulamada kullanılmış  $\alpha_0 = 4dB$  için  $s = 20/3$ 'tür. Bu eğri deneysel olarak belirlenmiştir (Loizou 2007).

### 3.4.4 Doğrusal Olmayan Spektral Çıkarma

Gürültünün, işareti spektral boyutta her bileşen için aynı şekilde etkilemediği düşünülerek tasarlanmış bir yaklaşımdır. Doğrusal olmayan spektral çıkarma, aşırı çıkarma kullanılan spektral çıkarmının aşırı çıkarma çarpanının frekansla değişen değerler alan hali olarak nitelendirilebilir.

Doğrusal olmayan spektral çıkarım kuralı

$$|\hat{Y}[\omega]| = \begin{cases} |\bar{X}[\omega]| - \alpha[\omega]N[\omega] & \text{eğer } |\bar{X}[\omega]| > \alpha[\omega]N[\omega] + \beta \cdot |\bar{D}[\omega]| \\ \beta |\bar{X}[\omega]| & \text{diğer} \end{cases} \quad (3.7)$$

ile verilmiştir. (3.7) denklemindeki  $\bar{D}$  ve  $\bar{Y}$  yumuşatılmış gürültü ve yumuşatılmış işaret kestirimine karşılık gelmektedir. Yumuşatma işlemi  $i$ . çerçeve için

$$|\bar{X}_i| = \mu_x |\bar{X}_{i-1}| + (1 - \mu_x) |\bar{X}_i| \quad (3.8)$$

$$|\bar{D}_i| = \mu_d |\bar{D}_{i-1}| + (1 - \mu_d) |\bar{D}_i| \quad (3.9)$$

ile yapılmıştır.

Buradaki yumuşatma katsayıları için  $0.1 \leq \mu_x \leq 0.5$  ve  $0.5 \leq \mu_d \leq 0.9$  eşitsizliklerini sağlayan değerler atanması önerilmiştir (Loizou 2007). Çalışan algoritmada iki yumuşatma çarpanı da 0.5 seçilmiştir.

(3.7) ifadesindeki  $N[\omega]$ , en son 40 gürültü çerçevesindeki maksimum bileşenlerle güncellenen gürültü genlik spektrumudur. Matematiksel ifadesi;

$$|N[\omega]| = \max_{i-40 \leq j < i} (|\hat{D}_j[\omega]|) \quad (3.10)$$

ile verilebilir.

Yöntemin temelini oluşturan  $\alpha[\omega]$  ise yüksek İGO durumunda küçük değer alacak düşük İGO durumlarında yüksek değer alacak şekilde oluşturmak üzere

$$\alpha[\omega] = \frac{1}{1 + \rho \frac{|\bar{X}[\omega]|}{|\bar{D}[\omega]|}} \quad (3.11)$$

ifadesi kullanılmıştır. Bu ifadede  $\rho$ , İGO değerinden elde edilecek bilgiyi kuvvetlendirme veya azaltmak üzere denkleme eklenmiştir. Çalıştırılan algoritmada  $\rho = 1$  alınmıştır.

### 3.4.5 Çok Bantlı Spektral Çıkarma

Doğrusal olmayan spektral çıkarım, gürültünün frekans uzayında birörnek dağılmadığından yola çıkarak oluşturulmuş bir yöntemdir ve çıkarım her frekans birleşimindeki İGO'yla değişen şekilde gerçekleştirilir. Bu yaklaşım ani gürültü değişikliklerinde sesin bozulmasına neden olabileceğinden bu çıkartım değerini bantlar için belirlemek üzere bu yöntem geliştirilmiştir (Loizou 2007).

Bu yöntem için gürültüsüz işaretin ifadesi  $i$ . bant için  $b_i$  bant başlangıç frekansı,  $e_i$  bant

sonlanma frekansı olsun

$$|\hat{Y}_i[\omega_k]|^2 = |\bar{X}_i[\omega_k]|^2 - \alpha_i \delta_i |\hat{D}_i[\omega_k]|^2 \quad b_i \leq \omega_k \leq e_i \quad (3.12)$$

şeklinde verilmiştir. Bu ifadedeki  $\alpha_i$  bandın aşırı çıkarım katsayısı,  $\delta_i$  çıkarımda ek kontrol sağlayan bant çıkarım katsayısı ve  $\bar{X}_i$  ise gürültülü işaretin  $i$ . bant için zamanda ağırlıklı spektral ortalamasıdır.  $\bar{X}_i$  bant için frekanslar  $\omega_k$  ile ifade edilsin ve  $X$ 'nin alt indisi zaman penceresi gösteriyor olsun.  $j$ . pencere için

$$|\bar{X}_j[\omega_k]| = \sum_{k=-M}^M W_k X_{j-k}[\omega_k] \quad (3.13)$$

şeklinde ifade edilir. Buradaki ağırlıklar deneyle [0.09, 0.25, 0.32, 0.25, 0.09] olarak belirlenmiştir (Loizou 2007).

Her bant için İGO hesabı

$$IGO_i(dB) = 10 \log_{10} \left( \frac{\sum_{\omega_k=b_i}^{e_i} |\bar{X}_i[\omega_k]|^2}{\sum_{\omega_k=b_i}^{e_i} |\hat{D}_i[\omega_k]|^2} \right) \quad (3.14)$$

ayrı ayrı yapılarak bant bazında aşırı çıkarım katsayısı,  $\alpha$ 'lar;

$$\alpha_i = \begin{cases} 4.75 & \text{İGO}_i < -5 \\ 4 - \frac{3}{20} \text{İGO}_i & -5 \leq \text{İGO}_i \leq 20 \\ 1 & \text{İGO}_i > 20 \end{cases} \quad (3.15)$$

ifadesiyle hesaplanır. Bantlar üstü frekans bazında ek kontrol sağlayan  $\delta_i$  katsayıları;

$$\delta_i = \begin{cases} 1 & f_i \leq 1kHz \\ 2.5 & 1kHz \leq f_i \leq \frac{F_s}{2} - 2kHz \\ 1.5 & f_i > \frac{F_s}{2} - 2kHz \end{cases} \quad (3.16)$$

ifadesiyle verilmiştir. Bu değerler de deneysel olarak elde edilmiş değerlerdir (Loizou 2007).

