



İnternet Protokolü Üzerinden Ses Taşıma Uygulamaları İçin Dalgacık Dönüşümü Kullanılarak Basit ve Etkili Bir Ses Sıkıştırma Uygulaması

Gürkan TUNA

Trakya Üniversitesi, Edirne Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Teknolojisi ve Programlama Bölümü, EDİRNE

ÖZET

Bilgisayar ağları ve İnternet son yıllarda hızla yaygınlaşmış ve bunun sonucunda yeni uygulamalar ortaya çıkmaya başlamıştır. Bilgisayar ağları ve İnternet üzerinden ses, video ve görüntülerin taşınması geleneksel yöntemlere göre büyük maliyet avantajı sağlamıştır [1]. İnternet üzerinde yapılan bu tür servislerde dikkat edilmesi gereken en önemli nokta uçtan uca olan taşıma süresidir. ITU (International Telecommunication Union) G.114 standardına göre ses taşıma için uçtan uca olan gecikme süresi 0-150 ms arasında olmalıdır [2]. İnternet protokolü üzerinden ses taşıma anlamına gelen VOIP (Voice over IP) teknolojisinde uçtan uca gecikmeyi kontrol altında tutacak bir servis kalitesi sistemi kullanılması kaliteli bir servis için zorunludur [3]. Ancak belirli servisler için garanti edilmiş bant genişliği ve öncelik sağlayan bir servis kalitesi uygulaması tek başına yeterli değildir. Taşınacak sesin daha az yer kaplaması için genellikle sıkıştırma algoritmaları ile önce ses sıkıştırılmakta ve daha sonra taşınmaktadır. Bu makalede VOIP uygulamaları için en uygun genel eşikli ayrık dalgacık dönüşümü kullanılarak basit ve etkili bir ses sıkıştırması sunulmaktadır. Uygulamada kullanılan metot basit ve hızlı bir sıkıştırma işlemi gerçekleştirmekte olup VOIP uygulamalarının gereksinimlerini karşılamaktadır. Uygulama ses sıkıştırma işlemi için işlemci gücü kullanmadan hızlı bir şekilde gerçekleştirmekte, tatmin edici ses kalitesi ve yüksek sıkıştırma oranları elde edilmektedir. Bu yöntem ses sıkıştırma işlemi gerçekleştiren yönlendirici ve VOIP geçitleri için uygundur.

Anahtar Kelimeler

Ayrık dalgacık dönüşümü,
Ses Sıkıştırma,
En uygun genel eşik,
İnternet protokolü üzerinden ses taşıma.

A Simple and Effective Audio Compression Application For Voice Over Internet Protocol Applications By Using Discrete Wavelet Transform

ABSTRACT

In recent years, computer networks and Internet have become widespread quickly, and as a result of this, new applications have started to emerge. The transmission of voice, video and images over computer networks and Internet has brought cost advantages over traditional methods [1]. The most important point in this kind of services over Internet is end-to-end transmission time. According to ITU G.114 standard, for carrying voice, end-to-end delay is to be 0 to 150 ms [2]. A quality of service system that keeps end-to-end delay under control in VOIP technology which means carrying voice over Internet protocol is required for a quality service [3]. But a quality of service application alone that provides guaranteed bandwidth and priority to specific services is not sufficient. Usually voice is compressed by the use of compression algorithms and then carried, to make it occupy less space. This paper presents a simple and effective lossless audio compression using discrete wavelet transform with an optimal global threshold for voice over IP applications. The method used in the application performs a simple and fast compression and satisfies the requirements of VOIP applications. The application performs voice compression quickly without requiring much processing power, and satisfactory audio quality and high compression rates are obtained. This method is suitable for routers and VOIP gateways which perform voice compression.

Keywords

Discrete wavelet transform,
Audio compression,
Optimal global threshold,
Voice over IP

* Sorumlu yazar (Corresponding author) e-posta: gurkantuna@trakya.edu.tr

1. GİRİŞ

Bilgisayar ağları ve İnternet son yıllarda hızla yaygınlaşmış ve bunun sonucunda yeni uygulamalar ortaya çıkmaya başlamıştır. İnternet'in bu yeni uygulama alanları kurumlara büyük maliyet avantajları sağlamıştır. İnternet artık ilk çıkış noktası olan veri iletişiminin dışına çıkmış ve görüntü, ses ve videonun taşınması için de bir platform olmuştur. İnternet'in en önemli iletişim ortamı olmaya başlamasının nedeni sağladığı maliyet avantajları ve farklı servisleri tek bir noktada buluşturmasıdır. İnternet ayrıca farklı servisler için ilave kablolama yapılması gereksinimini de ortadan kaldırmaktadır [4].

