

**YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**KARMA UYARIM DOĞRUSAL ÖNGÖRÜM
KODLAMASI YÖNTEMİ İLE DİSFONİK KONUŞMADAN
NORMAL KONUŞMA ELDE EDİLMESİ**

Bilgisayar Mühendisi H. İrem TÜRKMEN

**FBE Bilgisayar Mühendisliği Anabilim Dalında
Hazırlanan**

YÜKSEK LİSANS TEZİ

Tez Danışmanı:

Yrd. Doç. Dr. M. Elif Karslıgil

İSTANBUL, 2008

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	iii
KISALTMA LİSTESİ	iv
ŞEKİL LİSTESİ	v
ÇİZELGE LİSTESİ	vii
ÖNSÖZ	viii
ÖZET	ix
ABSTRACT	x
1. GİRİŞ	1
1.1 Ses Üretim Mekanizması ve Ses Bozuklukları	2
1.2 Literatürdeki Ses Değiştirme Sistemleri	3
1.3 Tez Çalışmasında Tasarlanan Sistem	4
2. KONUŞMA KODLAMA ALGORİTMALARINA GENEL BAKIŞ	6
2.1 Karma Uyarım Doğrusal Öngörüm Kodlaması	7
2.1.1 KUDÖK ile Konuşma Analizi	8
2.1.2 KUDÖK ile Konuşma Sentezi	22
3. DİSFONİK SESİN AKUSTİK ÖZELLİKLERİ	27
3.1 Sesliler ve Ötümlü Sessizlerin Akustik Farklılıkları	28
3.2 Ötümsüz Sessizlerin Akustik Farklılıkları	31
4. KARMA UYARIMLI DOĞRUSAL ÖNGÜRÜM KODLAMASI İLE DİSFONİK KONUŞMADAN NORMAL KONUŞMA ELDE EDİLMESİ	33
4.1 Sistem Kısıtları	33
4.2 Veri Tabanı Oluşturma	33
4.3 Sistemin Genel Yapısı	34
4.4 Ötümsüz Sessiz İçeren Çerçevelerin Tespiti	34
4.5 Disfonik Ses Düzeltme Sisteminin Genel Yapısı	34
4.5.1 Bant İçi Seslilik Değişikliği	35
4.5.2 Perde Değişikliği	36
4.5.3 Kazanç Değişikliği	38
4.5.4 Formant Değişikliği	40
4.5.4.1 Spektral Yumuşatma	40
4.5.4.2 Formant Yapısı Değişikliği	43
5. DENEYSEL SONUÇLAR	49
5.1 Spektral Uzaklık Hesabı	49

5.2	Öznel Dinleme Testleri ve Sentetik Ses Spektrogramları	51
6.	SONUÇLAR VE ÖNERİLER	56
KAYNAKLAR.....		58
ÖZGEÇMİŞ.....		60

SİMGE LİSTESİ

a	Doğrusal Öngörü Parametreleri
ab	Aperiyodik Bayrak
B	Bant Genişliği
F	Formant Frekansı
g	Kazanç Değeri
p	Tepesellik
vs	Bant Geçişli Sessizlik Değeri
β	Dinamik Erim Kontrol Katsayısı
ω_i	i. Çizgisel Spektral Frekansı

KISALTMA LİSTESİ

AFD	Ayrık Fourier Dönüşümü
ATFD	Ayrık Ters Fourier Dönüşümü
ÇAVN	Çok Amaçlı Vektör Nicemleme
ÇSF	Çizgisel Spektral Frekansları
DÇDR	Doğrusal Çok Değişkenli Regresyon
DÖK	Doğrusal Öngörüm Kodlaması
EODR	En Fazla Olabilir Doğrusal Regresyon
HFD	Hızlı Fourier Dönüşümü
KUDÖK	Karma Uyarım Doğrusal Öngörüm Kodlaması
VN	Vektör Nicemleme

ŞEKİL LİSTESİ

	Sayfa
Şekil 1.1 Gırtlığın Yapısı	2
Şekil 1.2 Sistemin Blok Diyagramı	5
Şekil 2.1 Konuşma Kodlama Sistemlerinin Temel Yapısı	6
Şekil 2.2 KUDÖ Kodlayıcı Blok Diyagramı	7
Şekil 2.3 KUDÖ Kod Çözücü Blok Diyagramı	7
Şekil 2.4. Fourier Büyüklük Hesabı Adımları	19
Şekil 3.1a Disfonik Sesle Söylenmiş /Otobus/ Kelimesi Spektrogramları.....	27
Şekil 3.1b Normal Sesle Söylenmiş /Otobus/ Kelimesi Spektrogramları	27
Şekil 3.2a Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /a/ Fonemi Spektrumları	28
Şekil 3.2b Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /e/ Fonemi Spektrumları	29
Şekil 3.3a Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /l/ Fonemi Spektrumları	30
Şekil 3.3b Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /m/ Fonemi Spektrumları	30
Şekil 3.3c Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /r/ Fonemi Spektrumları	30
Şekil 3.3d Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /v/ Fonemi Spektrumları	30
Şekil 3.4a Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /ç/ Fonemi Spektrumları	31
Şekil 3.4b Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /t/ Fonemi Spektrumları	31
Şekil 3.4c Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /k/ Fonemi Spektrumları	31
Şekil 3.4d Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /s/ Fonemi Spektrumları	31
Şekil 3.4e Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /ş/ Fonemi Spektrumları	31
Şekil 3.4f Disfoni Hastalarına ve Normal Sese Sahip Konuşmacıya Ait /f/ Fonemi Spektrumları	31
Şekil 4.1 Disfonik Ses Değiştirme Sistemi Blok Diyagramı	35
Şekil 4.2 Normal Sesle Söylenmiş /a/ Fonemi Spektrumları	37
Şekil 4.3 Fısıltılı Sesle Söylenmiş /a/ Fonemi Spektrumları	37
Şekil 4.4a /elma ağacı/ Cümlesi İçin Disfonik Ses.....	39
Şekil 4.4b /elma ağacı/ Cümlesi İçin Sessizlik ve Perde Değişikliği Yapılmış Disfonik Ses ..	39
Şekil 4.4c /elma ağacı/ Cümlesi İçin Kazanç Değişikliği ile Yükseltilmiş Ses	39
Şekil 4.4d /elma ağacı/ Cümlesi İçin Gürültü Eliminasyonu Yapılmış Ses	39
Şekil 4.5 Normal /a/ Fonemi Ses Sinyaline Ait ÇSF Gezingeleri	40
Şekil 4.6 Disfonik /a/ Fonemi Ses Sinyaline Ait ÇSF Gezingeleri	41
Şekil 4.7 Disfonik /a/ Fonemi Ses Sinyaline Ait Yumuşatılmış ÇSF Gezingeleri	42
Şekil 4.8 Disfonik /a/ Fonemi Ses Sinyaline Ait Yumuşatılmış ÇSF Gezingeleri	42
Şekil 4.9 Disfonik /a/ Fonemi Spektrumu ve Çizgisel İzge Frekansları.....	43

Şekil 4.10	Disfonik /a/ Fonemi Değiştirilmiş Spektrumu.....	45
Şekil 4.11	Disfonik /a/ Fonemi ve Chip z Dönüşümü ile Değiştirilmiş Spektrumu.....	46
Şekil 4.12	Disfonik /a/ Fonemi ve Formant Frekansları Değiştirilmiş Spektrum	47
Şekil 4.13	Disfonik /a/ Fonemi Spektrumu, Formant Frekansları ve Bant Genişlikleri Değiştirilmiş Spektrum ve Normal Ses Spektrumu.....	48
Şekil 5.1a	Sesliler ve ötümlü sessizler için akustik özellikler uyarlanarak sentezlenmiş seslerin, fısıltılı sesi normal sese yaklaştırma oranları	50
Şekil 5.1b	Ötümsüz sessizler için fonemin akustik özelliklerinin korunması ile sentezlenmiş sesin, akustik özellikleri uyarlanarak sentezlenmiş sese göre normal sese yakınlık oranı	50
Şekil 5.2a	Sesliler ve ötümlü sessizler için akustik özellikler uyarlanarak elde sentezlenmiş seslerin, disfonik sesi normal sese yaklaştırma oranları	50
Şekil 5.2b	Ötümsüz sessizler için fonemin akustik özelliklerinin korunması ile sentezlenmiş sesin, akustik özellikleri uyarlanarak sentezlenmiş sese göre normal sese yakınlık oranı	50
Şekil 5.3a	/ağrı dağı/ Tamlaması için Disfonik Ses Spektrogramı.....	51
Şekil 5.3b	/ağrı dağı/ Tamlaması için Sentetik Ses Spektrogramı.....	51
Şekil 5.3c	/ağrı dağı/ Tamlaması için Normal Ses Spektrogramı.....	51
Şekil 5.4a	/hava durumu/ Tamlaması için Disfonik Ses Spektrogramı	52
Şekil 5.4b	/hava durumu/ Tamlaması için Sentetik Ses Spektrogramı.....	52
Şekil 5.4c	/hava durumu/ Tamlaması için Normal Ses Spektrogramı.....	52
Şekil 5.5a	/taş/ Kelimesi için Normal Ses Spektrogramı.....	53
Şekil 5.5b	/taş/ Kelimesi için Disfonik Ses Spektrogramı.....	53
Şekil 5.5c	/taş/ kelimesi için Ötümsüz Sessizlerin Akustik Özellikleri Korunarak Sentezlenmiş Ses	53
Şekil 5.5d	/taş/ kelimesi için Tüm Seslerin Akustik Özellikleri Değiştirilerek Sentezlenmiş Ses.....	53
Şekil 5.6a	/çalışma/ Kelimesi için Normal Ses Spektrogramı.....	54
Şekil 5.6b	/çalışma/ Kelimesi için Disfonik Ses Spektrogramı	54
Şekil 5.6c	/çalışma/ kelimesi için Ötümsüz Sessizlerin Akustik Özellikleri Korunarak Sentezlenmiş Ses	54
Şekil 5.6d	/çalışma/ kelimesi için Tüm Seslerin Akustik Özellikleri Değiştirilerek Sentezlenmiş Ses	54

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2.1 KUDÖK Bit Paketi.....	21
Çizelge 4.1. Test İçin Kullanılan Örnekler.....	33
Çizelge 4.2. Fonem Gruplarının Sınıflandırılması Başarı Oranları.....	34
Çizelge 4.3 Disfonik /a/ Fonemi İçin Formant Frekansı Ve Bant Genişlikleri	44
Çizelge 4.4 Disfonik /a/ Fonemi İçin Çizgisel İzge Frekansları	44
Çizelge 4.5 Formant Bant Genişliği Değiştirilmiş Spektrum İçin Formant Frekansı Ve Bant Genişlikleri	45
Çizelge 4.6 Formant Frekansları Değiştirilmiş Spektrum İçin Formant Frekansı Ve Bant Genişlikleri	47

ÖNSÖZ

Günümüzde sağlık alanında pek çok problem disiplinler arası yaklaşımlarla çözülmektedir. Bu bakış açısıyla, sesini tamamen kaybetmiş insanların yaşam kalitesini arttırmak üzere Cerrahpaşa Tıp Fakültesi Kulak Burun Boğaz Ana Bilim Dalı'nda tedavi gören ses teli felci ve gırtlak kanseri hastalarından aldığım ses örnekleri üzerinde yaptığımız çalışmalar sonucunda tezimi tamamladım.

Tezimi gerçekleştirme sürecinde, bilgisini ve deneyimlerini bana aktaran değerli hocam ve tez danışmanım Yrd. Doç. Dr. M. Elif Karslıgil'e, tez çalışmama başlamam için beni yönlendiren Sayın Prof. M. Yahya Karslıgil'e, Y.T.Ü Bilgisayar Mühendisliği Bölümündeki bütün hocalarıma ve arkadaşlarıma teşekkür ederim.

Hayatım boyunca desteklerini hiçbir zaman eksik etmeyen, beni yüreklendiren ve bana güvenen sevgili anneme, babama ve kardeşlerime çok teşekkür ederim.

Çalışmama gösterdikleri ilgi ve katkılarından dolayı Cerrahpaşa Tıp Fakültesi Kulak Burun Boğaz Ana Bilim Dalı Öğretim Üyelerinden Prof. Dr. Özgün Enver'e, Doç. Dr. Ender İncir'e ve Dr. Yusuf Hacızade'ye teşekkür ederim.

KARMA UYARIM DOĞRUSAL ÖNGÖRÜM KODLAMASI YÖNTEMİ İLE DİSFONİK KONUŞMADAN NORMAL KONUŞMA ELDE EDİLMESİ

ÖZET

Konuşma, günlük etkileşimin en etkin ve verimli biçimlerinden biridir. Ses telleri, gırtlak içinden geçen havanın etkisiyle titrer ve ses çıkartılmasını sağlar. Damak, dil ve dudakların hareketiyle ses şekillendirilir ve konuşma halini alır. Oluşan sesin şiddetini akciğerler tarafından oluşturulan subglottik basıncın büyüklüğü belirler. Ses tellerinin gerginliğinin, şeklinin ve ses tellerine gelen hava basıncının değişmesiyle ses farklı şekillerde çıkar. Ses bozuklukları, gerek sinirsel, gerek yapısal, gerekse de ses telleri üzerinde oluşan patolojiler nedeniyle, konuşma için gırtlaktan çıkardığımız sesin tizliği, yüksekliği ve kalitesinde meydana gelen istenmedik değişikliklerdir.

Sesini tamamen kaybetmiş hastalar için elektro gırtlak ya da ses protezi gibi tıbbi ve mekanik çözümler önerilmiştir. Ancak bu teknikler enfeksiyon riski taşıyabilmekte ya da üretilen ses kalitesiz olabilmektedir. Bu tekniklerle üretilmiş seslerin kalitesinin yükseltilmesine yönelik çalışmalar son 15 yılda hızlanmış ve birçok sistem tasarlanmıştır. Ancak literatürde bu tekniklere alternatif olarak kullanılabilecek doyurucu bir çalışma mevcut değildir. Bu çalışmada, konuşma yeteneğini tamamen kaybetmiş hastalar için ses protezi gibi kullanılabilecek yazılım tabanlı bir sistem tasarlanmış ve gerçekleştirilmiştir.

Tasarlanan sistemde, iyileştirilmiş konuşmanın sentezi için KUDÖK (Karma Uyarımlı Doğrusal Öngörüm Kodlaması, MELP) kullanılmıştır. Disfonik konuşma için ötümsüz sessiz içeren çerçeveler tespit edilmiş ve bu çerçeveler dışındaki çerçeveler için formant frekansları ve perde bağlantısı kullanılarak perde oluşturulmuş, formant yapısı değişikliği yapılmış ve seslilik eklenmiştir. Perde üretimi için perde-formant frekansı ilişkisinden yararlanılmıştır.

Geliştirilen sistem ile elde edilen sentetik sesin kalitesinin incelenmesi için spektral uzaklık hesabı yapılmış ve öznel dinleyici testlerine başvurulmuştur. Testler sonunda sentetik sesin disfonik sese göre normal sese yaklaşık %20 oranında daha yakın olduğu saptanmıştır. İleriki çalışmalarda sentetik ses kalitesinin artırılması ve gerçek zamanlı bir gömülü sistem uygulaması haline getirilmesi ile sesini tamamen kaybetmiş hastaların yeniden etkin sesli iletişim kurabilmelerine yardımcı olacak alternatif bir metod geliştirilebilir.

Anahtar Kelimeler: Ses Bozuklukları, Ses Değiştirme Teknikleri, Karma Uyarım, Doğrusal Öngörüm Kodlaması

RECONSTRUCTION OF NORMAL SPEECH FROM DYSPHONIC SPEECH BY USING MIXED EXCITATION LINEAR PREDICTIVE CODING METHOD

ABSTRACT

Speech is one of the best effective ways of daily communication. While producing voice, vocal cords vibrate when air is forced through them and along the vocal tract. Voice shaped by palate, tongue and lips and gets form to speech. The intensity of the voice depends on the subglottic pressure produced by lungs. Voice can be adjusted by varying the shape and the tension in the vocal cords, and the pressure of the air behind them. Dysphony is the result of the neural, structural or pathological effects on the vocal cords or larynx and it causes undesirable changes in the pitch, amplitude and the quality of the speech.

Medical and mechanical solutions like electrolarynx and voice prosthesis are proposed for the patients who have lost ability to speak. However, these techniques have infection risk or are poor quality. Although studies about enhancing the speech that is produced by these techniques have begun to be seen in the literature for 15 years, there is not a comprehensive study on alternative solution of medical and mechanical techniques. The proposed system in this thesis can be used as speech prosthesis for the dysphony patients.

In the proposed system, MELP (Mixed Excitation Linear Predictive Coding) was used for synthesizing enhanced speech. Unvoiced phonemes were detected for dysphonic speech. Pitch and voicing were produced, formant modification was applied for the phonemes except unvoiced phonemes. Correlation between pitch and formant frequencies was used in order to produce pitch.

Spectral distance was calculated and subjective listening tests were applied in order to discuss the synthetic speech quality. It is observed that similarity between synthetic speech and normal speech is %20 higher compared to the similarity between dysphonic speech and normal speech. In the future, an alternative method that can help patients who are lack of ability to communicate effectively can be developed by increasing the synthetic speech quality and constructing a real time embedded system.

Keywords: Dysphony, Voice Modification Techniques, Mixed Excitation Linear Predictive Coding

1. GİRİŞ

İnsanı diğer canlı varlıklardan üstün kılan öğelerden birisi oldukça üst düzeyde gelişmiş olan sözel iletişim yeteneğidir. Yazılı, görsel ve sözel iletişim, sosyal iletişimin üç temel ögesini oluşturmaktadır. Sözlü iletişim bunların en efektif olanıdır. Konuşmacı, daha önce edinmiş olduğu bilgileri, duygu düşüncelerine göre korteksten alınacak komutlar ile önce sesi oluşturur, sonra da artikulatorler dediğimiz konuşma organları ile şekillendirilerek, en basit konuşma birimi olan fonemler oluşur. Yukarıda özetlenen konuşma zincirinin hangi halkasında bir problem oluşur ise, iletişimin ilgili elemanında sorunlar ortaya çıkmaktadır [2].

Sesin oluşumunda temel olarak üç sistem rol oynamaktadır. Bu sistemlerden birincisi jeneratör sistemdir. Jeneratör sistem basınçlı hava çıkışını sağlayan akciğerler tarafından oluşturulur. İkincisi vibratuar (titreşen) sistemdir ve bu görevi gırtlak üstlenir. Son olarak da oluşturulan sesin anlaşılabilir bir konuşma şeklinde ifade edilmesini sağlayan rezonatör sistem görev alır. Bu sistemi de yutak (burnun hemen gerisinden başlayarak boynuna inen ve özefagus (yemek borusu) olarak devam eden bir tüptür), dil, damak, dudaklar, burun ve paranasal sinüsler oluşturmaktadır. Jeneratör sistemin oluşturduğu basınçlı hava, ses tellerinin orta hatta birleşip titreşmesine ve sesin oluşumuna neden olmaktadır. Oluşan sesin şiddetini ses telleri altında akciğerler tarafından oluşturulan subglottik basıncın büyüklüğü belirlemektedir. Bu basınç ne kadar yüksek ise sesin şiddeti de o kadar büyük olur. Ses tellerinin gerginliği ve boyu arttıkça sesin saniyedeki titreşim sayısı artar, böylelikle ses inceler.

Ses kısıklığı, ses kalitesindeki değişimi tanımlamada kullanılan bir terimdir. Esas olarak ses tellerindeki patolojik titreşim sonucu oluşan, gırtlak fonksiyon bozukluğudur. Kronik ses kısıklıkları, ses tellerindeki polip, nodül ve reinke ödemi gibi organik lezyon oluşumları, gırtlak kötü hastalıkları, gırtlak kanseri, nörolojik bozukluklar, sigara kullanımına bağlı kronik irritasyon sebebiyle oluşur. Ses tellerinde patolojik oluşumlar bulunan hastalarda, nefesle alınan hava, nefes borusuna gönderilerek ses kaynağı oluşturulur, ses telleri çok az titreşir veya hiç titreşmez. Ses fısıltılı halde, normalden daha kalın veya çatallı çıkar [4].

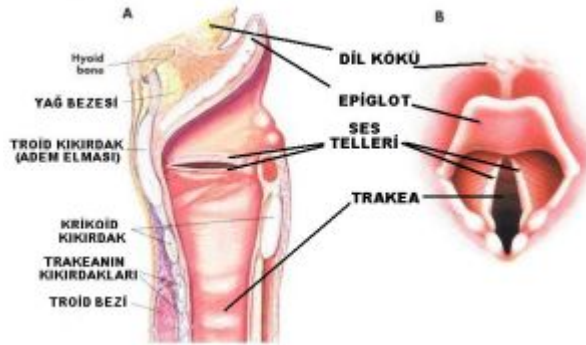
Bu çalışmada, başlıca nedenleri gırtlakta oluşan enfeksiyonlar, gırtlak kanseri ve ses tellerindeki organik lezyonlar olan kronik ses bozukluklarında, hastaların seslerinin düzeltilmesine yönelik bir sistem tasarlanmıştır. Geliştirilen sistem, kronik ses bozukluğu olan hastalar için sosyal hayatta daha etkili iletişim kurmalarına yönelik önemli bir adım olacaktır.

1.1 Ses Üretim Mekanizması ve Ses Bozuklukları

Sağlıklı bir insanda ses üretim mekanizması, aşağıdaki adımları içerir;

- Ses, akciğerlerden nefes borusu ile yukarı doğru çıkan havanın, gırtlakta bulunan ses tellerini titreştirmesi sonucu oluşmaktadır.
- Ses telleri, konuşulmadığı zamanlarda 'V' şeklindedir ve her iki yana açıktır. Konuşma sırasında orta hatta bir araya gelirler ve 'II' şeklini alırlar.
- Bu teller, ince seslerin çıkartılması için daha çok gerilir ve titreşir. Kalın seslerde ise bunun tam tersi bir mekanizma çalışmaktadır.
- Ses; gırtlaktan çıktıktan sonra ise, dil kökü, dil, dişler, dudaklar gibi organlarımızda son şeklini alır. Bu sırada geniz, burun boşluğu ve sinüslerde titreşerek kişiye özgü ses karakterleri ortaya çıkmaktadır.

Şekil 1.1' de gırtlak yapısı görülmektedir.



Şekil 1.1 Gırtlığın Yapısı

Bu çalışma kapsamında incelenen fonksiyonel kronik ses bozuklukları temel olarak 3 başlık altında toplanabilir. Birinci tip ses bozukluklarında, ses tellerinde oluşan nodül, polip veya ödem gibi organik lezyonlar nedeniyle ses telleri esnekliklerini kaybeder. İkinci tip ses bozukluklarında ses teli felci veya, gırtlakta oluşan enfeksiyonlar nedeniyle ses tellerinin biri veya ikisi işlevini yitirir. Bu durumda ses hiç çıkmaz. Gırtlak kanserine bağlı ses bozuklukları, üçüncü tipi oluşturmaktadır. Gırtlak kanserinin tedavisinde cerrahi yöntemler kullanılmışsa, ses tellerinin biri, gırtlığın bir bölümü veya tamamı alınmış olabilir. Üç tip cerrahi müdahalede de gırtlığa nefes alınması için bir delik açılmıştır [5].

1.2 Literatürdeki Ses Değiştirme Sistemleri

Gırtlak kanserinin tedavisinde uygulanan laryngectomy operasyonlarında gırtlak ve ses kablosu (vocal cord) tamamen alınabilir. Bu operasyondan sonra, boynun alt kısmına hastanın doğrudan nefes alabilmesi için bir delik açılır. Sesin düzeltilmesi için kullanılan üç yöntem vardır. Birincisi yapay gırtlak kullanmaktır (alaryngeal konuşma). Hasta bu elektro gırtlak kullanarak konuşabilir. Ancak elde edilen ses yapaydır ve elektrondur. İkinci yöntemde, hasta aldığı eğitimden sonra özefageal konuşabilmektedir. Özefageal konuşmada hasta boynun alt kısmındaki delikten aldığı havayı, ağızdan çıkarmayı öğrenerek ses çıkartır. Bu tarz konuşmada ses geçiriyormuşçasına çıkar. Üçüncü yöntem ise Trakeo-Özefageal konuşmadır. Bu yöntemde, laryngectomy sonrası, gırtlakta bir protez yerleştirilir. Diğer iki yönteme göre daha başarılı sonuçlar verebilmektedir ancak protez kullanımına bağlı enfeksiyon riski vardır ve protezin bakımı hasta için sıkıntı verici olmaktadır [3].