Bu değerler pencere ve bant başına bulunduktan sonra temel çıkarım denklemi (3.12) uygulandığında sıfırdan düşük çıkabilecek genliklerin önüne geçmek için;

$$|\hat{Y}_i[\omega_k]|^2 = \begin{cases} |\hat{Y}_i[\omega_k]|^2 & \text{eğer } |\hat{Y}_i[\omega_k]|^2 > \beta |\bar{X}_i[\omega_k]|^2 \\ \beta |\bar{X}_i[\omega_k]|^2 & \text{diğer} \end{cases} \quad (3.17)$$

uygulanır. Bu ifadede  $\beta = 0.002$  alınmıştır. Ortaya çıkabilecek müzikal gürültüyü bastırmak

için iyileştirilmiş spektruma yeniden gürültü

$$|\bar{Y}_i[\omega_k]|^2 = |\hat{Y}_i[\omega_k]|^2 + 0.05|\bar{Y}_i(\omega_k)|^2 \quad (3.18)$$

ile eklenerek iyileştirilmiş işaret elde edilir..

### 3.4.6 Algısal Spektral Çıkarma

Diğer algoritmalar ya sabit çıkarım yapıyorlardı ya da İGO'ya göre adaptif çıkarım yapıyorlardı. Bu algoritma Virag'ın 1999'da olaya insan algısını da katmasıyla duyma maskelemesi kullanarak çıkarım yapmaktadır (Virag 1999).

Duyma maskelemesi, insanların algısıyla ilgilidir. İşaretin yüksek enerjili bileşenlerin yanında algısal gürültü eşiği yükselirken işaretin düşük enerjili bileşenleri yanında algısal gürültü eşiği düşer. Bu yöntem işareten bu gürültü eşiğini belirlemeye ve buna göre aşırı çıkarım katsayısını ve spektral taban katsayısını çıkararak çalışır.

Yüksek enerjili kısımlar, formant frekansları çevresi iken düşük enerjili kısımlar formant frekansları arasında kalan çukur kısımlardır. Çıkarmayı buna bağlı yapmak için Liu ve arkadaşları doğrusal öngörüm kodlamasına dayanan bir algoritma önermişlerdir (Liu ve ark. 2006).

Her pencerede

$$P(z) = \frac{1 - \sum_{k=1}^p a_k \sigma_1^k z^{-k}}{1 - \sum_{k=1}^p a_k \sigma_2^k z^{-k}} \quad (3.19)$$

ifadesinden  $T(\omega) = P(z)|_{e^{j\omega}}$  dönüşümü yapılarak algısal filtrenin frekans bileşenleri ,  $0 \leq \sigma_2 \leq \sigma_1 \leq 1$  eşitsizliğine uygun olarak bulunur. Bu ifadede  $a$ 'lar doğrusal öngörüm kodlaması katsayılarıdır. Uygulamamızda  $\sigma_1 = 1$ ,  $\sigma_2 = 0.8$  alınmıştır (Liu ve ark. 2006).

Algı eşiği belirlendikten sonra aşırı çıkarım katsayısı ve spektral taban katsayısı

$$\alpha(\omega) = \alpha_{max} \left( \frac{T_{max}(\omega) - T(\omega)}{T_{max}(\omega) - T_{min}(\omega)} \right) + \alpha_{min} \left( \frac{T(\omega) - T_{min}(\omega)}{T_{max}(\omega) - T_{min}(\omega)} \right) \quad (3.20)$$

$$\beta(\omega) = \beta_{max} \left( \frac{T_{max}(\omega) - T(\omega)}{T_{max}(\omega) - T_{min}(\omega)} \right) + \beta_{min} \left( \frac{T(\omega) - T_{min}(\omega)}{T_{max}(\omega) - T_{min}(\omega)} \right) \quad (3.21)$$

ile belirlendi. Burada  $\alpha_{max} = 6$ ,  $\alpha_{min} = 1$ ,  $\beta_{min} = 0$ ,  $\beta_{max} = 0.02$  olarak alınmıştır (Liu ve ark. 2006).

Bu katsayılar belirlendikten sonra aşırı çıkarım kullanan spektral çıkarım uygulanarak iyileştirilmiş sesin frekans bileşenleri elde edilmiştir.

### 3.5 Frekans Uzayından Zaman Uzayına Geiř

Bütün yontemlerin ıktısı iyileřtirilmiř genlik spektrumudur. Buradan zaman uzayındaki iřarete dnüş yapmak için her pencereye iyileřtirilmemiř iřaretin faz bilgisi eklendi. Daha sonra her pencereye ters Fourier Dnüşümü uygulandı.

Ters Fourier Dnüşümü uygulanmıř pencerelerde baskın gerel byüklüğün yanında az da olsa karmařık bir byüklük belirdi. İřaret gerek bir iřaret olmasına raėmen bu byüklüğün belirmesinin nedeninin sınırlı pencere ve iřaret uzunluėu olduėu dřünülererek pencere ierisindeki karmařık deėer byüklükleri ihmal edildi.

Pencereler bazındaki örtüşmeden sakınmak için her pencerenin ortasındaki %25'lik kısım alınıp ard arda eklenerek zaman uzayındaki tek boyutlu ses iřareti elde edildi.



#### 4 SONUÇLAR ve VARGILAR

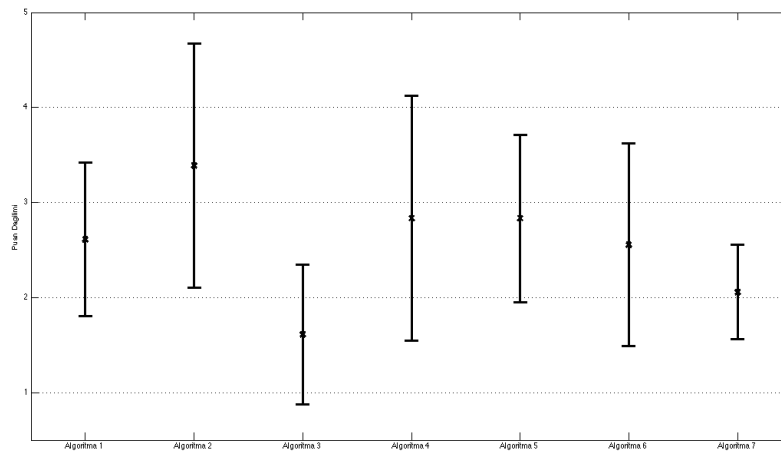
Algoritma çıktılarının kaliteleri hakkında değerlendirme yapmak için fikir puanı tercih edildi. Fikir puanı öznel bir sına olup iyi bir değerlendirme için uzmanların belirli şartları sağlayan mekanlarda bu sınamaya katılmasına ihtiyaç duyulur.