İnternet üzerinde yapılan bu tür servislerde dikkat edilmesi gereken en önemli nokta uçtan uca olan taşıma süresidir. Video ve ses taşımada dikkat edilmesi gereken hem bu verilerin taşınacağı noktalar arası bant genişliği hem de iletim süresidir. Bu konuda ITU (International Telecommunication Union) tarafından belirlenmiş olan standartlar bulunmaktadır. ITU G.114 standardına göre ses taşıma için uçtan uca olan gecikme süresi 0-150 ms arasında olmalıdır [2].

İnternet protokolü üzerinden ses taşıma anlamına gelen VOIP (Voice over IP) teknolojisinde uçtan uca gecikmeyi kontrol altında tutacak bir servis kalitesi sistemi kullanılması kaliteli bir servis için zorunludur [3]. Ancak belirli servislere garanti edilmiş bant genişliği ve öncelik sağlayan bir servis kalitesi uygulaması tek başına yeterli değildir. Taşınacak sesin daha az yer kaplaması için genellikle sıkıştırma algoritmaları ile önce ses sıkıştırılmakta ve daha sonra taşınmaktadır. Bütün sıkıştırma algoritmalarının ortak dezavantajları sıkıştırma işlemi için harcanan zaman ve işlem yüküdür. VOIP uygulamalarında analog ses bilgisinin bilgisayar ağlarında taşınması için dijital veriye dönüştürülmesi işlemi yönlendiriciler, VOIP geçitleri veya dijital telefon santralleri üzerinde gerçekleşmektedir. Yönlendiricilerin asıl görevi ses verilerinin dijital veriye dönüştürülmesi değil paketlerin ağ üzerinde yönlendirilmesidir. Dolayısıyla ses dönüşüm ve sıkıştırma işlemlerini gerçekleştiren yönlendiriciler üzerinde basit ve hızlı bir ses sıkıştırma algoritması kullanılması performans artışı sağlayacaktır. Bu çalışmada VOIP uygulamaları için ayrık dalgacık dönüşümü ile ses sıkıştırma uygulaması geliştirilmiştir. Uygulama ses sıkıştırma işlemini hızlı bir şekilde gerçekleştirmekte, tatmin edici ses kalitesi ve sıkıştırma oranları elde edilmektedir. Elde edilen sonuçlar diğer kayıpsız ve kayıplı yöntemlerle karşılaştırılmalı olarak tablolar halinde gösterilmiştir.

Makale giriş bölümü ile başlamakta, sırasıyla Ses Sıkıştırma, Ses Sinyallerinin Gösterimi İçin Dalgacık Dönüşümünün Kullanılması, İşitsel Model, Sıkıştırma İşlemi, Uygulama ve Testler, bölümleriyle çalışma hakkında bilgiler verilmektedir. Makale Sonuçlar ve Gelecekteki Çalışmalar bölümüyle sonuçlandırılmaktadır.

2. SES SIKIŞTIRMA

Ses sıkıştırma daha az yer kaplaması ve iletim sırasında daha az bant genişliği kullanması için ses verisinin kodlanmasıdır. Bu amaca ulaşmak için farklı sıkıştırma yöntemleri geliştirilmiştir. Diğer her dijital veri sıkıştırma yöntemi gibi, bu yöntemler de iki kategoriye ayrılabilir. Kayıpsız sıkıştırma ve kayıplı sıkıştırma.