Beynin kasları kontrol eden kısmının hasar görmesi sonucu, Spasite (Cerebral Palsy) meydana gelir. Hasta, çene ve ağızını doğru kullanamadığı için sesleri düzgün çıkaramaz. Alaryngeal ve özefageal konuşmanın iyileştirilmesine yönelik çalışmalar son 15 yılda hızlanmış ve çoğunluğu ses değişikliği tekniklerine dayanan birçok sistem tasarlanmıştır.

Aguilar ve Nakano (2005), örüntü tanıma yaklaşımı temelli bir alaryngeal konuşma iyileştirme sistemi sunmuşlardır. Önerilen sistemde, konuşmanın sesli kısımları yapay sinir ağları yöntemleri ile bulunur ve normal konuşmanın ilgili segmentleri ile değiştirilir.

Bi ve Qi (1997), var olan konuşmacı dönüştürme algoritmalarını alaryngeal konuşmayı iyileştirmek için kullanmışlardır. Spektral bozulmaları vektör nicemleme-VN (vector quantization, VQ) ve doğrusal çok değişkenli regresyon, (linear multivariate regression-LMR) tabanlı konuşmacı dönüştürme metotlarını kullanarak düzeltmeyi hedeflemişlerdir.

Pozo ve Young (2006), gırtlaksal dalga formu kullanarak ve seğirmeyi azaltarak tracheözefageal konuşmayı yeniden sentezlemişlerdir. Qi ve Winberg (1995) ise kadın tracheözefageal konuşmacıların seslerini gırtlaksal dalga formu ile yeniden sentezlemiş ve perde yumuşatması yapmışlardır.

Sawada ve Takeuchi (2004), perde periyodu kestirimi ve spektral değişiklik yöntemlerini kullanarak, özefageal konuşmanın iyileştirilmesi ve sinirsel felçli hastaların konuşmalarının düzeltilmesi üzerine çalışmışlardır.

Ses tellerinin titreşim yeteneklerini kaybetmesine bağlı konuşma bozukluklarında, fısıltılı konuşmada olduğu gibi, ses telleri arasındaki açıklık artar. Nefesle verilen havanın gırtlak geçidinden geçmesiyle, bu açıklıkta meydana gelen türbülanslı hava akımı, ses kaynağını oluşturur. Genellikle bu ses kaynağı, ses yolunun daha aşağıdaki parçalarında dağılır. Bu sebeple oluşan uyarım (excitation) normal sesinkinden daha gürültülü ve 20 db daha alçaktır (F. Icat, H. G. İlk, 2004). Literatürde, fısıltılı ile normal ses arasındaki yapısal farklılıkların analizi ve fısıltılı konuşmanın tanınması üzerine yapılmış az sayıda çalışma mevcuttur.

F. Icat ve H. G. İlk (2004), sürdürülebilir Türkçe fonemler için, fısıltılı ve normal konuşma arasındaki farklılıkları incelemişlerdir.

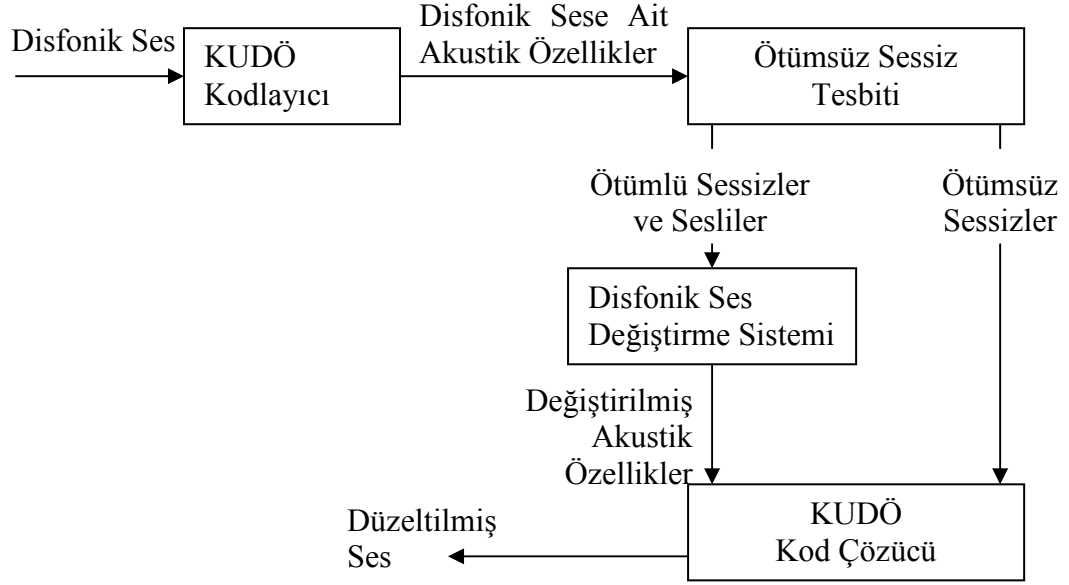
Itoh, T. Takeda (2001), fısıltılı konuşmanın akustik analizi ve tanınması üzerine çalışmışlardır. Yaptıkları çalışmada, fısıltılı konuşmanın tanınması için Saklı Markov Modeli tabanlı, En fazla Olabilir Doğrusal Regrasyon (Maximum likelihood linear regression) konuşmacı adaptasyonu kullanmışlar ve %78 oranında bir tanıma başarısı elde etmişlerdir.

Morris ve Clements (2002), fısıltının normal sese dönüştürülmesi üzerine çalışmışlardır. Çalışmalarında, normal ses elde edebilmek için, perde ve ses özelliklerinin alabileceği değerleri tahmin etmişler ve çizgisel spektrum frekansı değişikliği yapmışlardır.

1.3 Tez Çalışmasında Tasarlanan Sistem

Bu çalışmada, organik lezyonlar, ses teli felçleri ve gırtlak kanseri gibi insan ses üretim sisteminde fonksiyonel bozukluğa neden olan rahatsızlıklarda, hastanın sosyal hayatta etkili iletişiminde katkıda bulunabilecek Disfonik Konuşmadan Normal Konuşma Elde Etme Sistemi tasarlanmış ve gerçekleştirilmiştir.

Geliştirilen sistemde, sesin analizi ve düzeltilmesi için KUDÖK algoritması kullanılmıştır. KUDÖK giriş sinyalini, çerçeve çerçeve kodlar. Sistem, temel olarak dört adımdan oluşmaktadır. Sisteme ait blok diyagram Şekil 1.3.1’de görülmektedir. İlk adımda disfonik hastaya ait konuşma 8kHz örnekleme frekansında kaydedilir ve kaydedilen ses KUDÖK ile analiz edilir. İkinci adımda, konuşma sinyalinin ilgili çerçevesini içerdiği fonemin, ötümlü bir sese mi yoksa ötümsüz bir sese mi ait olduğu tespit edilir. Üçüncü adımda, konuşma sinyalinin ötümlü ses içeren her çerçevesi için analiz sonunda elde edilen, perde, ses borusu ve seslilik parametreleri, disfonik sesi, normal sese yakınlatacak şekilde değiştirilir. Son adımda ise, değiştirilmiş parametreler kullanılarak yeni ses sentezlenir.

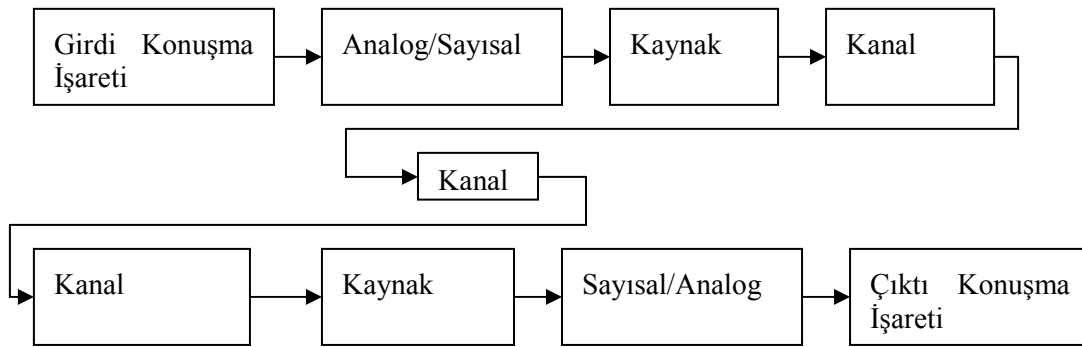


Şekil 1.2 Sistemin Blok Diyagramı

Çalışmanın 2. bölümünde, konuşma kodlama algoritmalarının temel prensiplerinden ve KUDÖK yönteminden, 3. bölümünde disfonik sesin akustik özelliklerinden, 4. bölümünde Disfonik Konuşmadan Normal Konuşma Elde Etme Sisteminin detaylarından bahsedilmiş, 5. bölümünde deneysel sonuçlar verilmiş, sonuç bölümünde çalışmanın başarısının değerlendirilmesi yapılmıştır.

2. KONUŞMA KODLAMA ALGORİTMALARINA GENEL BAKIŞ

Analog kaynaktan çıkan konuşma sinyali, dijital ortamda çok fazla yer tutmaktadır. Veri iletim hızının ve veri saklama kapasitelerinin en efektif şekilde kullanılabilmesi için ses sinyalinin kayıplı bir takım metotlarla sıkıştırılması gerekmektedir. Konuşma kodlama uygulamaları, sesin kalitesini bozmadan, konuşma işaretini en az bit ile ifade edebilmeyi hedefler. Konuşma kodlama sistemlerinin temel yapısı, Şekil 2.1’de gösterilmiştir.



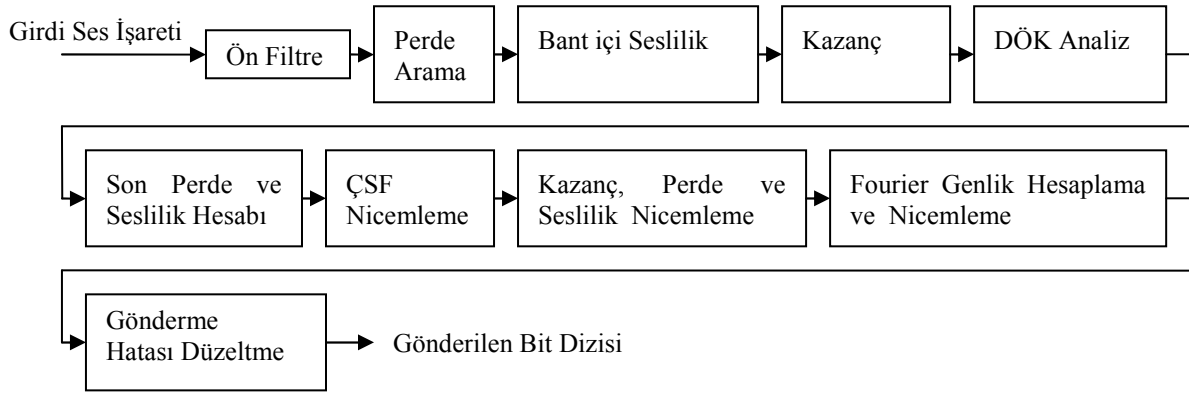
Şekil 2.1 Konuşma Kodlama Sistemlerinin Temel Yapısı

Konuşma kodlama algoritmaları kodlama tekniklerine göre, dalga biçimi kodlaması, parametrik kodlama ve karma kodlama metotları olarak sınıflandırılabilir.

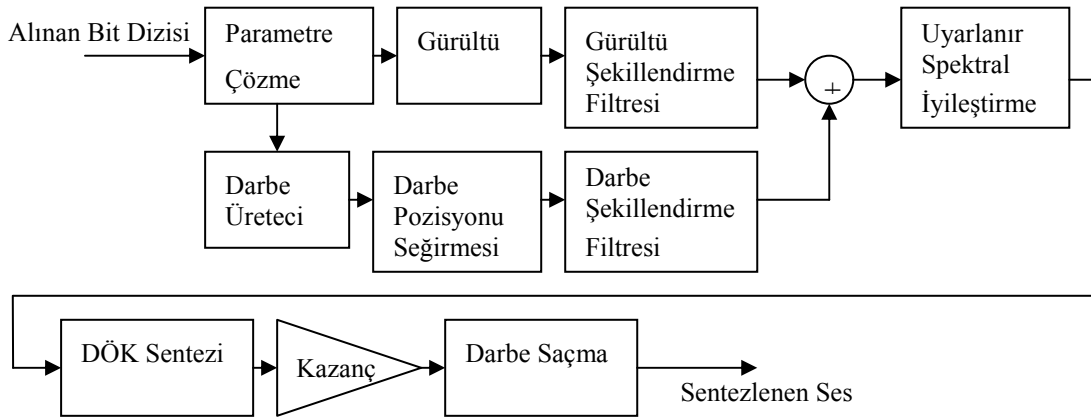
En efektif konuşma kodlama algoritmalarından biri, parametrik kodlama tekniği kullanan, Doğrusal Öngörüm Kodlaması-DÖK (Linear Predictive Coding-LPC)’dır. DÖK, insan sesini, temsil edebilmek için, tamamen parametrik bir model kullanır. Bu yaklaşımda kodlayıcı, sadece konuşma modelinin parametrelerini iletişim kanalından aktarır. Kod çözücü ise, giriş sinyalinin algısal karakteristiğine sahip bir konuşma sinyali sentezlemeyi hedefler. Model parametrelerinin, periyodik olarak kodlanması, konuşma sinyalinin doğrudan saklanmasına göre daha az bit gerektirdiğinden, DÖK, düşük bit hızlarında çalışabilmektedir. (3000 b/s den daha az). Düşük bit hızlarında çalışabilir olmasına rağmen, özellikle arka plan gürültülerinin bulunduğu ortamlardan alınan konuşma sinyallerinin kodlanmasında, çok iyi bir ses kalitesi elde edilememektedir. Böyle örneklerde, sentezlenen ses insan kulağına mekanik veya vızılı gelir. En basit anlamda DÖK, uyarım için basit darbe dizisi kullandığından yüksek kalitede konuşma üretememektedir.

2.1 Karma Uyarım Doğrusal Öngörüm Kodlaması

KUDÖK, DÖK'nın ürettiği ses kalitesini iyileştirmek için 1995 yılında, McCree tarafından geliştirilmiştir. Daha sonra iyileştirilerek, 1997'de resmi U.S federal standardı haline gelmiştir. KUDÖ kodlayıcı ve kod çözücü blok diyagramı Şekil 2.2 ve Şekil 2.3'de verilmiştir (Texas Instruments, 1998). Çalışmanın bu bölümünde, KUDÖ kodlayıcıya ve kod çözücüye ait işlem adımları anlatılırken Wai C. Chu'nun, 'Speech Coding Algorithms' adlı çalışmasından (Wai. C. Chu) ve Texas Instruments tarafından hazırlanan 'Final Report for Fixed-Point Implementation of the 2.4 kbps MELP Federal Standard Speech Coder' adlı çalışmasından (1998) faydalanılmıştır.



Şekil 2.2 KUDÖ Kodlayıcı Blok Diyagramı



Şekil 2.3 KUDÖ Kod Çözücü Blok Diyagramı

KUDÖK, basit darbe dizisi yerine, karma uyarım kullanarak, daha fazla sayıda ses çeşidi üretebilmekte böylece sentezlenen ses kulağa daha doğal gelmektedir. KUDÖK algoritması, geleneksel DÖK ses kodlayıcılara, karma uyarım (mixed excitation), periyodik olmayan darbe (aperiodic pulses), uyarlanır spektral iyileştirme (adaptive spectral enhancement), darbe

saçılımı (pulse dispersion) ve fourier genlik modelleme(fourier magnitude modeling) özelliklerinin eklenmesiyle geliştirilmiş tamamen parametrik çalışan bir kodlama algoritmasıdır.

Karma uyarım, çok bantlı karma model kullanılarak gerçekleştirir. Bu model, frekans bağımlı seslendirme (voicing) gücünü, uyarlanır filtre yapısı ile simüle eder. Karma uyarımın en önemli getirisi, DÖK ses kodlayıcılarda karşılaşılan vızıltı problemini azaltmasıdır.

DÖK kodlayıcısının en önemli kısıtlarından biri, konuşma sinyali çerçevesinin, sesli veya sessiz olarak sınıflandırılmasıdır. KUDÖK ise sesli, sessiz ve seğirme olmak üzere 3 sınıf kullanır. Seğirme durumunda, uyarım aperiodyiktir fakat tam olarak rasgele değildir. Genellikle konuşma sinyalinin, sesli ve sessiz parçaları arasındaki geçişlerde kullanılır. KUDÖK, sesin seğirme durumu modellemesi için, rasgele bir sayı olan perde seğirme(pitch jitter) parametresini kullanır. Bu özellik, KUDÖK kod çözücünün, hızlı değişebilen darbeleri, tonal sesler oluşmadan üretebilmesini sağlar (Wai. C. Chu).

Uyarlanır spektral iyileştirme filtresi, doğrusal öngörüm filtresinin kutupları kullanılarak oluşturulur. Sentezlenen konuşmanın, formant yapısını iyileştirmek için kullanılır. Darbe saçma, formant içermeyen frekans bölgelerinde, yapay sesin dalga biçiminin, doğal ses ile daha uyumlu olmasını sağlar. Böylece daha kaliteli bir ses sentezlenir

Fourier genlik modellemede, konuşma sinyalinin öngörüm kalıntısının ilk on Fourier genliği hesaplanır. Fourier genlikleri, kod çözücüye aktarılacak parametreler arasına alınmıştır . Bu sayede, özellikle alçak frekanslarda, sesin algısal kalitesi arttırılmış olur.

KUDÖK, 8kHz örnekleme hızında çalışmaktadır. Her bir çerçeve, 22.5 ms uzunluğundadır. Her bir çerçeve için, son perde, bant içi seslilik değerleri, kazanç değerleri, doğrusal öngörüm katsayıları, Fourier genlikleri ve aperiodyik bayrak (flag) parametrelerinin nicemlenmesinden oluşan 54 bitlik bir katar gönderilmektedir.

2.1.1 KUDÖK ile Konuşma Analizi

KUDÖK ile konuşma analizi, aşağıdaki adımlardan oluşmaktadır.

- Alçak Frekansların Çıkarılması (DC Removal)
- Birinci Aşama Perde Değerinin Hesaplanması
- Bant Geçişli Seslilik Analizi (Band Pass Voicing Analysis)
- Alçak Bant Seslilik Değeri Kararı

- Aperiiodik Bayrak
- Dört Yüksek Bant için Seslilik Değeri Kararı
- Kazanç Hesabı
- Doğrusal Öngörüm Analizi
- Tepesellik Hesaplanması
- Son Perde Değerinin Hesaplanması
- Doğrusal Öngörüm Katsayılarından Çizgisel İzge Frekanslarının Elde Edilmesi ve Nicemlenmesi
- Bant İçi Seslilik Nicemleme
- Perde ve Alçak - Bant Seslilik Nicemleme
- Kazanç Nicemleme
- Fourier Genliklerinin Hesaplanması ve Nicemlenmesi
- Gönderme Hatası Düzeltme ve Bit Paketleme

Alçak Frekansların Çıkarılması (DC Removal):

Kodlama işleminin ilk adımı, ses sinyalinden, çok alçak frekanslı enerjinin çıkartılmasıdır. Bu işlem 60 Hz kesim frekanslı, 4. dereceden Chebychev yüksek geçirgen filtre ile yapılır. Bundan sonraki işlemlerde, bu filtre çıktısı, girdi sinyali olarak kullanılacaktır. n. dereceden Chebychev yüksek geçirgen filtrenin, ω açısal frekansı yanıtı, (2.1) eşitliğindeki fonksiyon kullanılarak hesaplanır [1].

$$G_n(\omega, \omega_c) = \frac{1}{\sqrt{1 + (1/\varepsilon^2 T_n^2(\omega_c/\omega))}} \quad (2.1)$$

Burada ε , dalgacık (ripple) faktörü, ω_c , kesim frekansı ve $T_n(\cdot)$, n. Dereceden Chebychev polinomudur.

Birinci Aşama Perde Değerinin Hesaplanması:

Birinci aşama perde değerinin hesaplanması için, tam sayılı ve kesirli perde değerlerinin hesaplanması gerekmektedir.

Tam Sayılı Perde Değerinin Hesaplanması:

Tamsayılı Perde Değeri Hesaplaması, alçak geçiren filtre ile filtrelenmiş ses sinyali üzerinde yapılır. Filtreleme için, frekans yanıtı fonksiyonu (2.2) eşitliğinde verilen, 1 kHz kesim frekanslı 6. dereceden alçak geçiren Butterworth filtresi kullanılır.

$$G_n^2(\omega) = \frac{G_0^2}{1 + (\omega/\omega_c)^{2n}} \quad (2.2)$$

Burada, n , filtre derecesi, ω_c kesim frekansı, G_0 ise, 0 frekansındaki enerjidir. Tamsayılı perde değerinin hesaplanması için s , alçak geçiren filtre ile filtrelenmiş ses sinyali olmak üzere, her 180 örneklilik çerçeve için, eşitlik (2.3)'de verilen, normalize edilmiş özilinti fonksiyonu (normalized autocorrelation function) kullanılır.

$$r(l) = \frac{c[0, l, l]}{\sqrt{c[0, 0, l]c[l, l, l]}} \quad (2.3)$$

Burada $c[l, m, k]$, özilinti işlemidir ve eşitlik (2.4) ile ifade edilir.

$$c[l, m, k] = \sum_{n=-\lfloor k/2 \rfloor - 80}^{-\lfloor k/2 \rfloor + 79} s[n+1]s[n+m] \quad (2.4)$$

Periyodik sinyaller için, tam sayılı perde periyodu, $r[l]$ 'yi enbüyükten l değerine karşılık gelir. tam sayılı perde periyodu araması için, $l = 40, 41, \dots, 160$ değerleri kullanılmıştır.

Kesirli Perde Değeri Hesaplanması ve İyileştirilmesi:

Kesirli perde değerinin hesaplanması ve iyileştirilmesi için; T bölüm 1.2.1'de anlatılan tam sayılı perde değeri olmak üzere, aşağıdaki adımlar izlenmelidir.

1. $c[0, T+1, T]$ ve $c[0, T-1, T]$ hesaplanır. Eğer, $c[0, T-1, T] > c[0, T+1, T]$ ise, 2. adıma geçmeden önce T bir azaltılır. Bu adımın amacı, en büyük özilinti noktasının, $[T, T+1]$ ve $[T-1, T]$ aralıklarından hangisinde olduğunu bulmaktır.
2. Eşitlik (2.5) kullanılarak, kesirli perde periyodu hesaplanır.

$$\eta = \frac{c[0, T+1, T]c[T, T, T] - c[0, T, T]c[T, T+1, T]}{c[0, T+1, T](c[T, T, T] - c[T, T+1, T]) + c[0, T, T](c[T+1, T+1, T] - c[T, T+1, T])} \quad (2.5)$$

3. Eşitlik (2.6) kullanılarak, normalize edilmiş özilinti değeri hesaplanır.

$$r[T+\eta] = \frac{(1-\eta)c[0, T, T] + \eta c[0, T+1, T]}{\sqrt{c[0, 0, T]((1-\eta)^2 c[T, T, T] + 2\eta(1-\eta)c[T, T+1, T] + \eta^2 c[T+1, T+1, T])}} \quad (2.6)$$

Birinci aşama perde değeri $T^{(1)} = T + \eta$ olarak ifade edilir.

Bant Geçişli Seslilik Analizi (Band Pass Voicing Analysis):

Bu aşamada, 0-500, 500-1000, 1000-2000, 2000-3000 ve 3000-4000 Hz kesim frekanslarında, 6. derece Butterworth filtreler kullanılarak, beş frekans bandında seslilik analizi yapılır.