Bu çalışmada fikir puanlamasında uzmanlar yerine konuda uzman olan olmayan 10 kişilik bir topluluğa çıktılar değerlendirildi. Bu topluluktan çıktıların ne olduğunu bilmeden anlaşılabilirliklerine göre 1(çok kötü) ile 5(çok iyi, net) arasında kendi belirledikleri ortamda değerlendirme yapmaları istendi. 9 kişinin dönüş yaptığı değerlendirmede, kullanılan algoritmaların aldığı fikir puanlarının ortalaması Çizelge 4.1’de verilmiştir. Bu çıktılara verilen puanların dağılımını göstermek için ortalama puanlar, standart sapmalarıyla birlikte Şekil 4.1’de gösterilmiştir.

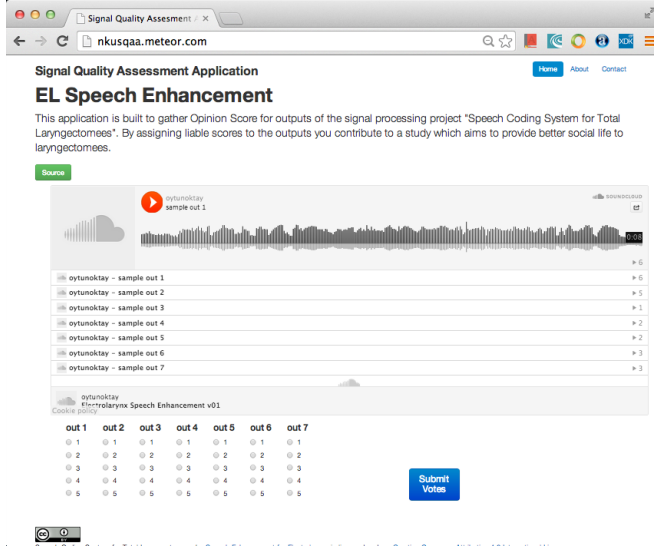
Çizelge 4.1: Ortalama Fikir Değerlendirme Çizelgesi

Algoritma	Ortalama Puan	Standart Sapma
Genlik Çıkarma	2.643	0.81
Güç Çıkarma	3.786	1.29
Aşırı Çıkarma	1.643	0.74
Doğrusal Olmayan Çıkarma	2.786	1.29
Çoklu Bant Çıkarma	2.643	0.88
Algısal Çıkarma	2.143	1.07
İşlenmemiş Ses	1.929	0.5

Bu puanlamanın sonucunda en iyi sonuç veren algoritma güç çıkarma algoritması olarak belirlenmiştir.



Şekil 4.1: Fikir puanlarının dağılımları



Şekil 4.2: İşaret Kalite Değerlendirme Uygulaması

## 4.1 Vargılar

Yöntemlerin değerlendirilmesi sonucu işlenmemiş sesin aşırı çıkarma yöntemi dışındaki yöntemlerden düşük puan alması işlemenin girdiyi az da olsa istenilen yönde iyileştirdiğini göstermektedir. Bütün çıktılara verilen puanların, girdiye verilen puandan daha fazla değışintiye sahip olması, puanlayanların girdinin anlaşılabilirliğinin yöntemlerin çıktularından kötü olduğunda hemfikir olduklarını gösterirken metotların çıktularının ne kadar iyi olduğunda hemfikir olmadıklarını göstermektedir.

Güç çıkarım yöntemi diğer yöntemlerden açık ara fazla puan almıştır. Genlik ve güç çıkarım yöntemleri bütün yöntemler içinde parametresiz yöntemler olarak nitelendirilebilir. Uygulanmasında parametre kullanılmayan bu iki yöntemde gürültünün tipi nedeniyle yüksek başarı sağlanmış olabilir.

Diğer yöntemler, genlik ve güç çıkarım yönteminin iyileştirilmiş sürümleridir. Buna rağmen görece düşük başarılarının nedeni barındırdıkları parametrelerin bu özel probleme göre ayarlanmaması olabilir.

## 4.2 İleri Çalışmalar

Bu çalışmada çıktuların değerlendirilmesi öznel ölçütlere bırakılmıştır. Öznel ölçütlerin istatistiksel olarak anlamlı çıktular verebilmesi için çok kişinin değerlendirmesine ihtiyaç vardır. Bunu sağlamak üzere bir sanal doku yöre uygulaması tasarlandı. Geliştirilmeye devam edilen uygulamanın an itibariyle kullanıcı arayüzü Şekil 4.2’te verilmiştir.

Bu çalışma birkaç yönde ilerletilebilir. Temel yaklaşım gürültü çıkarımı olduğundan

öncelikle gürültü kestiriminin üzerine gidilebilir.

Gürültü kestiriminde konuşmadaki duraklar işin içine katılabileceği gibi algılayıcı sayısı artırılarak da bu probleme yaklaşılabilir.

Yöntemlerin barındırdığı parametrelerin belirlenmesinde de iyileştirme yapılabilir. Çıktıların değerlendirilmesinde nesnel bir kalite ölçütü tanımlanırsa, bu kalite ölçütü yardımıyla parametrelerin en uygun değerleri en iyileme yöntemleriyle belirlenebilir. Hatta bu sayede farklı gürültüye sahip ortamlarda da başarımın sürekliliği sağlanmış olur.

Bu çalışma konuşmanın cihazlar üzerinden yapıldığı durumlara daha uygun olduğundan sabit ve/veya taşınabilir haberleşme cihazlarına nasıl uygulanabileceği üzerinde çalışılabilir. Birçok cihazda çalışabilen bir yazılım olarak veya cihazlara eklenebilen bir eklenti olarak uygulanabilir.