Kayıpsız ses sıkıştırmada kullanılan kodlama teknikleri analog ses sinyalinin gerçek şeklini sinyalin her bir örneğini farklı kuantalama yöntemleri ile kuantalayarak elde etmeye çalışırlar. Bu teknikler ile ses dalga biçimine yaklaşılmaya çalışılır ve eğer yeterli büyüklükte bir bit oranı kullanılırsa oldukça yakın sonuçlar elde edilir. Analog işaretlerin genlikleri sürekli olduğu için örneklenen analog işaret kuantalama işlemine tabi tutulur ve böylece örneklenmiş işaret değerlerinin önceden belirlenmiş sınırlı sayıdaki seviyeden en yakın olana yuvarlanması yoluyla işaretin alabileceği genlik değerleri sınırlandırılır. Örnekleme ve kuantalama işlemlerinden sonra analog işaretler sayısal işaretlere dönüşür. Örneğin müzik CD'lerinde kullanılan PCM (Pulse Code Modulation: Darbe Kod Modülasyonu) sıkıştırılmamış bir ses formatıdır. Müzik CD'lerinde kullanılan ses sinyalleri 44.1 kHz'de örneklenerek örnek başına 16 bit ile kuantalanır ve sonuçta tek kanal için 705 kbps, iki kanal stereo için 1411 kbps gibi çok yüksek bit oranları elde edilir. Bu sonuçlar elde edilirken kullanılan hesaplama yöntemi aşağıda gösterilmiştir.

Bit oranı = (bit derinliği) x (örnekleme oranı) x (kanal sayısı)

Örneğin 44.1 kHz örnekleme oranı, 2 kanal (stereo) ve 16 bit derinliğinde bir kayıt için:

1 Saniyede $44100 \times 2 \times 16 = 1411200$ bit diğer bir ifadeyle 1411.2 kbit/s

Ses sıkıştırma için kullanılan diğer kayıpsız tekniklerin çoğu artıklık kısımları bularak onları uzaklaştırma yöntemini kullanır. Bazı kayıpsız sıkıştırma teknikleri ise kuantalama işlemini optimize ederek sıkıştırma işlemini gerçekleştirir.

Kayıplı sıkıştırmada ise algısal artıklık niteliğindeki bölümler azaltılır. Bir ses dalgasında insan kulağı tarafından algılanamayacak olan bölümler çıkartılır. Çok yüksek frekanslar, çok düşük frekanslar ve yüksek sesler tarafından bastırılan alçak sesler çıkartılır. Kayıplı sıkıştırma algoritmalarının çoğunda zaman bölgesinde örneklenmiş dalga biçimleri MDCT (Modified Discrete Cosine Transform: Değiştirilmiş Ayrık Kosinüs Dönüşümü) gibi dönüşümler kullanılarak frekans bölgesine dönüştürülür. Frekans bölgesine dönüştürüldükten sonra frekans bileşenleri ne kadar duyulabileceklerine göre ayrıştırılır. Spektral bileşenlerin duyulabilirliğini belirlemek için insan kulağının duyamayacağı sesleri gösteren bir maskeleme eşik seviyesi belirlenir. Bu maskeleme seviyesinin üzerindeki seslerin insan kulağı tarafından algılanamayacağı kabul edilir. MDCT büyük bir veri setinin takip eden blokları üzerinde uygulanmak için tasarlanmış olup bir bloğun son yarısı ile takip eden bloğun ilk yarısını çakıştırmak için ardışık bloklar örtüştürülür. MDCT bu örtüştürme özelliği ve DCT (Discrete Cosine Transform: Ayrık Kosinüs Dönüşümü) dönüşümünden gelen özellikleriyle sinyal sıkıştırma uygulamaları için ideal bir tekniktir.

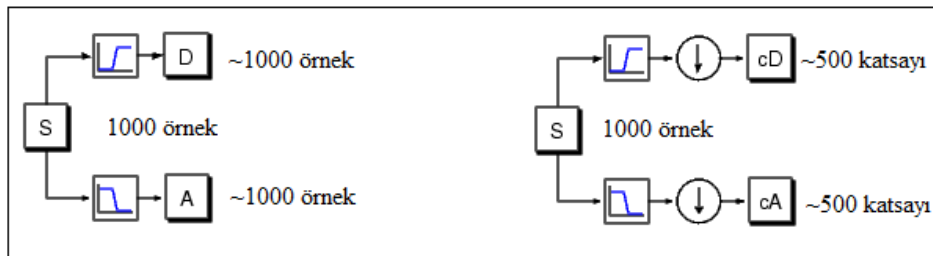
Ses sıkıştırma tekniklerinde dikkat edilesi gereken bir nokta da dinamik sözlük tabanlı sıkıştırma teknikleri uzun bir kodlama gecikmesine neden olmaktadır. Dolayısıyla bu tekniklerin gerçek zamanlı yada VOIP gibi gecikmeye duyarlı uygulamalarda kullanılması zordur.