Alçak Bant Seslilik Değeri Kararı:

En alçak frekans bandının (0-500 Hz) seslilik değeri (vs_1), (2.6) eşitliği ile hesaplanan normalize edilmiş özilinti değerine eşittir.

$$vs_1 = r(T^{(1)}) \quad (2.7)$$

Aperiyodik Bayrak:

Aperiyodik bayrak-ab (aperiodic flag-af), alçak bant seslilik değeri ile ilişkilidir.

$$af = \begin{cases} 1, & vs_1 < 0.5 \\ 0, & vs_1 \geq 0.5 \end{cases} \quad (2.8)$$

0.5 'in üstünde yüksek seslilik değerlerine sahip çerçeveler yüksek olasılıkla periyodiktir ve bayrak değeri 0 olarak atanır. Zayıf periyodiklik için kod çözücünün sesli uyarımlı, aperiyodik darbeler üretmesi için (seğirmeli seslilik) bayrak değeri 1 olarak atanır. Bu bayrak için bit katarında 1 bit ayrılmıştır.

Dört Yüksek Bant için Seslilik Değeri Kararı:

Dört Yüksek Bant için Seslilik Değeri Kararı için gerçekleşmesi gereken adımlar aşağıdaki gibidir:

1. $r(\)$, (2.6) eşitliği ile hesaplanan normalize edilmiş özilinti ve $T^{(1)}$, birinci aşama gerçek perde periyodu olmak üzere $r_1 = r(T^{(1)})$ hesaplanır. Özilintinin hesaplanması için kullanılan sinyal, ilgili bant geçiren filtre çıkışıdır.
2. (2.6) eşitliği ile, $r_2 = r(T^{(1)})$ hesaplanır. Fakat bu defa özilintinin hesaplanması için kullanılan zaman zarfı, işaretin önce tam dalgalı doğrultucudan daha sonra da alçak geçiren filtreden geçirilmesi ile elde edilir. Alçak geçiren filtrenin frekans yanıtı, $\omega = 0$ için 0'dır. Böylece tam dalgalı doğrultulmuş sinyal düzgünleştirilir.
3. İlgili bandın seslilik değeri eşitlik (2.9) ile hesaplanır.

$$vs = \max(r_1, r_2) \quad (2.9)$$

Yukarıda verilen adımlar, dört filtre için de tekrarlanarak, dört bant için de vs_2, vs_3, vs_4, vs_5 seslilik değerleri hesaplanır.

Kazanç Hesabı:

Girdi ses işaretinin kazancı, perde uyarlanır pencere boyu kullanılarak, her çerçeve için iki kez hesaplanır. Pencere boyu, her iki kazanç hesabı için aynıdır ve aşağıdaki gibi hesaplanır:

1. Eğer $vs_1 > 0.6$ ise, pencere boyu $T^{(1)}$ (birinci aşama son perde periyodu)'in 120 örnekten büyük olan en küçük katıdır. Eğer boy 320 örneğe çıkarsa, ikiye bölünür.
2. Eğer $vs_1 \leq 0.6$ ise, pencere boyu 120 örnek olarak alınır. Böylece sessiz ve seğirmeli çerçeveler için varsayılan pencere boyu kullanılır.

İlk pencere için kazanç değeri g_1 olarak adlandırılır. Bu pencerenin merkezi, o anki çerçevenin son örneğinden 90 örnek önceki örnektir. İkinci pencere için hesaplanan kazanç değeri g_2 olarak adlandırılır. İkinci pencerenin merkezi, o anki çerçevenin son örneğidir. Kazanç hesabı, N pencere boyu, $s[n]$ girdi sesi sinyali olmak üzere, eşitlik (2.10) ile yapılır.

$$g = 10 \log_{10} \left(0.01 + \frac{1}{N} \sum_n s^2[n] \right) \quad (2.10)$$

Eğer kazanç 0'dan küçük çıkarsa, 0 olarak kabul edilir. Kazanç hesabında, girdi sinyalinin 16 bit ile örneklendiği varsayılır.

Doğrusal Öngörüm Analizi:

Doğrusal Öngörüm Analizi, günümüzde kullanılan konuşma kodlama algoritmalarının çoğunun, temelini oluşturmaktadır. Temel düşünce, bir konuşma sinyali örneğinin, daha önceki örneklerin doğrusal birleşimi ile ifade edilebilir olduğudur. Bir sinyal çerçevesi içinde, doğrusal birleşimi hesaplamak için kullanılan ağırlıklar, öngörüm hatasının ortalama karesi enküçültülerek bulunur. Elde edilen ağırlıklar ya da başka bir deyişle doğrusal öngörüm katsayıları, ilgili çerçeveyi ifade etmek için kullanılır. KUDÖK'te giriş sinyaline, merkezi güncel çerçevenin son örneği olan 200 örneklik bir Hamming penceresi uygulanarak, 10. dereceden doğrusal öngörü analizi yapılır.

a_i , doğrusal öngörücü parametreler olmak üzere, güncel ses örneği $s[n]$ 'i, M adet geçmiş örneğin doğrusal birleşimi ile elde eden M. dereceden bir doğrusal öngörücü, bu işlemi (2.11)

eşitliği ile yapar.

$$\hat{s}[n] = -\sum_{i=1}^M a_i s[n-i] \quad (2.11)$$

Öngörüm, doğrusal olduğundan, ileri öngörü hatası aşağıdaki eşitlik ile hesaplanır.

$$e[n] = s[n] - \hat{s}[n] \quad (2.12)$$

Öngörü parametreleri, ileri öngörü hatasının ortalama karesi en küçük olacak şekilde belirlenir. Bu enküçültme sonucu, aşağıdaki eşitlik elde edilir.

$$\sum_{i=1}^M a_i R_s(i-k) = -R_s[k] \quad k = 1, 2, \dots, M \quad (2.13)$$

Burada, $E\{\}$ istatistiksel beklenti işleci olmak üzere, R_s eşitlik (2.14) ile ifade edilir.

$$R_s[m] = E\{s[n]s[n+m]\} \quad (2.14)$$

$R_s[l]$, $s[n]$ 'nin özilinti işlemini ifade etmektedir. Eşitlik (2.13), $R_s[l]$ açısından, optimum doğrusal öngörü katsayılarını ifade etmektedir. Bu matris formunda, aşağıdaki gibi yazılabilir.

$$R_s a = -r_s \quad (2.15)$$

Burada,

$$R_s = \begin{pmatrix} R_s[0] & R_s[1] & \dots & R_s[M-1] \\ R_s[1] & R_s[0] & \dots & R_s[M-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_s[M-1] & R_s[M-2] & \dots & R_s[0] \end{pmatrix} \quad (2.16)$$

$$a = [a_1 \quad a_2 \quad \dots \quad a_M]^T \quad (2.17)$$

$$r_s = [R_s[1] \quad R_s[2] \quad \dots \quad R_s[M]]^T \quad (2.18)$$

olarak gösterilir.

Doğrusal öngörü parametrelerinin veri örnekleri dizisinden elde edilmesinde en çok kullanılan yöntem özilinti yöntemidir. Bu yöntemde özilintiler ($R_s[l]$), orijinal ses örneklerinden kestirilir ve (2.16) eşitliğinde yerine koyulur. Özilinti yönteminde, işaret önce, Hamming penceresinden geçirilir. $R_s[m]$ özilinti dizisi, N adet ses örneği ile aşağıda verilen eşitlik ile

kestirilebilir.

$$R_s[m] = \frac{1}{N-|m|} \sum_{i=0}^{N-|m|-1} s[n+|m|]s[n] \quad (2.19)$$

Eşitlik (2.16)'daki matris simetriktir. Köşegen elemanları birbirine eşittir. Bu eşitliği çözmek için kullanılan en efektif yöntemlerden biri Levinson-Durbin yöntemidir. m. dereceden bir öngörücü için Levinson-Durbin yöntemi aşağıdaki gibi özetlenebilir.

1. $l = 0$ için ilklendirme

$$j_0 = R_s[0] \quad (2.20)$$

2. $l = 1, 2, \dots, M$ için özyineleme

1.Adım:

$$k_l = \frac{R_s[l] + \sum_{i=1}^{l-1} a_i^{(l-1)} R_s[l-i]}{j_{l-1}} \quad (2.21)$$

$$a_l^{(l)} = -k_l \quad (2.22)$$

$$a_i^{(l)} = a_i^{(l-1)} - k_l a_{l-i}^{(l-1)} \quad i = 1, 2, \dots, l-1 \quad (2.23)$$

Eğer $l = M$ ise dur.

3.Adım:

$$j_1 = j_{l-1} (1 - k_l^2) \quad (2.24)$$

$$l \leftarrow l + 1 \quad (2.25)$$

1.Adıma dön.

Eşitlik (2.20) ile eşitlik (2.25) arasındaki eşitlikler özyinelemeli olarak çözüldükten sonra, doğrusal öngörüm katsayıları, eşitlik (2.26) ile bulunur.

$$a_i = a_i^{(M)} \quad i = 1, 2, \dots, M \quad (2.26)$$

Doğrusal öngörü uygulamalarında, nicemleme ya da işlemler sırasındaki duyarlılık kaybı

nedeniyle filtrelenen ses stabil olmayabilir. Doğrusal öngörü katsayıları, eşitlik (2.27)'ye göre değiştirilerek stabilite sağlanır.

$$a_i^{new} = \gamma^i a_i \quad i = 1, 2, \dots, M \quad (2.27)$$

γ katsayısı 1'den küçük olduğunda, orijinal spektrum bant genişlikleri artar. KUDÖK'te γ katsayısı 0.994 olarak alınmıştır.

Ses sinyalinin, bulunan doğrusal öngörü katsayıları ile ters filtrelenmesi ile, kalıntı işareti elde edilir. Kalıntı işareti, son perde değeri ve tepesellik hesaplamasında kullanılır.

Tepesellik Hesaplanması:

Kalıntı işaretinin tepesellik değeri, merkezi güncel çerçevenin son örneği olan 160 örneklilik bir pencere kullanılarak hesaplanır. Tepesellik değeri, eşitlik (2.28) ile ifade edilir.

$$p = \frac{\sqrt{\frac{1}{160} \sum_{n=-80}^{79} e^2[n]}}{\frac{1}{160} \sum_{n=-80}^{79} |e[n]|} \quad (2.28)$$

Tepesellik değerine göre, en alçak üç bandın seslilik değerleri aşağıdaki şekilde değiştirilir.

$$p > 1.34 \text{ ise } vs_1 \leftarrow 1$$

$$p > 1.60 \text{ ise } vs_2 \leftarrow 1 \text{ ve } vs_3 \leftarrow 1$$

Son Perde Değerinin Hesaplanması:

Son perde değerinin hesaplanması için, doğrusal öngörüm analizi sonunda elde edilen kalıntı işareti, kesim frekansı 1 kHz olan 6. dereceden Butterworth alçak geçiren filtreden geçirilir. Birinci aşama perde periyodu değerinin ($T^{(1)}$) beş örnek öncesinden beş örnek sonrasına kadar, tamsayı perde araması yapılır. Sinyal kalıntısı üzerinde yeniden perde periyodu araması yapılması, kalıntı sinyali, üzerinde formant bilgisi taşımadığından, daha doğru sonuç verir.

Doğrusal Öngörüm Katsayılarından Çizgisel İzge Frekanslarının Elde Edilmesi ve Nicemlenmesi:

Çizgisel Spektral Frekansları, ÇSF (Line Spectral Frequency, LSF), ilk olarak Itakura tarafından, 1975 yılında, DÖ katsayılarının alternatif bir betimlemesi olarak kullanılmıştır. Getirdiği avantajlar sayesinde, konuşma kodlama uygulamalarında çok büyük bir kullanım

oranına sahiptir. KUDÖK algoritması da DÖ katsayılarını, ÇSF'ları şeklinde ifade eder ve nicemler.

M. dereceden, DÖ katsayıları a_i , $i = 1, 2, \dots, M$ olmak üzere, DÖ katsayılarına denk düşen ÇSF'lar $0 < \omega_i < \pi$ olmak üzere $\omega_i, i = 1, \dots, M$ şeklinde ifade edilir.

DÖ katsayılarından, ÇSF'larına geçiş için izlenmesi gereken adımlar aşağıdaki gibidir;

M çift ise; $M_1 = M_2 = M/2$,

M tek ise; $M_1 = (M+1)/2, M_2 = (M-1)/2$,

$$p_i = a_i + a_{M-i+1}; \quad i = 1, \dots, M_1$$

$$q_i = a_i + a_{M-i+1}; \quad i = 1, \dots, M_2$$

$$p'_0 = 1$$

$$p'_i = p_i - p'_{i-1}; \quad i = 1, \dots, M_1$$

M çift ise;

$$q'_0 = 1$$

$$q'_i = q_i + q'_{i-1}; \quad i = 1, \dots, M_2$$

M tek ise;

$$q'_0 = 1, \quad q'_1 = q_1$$

$$q'_i = q_i + q'_{i-2}; \quad i = 1, \dots, M_2$$

Yukarıdaki polinom katsayıları kullanılarak, aşağıdaki fonksiyonlar gerçekleştirir;

$$P_0(\omega) = 2 \cos(M_1 \omega) + 2p'_1 \cos((M_1 - 1) \omega) + \dots + 2p'_{M_1-1} \cos \omega + p'_{M_1}, \quad (2.29)$$

$$Q_0(\omega) = 2 \cos(M_2 \omega) + 2q'_1 \cos((M_2 - 1) \omega) + \dots + 2q'_{M_2-1} \cos \omega + q'_{M_2}, \quad (2.30)$$

$0 < \omega < \pi$ aralığı için, $P_0(\omega)$ ve $Q_0(\omega)$ fonksiyonlarının kökleri, ÇSFlarıdır. Daha önce de

belirtildiği gibi, $P_0(\omega)$ ve $Q_0(\omega)$ fonksiyonları M 'e bağlıdır. Pek çok uygulamada olduğu gibi KUDÖK algoritmasında da $M=10$ olarak seçilmiştir. Kosünüs hesaplama, algoritma hızı ve karmaşıklığı açısından pratik olmadığı için, kök hesaplamaları boyunca, $\cos(\omega) = x$ eşitliği kullanılır. Bu eşitlik ile hesaplanan kökler, x_i ile gösterilmek üzere,

$$\omega_i = \cos^{-1}(x_i) \quad (2.31)$$

formülü ile ÇSF'ları elde edilir.

$M=10$ için, $P_0(x)$ ve $Q_0(x)$ fonksiyon, $\cos(\omega) = x$ eşitliği kullanılarak, aşağıdaki şekilde tekrar yazılabilir;

$$P_0(x) = 32x^5 + 16p'_1x^4 + 8(p'_2 - 5)x^3 + 4(p'_3 - 4p'_1)x^2 + 2(p'_4 - 3p'_2 + 5)x + p'_5 - 2p'_3 + 2p'_1 \quad (2.32)$$

$$Q_0(x) = 32x^5 + 16q'_1x^4 + 8(q'_2 - 5)x^3 + 4(q'_3 - 4q'_1)x^2 + 2(q'_4 - 3q'_2 + 5)x + q'_5 - 2q'_3 + 2q'_1 \quad (2.33)$$

Vektör nicemleme, VN, (Vector Quantization, VQ) sürekli bir bilgi kaynağı topluluğu ile ayrık bir bilgi topluluğu arasında, çok boyutlu bir uzayda eşleme yapmayı hedefler. VN'nin doğrudan gerçekleşmesi yer ve zaman açısından çok maliyetlidir. Çok Aşamalı Vektör Nicemleme, ÇAVN, (Multi Stage Vector Quantization, MSVQ), VN 'nin karmaşıklığı daha az olan özel bir şeklidir. ÇAVN, gerçekleşme maliyetlerinin azalmasını sağlamıştır ancak pratik uygulamalarda, performans düşüşüne de sebep olmaktadır. KUDÖ kodlayıcısı, ÇSF larını, ÇAVN ile nicemler. ÇAVN, 128, 64, 64, 64 seviyeli 4 aşamadan ve 64 kod sözcüğünden oluşur. ÇAVN, ağırlıklandırılmış Euclidian uzaklığını en küçükten kod vektörlerini hesaplar. Ağırlıklandırılmış uzaklık hesabı, eşitlik (2.34) ile hesaplanır.

$$d(x, \hat{x}) = (x - \hat{x})^T W (x - \hat{x}) = \sum_{i=1}^M w_i (\omega_i - \hat{\omega}_i)^2 \quad (2.34)$$

Burada W , diyagonal elemanları w_i olmak üzere bir diogonal matristir. Bu elemanlar x 'e bağlı olacaktır. $S(e^{j\omega_i})$, ÇSF'ları seti $\{\omega_i\}$ ile tanımlanan güç spektrumunu göstermek üzere, optimum ÇSF nicemleme için, ağırlık parametreleri olan w_i 'ler aşağıdaki gibi seçilmelidir.

$$w_i = \begin{cases} S(e^{j\omega_i})^{0.3}, & i \leq 8, \\ 0.64[S(e^{j\omega_i})]^{0.3}, & i = 9, \\ 0.16[S(e^{j\omega_i})]^{0.3}, & i = 10, \end{cases} \quad (2.35)$$

Ağırlıklandırılmış Euclidian uzaklıklarını en küçükten kod vektörlerinin bulunması için arama işlemi, her aşamadan en iyi sonuç veren 8 kod vektörünün bulunması ve sonraki aşamaya aktarılacak üzere kaydedilmesi ile yapılır.

Bant İçi Seslilik Nicemleme:

Bölüm 1.3'te anlatıldığı şekilde hesaplanan bant içi seslilik değerleri, KUDÖ kod çözücüyeye gönderilmeden önce aşağıdaki yarı-kodlamada anlatıldığı gibi nicemlenmelidir.

```

NICEMLE_VS (vs1, ..., vs5)
1.  if vs1 ≤ 0.6           // Sessiz
2.    for i ← 2 to 5
3.      qvsi ← 0
4.  else                   // Sesli
5.    for i ← 2 to 5
6.      if qvsi > 0.6
7.        qvsi ← 1
8.      else
9.        qvsi ← 0
10.   if qvs2 = 0 and qvs3 = 0 and qvs4 = 0
11.     qvs5 ← 0
12.   return qvs2, ..., qvs5

```

Bu yarı-kod, giriş olarak 5 bant içi seslilik değerini alır. Sesli Olmama durumunda, (vs₁) değerine göre karar verilir) diğer dört bant içi seslilik değerine sıfır atanır. Diğer durumlarda, bu değerlere, büyüklüklerine göre sıfır ya da bir atanır.

Perde ve Alçak - Bant Seslilik Nicemleme:

Son perde periyodu değeri T ve alçak-bant seslilik değeri vs₁, birlikte yedi bit ile kodlanır. Eğer vs₁ ≤ 0.6 ise, çerçeve sessizdir ve tüm bitleri sıfır olan bir kod gönderilir. Aksi durumda, log T, log 20 ve log 160 aralığında 99 seviyeli düzgün nicemleme ile nicemlenir. Elde edilen yedi bit, kod çözücüyeye gönderilir.

Kazanç Nicemleme:

Kazanç g_2 , 10 ile 77 dB aralığında, beş bitlik, düzgün nicemleme ile nicemlenir. Kazanç g_1 ise aşağıdaki yarı-kodda anlatıldığı gibi üç bit ile nicemlenir.

```

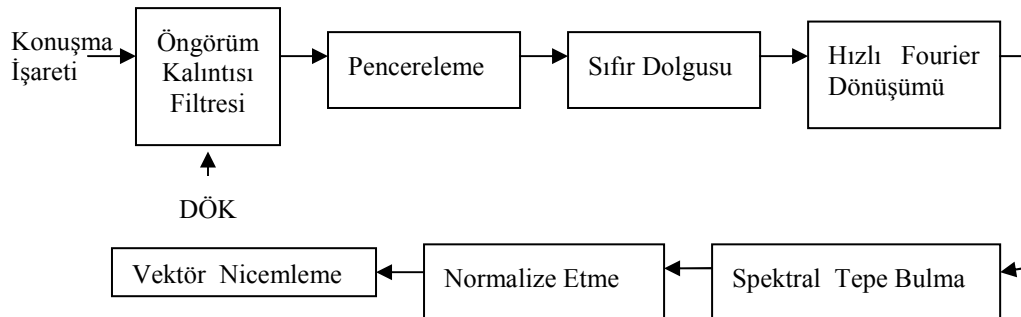
NICEMLE_g1( $g_1, g_2, g_{2,past}$ )
1.  if  $|g_2 - g_{2,past}| < 5$  and  $|g_1 - (g_2 + g_{2,past})/2| < 3$ 
2.    index  $\leftarrow 0$ 
3.  else
4.    g max  $\leftarrow MAX(g_{2,past}, g_2) + 6$ 
5.    g min  $\leftarrow MIN(g_{2,past}, g_2) - 6$ 
6.    if g min  $< 10$ 
7.      g min  $\leftarrow 10$ 
8.    if g max  $> 77$ 
9.      g max  $\leftarrow 77$ 
10.  index  $\leftarrow DUZGUN\_NICEMLE(g_1, g_{min}, g_{max}, 7)$ 
11.  return index

```

Bu yarı-kod, güncel çerçevenin g_1 ve g_2 kazançları ile bir önceki çevrenin g_2 kazancını ($g_{2,past}$) alır. İlk satırdaki koşulların sağlanması halinde, çerçeve kalıcı durumdadır ve nicemleme sonucunda tüm bitleri sıfır olan bir kod gönderilir. Aksi durumda ise, çerçeve geçicidir ve yedi seviyeli düzgün nicemleyici ile nicemlenir.

Fourier Genliklerinin Hesaplanması ve Nicemlenmesi:

KUDÖ kodlayıcı, uyarım darbesinin şeklini yakalamak için, öngörüm kalıntısının Fourier büyüklüklerini kullanır. Fourier büyüklüklerinin hesaplanması için izlenen adımlar, aşağıda gösterilmiştir.



Şekil 2.4. Fourier Büyüklük Hesabı Adımları

Fourier büyüklüklerinin hesaplanması öncesinde, DÖ katsayıları ve perde periyodu değerleri hesaplanmıştır. Konuşma işareti, öngörüm kalıntısı filtresinden geçirilerek, öngörüm kalıntısı hesaplanmıştır. Kalıntıya 200 örneklik Hamming penceresi uygulanır. Pencereleme işlemi, Hızlı Fourier Dönüşümü, HFD, (Fast Fourier Transform, FFT) için standart bir ön işlemdir. Sonuçta elde edilen 200 örneklik bileşkenin, 512 örneğe kadar sıfır ile dolgulanması ve HFD uygulaması ile 512 karmaşık değer elde edilir. HFD, Ayrık Fourier Dönüşümü, AFD, (Discrete Fourier Transform, DFT) hesaplamının verimli bir yoludur. DFT, aşağıdaki gibi tanımlanabilir;

$$X[k] = \sum_{n=0}^{N_0-1} x[n] e^{-j(2\pi k n / N_0)}; \quad k = 0, \dots, N_0 \quad (2.36)$$

Fourier genlikleri hesaplandıktan sonra, fourier büyüklüklerinin, tepe büyüklüklerinin bulunması için, aşağıdaki yarı-kod işletilmelidir;

```

TEPE _ BUYUKLUGU _ ARAMA(X[0...255], T, Fmag[1...10] )
1.  for i ← 1 to 10
2.    freq ← round(512 * i / T)
3.    if freq > 255
4.      Fmag[i] ← ENKUCUK _ DEGER
5.    else
6.      peak ← 0
7.      for j ← freq - 5 to freq + 5
8.        if j > 255 break
9.        if |X[j]| > peak
10.         peak ← |X[j]|
11.      Fmag[i] ← peak
12.  return Fmag[1...10]

```

Bu yarı-kodun amacı, T, perde periyodu olmak üzere, $512i/T$, $i = 1, \dots, 10$ frekans örnekleri için, tepe büyüklüğü araması yapmaktır. Eğer frekans değeri 255'in üstündeyse, harmonikler değerlendirme aralığı dışındadır ve ona karşılık gelen fourier büyüklüğüne, sabit değer olarak belirlenen ENKUCUK_DEGER değeri atanır. Bu durum, perde periyodu değeri çok büyük olduğunda ortaya çıkar. Eğer frekans değeri 255'e eşit veya küçükse, $(512i/T) \pm 5$, $i = 1, \dots, 10$ frekans örnekleri için arama yapılır. Bu yarı-kod sonunda 10 tepe büyüklüğü ($Fmag[i]$, $i = 1$ to 10) elde edilir.