## KAYNAKLAR

- Aksoy, F., Veyseller, B., Yıldırım, Y. S., Demirhan, H., Özturan, O. (2010). Total larenjektomi sonrası ses restorasyonunda provox konuşma protezi tecrübemiz. *Acta Oncologica Turcica*, (43):65–69.
- Anadolu, Y., Akbaş, Y., Uzun, L. (1998). Total larenjektomi sonrası provox ses protezi ile ses rehabilitasyonu. *K.B.B. ve Baş Boyun Cerrahisi Dergisi*, 6(1):26–29.
- Barkmeier, J., Bunting, G. W., Hicks, D. M., Karnell, M. P., McFarlane, S. C., Stone, R. E., Berg, S. V., Watterson, T. L. (2003). Evaluation and treatment for tracheoesophageal puncture and prosthesis. Technical report, American Speech-Language-Hearing Association.
- Berouti, M., Schwartz, M., Makhoul, J. (1979). Enhancement of speech corrupted by acoustic noise. *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pages 208–211.
- Cole, D., Sridharan, S., Moody, M., Geva, S. (1997). Application of noise reduction techniques for alaryngeal speech enhancement. In *IEEE TENCON Speech and Image Technologies for Computing and Telecommunications*.
- Ergenç, I. (2002). *Konuşma Dili ve Türkçenin Söyleyiş Sözlüğü*. Multilingual Yabancı Dil Yayınları / Dilbilim Dizisi, 486 s, Ankara, Türkiye.
- Erişir, F. İnci, E. (2001). Total larenjektomiden sonra vokal rehabilitasyon. *Cerrahpaşa Tıp Dergisi*, 32(2):80–85.
- Flanagan, J., Allen, J., Hasegawa, M. (2008). *Speech Analysis, Synthesis and Perception*. J. Allen ECE537 Speech Recognition Course Material, 486 p, USA.
- Genden, E. M., Ferlito, A., Silver, C. E., Jacobson, A. S., Werner, J. A., Suárez, C., Leemans, C. R., Bradley, P. J., Rinaldo, A. (2007). Evolution of the management of laryngeal cancer. *Oral Oncology*, 43(5):431–439. <http://dx.doi.org/10.1016/j.oraloncology.2006.08.007>.
- Griffin Labs, U. (2008). *Trutone users manual*. <http://www.griffinlab.com/Manuals/TruTone-Users-Manual-English-Spanish-v20080214.pdf>.
- Liu, H., Zhao, Q., Wan, M., Wang, S. (2006). Enhancement of electrolarynx speech based on auditory masking. *IEEE Transactions on Biomedical Engineering*, 53(5):864–875.
- Loizou, P. C. (2007). *Speech Enhancement, Theory and Practice*. Taylor-Francis, 632 p, USA.
- Meltzner, G. S. (2003). *Perceptual and Acoustic Impacts of Aberrant Properties of Electrolaryngeal Speech*. Doktora Tezi, Massachusetts Institute of Technology.
- Niu, H.-J., Wan, M.-X., Wang, S.-P., Liu, H.-J. (2003). Enhancement of electrolarynx speech using adaptive noise cancelling based on independent component analysis. *Medical and Biological Engineering and Computing*, 41:670–678.
- Rosenberg, A. E. (1971). Effect of glottal pulse shape on the quality of natural vowels. *The Journal of Acoustical Society of America*, 49(1A):583–590.

- Roy, D. (2011). The birth of a word, TED2011. [http://www.ted.com/talks/deb\\_roy\\_the\\_birth\\_of\\_a\\_word.html](http://www.ted.com/talks/deb_roy_the_birth_of_a_word.html).
- Sharifzadeh, H. R., Ahmadi, F., McLoughlin, I. (2009). Speech reconstruction in post-laryngectomised patients by formant manipulation and pitch profile generation. In Proceedings of the World Congress on Engineering, volume II, London, UK.
- Tarakçıođlu, G. S. (2010). Voice conversion for reconstruction of dysphonic speech. Y.Lisans Tezi, Fen Bilimleri Enstitüsü, Bođaziçi Üniversitesi.
- Türkmen, H. I. (2008). Karma uyarım doğrusal Öngörüm kodlaması yöntemi ile disfonik konuşmadan normal konuşma elde edilmesi. Y.Lisans Tezi, Fen Bilimleri Enstitüsü, Yıldız Teknik Üniversitesi.
- Virag, N. (1999). Single channel speech enhancement based on masking properties of the human auditory system. IEEE Transactions on Speech and Audio Processing, 7(2).

## EKLER

### Konuşma Bilgi Aktarım Hızı

Türkçe'deki ses birimleri Çizelge 2.1'de gösterilmiştir. Bunları baz alırsak toplamda 42 tanedir.

Bir ses biriminin aktardığı bilginin hesaplanmasında bilgi miktarının üst sınırına, ses birimlerinin bir örnek dağıldığı durumda ulaşılır. Ses birimlerinin bir örnek dağıldığı durumda bir ses biriminin taşıdığı bilgi miktarı Shannon Entropisi'yle

$$H(m) = - \sum_{i=1}^{42} \frac{1}{42} \log_2 \left( \frac{1}{42} \right) = 5.3923 \text{ bit} \quad (4.1)$$

olarak bulunur.

Saniyede 25-30 ses birimi normal hızda konuşmaya karşılık gelmektedir. Buradan saniyede aktarılan bilgi miktarının üst sınırı 161.869 bit olarak hesaplanır.

## Konuşma İşlemede Kullanılan Betikler

Çalışmada kullanılan Matlab betikleri bu bölümde verilmiştir.