Sonraki adımda ise, fourier büyüklükleri dizisi, aşağıdaki eşitliğe göre normalize edilir.

$Fmag'[i]$ normalize edilmiş fourier büyüklükleri olmak üzere;

$$Fmag'[i] = \alpha * Fmag[i] \quad (2.37)$$

Burada, α aşağıdaki gibi ifade edilebilir.

$$\alpha = \left(\frac{1}{10} \sum (Fmag[i])^2 \right)^{-1/2} \quad (2.38)$$

Normalize edilmiş fourier büyüklüğü dizisi, 8 bitlik vektör nicemleme ile nicemlenir.

Gönderme Hatası Düzeltme ve Bit Paketleme:

Bu aşamada, nicemlenmiş bitler kullanılarak, bit paketi oluşturulur. Sessiz çerçeveler için, fourier büyüklükleri, bant içi seslilik değerleri ve periyodik olmayan bayrak parametrelerinin aktarılması gerekli olmadığından, bu parametreler, hata düzeltimi bitleri ile değiştirilir. Hata düzeltimi için Hamming kodu kullanılır.

Aşağıdaki tabloda, KUDÖ kodlayıcının kullandığı bit paketi yapısı görülmektedir. Senkronizasyonu sağlamak için, ardışık çerçeveler için sıfır ve bir değeri alan senkronizasyon biti kullanılır. Sesiz çerçevelerde, 13 bit hata düzeltme için ayrılmış olup, 22.5 ms'lik her çerçeve için 54 bit gönderilir. Sonuç olarak KUDÖK, 2400 bit/saniye'lik bir veri aktarım hızına sahiptir.

Çizelge 2.1 KUDÖK Bit Paketi

Parametre	Sesli	Sessiz
ÇSF	25	25
Perde Periyodu/Alçak Bant Seslilik	7	7
Bant İçi Seslilik	4	-
Kazanç 1	3	3
Kazanç 2	5	5
Periyodik Olmayan Bayrak	1	-
Fourier Büyüklükleri	8	-
Senkronizasyon	1	1
Hata Koruma	-	13
Toplam	54	54

2.1.2 KUDÖK ile Konuşma Sentezi

KUDÖK ile konuşma sentezinde, kodlayıcıdan gelen bit paketi açılır ve parametreler kullanılarak, konuşma sentezi yapılır. Klasik konuşma sentez algoritmalarına ek olarak, KUDÖK kod çözücü spektral iyileştirme ve darbe saçılım filtresi olmak üzere, iki yeni filtreye sahiptir. KUDÖK ile konuşma sentezi aşağıdaki adımlardan oluşmaktadır.

- Parametre Çözme ve Aradeğerleme
- Karma Uyarımın Oluşturulması
- Darbe Uyarımın Oluşturulması
- Gürültü Uyarımın Oluşturulması
- Biçimlendirme Süzgeçleri
- Uyarlanır Spektral İyileştirme Filtresi
- DÖK Sentez Filtresi
- Ölçekleme Katsayısı Hesaplama
- Darbe Saçma Filtresi

Parametre Çözme ve Aradeğerleme:

KUDÖK kod çözücü, kodlayıcıdan gelen bit paketini açarak, ÇSF, perde periyodu, alçak bant seslilik ve bant içi seslilik değeri, kazanç₁, kazanç₂, periyodik olmayan bayrak ile fourier büyüklükleri parametrelerini çözer. Bu parametreler, ilgili çerçeve için akustik özellikleri ifade eder ve sentez sırasında, doğrusal olarak aradeğerlenir (interpolation).

Sessiz çerçeveler için, yukarda bahsedilen parametrelere varsayılan bir takım sabit değerler atanır. Bu sabit değerler, perde periyodu için 50, seğirme için 0.25, tüm Fourier büyüklükleri için 1 ve tüm seslilik değerleri için 0'dır. Sesli çerçeveler için seğirme parametresi, periyodik olmayan bayrağa göre aşağıdaki gibi belirlenir;

$$\text{periyodik olmayan bayrak} = 1 \text{ ise } \text{seğirme} \leftarrow 0.25$$

$$\text{periyodik olmayan bayrak} \neq 1 \text{ ise } \text{seğirme} \leftarrow 0$$

Seğirmenin 0 olması durumunda, perde periyodu, bit paketinden çıkarılır. Gerçek perde periyodu aşağıdaki eşitlik ile bulunur;

$$T = T_0 * (1 + \text{seğirme} * x) \quad (2.39)$$

Burada, T_0 , bit paketinden çıkarılan ve aradeğerlenen perde periyodu, x ise, $[-1,1]$ sınırları

arasında eşit olasılıklı rastgele bir sayıdır.

Konuşma sentezi boyunca, sinyal, bir perde periyodundan, diğer perde periyoduna hesaplanır. Sentez için gerekli parametreler, geçmiş çerçeve ve güncel çerçeve verisinden elde edilir. Aradeğerleme etmeni α , aşağıdaki gibi ifade edilebilir;

$$\alpha = n_0 / 180 \quad (2.40)$$

Aradeğerleme etmeni, ÇSF, perde periyodu, fourier büyüklükleri ve biçim filtresi katsayılarına uygulanır. Seslilik değerlerine de aradeğerleme etmeni uygulanabilse de, bunun yerine biçim filtresi katsayılarına uygulanması, hesaplama karmaşıklığını düşürür. Perde periyodu için aradeğerleme etmeni aşağıdaki gibi uygulanır;

$$T = (1 - \alpha) * T_{past} + \alpha * T_{present} \quad (2.41)$$

Elde edilen sonuç en yakın tam sayıya yuvarlanır. Kullanılacak kazanç parametresini elde etmek için aşağıdaki aradeğerleme formülü kullanılır;

$$g = \begin{cases} (1 - \alpha) * g_{2,past} & n_0 < 90 \\ (1 - \alpha) * g_{1,present} & 90 \leq n_0 < 180 \end{cases} \quad (2.42)$$

Karma Uyarımın Oluşturulması:

KUDÖK kod çözücü, fourier büyüklüklerini kullanarak, T örneklik darbe uyarımını hesaplar. Biçimlendirici filtre ile filtrelenmiş darbe uyarım dizisi ve filtrelenmiş gürültü uyarımlarının toplanması ile karma uyarım oluşturulur.

Darbe Uyarımın Oluşturulması:

Önceki bölümlerde belirtildiği gibi, konuşma sentezi boyunca, sinyal, perde periyodundan, perde periyoduna hesaplanır. Dolayısıyla kod çözme sırasında, T örneklik darbe uyarımı hesaplanır. Darbe uyarımı hesaplanmasına ilişkin yarı-kodda aşağıda verilmiştir.

```

DARBE_HESAPLAMA(Fmag[1...10], T, y[0...T-1])
1.  y[0] ← 0
2.  for k ← 11 to T-11
3.    y[k] ← 1
4.  for k ← 1 to 10
5.    y[k] ← Fmag[k]
6.    y[T-k] ← Fmag[k]
7.  ATFD(y[0...T-1], y[0...T-1], T)
8.  DAIRESEL_KAYDIRMA(y[0...T-1], T, 10)
9.  NORMALIZE_ET(y[0...T-1], T)
10. return y[0...T-1]

```

Darbe hesaplama yarı-kodu, T örneklilik dizi y'yi oluşturmak üzere, fourier büyüklüklerini ve perde periyodunu girdi olarak alır. y dizisinin ilk elemanına 0 ve diğerlerine 1 atar. 10 fourier büyüklüğü y dizisine $y[1...10]$ ve $y[T-10...T-1]$ aralıklarına yerleştirilir ve böylece simetrik bir dizi elde edilir. 7. satırda, aşağıdaki formül kullanılarak ayrık ters fourier dönüşümü, ATFD (Inverse Discrete Fourier Transform, IDFT) uygulanır.

$$y[n] = \frac{1}{T} \sum_{k=0}^{T-1} y[k] e^{j2\pi nk/T}; \quad T-1 \quad (2.43)$$

y dizisinin simetrik olması, ATFD sonrasında gerçek değerli bir dizi elde edilebilmesi için önemlidir.

8. satırda, y[n] dizisinin son hali, 10 örnek dairesel kaydırılır ve 9. satırda normalize edilir. Böylece, y dizisi birim etkin değere (unit root mean square) sahip olur.

Gürültü Uyarımın Oluşturulması:

Gürültü uyarımı, (-1732) – (+1732) aralığında rasgele sayılardan oluşur.

Biçimlendirme Süzgeçleri:

KUDÖK konuşma sentezi modeli, darbe uyarımı ile gürültü uyarımını birleştirmek için, iki biçimlendirme filtresi kullanır. Bu filtrelerin yanıtları, seslilik değerleri ile kontrol edilir. Bu filtreler, uyarım içindeki darbe ve gürültü miktarını ayarlar.

Her biçimlendirici filtre, kod çözme esnasında karma uyarım sinyali oluşturmak için kullanıldıklarından sentez filtreleri olarak adlandırılan beş filtreden oluşmaktadır. Her sentez filtresi, frekans aralıkları 0-500, 500-1000, 1000-2000, 2000-3000, 3000-4000 Hz olan bir

frekans bandını kontrol etmektedir. Sentez filtresinin dürtü yanıtı (impulse response), $h_i[n]$, $i = 1 \text{ to } 5$ olarak ifade edilmek üzere, darbe biçimlendirici filtrenin toplam yanıtı aşağıdaki gibi hesaplanır.

$$h_p[n] = \sum_{i=1}^5 v s_i h_i[n] \quad (2.44)$$

Gürültü biçimlendirici filtre yanıtı ise aşağıdaki gibi ifade edilir.

$$h_n[n] = \sum_{i=1}^5 (1 - v s_i) h_i[n] \quad (2.45)$$

İki filtre birbirinin tümleyeni olduğu için, birinin kazancı düşerken diğerininki artar ve iki filtrenin toplam kazancı sabit kalır.

Uyarlanır Spektral İyileştirme Filtresi:

Bu filtre, birinci derece eğim dengelemeli 10. derece kutup/sıfır filtresidir. Karma uyarım işaretine uygulanır. Uyarlanır spektral iyileştirme filtresinin sistem fonksiyonu aşağıdaki gibidir;

$$H(z) = (1 - \mu z^{-1}) \frac{1 + \sum_{i=1}^{10} a_i \beta^i z^{-i}}{1 + \sum_{i=1}^{10} a_i \alpha^i z^{-i}} \quad (2.46)$$

Burada a_i doğrusal öngörüm katsayılarını ifade eder. μ, α ve β işaret durumuna göre uyarlanır.

DÖK Sentez Filtresi:

Katsayıları, aradeğerlenmiş ÇSF'na karşılık gelen DÖK sentez filtresi kullanılır. a_i , orjinal sinyalin DÖ katsayıları olmak üzere, sentez filtresinin sistem fonksiyonu aşağıdaki gibidir;

$$H(z) = \frac{1}{1 + \sum_{i=1}^M a_i z^{-i}} \quad (2.47)$$

Ölçekleme Katsayısı Hesaplama:

Sentezlenen sesin gücü, kodlayıcıdan gelen aradeğerlenmiş kazanca eşit olmalıdır. Uyarım belirli bir seviyede üretildiğinden, sentez filtresi çıkışının gücü için ölçekleme katsayısı g_0 ,

hesaplanmalıdır. Böylece, uygun seviyede bir ses üretilebilir. Ölçekleme katsayısı hesabı aşağıda gösterilmiştir.

$$g_0 = \frac{10^{g/20}}{\sqrt{\frac{1}{T} \sum_n y^2[n]}} \quad (2.48)$$

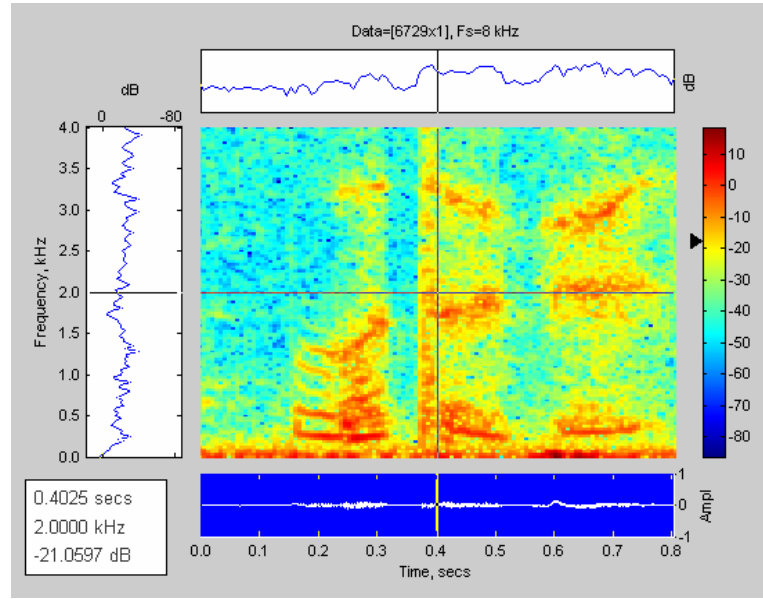
Sentez filtresi çıkışı $y[n]$ 'nin, g_0 ile çarpılması ile, T örneklilik dizinin gücü, g dB olur.

Darbe Saçma Filtresi:

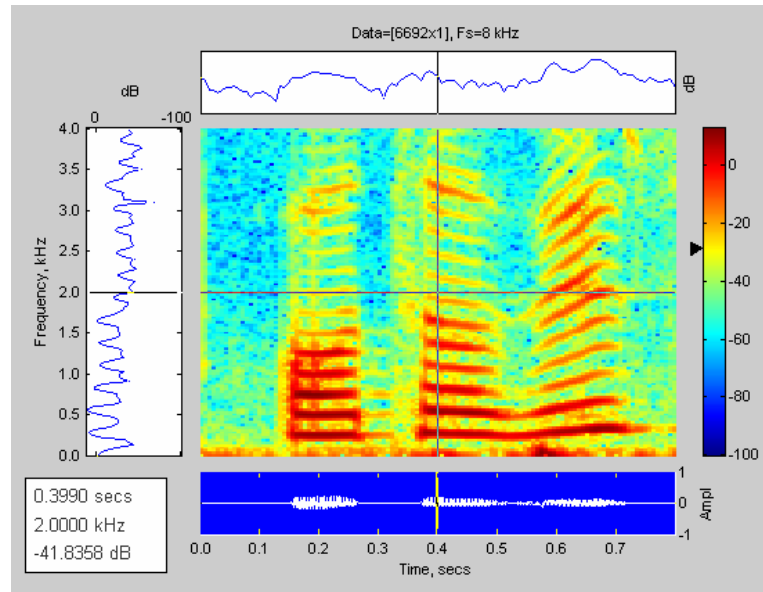
Bu işlem, kod çözücünün son işlemidir. Bu filtre aslında genlik yanıtı değişiklikleri oldukça küçük olan bir tür tüm geçiren süzgeçtir. McCree ve Barnwell (1995)'a göre darbe saçma filtresi, spektral rezonansın bulunmadığı bölgelerde, bant geçiren süzgeç ile süzgeçlenmiş sentetik konuşma ile, doğal konuşmanın uyumunu artırır. Böylece daha doğal bir konuşma elde edilir.

3. DİSFONİK SESİN AKUSTİK ÖZELLİKLERİ

Birinci bölümde anlatıldığı gibi kronik disfoni temel olarak ses tellerinin işlevsel bozukluğu sonucu ortaya çıkmaktadır. Dolayısıyla oluşan ses, fısıltılı konuşma özelliklerine benzer özellikler göstermektedir. Disfonik sesin, seslilik, perde ve formant yapısı özellikleri, normal sese göre farklılık göstermektedir. Disfonik seste algılanabilir bir perde periyodu bulunmaz ve tamamen gürültülüdür. /Otobüs/ kelimesi için normal ve disfonik sese ait spektrogramlar, Şekil 3.1’de görülmektedir.



(a)



(b)

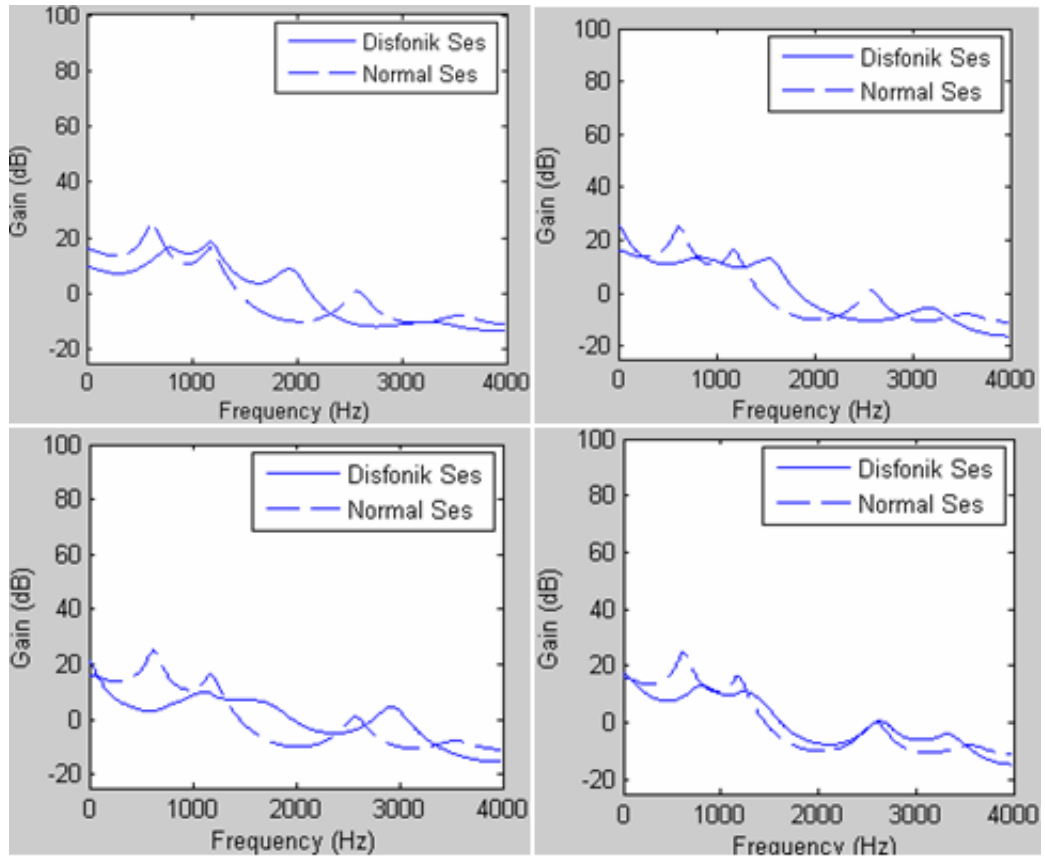
Şekil 3.1 (a) Disfonik Sesle Söylenmiş, (b) Normal Sesle Söylenmiş /Otobüs/ Kelimesi Spektrogramları

Türkçe’de sessizler ses tellerinin titreşimine göre ötümlü ve ötümsüz olarak ikiye ayrılmaktadır. Ötümlü sessizler (b, c, d, g, ğ, j, l, m, n, r, v, y, z), ses tellerinin titremesi ile, ötümsüz sessizler ise (p, ç, t, k, s, ş, f, h) ses telleri titreşimi olmadan oluşmaktadır.

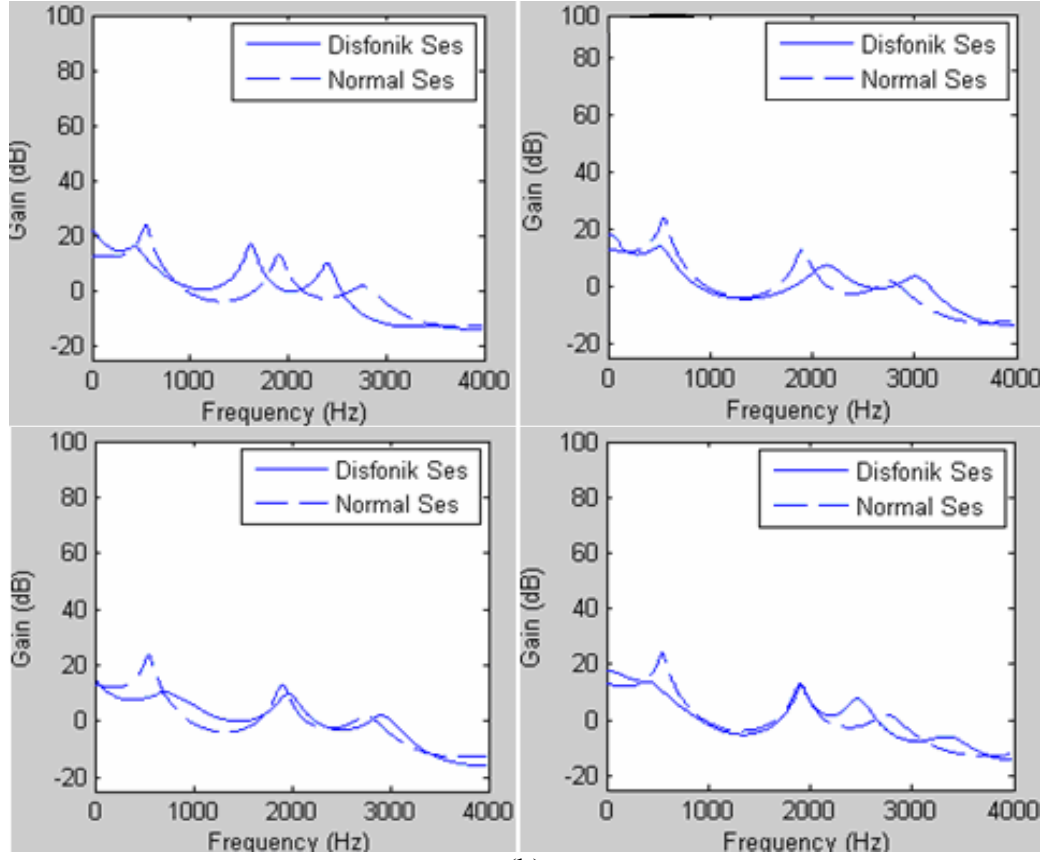
Yapılan analizlerde disfonik sese ait sesliler ve ötümlü sessizler ile ötümsüz sessizlerin yapısında meydana gelen değişiklikler ayrı ayrı incelenmiştir.

3.1 Sesliler ve Ötümlü Sessizlerin Akustik Farklılıkları

İncelenen disfonik seslerde, sesli harfler için özellikle ilk formant tepelerine ait formant bant genişliklerinin genellikle daha geniş ve formant frekansların daha büyük olduğu gözlenmiştir. /a/ fonemi için normal ses ve disfonik ses spektrumları Şekil 3.2’de görülmektedir.



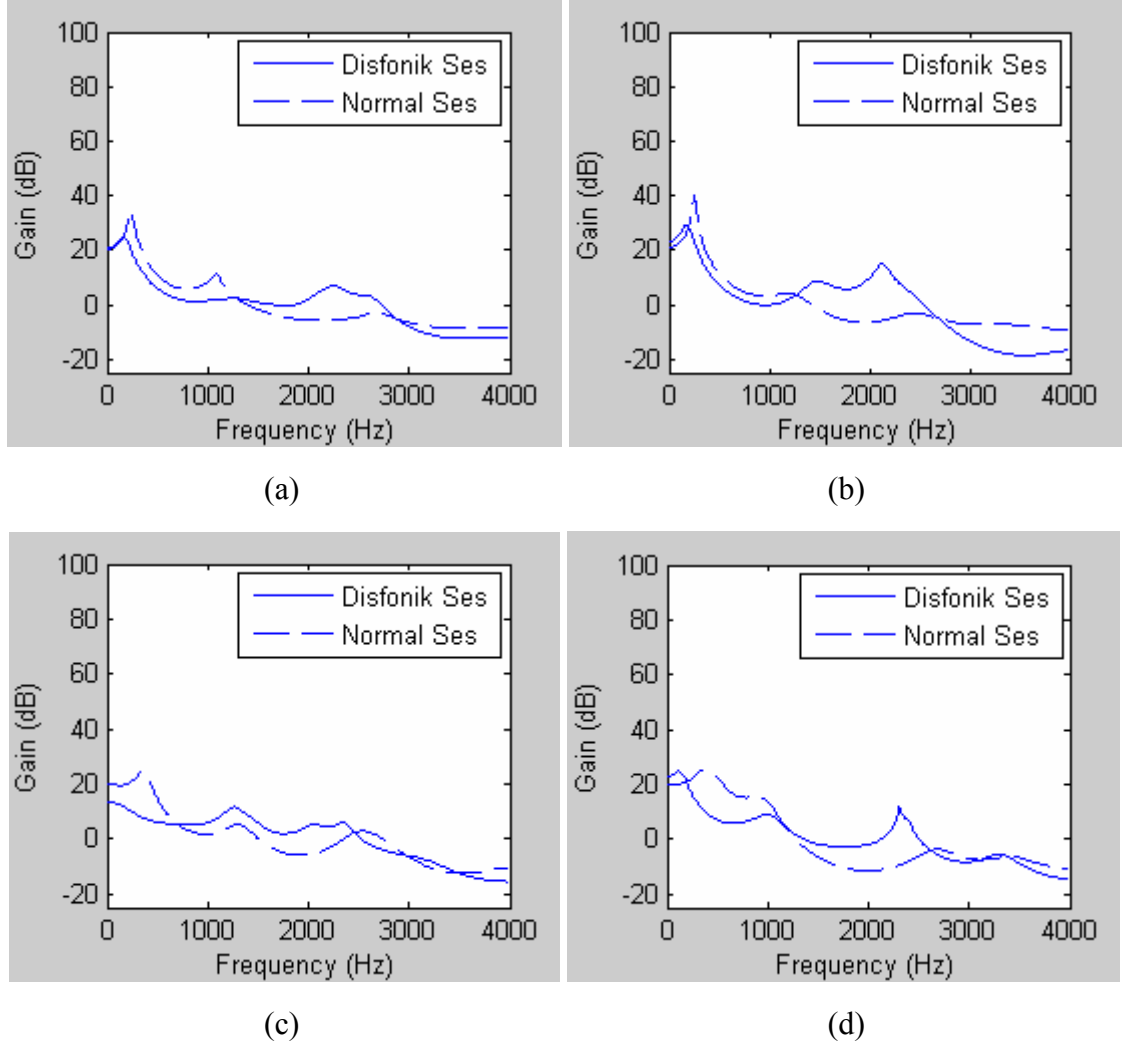
(a)



(b)

Şekil 3.2 Ses Teli Felci, Gırtlak Kanseri Hastalarına ve Normal Ses Sahip Konuşmacıya Ait
(a) /a/ Fonemi Spektrumları (b) /e/ Fonemi Spektrumları

Disfoni hastalarından alınmış ötümlü sessiz örneklerine ait spektrumlar ve normal sese sahip konuşmacıdan alınmış aynı seslere ait spektrumlar ise Şekil 3.3'de görülmektedir.

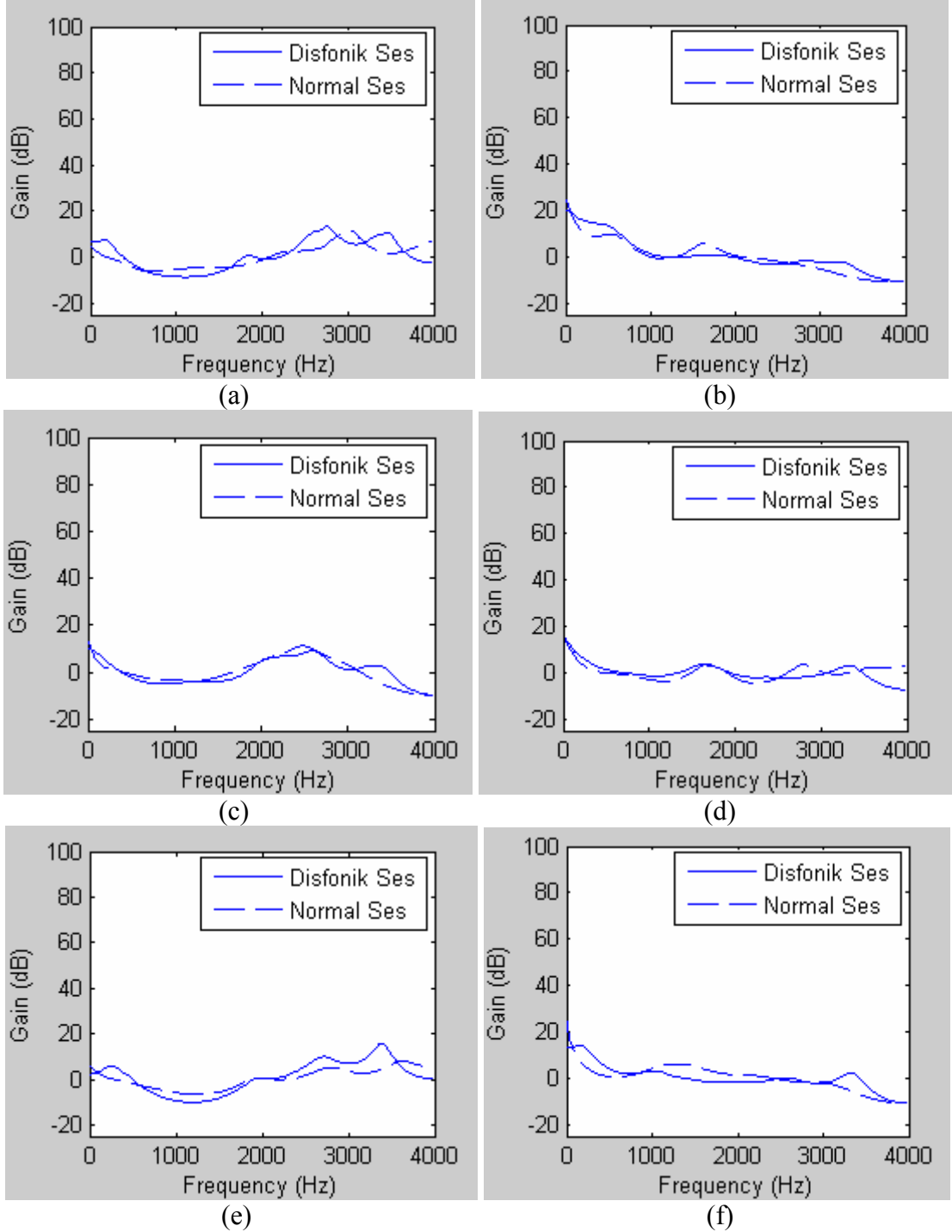


Şekil 3.3 Ses Teli Felci Hastasına ve Normal Sese Sahip Konuşmacıya Ait (a) /l/ (b) /m/ (c) /r/ (d) /v/ Fonemi Spektrumları

Şekil 3.3'de açıkça görüldüğü gibi, disfonik sese ait ötümlü sessizler de sesliler gibi, formant yapısı bakımından, normal sese göre farklılık göstermektedir. Bununla birlikte disfonik sese ait ötümlü sessizlerde ve seslilerde, normal sese ait ötümlü sessizler ve seslilerin aksine algılanabilir perde periyodu ve seslilik gözlenmemiştir.

3.2 Ötümsüz Sessizlerin Akustik Farklılıkları

Disfoni hastalarından alınmış ötümsüz ses örneklerine ait spektrumlar ve normal sese sahip konuşmacıdan alınmış aynı seslere ait spektrumlar Şekil 3.4 'de görülmektedir.



Şekil 3.4 Ses Teli Felci Hastasına ve Normal Sese Sahip Konuşmacıya Ait (a) /ç/ (b) /t/ (c) /k/ (d) /s/ (e) /ş/ (f) /f/ Fonemi Spektrumları

Şekil 3.4'te görüldüğü gibi, ötümsüz sesler için disfonik ve normal ses spektrumlarında belirgin değişiklik gözlenmemektedir. Bununla birlikte, normal sese sahip konuşmacıya ait ötümsüz sesler ve disfoni hastası konuşmacıya ait ötümsüz sesler için seslilik analizi yapıldığında, beş frekans bandı için de belirgin bir fark olmadığı görülmüştür. Ayrıca normal konuşmacıya ait ötümsüz sesler için algılanabilir bir perde periyodu değeri de görülmemiştir.

Şekil 3.2, 3.3 ve 3.4'te görüldüğü gibi, formant yapısındaki bozulmalar hastadan hastaya değişiklik göstermektedir. Sistemin formant yapısı değiştirme adımı için genel geçerli kurallar belirlemek yerine, hastanın ses özelliklerine göre, formant yapısı değişikliği yapmak faydalı olacaktır.

4. KARMA UYARIMLI DOĞRUSAL ÖNGÜRÜM KODLAMASI İLE DİSFONİK KONUŞMADAN NORMAL KONUŞMA ELDE EDİLMESİ

Bu bölümde, bu tez çalışmasında geliştirilen Disfonik Konuşmadan Normal Konuşma Elde Etme Sisteminin detayları anlatılmıştır.

4.1 Sistem Kısıtları

Tasarlanan sisteminde düzeltilecek ses kaydı, bir baş mikrofonu ile alınmaktadır. Sistemin düzeltme başarısının yeterli olabilmesi için sistem bazı kısıtları beraberinde getirmektedir. Sesin alındığı ortamda, arka plan gürültüsü olmamalı, gırtlığında delik bulunan hastalarda, delikten çıkan havanın, ses kaydını daha az etkileyebilmesi için mikrofon ağız hizasından en az 10 cm yuksekte bulunmalıdır.

4.2 Veri Tabanı Oluşturma

Bu çalışmada, 3 ses teli felçli hastaya ait ve 4 gırtlak kanserli hastaya ait sesler incelenmiştir. Her hastadan alfabedeki sesli harfler ile farklı uzunluklarda, fonetik olarak düzgün dağılımlı, toplamda 50 kelimeden oluşan cümle ve isim tamlaması kaydı alınmıştır.

Sistemin testi için kullanılan cümleler ve sesler beş grupta incelenebilir. Disfonik sesleri içeren, mevcut bir veritabanı bulunmadığından, sınırlı sayıda hasta için test yapılabilmiş ve sistem başarısının daha iyi gözlenebilmesi için her gruptan farklı cümlelere ait ses kayıtları, farklı hastalardan alınmıştır. Çizelge 4.1’de bu gruplara ait birkaç örnek görülmektedir.

Çizelge 4.1. Test İçin Kullanılan Örnekler

Örnekler	Ünlüler		Ünsüzler	
	Yuvarlak Geniş	Düz Dar	Ötümsüz	Ötümlü
Edirne		*		*
Japonya	*			*
Takım Çalışması		*	*	
Boğaz Köprüsü	*		*	*
Bafra Ovası	*			*
Vapur Seferleri		*	*	
Doğu Anadolu Bölgesi	*			*
Çocuk Masalları		*	*	
Elma Ağacı		*		*

4.3 Sistemin Genel Yapısı

Disfonik ses düzeltme sistemi, giriş sinyali olan disfonik sesin, KUDÖ kodlayıcı kullanılarak analiz edilmesi ile başlar. Analiz sonunda elde edilen ÇSF kullanılarak, giriş sinyalindeki ötümlü sessiz ve sesli içeren çerçeveler tespit edilir. Bu çerçeveler için kazanç, seslilik, perde ve formant yapısı Değişikliği yapılır. Son adımda ise ötümlü sessiz ve sesli içeren çerçevelerden, uyarlanmış parametreleri kullanılarak, ötümsüz sessiz içeren çerçevelerden ise orijinal parametreleri koruyarak sentetik ses elde edilir.

4.4 Ötümsüz Sessiz İçeren Çerçevelerin Tespiti

Ötümsüz sessiz tespiti için, şablon eşleştirme yöntemi kullanılmıştır. Eğitim seti için, farklı kişilere ait ötümlü sessiz ile sesli içeren fonemler ve ötümsüz sessiz içeren fonemler olmak üzere iki sınıf oluşturulmuştur. Oluşturulan sistemin testi için hazırlanan disfonik ses veritabanı kısıtlı sayıda örnek içerdiğinden, eğitim seti, disfonik sesler yerine özellik olarak benzerlik gösteren fısıltılı seslerden oluşturulmuştur. Eğitim seti, bu sınıflara ait her ses için ortalama ÇSF 'nı içerir. Sınıflandırılacak çerçeveye ait ÇSF kullanılarak, k-en yakın komşuluk yöntemi ile çerçevenin ait olduğu sınıf tespit edilir. Çizelge 4.2'de k=3 için yapılan sınıflandırma başarıları görülmektedir.

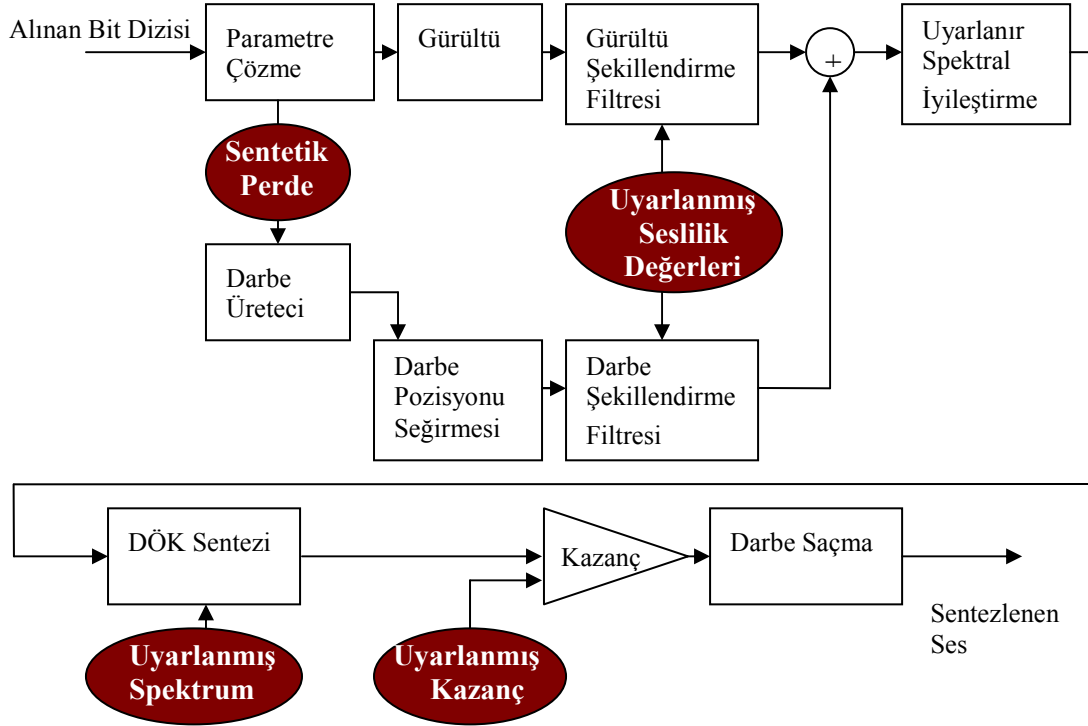
Çizelge 4.2. Fonem Gruplarının Sınıflandırılması Başarı Oranları

	Sesliler	Ötümlü Sessizler	Ötümsüz Sessizler
Ötümsüz sessiz olarak sınıflandırılma	5,12%	17,23%	74,38%
Ötümlü sessiz veya sesli olarak sınıflandırılma	94,88%	82,77%	25,62%

Sınıflandırma hataları incelendiğinde, ötümlü sessizlerden z, r, j ve g'nin %48 oranında yanlış sınıflandırıldıkları bunun yanında v, y, m, n, l sessizlerinin ortalama %95 'in üzerinde doğru sınıflandırıldığı gözlenmiştir. Ötümsüz sessizlerin sınıflandırılma hatası incelendiğinde h ve p sessizlerinin yaklaşık %42 oranında doğru sınıflandırıldığını, ç, ş, s, f, t, k sessizlerinin ise ortalama %85 oranında doğru sınıflandırıldığı görülmüştür.

4.5 Disfonik Ses Düzeltme Sisteminin Genel Yapısı

Tasarlanan düzeltme sistemi KUDÖ kod çözücünün kırmızı kutucuklar ile gösterilen adımlarında, gerekli değişiklikleri yaparak, sentetik ses üretir. Sistemin blok diyagramı Şekil 4.1 'de görülmektedir.



Şekil 4.1 Disfonik Ses Düzeltme Sistemi Blok Diyagramı

4.5.1 Bant İçi Seslilik Değişikliği

KUDÖK’da, seslilik, beş frekans bandı ile kontrol edilir. Bölüm 2.1.2.5’de anlatıldığı gibi, KUDÖK konuşma sentezi modeli, darbe uyarımı ile gürültü uyarımını birleştirmek için, iki biçimlendirme filtresi kullanır. Bu filtrelerin yanıtları, seslilik değerleri ile kontrol edilir. Bu filtreler, uyarım içindeki darbe ve gürültü miktarını ayarlar.

Yapılan analizlerde normal konuşmanın sesli kısımlarında, en düşük dört veya beş frekans bandına (0–4 veya 0-5 kHz) ait seslilik değerlerinin 1, disfonik konuşmanın sesli kısımlarında ise 0 olarak hesaplandığı gözlenmiştir. Normal sese yakın bir sentetik ses elde edilebilmesi için, ötümlü seslerde (0 – 5 kHz) frekans bandı için darbe uyarımı üretilmelidir. Bu amaçla, disfonik sesler için hesaplanan bant içi seslilik değerleri uyarlanmıştır.

Önerilen sistemde, yukarıda anlatılan gözlemlere dayanılarak, ötümlü fonem bilgisi içeren çerçeveler için en düşük dört frekans bandına ait seslilik değeri 1 olarak atanmış, diğer çerçeveler için ise, tüm frekans bantlarına ait seslilik değerleri 0 olarak sabitlenmiştir.

4.5.2 Perde Değişikliği

Kronik disfonik konuşmada algılanabilir bir perde periyodu değeri görülmemektedir. Disfonik sestten normal ses elde edilebilmesi için, kişinin normal konuşması esnasında gözlenebilecek perde periyodu değerlerine uygun, sentetik perde değerleri üretilmelidir. Önerilen sistemde, perde ve kazanç değerleri arasında gözlemlenen ilinti kullanılarak, aşağıdaki eşitlik ile perde tahmini yapılmıştır (Morris R.W.1; Clements, 2002).

$$perde_{yeni}^n = (((kazanç_1^n + kazanç_2^n) / 2 - kazanç_{ortalama}) * \beta) + perde_{referans} \quad (4.1)$$

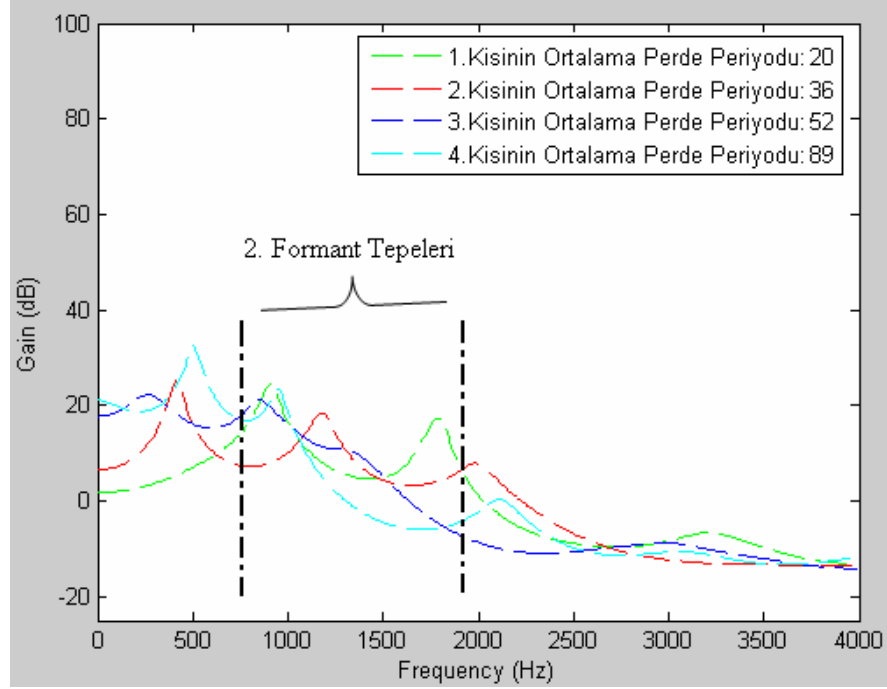
Burada $perde_{yeni}^n$, tahmin edilen perde değerini $kazanç_1^n$ ve $kazanç_2^n$, n numaralı çerçeveye ait kazanç değerlerini, $kazanç_{ortalama}$, ortalama kazanç değerini, $perde_{referans}$ ise referans perde değerini göstermektedir. $perde_{referans}$, sentezlenecek sesin tonunu ayarlamak için, β ise dinamik erimi kontrol etmek için kullanılır.

Önerilen sistem, referans perde değerini, otomatik olarak hesaplamaktadır. Her zaman, disfonik ses sahibi hastanın, hastalıktan önceki sesinin bulunması mümkün olmadığından ve fısıltılı seste de kronik disfonik seste olduğu gibi algılanabilir bir perde değeri bulunmadığından (Morris R.W.1; Clements, 2002), disfonik ses sahibine en uygun perde değerlerinin hesaplanması için fısıltılı konuşmadan faydalanılmıştır. Disfonik ses sahibine en uygun referans perde değerinin hesaplanması için, fısıltılı /a/ foneminin formant frekanslarından yararlanılmıştır.

Formant frekansları ile perde değerleri arasındaki bağıntıyı inceleyen pek çok çalışma bulunmaktadır (I. B. Thomas, 1969 ve M. Higashikawa, K. Nakai, 1996).

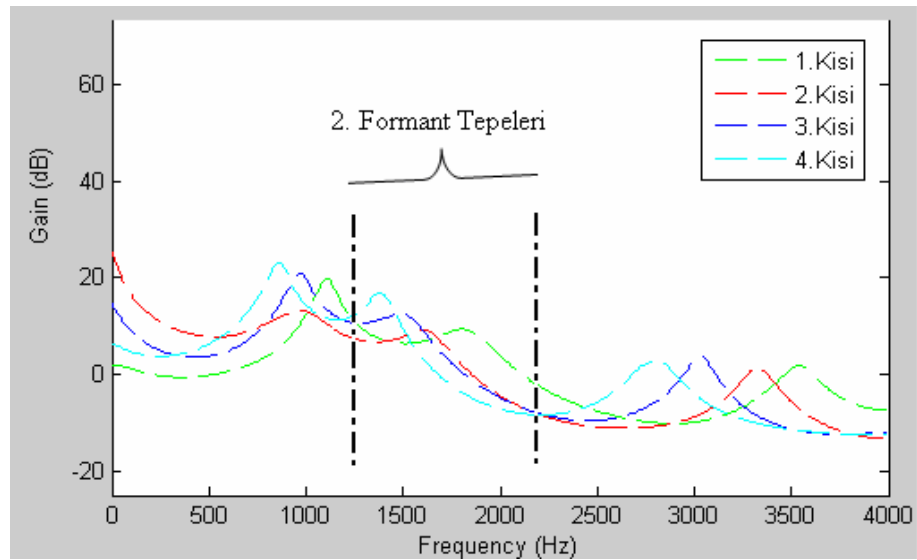
Formant frekansları ile perde arasında bağıntı bulunması için, farklı perde değerlerine sahip konuşmacılara ait /a/ fonemi formant frekansları incelenmiştir.

Farklı tonlarda sese sahip dört konuşmacının, normal bir sesle söyledikleri /a/ fonemi spektrumları Şekil 4.2’de görülmektedir. Bu konuşmacıların perde periyodu değerleri normalize edilmiş özilinti fonksiyonu ile bulunmuştur. Konuşmacıların /a/ fonemi için, ortalama perde periyodu değerleri sırayla 20, 36, 52 ve 89 olarak hesaplanmıştır. Şekil 4.2’ de görüldüğü gibi, perde periyodu arttıkça, özellikle 2.formant tepesinin frekansı azalmaktadır.



Şekil 4.2 Normal Sesle Söylenmiş /a/ Fonemi Spektrumları

Aynı dört konuşmacıdan alınan fısıltılı /a/ seslerinin, spektrumları ise Şekil 4.3’te görülmektedir.