### Ana Betik

Ana betik, deneme dosyalarını yükleyip ön işlemeye tabi tuttukten sonra, sanki gerçek zamanlı işliyormuş gibi pencere pencere yöntemleri uygulayıp çıktığı kaydedecek şekilde tasarlanmıştır.

```
% VERIYLE ILGILI GIRDILER
% Noise Periyotlari
% BB01_04_electrolarynxSpeech.wav
% yalin el sesi
% 9.0 – 11.7s & 20.6 – 23.6s
[sub01data,fs01] = wavread(...
'../../../../data/electrolarynx_01/signed16bit/
BB01_04_electrolarynxSpeech.wav');
noiSub01Markers = [9.0 11.7; 20.6 23.6];
% [gurultubasi1 gurultusonul; gurultubasi2 gurultusonul2]
% MO01_04_electrolarynxDpeech.wav
% yalin el sesi
% 9.20 – 9.60s & 16.9 – 17.3s
[sub02data,fs02] = wavread(...
'../../../../data/electrolarynx_01/signed16bit/
MO01_04_electrolarynxSpeech.wav');
noiSub02Markers = [9.2 9.6; 16.9 17.3];
% [gurultubasi1 gurultusonul; gurultubasi2 gurultusonul2]

% kullanıcı girdileri
winDuration = 0.030; % pencere uzunlugu s
winOverPercent = 75; % pencere ortusme yuzde
dataInPro.rawData = sub01data;
dataInPro.fs = fs01;
dataInPro.preProLPFilterFile = 'LPfilt01.mat'; %FIR Filtre AG
dataInPro.preProHPfilterFile = 'HPfilt01.mat'; %FIR Filtre YG
dataInPro.windowType = 1; % 1 Hamming Pencerelemesi (sadece
Hamming)
dataInPro.noiseMarker = noiSub01Markers;

% onisleme
```

```

% GIRDI:
% {1:isaret dizisi (1D),2:ornekleme frekansi ,
%      3:AGF dosya , 4:YGF dosya , 5:pencere uzunlugu , 6:pencere
%      tipi ,
%      7:ortusme orani}
% CIKTI:
% {1:Filtrelenmis ->PencereLenmis Isaret(zaman uzayi) , 2:1
%      yapısında zaman
%      bilgisi , 3: fft(1) , 4:filtrelenmis isaret}
[dataInPro.winT , dataInPro.times , dataInPro.winF , dataInPro.
    filtered] =...
    preprocess(dataInPro.rawData , dataInPro.fs ,...
        dataInPro.preProLPFilterFile , dataInPro.preProHPfilterFile
        ,...
        winDuration , dataInPro.windowType , winOverPercent);
dataInPro.winTsize = size(dataInPro.winT,1);
dataInPro.winFsize = size(dataInPro.winF,1);
% gurultu kestirimi
dataInPro.noiseEstF = zeros(dataInPro.winFsize ,1);

% yalin EL sesi barindiran pencere numaralari
dataInPro.noiseWindowNos = [];
for i =1:size(dataInPro.noiseMarker ,1);
dataInPro.noiseWindowNos = [dataInPro.noiseWindowNos ...
    find((dataInPro.noiseMarker(i ,1)<dataInPro.times(1 ,:)) ...
        .*(dataInPro.noiseMarker(i ,2)>dataInPro.times(1 ,:))>0)];
end

% ciktilari ilk kullanima hazirlamak
dataInPro.method01.outWinF = zeros(size(dataInPro.winF));
dataInPro.method02.outWinF = zeros(size(dataInPro.winF));

% pencere bazinda isleme: yontem1 ve yontem2 icin
for win = 1:size(dataInPro.winF ,2)
    %gurultuyse , kestirimi yap
    if ismember(win , dataInPro.noiseWindowNos)
        dataInPro.noiseEstF = noiseEst(dataInPro.noiseEstF ,...
            dataInPro.winF(:, win));
    end
end

```



```

        dataInPro.method01.outWinF(:,win) = method01AmpSubt(...
            dataInPro.winF(:,win), dataInPro.noiseEstF);
        dataInPro.method02.outWinF(:,win) = method02PowSubt(...
            dataInPro.winF(:,win), dataInPro.noiseEstF);
    end

% method 1 ciktisi Spektral Genlik Cikarimi, zaman uzayina
donus
dataInPro.method01.outWinT = ifft(dataInPro.method01.outWinF);
dataInPro.method01.outT = real(matrix2array(dataInPro.method01.
    outWinT,...
        dataInPro.winTsize/2 , dataInPro.winTsize*(100-
            winOverPercent)/100));

% method 2 ciktisi Spektral Guc Cikarimi, zaman uzayina donus
dataInPro.method02.outWinT = ifft(dataInPro.method02.outWinF);
dataInPro.method02.outT = real(matrix2array(dataInPro.method02.
    outWinT,...
        dataInPro.winTsize/2 , dataInPro.winTsize*(100-
            winOverPercent)/100));

% method 3 Asiri Cikarma Kullanaral Spektral Cikarma
% gurultu bilgisi sifirlama
dataInPro.noiseEstF = zeros(size(dataInPro.noiseEstF));
% ciktinin ilk kullanima hazirlanmasi
dataInPro.method03.outWinF = zeros(size(dataInPro.winF));

% pencere bazinda isleme
for win = 1:size(dataInPro.winF,2)
    if ismember(win, dataInPro.noiseWindowNos)
        dataInPro.noiseEstF = noiseEst(dataInPro.noiseEstF ,
            dataInPro.winF(:,win));
    end
    [dataInPro.method03.outWinF(:,win), postSNRs(win), ...
        numBigInds(win), numElseInds(win)] = method03SpecSubOver
        (...

```

```

        dataInPro.winF(:,win), dataInPro.noiseEstF);

end
dataInPro.method03.outWinT = ifft(dataInPro.method03.outWinF);
dataInPro.method03.outT = real(matrix2array(dataInPro.method03.
    outWinT,...
        dataInPro.winTsize/2, dataInPro.winTsize*(100-
            winOverPercent)/100));

% method 4 dogrusal olmayan spektral cikarma
% degisken ilk degerlemesi
dataInPro.noiseEstF = zeros(size(dataInPro.noiseEstF));
dataInPro.method04.outWinF = zeros(size(dataInPro.winF));
sizeOfaWindow = size(dataInPro.noiseEstF,1);

% yontem degiskenleri
mux = 0.5;
mud = 0.5;
betas = 0.1;
gammas = 1;
Nomega = zeros(sizeOfaWindow);
Noind = 0;
NomegaData = zeros(size(dataInPro.noiseEstF,1),40);
XuFa = zeros(sizeOfaWindow,2);
DuFa = zeros(sizeOfaWindow,2);
sonindex = 1;

for win = 2: size(dataInPro.winF,2)

    indexs = mod(win,2)+1;
    indexsB = mod(win+1,2)+1;
    % gurultu ?
    if ismember(win, dataInPro.noiseWindowNos)
        % en yuksek gurultu guncellemesi
        Noind = mod(Noind, 41)+1; % N icinde dolasmak icin
            indis
        dataInPro.noiseEstF = noiseEst(dataInPro.noiseEstF, ...
            dataInPro.winF(:,win)); % gurultu kestirimi
    end
end