Şekil 4.3 Fısıltılı Sesle Söylenmiş /a/ Fonemi Spektrumları

Şekil 4.3' de açıkça görüldüğü gibi, en alçak perde periyodu değerine sahip olan 1. konuşmacıya ait fısıltılı /a/ fonemi spektrumunun 2. formant frekansı değeri en büyüktür. Perde değeri azaldıkça, formant frekansları artmaktadır.

En yüksek perde periyodu değerine sahip konuşmacının 2. formant frekansı ile perde periyodu değeri f^{buyuk} ve p^{buyuk} , en düşük perde periyodu değerine sahip konuşmacının 2. formant frekansı ile perde periyodu değeri ise f^{kucuk} ve p^{kucuk} olmak üzere, iyileştirilecek disfonik sese ait /a/ fonemi spektrumunun 2. formant frekansı f_2 ile ifade edilirse, bu sesin iyileştirilmesi için kullanılacak referans perde ($p_{referans}$) değeri aşağıdaki gibi bulunabilir.

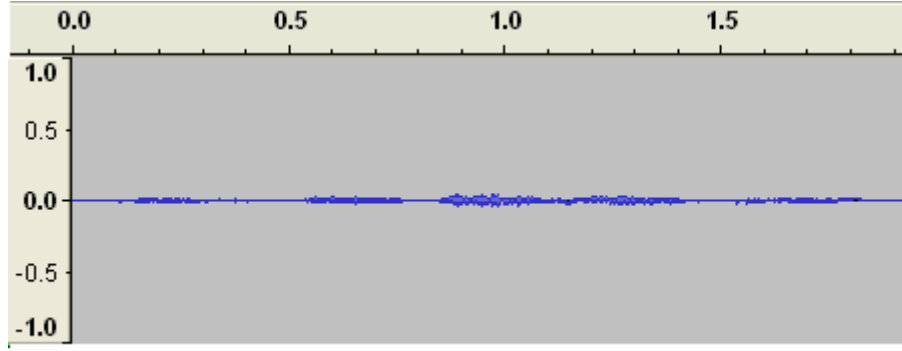
$$a = (p^{buyuk} - p^{kucuk}) / (f^{kucuk} - f^{buyuk}) \quad (4.2)$$

$$p_{referans} = (f^{kucuk} - f_2) * a + p^{kucuk} \quad (4.3)$$

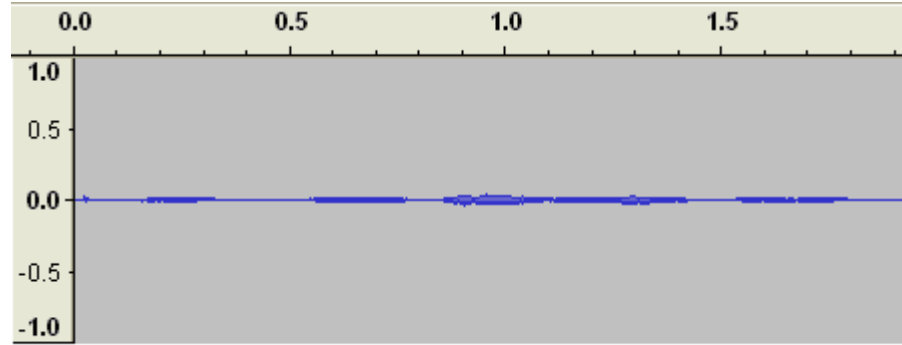
Önerilen sistemde, düzeltilecek disfonik için referans perde değeri, karşılaştırılması muhtemel en büyük ve en küçük perde değerine sahip konuşmacılara ait perde ve 2. formant frekansları kullanılarak hesaplanmıştır. Buna göre, $f^{buyuk}=897$, $p^{buyuk}=89$ ve $f^{kucuk}=1788$, $p^{kucuk}=20$ olarak alınmıştır.

4.5.3 Kazanç Değişikliği

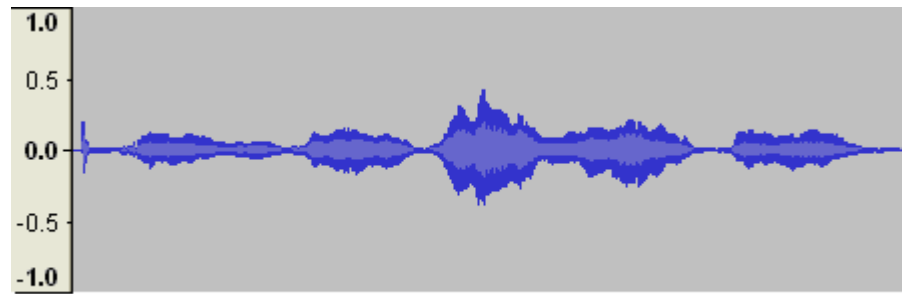
Normal konuşmada, sesin büyüklüğü, gırtlaktan gelen basıncın büyüklüğüne bağlıdır. Disfonik sesin büyüklüğü, normal sesinkinden yaklaşık 20 dB daha alçaktır. Bununla birlikte seslilik değişikliği sonrası, iyileştirilmiş disfonik seste, sessiz çerçevelerde gürültü gözlenmiştir. Bu etkiyi azaltabilmek için, bu çalışmada belirli bir eşik değerinin altında kalan konuşma parçaları elimine edilmiş, eliminasyondan sonra ilgili çerçeveye ait kazanç değeri arttırılmıştır. Eşik değeri, giriş sinyalinin ortalama kazancının sabit bir katsayı ile çarpılmasıyla dinamik olarak elde edilmektedir. Yapılan geçerlilik testleri sonucuna göre, bu katsayı 0.6 olarak belirlenmiştir. Disfonik ses ile söylenmiş /elma ağacı/ cümlesine ait ses sinyali ve bu sinyalin seslilik ve kazanç değişikliği adımlarından sonra aldığı hal, Şekil 4.4' te görülmektedir.



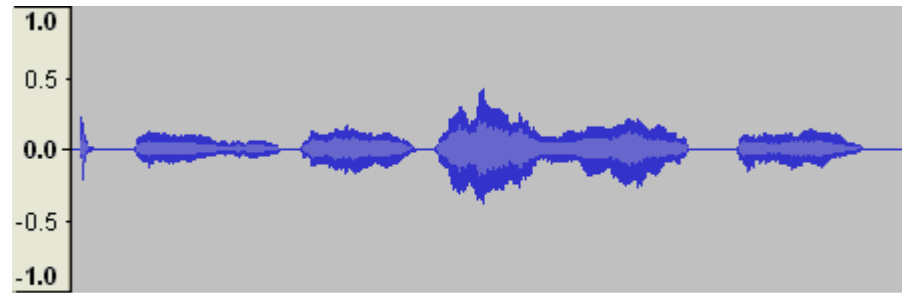
(a)



(b)



(c)



(d)

Şekil 4.4 /elma ağacı/ Cümlesi İçin (a) Disfonik Ses (b) Seslilik ve Perde Değişikliği Yapılmış Disfonik Ses (c) Kazanç Değişikliği ile Yükseltilmiş Ses (d) Gürültü Eliminasyonu Yapılmış Ses

Şekil 4.4 (c)'de kazanç Değişikliği ile yükseltilmiş ses sinyali görülmektedir. Açıkça görüldüğü gibi, heceler ve kelimeler arasında, gürültülü oluşumlar bulunmaktadır. Şekil 4.4 (d)'de, gürültü eliminasyonu yapıldıktan sonra, bu gürültülerin yok olduğu gözlenmiştir.

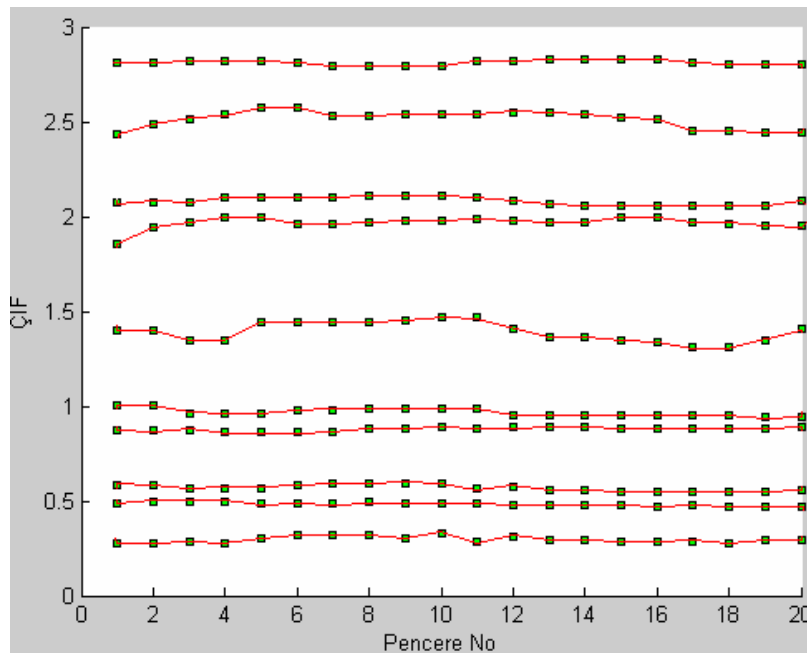
4.5.4 Formant Değişikliği

Normal konuşmaya göre disfonik konuşmanın formant yapısında bozulma olduğu gözlenmiştir. Bu çalışmada, konuşma kodlama esnasında spektrumda meydana gelen gürültüyü azaltabilmek için spektral yumuşatma ile disfonik sesin spektrumunu, normal sesine yakınlaştırmak için, formant bant genişliği ve frekansı değişikliği yapılmıştır.

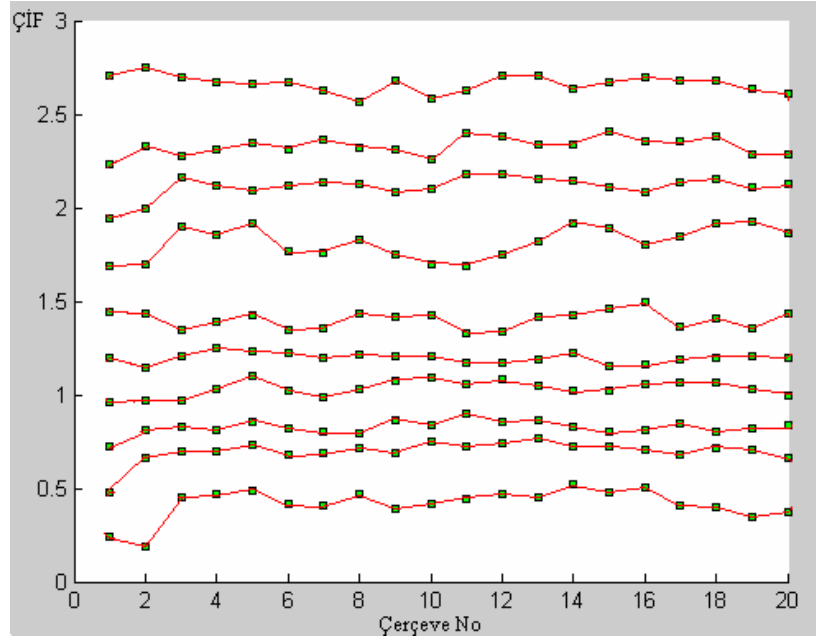
4.5.4.1 Spektral Yumuşatma

Pek çok konuşma kodlama algoritmasında olduğu gibi KUDÖK’nda da ÇSF, kuantalanmaktadır. Bu kuantalama, kuantalanan parametlerde, güç spektrumu zarfında ani değişiklikler olarak kendini gösteren gürültüye sebep olmaktadır. (H. P. Knagenhjelm ve W. B. Kleijn, 1995). Oluşan gürültünün azaltılabilmesi ve sentezlenen sesin kalitesinin artırılabilmesi için bu çalışmada ÇSF kuantalanmadan kullanılmıştır.

Disfonik sese ait uyarım gürültülü olduğundan, oluşan spektral zarf da gürültülü olmaktadır (Morris R.W.1 ve Clements, 2002). Bu gürültünün ÇSF’na etkisinin görülebilmesi için, normal sesle söylenmiş /a/ fonemine ait ÇSF gezingeleri ve disfonik ÇSF gezingeleri hesaplanmıştır. m numaralı çizgisel spektral frekansına ait gezinge, arka arkaya gelen n çerçeveye ait m. çizgisel spektral frekanslarından oluşmaktadır. Şekil 4.5’ te normal /a/ sesine ait ÇSF gezingeleri ve 4.6’da ise disfonik /a/ sesine ait ÇSF gezingeleri (LSF Trajectories) görülmektedir.



Şekil 4.5 Normal /a/ Fonemi Ses Sinyaline Ait ÇSF Gezingeleri

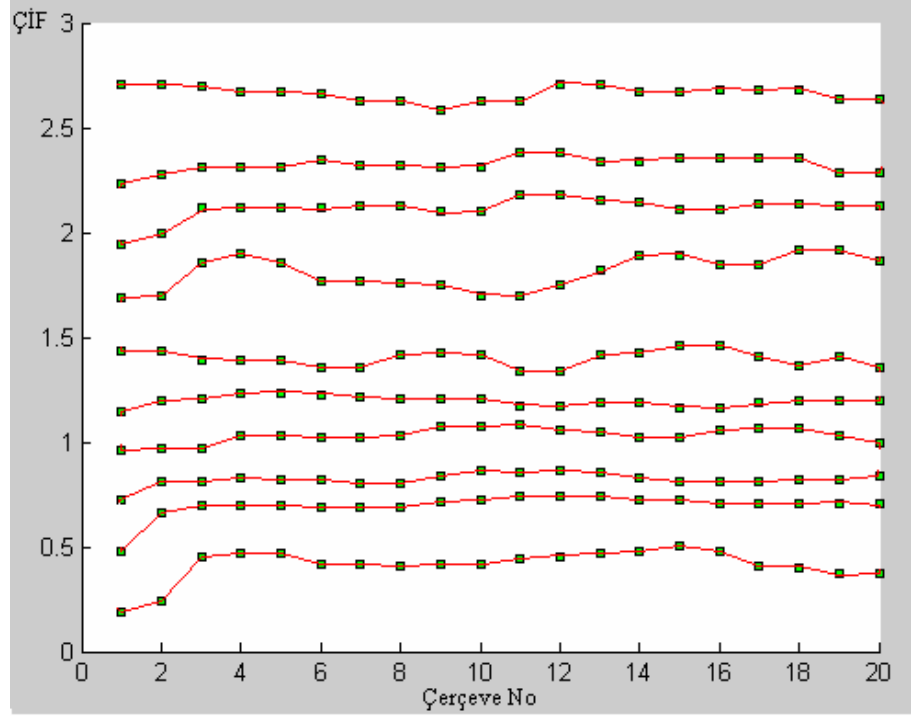


Şekil 4.6 Disfonik /a/ Fonemi Ses Sinyaline Ait ÇSF Gezingeleri

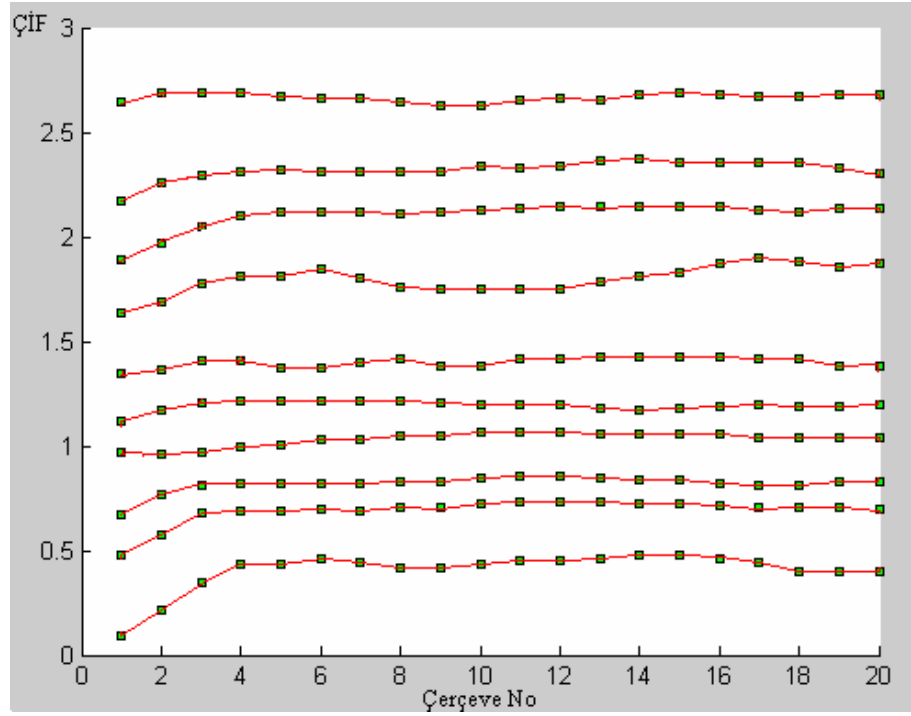
Geliştirilen sistemde, spektrumu yumuşatmak için ortanca filtresi kullanılmıştır. $lsf_n^{(m)}$, n numaralı çerçeveye ait m. ÇSF olmak üzere, yumuşatılmış çizgisel spektral frekansları Eşitlik 4.4 ile hesaplanır (Morris R.W.1 ve Clements, 2002). Ortanca filtresi için arka arkaya gelen k adet çerçeve için, percere boyutu p olan ortanca filtresi uygulanmıştır.

$$\hat{lsf}_n^{(m)} = median(lsf_n^{(m)}, lsf_{n-1}^{(m)}, \dots, lsf_{n-K+1}^{(m)}, p) \quad (4.4)$$

p=3 için ortanca filtresi ile yumuşatılmış ÇSF ları Şekil 4.7, p=5 için ortanca filtresi ile yumuşatılmış ÇSF ları ise 4.8'de görülmektedir.



Şekil 4.7 Disfonik /a/ Fonemi Ses Sinyaline Ait Yumuşatılmış ÇSF Gezingeleri (k=20 ve p=3)

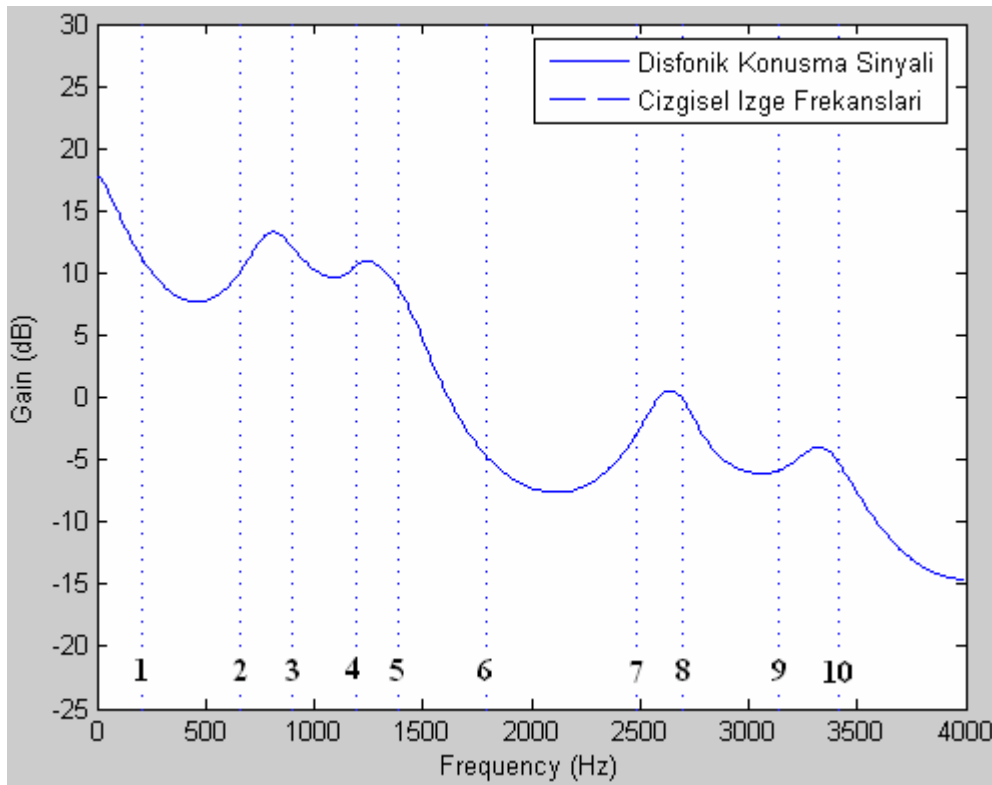


Şekil 4.8 Disfonik /a/ Fonemi Ses Sinyaline Ait Yumuşatılmış ÇSF Gezingeleri (k=20 ve p=5)

4.5.4.2 Formant Yapısı Değişikliği

Disfonik ve normal seslendirilmiş sesli harfler incelendiğinde, disfonik sese ait formant tepelerinin özellikle ilk iki tanesinin formant bant genişliklerinin daha geniş ve formant frekanslarının daha büyük olduğu gözlenmiştir. Disfonik sesle söylenmiş kelimelerin anlaşılabilirliğini arttırmak için formant değişikliği yapılmıştır.

Formant değişikliği için ÇSF'ları kullanılmıştır. Şekil 4.9'da /a/ fonemine ait spektrum ve ÇSF'ları görülmektedir.



Şekil 4.9 Disfonik /a/ Fonemi Spektrumu ve Çizgisel İzge Frekansları

Şekil 4.9'da görülen spektruma ait formant frekans ve bant genişlikleri, en sık kullanılan yöntemlerden biri olan polinomsal kök çözme yöntemi ile hesaplandığında, Çizelge 4.3'deki sonuçlar elde edilmiştir. Polinomsal kök yöntemi ile formant frekans ve bant genişlikleri, $z = re^{j\theta}$, eşlenik tepe polinomsal kökü ve f_0 örnekleme frekansı olmak üzere, formant frekansı F ve 3 dB bant genişliği B , Eşitlik 4.5 ve 4.6 ile hesaplanır (M. E. Lee ve M. J. T. Smith, 2003).

$$F = \frac{f_0}{2\pi} * \theta \quad \text{Hz} \quad (4.5)$$

$$B = -\frac{f_0}{\pi} * \ln r \quad \text{Hz} \quad (4.6)$$

Çizelge 4.3 Disfonik /a/ Fonemi İçin Formant Frekansı Ve Bant Genişlikleri

Formant Tepesi	Formant Frekansı	Formant Bant Genişliği
1	749.7	170.1
2	1313.0	237.6
3	2436.3	325.6
4	3252.1	418.0

Şekil 4.9’da da görüldüğü gibi aynı spektruma ait ÇSF’ları ise aşağıdaki gibidir

Çizelge 4.4 Disfonik /a/ Fonemi İçin Çizgisel İzge Frekansları

ÇSF	1	2	3	4	5	6	7	8	9	10
Frekans	209	661.7	896.9	1191.7	1390.9	1792.0	2480.9	2700.4	3134.2	3417.1

4.5.4.2.1 Formant Bant Genişliği Değişikliği

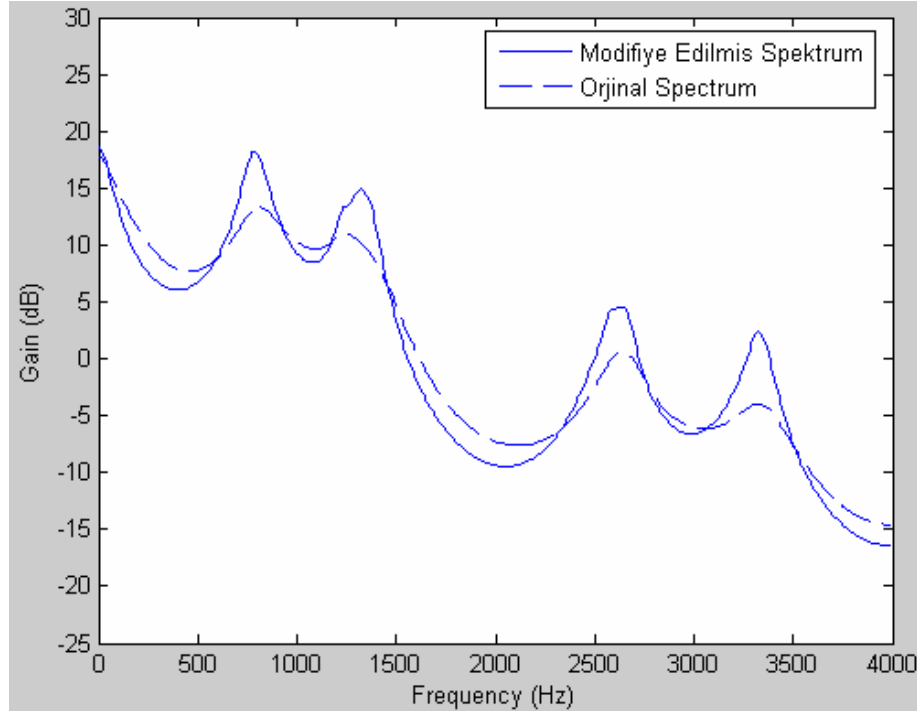
Formant bant genişliği değişikliği için, ω_i , $i=1...11$ olmak üzere ÇSF’nı ve ω'_i değiştirilmiş frekansları göstermek üzere, birinci ve ikinci ÇSF larına ait bant genişliği eşitlik 4.7 ve 4.8 ile a derecesinde daraltılır (McLoughlin ve Chance, 1997).

$$\omega'_1 = \omega_1 + a(\omega_2 - \omega_1) \quad (4.7)$$

$$\omega'_2 = \omega_2 - a(\omega_2 - \omega_1) \quad (4.8)$$

Bu yöntemin spektrumun tamamına uygulanabilmesi için, formant tepelerinin frekanslarının doğru tespiti çok önemlidir.