```

```

NomegaData(:,Noind) = dataInPro.noiseEstF(:); % gurultu
      kestirimi
% atamasi

DuFa(:,indexS) = mud*abs(DuFa(:,indexSB)) +...
      (1-mud)*abs(dataInPro.noiseEstF); % gurultu
      kestirimi yumusatma
sonindex = indexS;
end

Nomega = max(NomegaData, [], 2); % en yuksek gurultu
XuFa(:,indexS) = mux*abs(XuFa(:,indexSB)) + ...
      (1-mux)*abs(dataInPro.winF(:,win)); %! iyilestirilmis
      isaret yumusatma

roF = abs(XuFa(:,indexS))./abs(DuFa(:,sonindex)); %
      posterior IGO etki katsayisi

alphaF = 1./(1+gammas*roF);

dataInPro.method04.outWinF(:,win) = method04NonLinSS(...
      dataInPro.winF(:,win), XuFa(:,indexS), Nomega, alphaF
      ,...
      betas, DuFa(:,sonindex));

end

dataInPro.method04.outWinT = ifft(dataInPro.method04.outWinF);
dataInPro.method04.outT = real(matrix2array(dataInPro.method04.
      outWinT,...
      dataInPro.winTsize/2, dataInPro.winTsize*(100-
      winOverPercent)/100));

% yontem 5 coklu bant spektral cikarma

% degisken ilk degerleme
dataInPro.noiseEstF = zeros(size(dataInPro.noiseEstF));
dataInPro.method05.outWinF = zeros(size(dataInPro.winF));

```

```

% yontem degiskenleri
sizeOfaWindow = size(dataInPro.noiseEstF,1);
sizeOfAll = size(dataInPro.winF);

% yontem girdileri inputs
% frekans bantlari, satir tabanlı, baslangic bitis
% frekanslari
bands = [0 500;
         500 4000;
         4000 12000;
         12000 24000];
% spektral taban katsayisi
betam = 0.002;
% agirlikli spektral ortalama alma uzunlugu
M = 2;
% spektral agirliklar
Wi = [0.09 0.25 0.32 0.25 0.09];
% bantici cikarma kontrolu saglayan delta
deltaifrs = [0 1000;
            1000 dataInPro.fs/2-2000;
            dataInPro.fs/2-2000 dataInPro.fs/2];
% bantlara karsilik gelen delta katsayilari
deltaigain = [1.0 2.5 1.5];
% alpha IGO araliklari
alphaISNRs = [-inf -5; -5 20;20 inf];
% alpha_i = alpha degerleri islev icinde.

% frekans belirtecleri
frs = linspace(0, dataInPro.fs/2, sizeOfaWindow/2+1);
srf = -1*fliplr(frs);
frs = [frs srf(2:end-1)]';

for win = 1:sizeOfAll(2)
% gurlutu penceresi?
if ismember(win, dataInPro.noiseWindowNos)
    dataInPro.noiseEstF = noiseEst(dataInPro.noiseEstF, ...
    dataInPro.winF(:,win));
end
% agirlikli ortalama alinacak pencereler

```

```

    blockInds = win-M:win+M;
    % kenar kontrolleri
    inds = find((blockInds >=1) .* (blockInds <=sizeOfAll(2)));
    % agirlikli spektral ortalama
    XuFa = sum(repmat(Wi(inds),sizeOfaWindow,1) .* ...
        dataInPro.winF(:,blockInds(inds)),2);

    [dataInPro.method05.outWinF(:,win), SNrs(:,win)] =
        method05MultiSS(...
            dataInPro.winF(:,win), dataInPro.noiseEstF, XuFa, frs,
            ...
            bands, betam, alphasNRs, deltaifrs, deltaigain);
end

dataInPro.method05.outWinT = ifft(dataInPro.method05.outWinF);
dataInPro.method05.outT = real(matrix2array(dataInPro.method05.
    outWinT,...
    dataInPro.winTsize/2, dataInPro.winTsize*(100-
        winOverPercent)/100));

% yontem 6 algisal ozellikli cikarma
% degisken ilk degerlemesi
dataInPro.noiseEstF = zeros(size(dataInPro.noiseEstF));
dataInPro.method06.outWinF = zeros(size(dataInPro.winF));

% yontem degiskenleri
sizeOfaWindow = size(dataInPro.noiseEstF,1);
sizeOfAll = size(dataInPro.winF);
% dogrusal ongorum kodlamasi kutup sayisi
p = 24;
% P(z) pay sigma
sigma1 = 1;
% P(z) payda sigma
sigma2 = 0.8;

% alpha-beta uc degerleri
alphamin = 1;
alphamax = 6;

```

```

betamin = 0;
betamax = 0.02;

for win = 1:sizeofAll(2)
    % gurultu?
    if ismember(win, dataInPro.noiseWindowNos)
        dataInPro.noiseEstF = noiseEst(dataInPro.noiseEstF,
            dataInPro.winF(:, win));
    end
    [dataInPro.method06.outWinF(:, win), alphas06(:, win),
        betas06(:, win)] = ...
        method06PerceptSS(dataInPro.winT(:, win), dataInPro.winF
            (:, win), ...
            dataInPro.noiseEstF, p, sigma1, sigma2, alphamin,
            alphamax, ...
            betamin, betamax);
end

dataInPro.method06.outWinT = ifft(dataInPro.method06.outWinF);
dataInPro.method06.outT = real(matrix2array(dataInPro.method06.
    outWinT, ...
    dataInPro.winTsize/2, dataInPro.winTsize*(100-winOverPercent)
    /100));

```

## Ön İşleme İşlevi

Sesin filtrelenip, belirtilen örtüşme oranında pencerelere bölünüp, ilgili pencere fonksiyonuyla çarpıldığı fonksiyon.

```
function [ outputt , tt , outputf , dataFiltered ] = s01preprocess
    ( dataId , fsdata , ...
        filterLP , filterHP , winDur , winType , winOverlap )
% dataId : ham girdi , 1B
% fsdata : ornekleme frekansi
% filtre dosyaları filterHP , filterLP (.mat uzantili)
% winDur : pencere süresi
% winType : {1: 'HAMMING'}
% winOverlap : örtüşme yüzdesi
% outputt : Matris , her sütun bir pencere
% tt : Matris , zaman değerleri
% outputf : Matris , her sütun bir pencerenin fftsi
% dataFiltered : sadece filtrelenmiş veri 1B

ou = load(filterHP);
tempData = filter(ou.Num,1,dataId);
ou = load(filterLP);
dataFiltered = filter(ou.Num,1,tempData);
tempT = (0:length(dataFiltered)-1)*(1.0/fsdata);
winNop = winDur*fsdata;

if (winType==1)
    winData = hamming(winNop);
end

outputt = buffer(dataFiltered , winNop , 1.0*winOverlap/100*
    winNop);
tt = buffer(tempT , winNop , 1.0*winOverlap/100*winNop);
outputt = outputt.*repmat(winData , 1 , size(outputt , 2));

outputf = fft(outputt);
end
```

## Gürültü kestirimi

Gürültü penceresini alıp, gürültü kestirim işleviyle gürültü kestirimini tazeleyen işlevdir.

```
function [ updatedNoiseSpectrum ] = noiseEst( noiseEst ,  
    noisyDataWindowF )  
% INPUTs  
% noiseEst ; previous estimation of noise spectrum  
% noisyDataWindowF; new windows marked as noise  
% OUTPUTs  
% updatedNoiseSpectrum; new noise estimation  
  
    updatedNoiseSpectrum = (noiseEst + noisyDataWindowF)/2;  
  
end
```



## Genlik Çıkarma İşlevi

Genlik çıkarma yöntemini pencere bazında uygulayan işlev dosyasının içeriği aşağıda verilmiştir.

```
function [ enhancedWindowF ] = method01AmpSubt( noisyWindowF ,  
        noiseEstF )  
%METHOD01AMPSUBT Summary of this function  
% INPUTs  
% noisyWindowsF : a window of noisy signal  
% noiseEstF : estimated noise spectrum  
% OUTPUT  
% enhancedWindowsF : enhanced window! a window, vector  
  
        tempF = abs(noisyWindowF)-abs(noiseEstF);  
        tempF( find (tempF<0)) = 0;  
        enhancedWindowF = abs(tempF).*exp(1i*angle(noisyWindowF));  
end
```

## Güç Çıkarma İşlevi

Güç çıkarma yöntemini pencere bazında uygulayan işlev dosyasının içeriği aşağıda verilmiştir.

```
function [ enhancedWindowF ] = method02PowSubt( noisyWindowF ,
    noiseEstF )
%METHODO2POWSUBT Summary of this function goes here
% INPUTs
% noisyWindowsF : a window of noisy signal
% noiseEstF : estimated noise spectrum
% OUTPUT
% enhancedWindowsF : enhanced window! a window, vector

    tempF = noisyWindowF.*conj(noisyWindowF)-noiseEstF.*conj(
        noiseEstF);
    tempF(find(tempF<0)) = 0;
    enhancedWindowF = (tempF.^0.5).*exp(1i*angle(noisyWindowF))
        ;

end
```

## Aşırı Çıkarma İşlevi

Aşırı çıkarma yöntemini pencere bazında uygulayan işlev dosyasının içeriği aşağıda verilmiştir.

```
function [ enhF, postSNR, numBigInds, numElseInds ] =
    method03SpecSubOver( winF, noiF )
% method03SpecSubOver
% Girdi
% winF: gürültülü isaretin bir penceresi (f uzayı)
% noiF: bir pencerelik gürültü kestirimi (f uzayı)
% Çıktı
% enhF: iyileştirilmiş isaretin bir penceresi (f uzayı)
% postSNR: posterior SNR
%

% kullanıcı tanımlı parametreler
alpha0 = 4;
betaHsnr = 0.02;
betaLsnr = 0.06;

% çıktı ilk kullanıma hazırlama
enhF2 = zeros(size(winF));
% posterior SNR hesaplama
postSNR = 10*log10(abs(winF'*conj(winF))/abs(noiF'*conj(noiF)))
    ;

postSNRLowLimit = -5;
postSNRHighLimit = 20;

% alpha, beta belirleme
if (postSNRLowLimit < postSNR) && ( postSNR < postSNRHighLimit)
    alphas = alpha0 - (3.0/20)*postSNR; % !
                                                % ----->
        should depend on alpha0
    if postSNR > 0 % !
        betas = betaHsnr;
    else
        betas = betaLsnr;
```

```

        end

    else
        alphas = 0; %
        betas = 0; %
    end

    winF2 = winF.*conj(winF);
    noiF2 = noiF.*conj(noiF);

    % durumlara gore cikarmalar
    bigInds = find(winF2>(alphas+betas)*noiF2);
    elseInds = find(winF2<=(alphas+betas)*noiF2);
    numBigInds = length(bigInds);
    numElseInds = length(elseInds);
    enhF2(bigInds) = winF2(bigInds)-alphas*noiF2(bigInds);
    enhF2(elseInds) = betas*noiF2(elseInds);

    enhF = sqrt(enhF2).*exp(1i*angle(winF));

end

```

## Doğrusal Olmayan Çıkarma İşlevi

Doğrusal olmayan çıkarma yöntemini pencere bazında uygulayan işlev dosyasının içeriği aşağıda verilmiştir.

```
function [ enhF ] = method04NonLinSS( winF, XuFa, NF,...
    alphas, betas, DuFa )
% METHOD04NONLINSS
% Girdi :
% winF: Pencerelemis isaret(f)
% XuFa: Yumusatilmis EL isareti (f)
% NF: Son 40 penceredeki en yuksek gurultu (f)
% alphas: Pencere icin asiri cikarim katsayisi (f)
% betas: Pencere icin spektral taban katsayisi (f)
% DuFa: Yumusatilmis Gurultu

% iyilestirilmis isaret baslangic degeri atama
enhFa = zeros(size(winF));
% buyuk degerlerin indisleri
bigInds = find(XuFa > (alphas.*NF + betas.*DuFa));
% kucuk degerlerin indisleri
elseInds = find(XuFa <= (alphas.*NF + betas.*DuFa));
enhFa(bigInds) = XuFa(bigInds) - alphas(bigInds).*NF(bigInds);
enhFa(elseInds) = betas*XuFa(elseInds);
% faz bilgisi ekleme
enhF = enhFa.*exp(1i*angle(winF));

end
```