Çizelge 4.3 ve 4.4’ten yararlanarak 4.7 ve 4.8 eşitliklerinin hangi ÇSF için kullanılması gerektiği tespit edilebilir. Her formant tepesi için, o formant tepesine ait formant frekansını içine alan ÇSF çiftleri kullanılarak daha dar ve yüksek genlikli spektral tepeler elde edilmiştir. 2-3, 4-5, 7-8 ve 9-10 numaralı ÇSF çiftleri için daraltma yapılarak değiştirilen spektrum Şekil 4.10’da görülmektedir ($a=0.3$ olarak alınmıştır). Bu spektrum için hesaplanan formant frekansları ve bant genişlikleri ise Çizelge 4.5’de görülmektedir.



Şekil 4.10 Disfonik /a/ Fonemi Spektrumu ve Formant Bant Genişlikleri Değiştirilmiş Spektrum

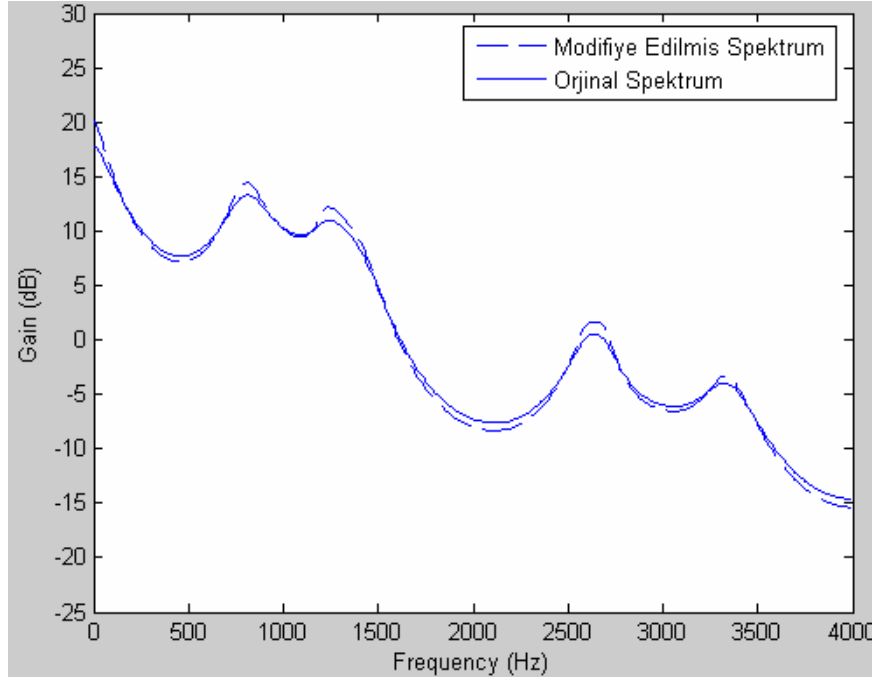
Çizelge 4.5 Formant Bant Genişliği Değiştirilmiş Spektrum İçin Formant Frekansı Ve Bant Genişlikleri

Formant Tepesi	Formant Frekansı	Formant Bant Genişliği
1	749.7	69.2
2	1331.5	100.1
3	2598.3	138.1
4	3326.2	167.1

Bu yöntem, formant tepelerinde istenen daralma ve yükselmeyi sağlamaktadır. Ancak tepe etkileşimi problemi nedeniyle, polinomsal kök çözme yöntemiyle hesaplanan formant polinomunda hatalar oluşabilmektedir (Y. S. Hsiao ve D. G. Childers, 1996). Değiştirilecek formantlar birbirlerine çok yakın olduğunda, tepe etkileşimi problemi nedeniyle, formantları yukarıda anlatıldığı gibi değiştirmek spektrumda bozulmaya ve anlaşılabilirliğin azalmasına yol açabilmektedir. Bunun yerine, formantlar daraltılırken, dengeli bir spektrum elde edebilmek için Chirp z dönüşümü kullanılabilir. (Wai. C. Chu, 2003). Buna göre η 1'den küçük pozitif bir sabit olmak üzere a_i $i=1...p$, değiştirilmiş p. dereceden DÖK, eşitlik 4.9 ile hesaplanır.

$$a_i = \eta^{-1*i} * a_i \quad i = 1 \dots p \quad (4.9)$$

Chirp z dönüşümü kullanılarak değiştirilmiş spektrum Şekil 4.11’de görülmektedir. Ancak bu yöntem ile ortalama 40 Hz lik bir bant genişliği daralması elde edilebilmektedir.



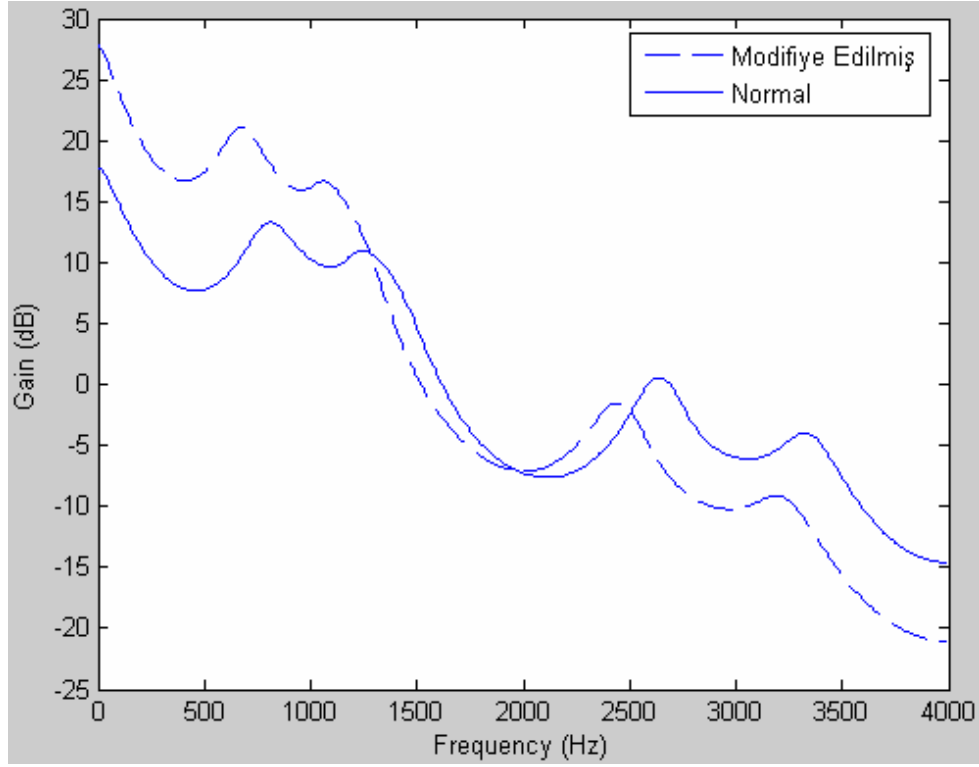
Şekil 4.11 Disfonik /a/ Fonemi Spektrumu ve Chip z Dönüşümü Kullanılarak Değiştirilmiş Spektrum

4.5.4.2.2 Formant Frekansı Değişikliği

Formant frekansı değişikliği için yine ÇSF’ndan yararlanılmıştır. k numaralı çizgisel spektral frekansını γ derecesinde ötelemek için (4.10) eşitliği kullanılır (McLoughlin ve Chance, 1997).

$$\omega'_k = \omega_k + \omega_k (\gamma - 1)(\pi - \omega_k) / \pi \quad (4.10)$$

Spektrumun dengeli olabilmesi için, önerilen yöntemde, tüm çizgisel spektral frekansları ötelenmiştir. Şekil 4.12’ de Eşitlik 4.10 kullanılarak değiştirilmiş spektrum görülmektedir ($\gamma = 0.8$ olarak alınmıştır).

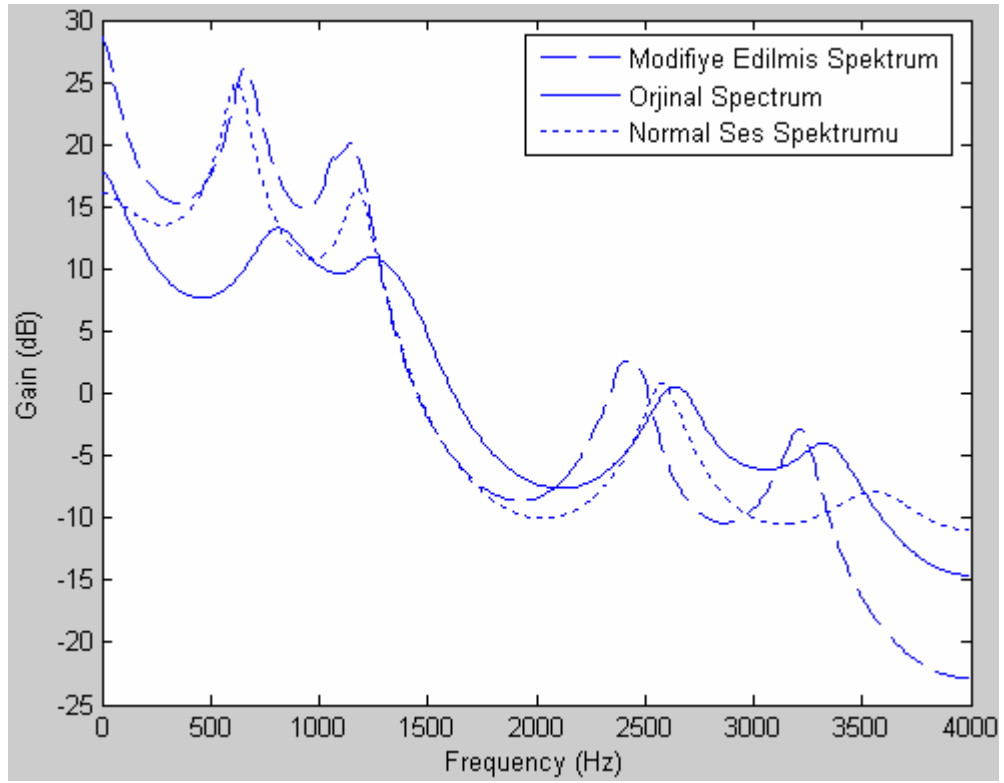


Şekil 4.12 Disfonik /a/ Fonemi Spektrumu ve Formant Frekansları Değiştirilmiş Spektrum
Değişiklik sonucu elde edilen Formant frekansları ve bant genişlikleri aşağıdaki gibidir.

Çizelge 4.6 Formant Frekansları Değiştirilmiş Spektrum İçin Formant Frekansı Ve Bant Genişlikleri

Formant Tepesi	Formant Frekansı	Formant Bant Genişliği
1	630.9	167.2
2	1136.2	222.3
3	2246.3	317.0
4	3126.3	441.6

Eşitlik 4.7, 4.8 ve 4.10 kullanılarak anlatılan formant bant genişliği ve frekansı değişiklikleri disfonik sese uygulandığında elde edilen spektrum, orijinal spektrum ve normal ses spektrumları Şekil 4.13’ da görülmektedir.



Şekil 4.13 Disfonik /a/ Fonemi Spektrumu, Formant Frekansları ve Bant Genişlikleri Değiştirilmiş Spektrum ve Normal Ses Spektrumu

5. DENEYSEL SONUÇLAR

Bu çalışmada Cerrahpaşa Tıp Fakültesi Kulak Burun Boğaz Ana Bilim Dalı'nda ses teli felci ve gırtlak kanseri teşhisi ile yatmakta olan 15 hastanın 20'şer adet trifon dengeli cümlesi, 8kHz örnekleme frekansında kaydedilmiştir. Ancak geçirdikleri tam laryngectomy operasyonlarında gırtlaklarına delik açılmış bazı hastalardan alınan konuşmaların, bu deliğin iyi kapatılamaması nedeniyle çok gürültülü olması ve ses teli felci teşhisi konulmuş bazı hastalardan alınmış konuşmaların dil ve dişlerindeki bir takım problemler nedeniyle anlaşılabilirliğinin çok düşük olması sebebiyle, incelenen tüm hastaların sesleri test için kullanılamamış ancak 5 erkek, 2 kadın olmak üzere 7 hastanın ses kayıtları kullanılabilmektedir. Sentezlenen ses spektrumlarının normal sese yakınlığının nesnel olarak ölçülebilmesi için log spektral uzaklık (log spectral distance) hesabı yapılmıştır. Ancak spektral farkın hesaplanması, sentezlenen sesin normal sese yakınlığının tespiti için tek başına yeterli değildir. Bunun eksikliğin giderilmesi için öznel dinleme testlerine başvurulmuş, sentetik ses ve normal ses spektrogramları incelenmiştir.

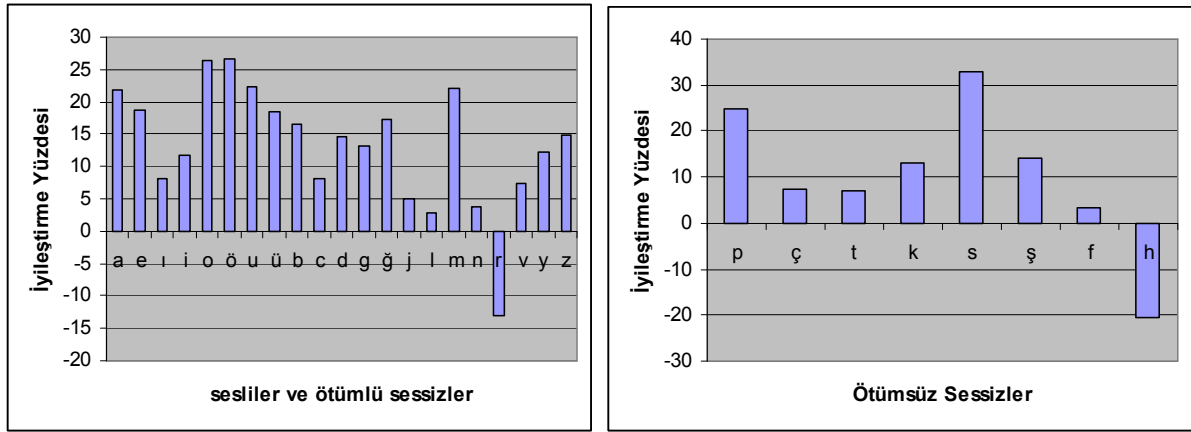
5.1 Spektral Uzaklık Hesabı

Sentezlenmiş ses spektrumunun, normal sese yakınlığının tespit edilebilmesi için, disfonik konuşmacılardan alınmış fonemler ile normal konuşmacılardan alınmış fonemlerin spektral farkları, disfonik konuşmacılar için sentezlenmiş fonemler ile normal konuşmacılardan alınmış fonemlerin spektral farklılıkları ile kıyaslanmıştır.

x ve y , x_i ve y_i sinyallerine ait ayrık spektrumları olmak üzere, x_i ve y_i sinyallerinin log spektral uzaklığı eşitlik 5.1 ile hesaplanmıştır (Elena Deza, Michel Marie Deza, 2006).

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (\ln x_i - \ln y_i)^2} \quad (5.1)$$

Veritabanında disfonik sesi bulunan hastalara ait normal sesler elimizde bulunmamaktadır. Türkçe'deki sesliler ve ötümlü sessizler için normal ses-fısıltılı ses ve normal ses-sentezlenmiş ses log spektral farkları hesaplanmıştır. Ötümsüz sessizler için ise fonemin akustik özelliklerinin korunması ile sentezlenmiş sesin, akustik özellikleri uyarlanarak sentezlenmiş sese göre ne oranda normal sese yakınlık sağladığı hesaplanmıştır. Normal ses sahip bir kullanıcıya ait fonem örnekleri için sentezlenen sesin, fısıltılı sesi ne oranda normal sese yaklaştırdığı Şekil 5,1'de görülmektedir.

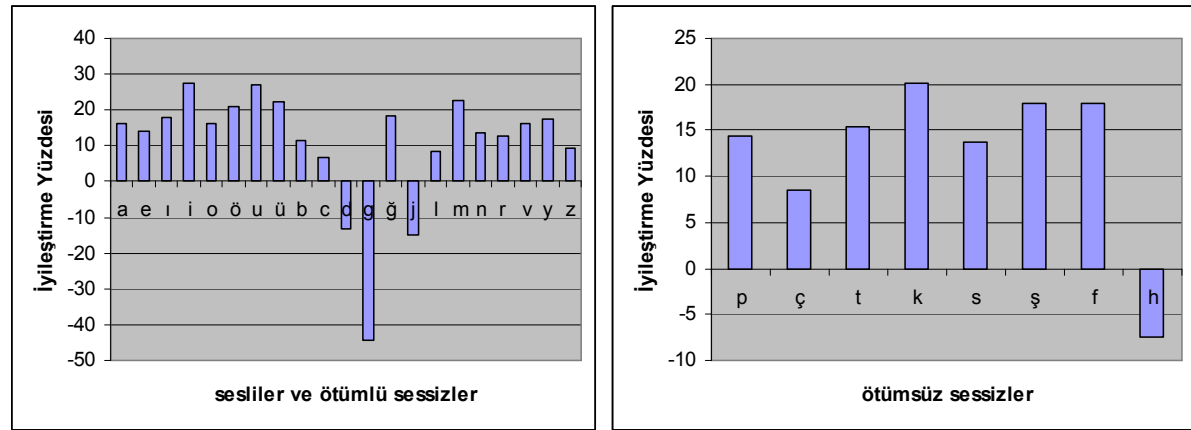


(a)

(b)

Şekil 5.1 (a) Sosliler ve ötümlü sessizler için akustik özellikler uyarlanarak sentezlenmiş seslerin, fısıltılı sesi normal sese yaklaştıırma oranları (b) Ötümsüz sessizler için fonemin akustik özelliklerinin korunması ile sentezlenmiş sesin, akustik özellikleri uyarlanarak sentezlenmiş sese göre normal sese yakınlık oranı

Ses teli felci hastasına ait fonem örnekleri için sentezlenmiş seslerin, disfonik sesi normal sese yakınlılaştıırma oranları şekil 5.2' de görölmektedir.



(a)

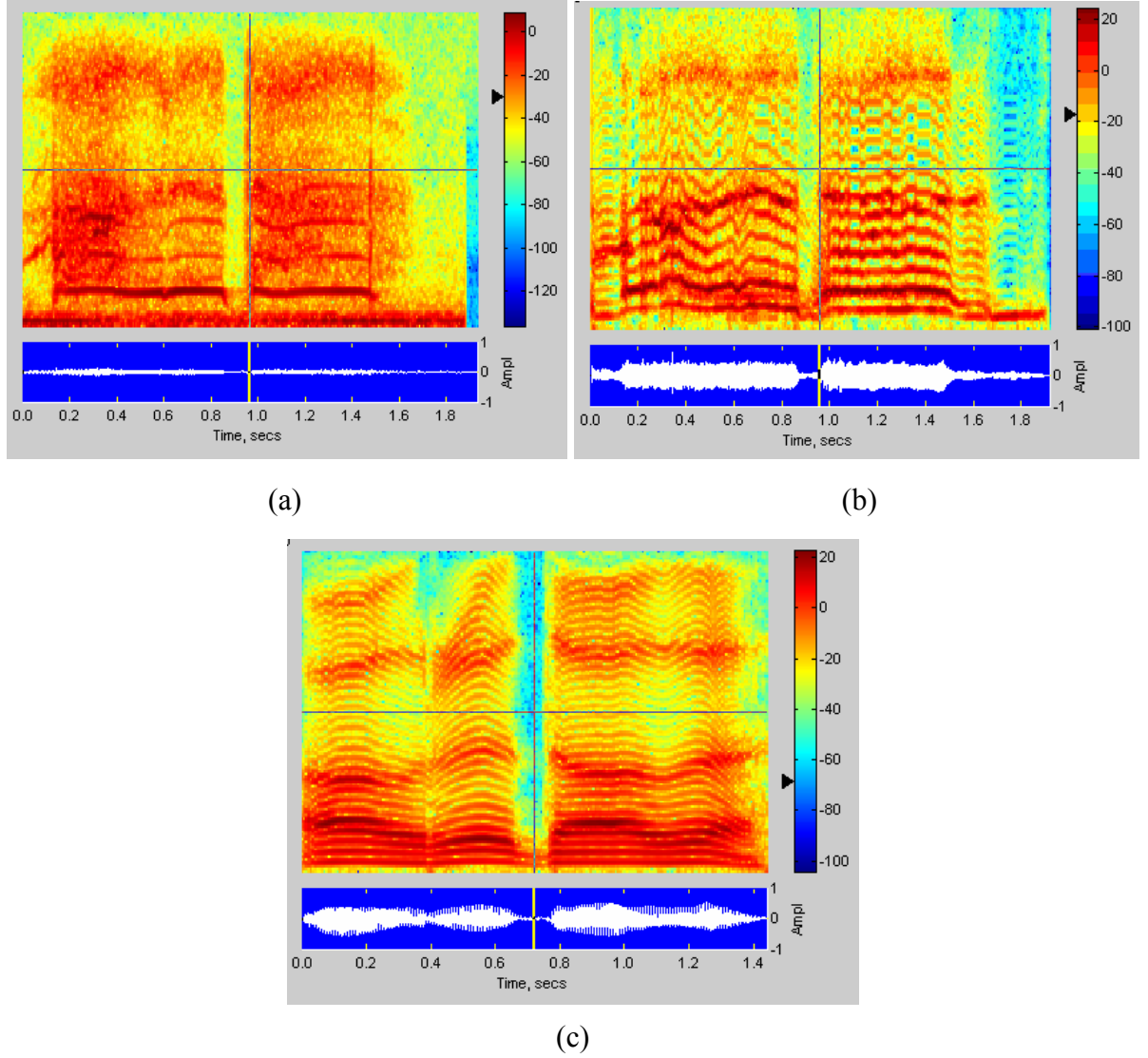
(b)

Şekil 5.2 (a) Sosliler ve ötümlü sessizler için akustik özellikler uyarlanarak elde sentezlenmiş seslerin, disfonik sesi normal sese yaklaştıırma oranları (b) Ötümsüz sessizler için fonemin akustik özelliklerinin korunması ile sentezlenmiş sesin, akustik özellikleri uyarlanarak sentezlenmiş sese göre normal sese yakınlık oranı

Şekil 5.1 ve 5.2'de göröldüğü gibi, özellikle sesli harfler için sentezlenen ses, normal sese yaklaşık %20 civarında daha yakındır. Ötümsüz seslerde ise akustik özellikler korunarak sentezlenen sesin, akustik özellikler uyarlanarak sentezlenen sese göre yaklaşık %11 oranında normal sese daha yakın olduğı gözlenmiştir.

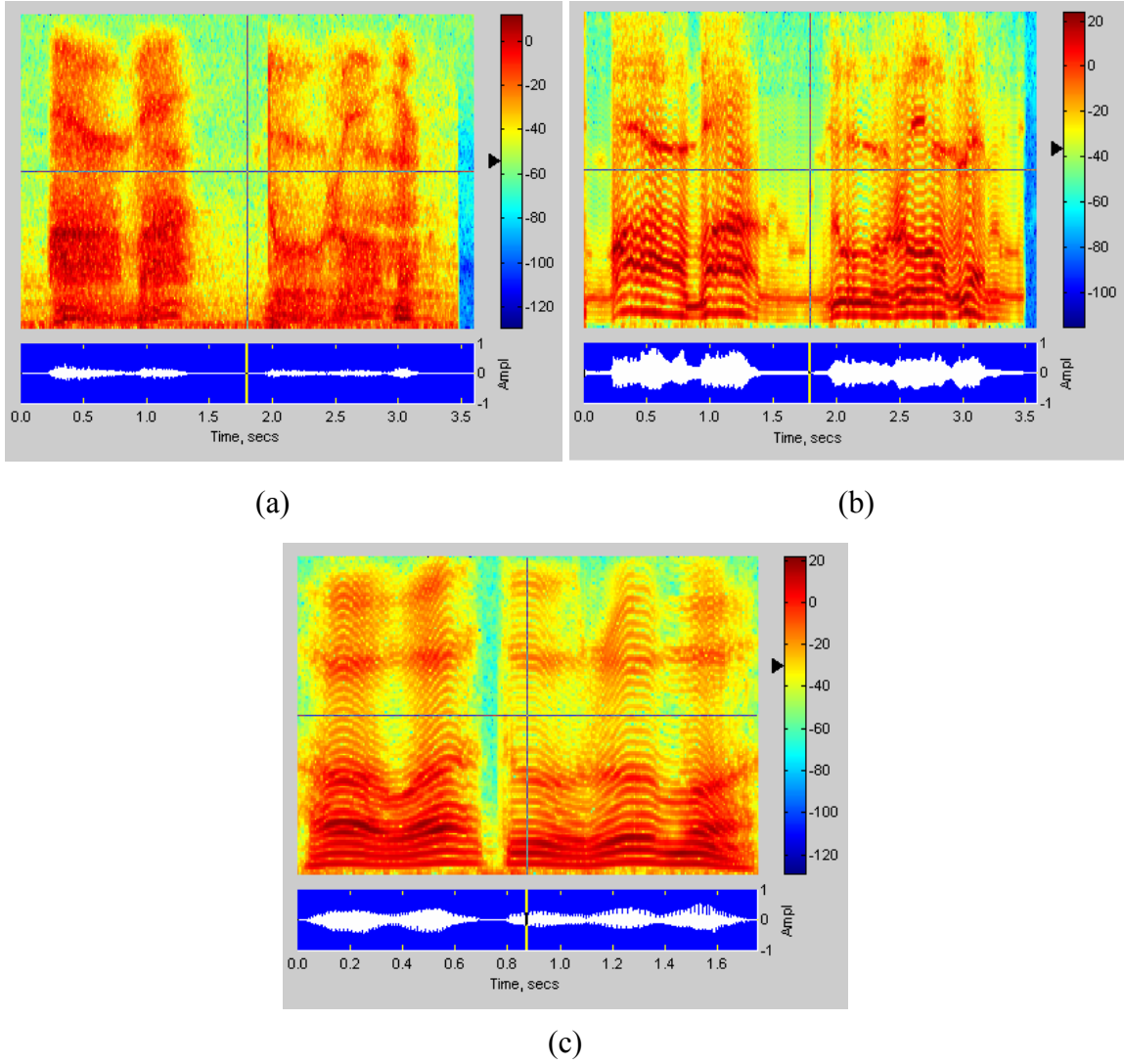
5.2 Özel Dinleme Testleri ve Sentetik Ses Spektrogramları

Disfonik sesin normal sesteki farklılıklarının görsel olarak algılanabilmesi için spektrogramlardan faydalanılmıştır. Şekil 5.3’de gırtlak kanseri hastasında alınmış /ağrı dağı/ tamlaması spektrogramı, bu tamlama için elde edilmiş sentetik konuşma spektrogramı ve normal sese sahip konuşmacıdan alınan örneğe ait spektrogram görülmektedir.



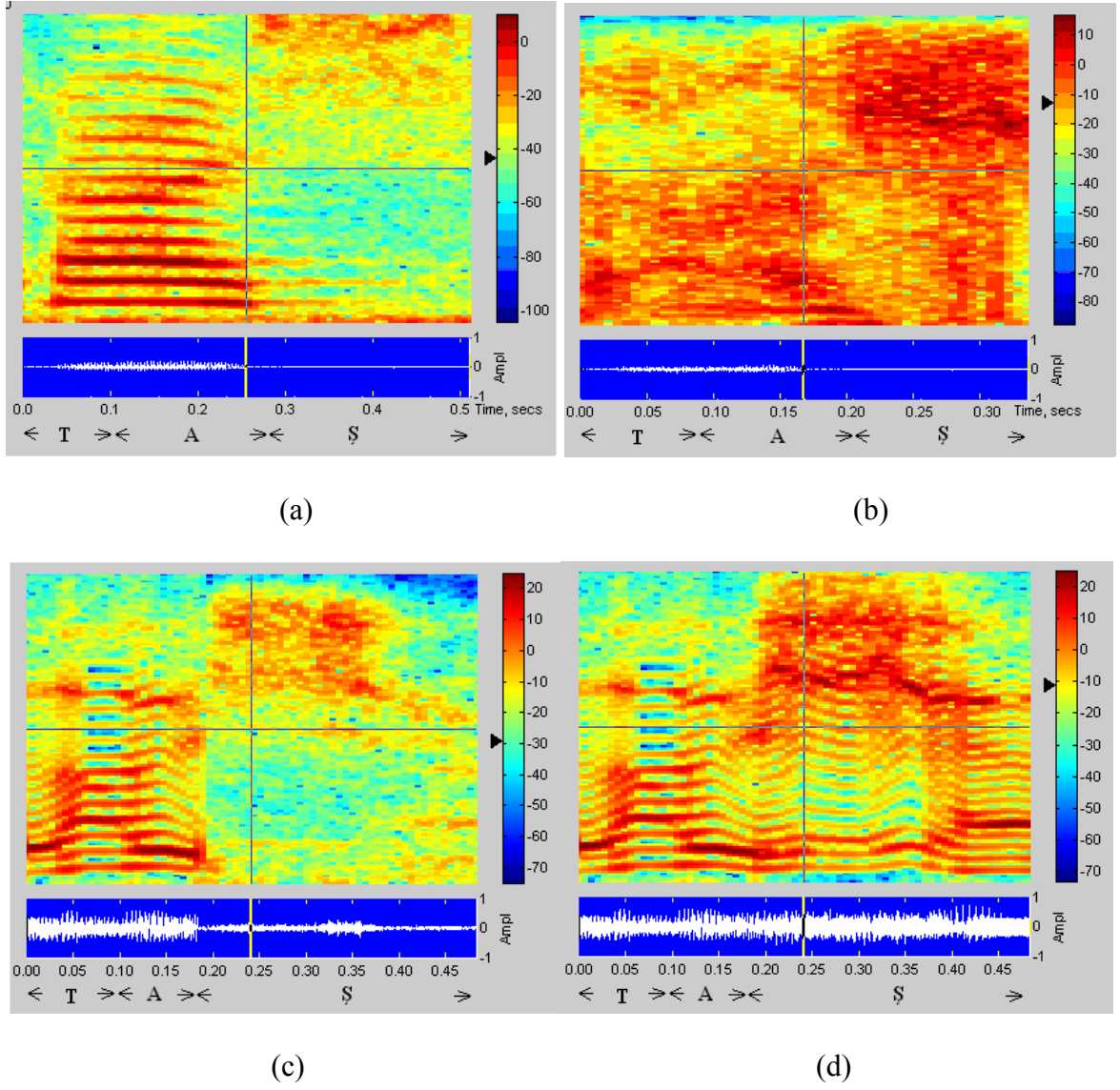
Şekil 5.3 /ağrı dağı/ tamlaması için (a) Disfonik Ses (b) Sentetik Ses (c) Normal Ses Spektrogramları

Şekil 5.4’de ise ses teli felci hastasından alınmış /hava durumu/ tamlaması spektrogramı, bu tamlama için elde edilmiş sentetik konuşma spektrogramı ve normal sese sahip konuşmacıdan alınan örneğe ait spektrogram görülmektedir.

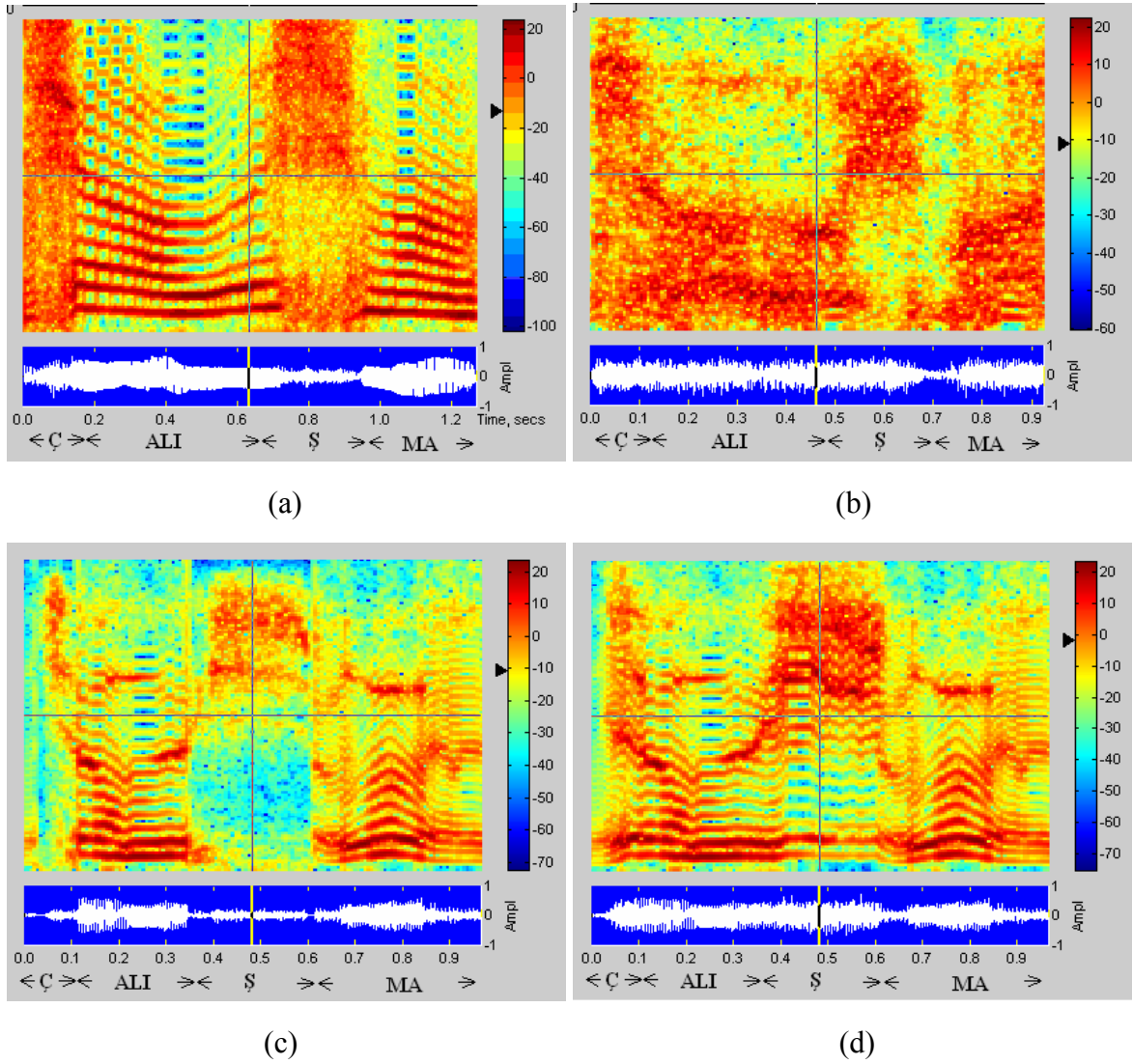


Şekil 5.4 /hava durumu/ tamlaması için (a) Disfonik Ses (b) Sentetik Ses (c) Normal Ses Spektrogramları

Şekil 5.3 ve 5.4’de görüldüğü gibi, sentezlenen ses, belirli frekansda, disfonik sese göre daha kuvvetlidir. Bu açıdan, normal sese yakınlık sağlamaktadır. Ancak, disfonik sesin gürültülü yapısı, sentetik ses kalitesinin belirgin ölçüde azalmasına sebep olmaktadır. Şekil 5.5 ve 5.6’de ise ses teli felci hastasından alınmış /taş/ ve /çalışma/ kelimelerinin, bu kelimeler için, ötümsüz harflerin akustik özellikleri korunarak ve korunmadan sentezlenmiş konuşmalar ile normal bir konuşmacıdan alınmış ses örneği spektrogramları görülmektedir.



Şekil 5.5 /taş/ kelimesi için (a) Normal Ses (b) Disfonik Ses (c) Ötümsüz Sessizlerin Akustik Özellikleri Korunarak Sentezlenmiş Ses (d) Tüm Seslerin Akustik Özellikleri Değiştirilerek Sentezlenmiş Ses



Şekil 5.6 /çalışma/ kelimesi için (a) Normal Ses (b) Disfonik Ses (c) Ötümsüz Sessizlerin Akustik Özellikleri Korunarak Sentezlenmiş Ses (d) Tüm Seslerin Akustik Özellikleri Değiştirilerek Sentezlenmiş Ses

Ötümlü sesler, ötümsüz seslere göre daha periyodik bir yapı göstermektedir. Şekil 5.6 a ve b'de açıkça görülebileceği gibi, disfonik ses ve normal sese ait ötümsüz seslerin yapısı farklılık göstermemektedir. Bu nedenle sentezlenen seste ötümsüz seslerin akustik özelliklerinin korunması, sentezlenen sesin anlaşılabilirliğini arttırmaktadır.

Disfonik ses düzeltme sisteminin başarısını ölçmek amacıyla öznel dinleme testlerine başvurulmuştur. Öznel dinleme testlerinde 5 dinleyiciye, disfoni hastalarına ait cümleler ve bu cümleler için üretilmiş sentetik sesler dinletilmiştir. Dinleyicilerden sentetik seslere, disfoni hastasının orijinal sesine göre anlaşılabilirlik ve normal sese yakınlık açısından 5 çok iyi ve 1 çok kötü olmak üzere 5 üzerinden puan vermesi istenmiştir. Testler sonucunda sentetik sesin anlaşılabilirliği ortalama 2.5 ve normal sese yakınlığı ortalama 2.4 olarak hesaplanmıştır.

Ötümsüz sessizlerin yüksek başarı oranlarıyla tanındığı ve yuvarlak geniş ünlüler içeren örneklerin diğerlerine göre daha yüksek anlaşılabilirlikte olduğu gözlenmiştir.

6. SONUÇLAR VE ÖNERİLER

Ses, konuşma ve lisan ile anlam kazanan, iletişimin en önemli ögesidir. Ancak, gırtlakın işlev bozukluğu ile bu iletişim ögesine gölge düşmektedir. Gırtlak kanseri, ses teli felci gibi rahatsızlıklar hastanın sesini tamamen kaybetmesi ile sonuçlanabilmektedir. Kronik ses kısıklıklarında hastanın tekrar konuşabilmesi için yapay gırtlak, özofageal konuşma ve gırtlak protezi yöntemleri kullanılmaktadır. Ancak bu yöntemler ile elde edilen konuşma ya doğallıktan uzaktır ya da enfeksiyon riski barındırmaktadır. Cerrahi operasyon gerektirmeden doğal sesle konuşmayı sağlayacak bir sistem kronik disfoni hastalarının hayatlarını önemli ölçüde kolaylaştıracaktır. Literatürde sesini tamamen kaybetmiş insanların konuşmalarını iyileştirmeye yönelik az sayıda çalışma mevcuttur. Bu çalışmada, kronik ses kısıklığı problemi olan hastaların konuşmalarını, konuşmanın sürekli yapısını bozmadan normal sese yaklaştıran KUDÖK tabanlı bir sistem tasarlanmıştır.

Literatürde disfonik ses yapısındaki bozulmaları inceleyen çalışmalar mevcuttur. Bu bozulmalar fonem bazında incelendiğinde, disfonik ve normal sese ötümsüz sessizlerin akustik özelliklerinin belirgin bir farklılık göstermediği ancak disfonik ötümlü sessizler ve seslilerde normal seslerin aksine algılanabilir perde periyodu ve seslilik olmadığı ve formant yapısında bozulma olduğu gözlenmiştir.

Geliştirilen sistem ile disfonik konuşma çerçeve çerçeve analiz edilmiş ve her çerçeve için, son perde, bant içi seslilik, kazanç, doğrusal öngörüm katsayıları, fourier genlikleri ve aperiodyik bayraktar parametreleri hesaplanmıştır. Doğrusal öngörüm katsayıları kullanılarak şablon eşleme yöntemi ile ilgili çerçevedeki sesin ötümlü bir sessiz ya da sesli olup olmadığı tespit edilmiştir. Ötümlü sessizler ve sesliler için disfonik sese seslilik ve perde eklenmiş, kazanç ve formant yapısı değişikliği yapılmıştır. Son adımda, değiştirilmiş akustik özellikler kullanılarak KUDÖK ile yeni ses sentezlenmiştir.

Sistemin testi için 5 erkek, 2 kadın olmak üzere, ses teli felci ve gırtlak kanseri teşhisi konmuş 7 hastanın 20'şer adet trifon dengeli cümlesi kullanılmıştır. Sistem başarısının değerlendirilmesi log spektral uzaklık hesabı yapılmış, öznel dinleme testlerine başvurulmuştur. Testler sonunda sentetik sesin, disfonik sese göre ortalama %20 oranında normal sese daha yakın olduğu gözlenmiştir. Dinleyici testlerinde sentetik sesin normal sese göre daha mekanik olduğu ve barındırdığı gürültünün anlaşılabilirliği azalttığı tespit edilmiştir.

Disfonik sesin gürültülü yapısının temizlenmesi ile sentetik sesin anlaşılabilirliği artırılabilir. Anlaşılabilirliğin artırılması için önerilebilecek bir diğer çözüm, formant bant genişliği ve

frekansında yapılan deęişiklięin fonem baęımlı hale getirilmesi olacaktır. Geliştirilen sistemde disfonik ses için üretilen perde deęerleri kazanç parametresine baęlı olarak deęişmektedir. Perde üretiminin sesin kazanç parametresi ile birlikte başka parametrelerinin de göz önünde bulundurularak yapılması sentetik sesin mekaniklięinin azaltılmasına yardımcı olabilir.

Sistem eksikliklerinin giderilmesi disiplinler arası bir yaklaşım gerekmektedir. Sentetik ses kalitesinin arttırılması ve sistemin gerçek zamanlı hale getirilmesi ile geliştirilen sistem telekomünikasyon alanında disfoni hastaları için oldukça faydalı olacaktır. Sistemin gömülü sistem uygulaması haline getirilmesi ise disfoni hastalarının günlük yaşantılarında normal konuşabilmeleri için alternatif bir çözüm olacaktır. Sistemin mevcut yapısı ile ileride yapılacak çalışmalara yön vereceęi umut edilmektedir.

KAYNAKLAR

- Aguilar G., Nakano-Miyatake M., (2005), "Alaryngeal Speech Enhancement Using Pattern Recognition Techniques", IEICE - Transactions on Information and Systems, Volume E88-D, Issue 7, pp. 1618-1622
- Ali R.H. and Jebara S.B., (2006), "Esophageal speech enhancement using excitation source synthesis and formant structure modification".
- Bi N. and Qi Y., (1997), "Speech conversion and its application to alaryngeal speech enhancement", Proc. ICSP96, pp. 1586-1589.
- Deza E., Deza M. M., (2006), "Dictionary of Distances", Elsevier Publications.
- Hansen J.H.L., (1991), "Constrained Iterative Speech Enhancement with Application to Speech Recognition" IEEE Transactions On Signal Processing, 1991, vol 39
- Higashikawa M., Nakai K., Sakakura A. and Takahashi H., (1996), "Perceived pitch of whispered vowels-relationship with formant frequencies: A preliminary study", Journal of Voice, pp 155-158
- Hsiao Y. S. and Childers D. G., (1996), "A New Approach to Formant Estimation and Modification Based on Pole Interaction", Signals, Systems and Computers, Volume 1, 3-6 Nov. 1996 pp 783 – 787.
- Icat F., İlk H. G., (2004), "Sürdürülebilir Türkçe Ünlüler İçin Fisilti ile Normal Konuşma Arasındaki Farklılıkların İncelenmesi", Sinyal İşleme ve Uygulamaları Kurultayı, SIU04, 564-566, Kusadasi.
- Itoh, T. Takeda, K. Itakura, (2001), "Acoustic analysis and recognition of whispered speech", Automatic Speech Recognition and Understanding, 2001. ASRU '01. IEEE Workshop on, pp. 429- 432.
- Jovicic S. T. and Dordevic M. M., (1996), "Acoustic features of whispered speech," Acustica- acta acustica, vol. 82, pp. S228.
- Knagenhjelm H. P. and Kleijn W. B., (1995), "Spectral Dynamics is More Important Than Spectral Distortion", Proc. IEEE Internat. Conf. Acoust. Speech Signal Process., pp. 732-735.
- Lee M. E. and Smith M. J. T., (2003), "Spectral Modification For Digital Singing Voice Synthesis Using Asymmetric Generalized Gaussians", ICASSP.
- McLoughlin I. V. and Chance R. J., (1997), "LSP-based speech modification for intelligibility enhancement," in Proceedings 13th International Conference on DSP, 1997, vol. 2, pp. 591-594.
- Morris R.W.1; Clements M.A., (2002), "Reconstruction of speech from whispers", Medical Engineering and Physics, Volume 24, Number 7, pp. 515-520(6), September.
- Pozo A. and Young S., (2006), "Continuous Tracheoesophageal Speech Repair", EUSIPCO.

Qi Y., Weinberg B. and Bi N., (1995), "Enhancement of female esophageal and tracheoesophageal speech", Journal of the Acoustical Society of America, vol. 98, pp. 2461-2465.

Sawada H. , Takeuchi N., Hisada A. , (2004), "A Real-time Clarification Filter of a Dysphonic Speech and Its Evaluation by Listening Experiments", International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT2004), pp. 239-246.

Texas Instruments DSPS R&D Center Media Technologies Laboratory, (1998), Final Report for Fixed-Point Implementation of the 2.4 kbps MELP Federal Standard Speech Coder.

Thomas I. B., (1969), "Perceived pitch of whispered vowels," J. Acoust. Soc. Am., vol. 46, no.2, pp. 468.

Wai. C. Chu, (2003), "Speech coding algorithms- foundation and evolution of standardized coders", Wiley-interscience publications.

INTERNET KAYNAKLARI

[1] http://en.wikipedia.org/wiki/Chebyshev_filter

[2] http://geriatri.dergisi.org/pdf/pdf_TJG_111.pdf

[3] <http://www.cancerbackup.org.uk/Cancertype/Larynx/Aftertreatment/Voicerestoration>

[4] <http://www.mehmetomur.com/tr/article.asp?ID=107&C=3>

[5] <http://www.zibec.org/index.php?lang=tr&action=show&id=10>

ÖZGEÇMİŞ

Doğum tarihi 22.05.1983

Doğum yeri İstanbul

Lise 1998-2000 Isparta Süleyman Demirel Fen Lisesi
2000-2001 Isparta Gazi Lisesi

Lisans 2001-2005 Yıldız Teknik Üniversitesi Elektrik – Elektronik Fak.
Bilgisayar Mühendisliği Bölümü

Çalıştığı kurumlar

07.2005-10.2005 Yetkinel Danışmanlık Hizmetleri

2005-Devam ediyor YTÜ Fen Bilimleri Enstitüsü Araştırma Görevlisi