## Çoklubant Çıkarma İşlevi

Çoklubant çıkarma yöntemini pencere bazında uygulayan işlev dosyasının içeriği aşağıda verilmiştir.

```
function [ enhF, SNRi ] = method05MultiSS( winF, noiEstF, XuFa,
    ...
    freqScope, bands, betam, alphaSNRs, deltaFs, deltaG )
% METHOD05MULTISS
% Girdiler:
% winF: EL konuşma penceresi (f)
% noiEstF: Gurultu penceresi (f)
% XuFa: agirlikli spektral ortalama (f)
% freqScope: fft ciktisinin frekans bileşenleri
% bands: bantlar, genel, baslangic, bitis
% betam: spektral taban katsayisi
% alphaSNRs: alpha katsayisi sinirlari
% deltaFs: delta katsayisi sinirlari
% deltaG: delta katsayilari

sizeOfaWindow = size(winF,1);

% band belirtec vektoru, her bant icin
bandInd = zeros(size(winF,1), size(bands,1));
for i = 1:size(bands,1)
    bandInd(:,i) = (freqScope >= bands(i,1)) .* (freqScope <
        bands(i,2)) + ...
        (freqScope <= -bands(i,1)) .* (freqScope > -bands(i,2))
    ;
    bandInd(:,i) = bandInd(:,i)>0;
end

% delta belirtec vektoru .* delta degeri
deltai = zeros(sizeOfaWindow, 1);
for i = 1:size(deltaFs,1)
    deltaInd(:,1) = (freqScope >= deltaFs(i,1)) .* (freqScope <=
        deltaFs(i,2)) + ...
        (freqScope <= -deltaFs(i,1)) .* (freqScope >= -deltaFs(i,2));
    gainInds = find(deltaInd(:,1)>0);
    deltai(gainInds) = deltaG(i);
end
```

```

end

% bant IGO hesabi
SNRi = zeros(size(bands,1),1);
for i = 1:size(bands,1)
    bwinF = XuFa.*bandInd(:,i);
    bnoiF = noiEstF.*bandInd(:,i);
    SNRi(i) = 10*log10(abs(bwinF'*conj(bwinF))/abs(bnoiF'*conj(
        bnoiF)));
end

% IGO'ya gore alpha carpan vektorleri olusturma
alphai = zeros(sizeOfaWindow, size(alphaSNRs,1));
for i = 1:size(bands,1)
    if alphaSNRs(1,1) <= SNRi(i) && SNRi(i) <= alphaSNRs(1,2)
        alphai(:,i) = repmat([4.75], sizeOfaWindow, 1);
    elseif alphaSNRs(2,1) <= SNRi(i) && SNRi(i) <= alphaSNRs
        (2,2)
        alphai(:,i) = repmat([4-3.0/20*SNRi(i)], sizeOfaWindow
            ,1);
    elseif alphaSNRs(3,1) <= SNRi(i) && SNRi(i) <= alphaSNRs
        (3,2)
        alphai(:,i) = repmat([1], sizeOfaWindow,1);
    end
end

% Pencere guc bilesenleri
XuFa2 = abs(XuFa.*conj(XuFa));
% Bant bazinda pencere gucleri
xi2u = zeros(sizeOfaWindow, size(bands,1));
for i = 1:size(bands,1)
    XuFa2band = XuFa2.*bandInd(:,i);
    noiband = abs(noiEstF.*conj(noiEstF)).*bandInd(:,i);
    % alpha + delta cikarimi
    temp2 = XuFa2band - alphai(:,i).*deltai.*noiband;
    % beta eklemesi
    bigInds = find(temp2>(betam*XuFa2band));
    elseInds = find(temp2<=(betam*XuFa2band));
    temp2(elseInds) = betam*XuFa2band(elseInds);
end

```

```
        xi2u(:,i) = temp2 + 0.05*XuFa2band;  
end  
  
enhFa2 = sum(xi2u,2);  
% faz bilgisi ekleme  
enhF = sqrt(enhFa2).*exp(1i*angle(winF));  
  
end
```



## Algısal Çıkarma İşlevi

Algısal çıkarma yöntemini pencere bazında uygulayan işlev dosyasının içeriği aşağıda verilmiştir.

```
function [ enhF, alphas, betas ] = method06PerceptSS( winT,
    winF, noiF, ...
    lpcP, sig1, sig2, alphamin, alphamax, betamin, betamax )
%METHODO6PERCEPTSS
% Girdiler:
%   winT: EL isareti (zaman)
%   winF: EL isareti (frekans)
%   noiF: Gurultu kestirimi
%   lpcP: dogrusal ongorum kodlamasi kutup sayisi
%   sig1: P(z) pay carpani
%   sig2: P(z) payda carpani
%   alphamin: alpha alt sinir
%   alphamax: alpha ust sinir
%   betamin: beta alt sinir
%   betamax: beta ust sinir

% zaman isaretinden DOK katsayilari
[a,q] = lpc(winT,lpcP);
% sigma carpanlari
sigma1 = ones(lpcP+1,1);
sigma2 = ones(lpcP+1,1);
% pay ve payda ifadeleri
sigma1(2:end) = a(2:end).*(sig1.^sigma1(2:end)');
sigma2(2:end) = a(2:end).*(sig2.^sigma2(2:end)');
% P(z) pay ve payda
numerator = [sigma1(1); -sigma1(2:end)];
denominator = [sigma2(1); -sigma2(2:end)];
% P(z)'den frekans uzayina gecis
[h,w] = freqz(numerator, denominator, length(winT)/2+1);
% fft frekanslariyla uyusturma
pwfc = [h(1:end); flipud(h(2:end-1))];
% T(omega) sinirlarini belirleme
Tmax = max(pwfc);
Tmin = min(pwfc);
% alpha beta ifadeleri
```

```

alphas = ((Tmax-pwfc)/(Tmax-Tmin))*alphamax + ...
         ((pwfc-Tmin)/(Tmax-Tmin))*alphamin;
betas = ((Tmax-pwfc)/(Tmax-Tmin))*betamax + ...
        ((pwfc-Tmin)/(Tmax-Tmin))*betamin;

Y2f = abs(winF.*conj(winF));
D2f = abs(noiF.*conj(noiF));
X2f = zeros(size(Y2f));
bigInds = find(Y2f>(alphas+betas).*D2f);
elseInds = find(Y2f<=(alphas+betas).*D2f);

alphabetabeta = alphas(bigInds)+betas(bigInds);
X2f(bigInds) = Y2f(bigInds)-alphabetabeta.*D2f(bigInds);
X2f(elseInds) = betas(elseInds).*D2f(elseInds);
% faz bilgisi ekleme
enhF = sqrt(X2f).*exp(1i*angle(winF));

```

```
end
```

## ÖZGEÇMİŞ

2008 yılında Gebze Yüksek Teknoloji Enstitüsü, Mühendislik Fakültesi, Elektronik Mühendisliği Bölümü'nden lisans diplomasını alan Mesut Oytun Oktay, gene aynı yıl İstanbul Teknik Üniversitesi, Bilişim Enstitüsü, Hesaplamalı Bilim ve Mühendislik Anabilim Dalı'nda yüksek lisans eğitimine başladı. 2010 yılında Namık Kemal Üniversitesi, Çorlu Mühendislik Fakültesi, Elektronik ve Haberleşme Bölümü'nde araştırma görevlisi olarak çalışmaya başladığından İstanbul Teknik Üniversitesi'ndeki yüksek lisans eğitimini yarıda bıraktı. Kendi bölümünde yüksek lisans eğitimini 2014 yılında bitirdi.