3. SES SİNYALLERİNİN GÖSTERİMİ İÇİN DALGACIK DÖNÜŞÜMÜNÜN KULLANILMASI

Dalgacık dönüşümü son yıllarda özellikle ses ve görüntü işlemede yaygın olarak kullanılmaya başlanmıştır. Dalgacık dönüşümü iki farklı yöntemle uygulanabilir. Bu yöntemler CWT (Continuous Wavelet Transform: Sürekli Dalgacık Dönüşümü) ve DWT (Discrete Wavelet Transform: Ayrık Dalgacık Dönüşümü)'dir. DWT sinyallerin ve görüntülerin analizi ve ayrıştırılmasında kullanılan bir tekniktir [6].

Dalgacık dönüşümü sırasında dalgacık katsayıları oluşur. Mümkün olan her ölçekte dalgacık katsayılarının hesaplanması iş yükü gerektirir ve çok fazla veri oluşturur [7]. Sadece ikinin katları olan ölçeklerin ve konumların seçilmesi durumunda elde edilen analiz yeterli olacaktır. DWT bu analiz için uygun yöntem olacaktır. Birçok sinyal için düşük frekans içeriği en önemli veridir. Yüksek frekans içeriği ise ayrıntıları oluşturan ve düşük frekansa değer katan verilerdir. İnsan sesi örneği ele alınırsa yüksek frekans bileşenlerinin yok edilmesi sesi farklılaştıracak ancak ne söylendiği yine de anlaşılacaktır [7]. Ancak alçak frekans bileşenlerinin yok edilmesi sesi anlamsızlaştıracaktır.

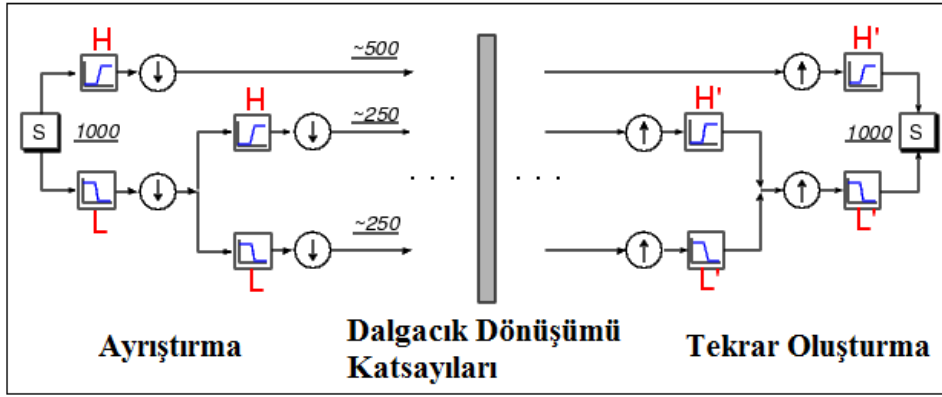
Dalgacık dönüşümünden sonra A (approximations: yaklaşıklıklar) ve D (details: detaylar) katsayıları oluşur. A sinyalin yüksek ölçekli düşük frekans bileşenlerini, D ise sinyalin düşük ölçekli yüksek frekans bileşenlerini gösterir. Dalgacık dönüşümünde giriş sinyali S iki ayrı filtreden geçer ve sonuçta iki ayrı sinyal oluşur. Giriş sinyali 1000 örnekten oluşuyorsa sonuçta iki adet 1000 örnekli sinyal olacaktır. Bunun yerine alt-örnekleme denilen işlem uygulanır ve çıkış sinyalinde her iki noktadan bir tanesi alınır [6]. Bu işlem şekil 1'de gösterilmektedir.



Şekil 1. Ayrık Dalgacık Dönüşümü ve Alt-Örnekleme İşlemi

Alt-örnekleme işlemi uygulanmış bir sinyal direkt olarak birleştirilemez. Öncelikle üst-örnekleme işlemi uygulanması gereklidir [8]. Dalgacık

dönüşümü çok kademeli olarak uygulanabilir. Şekil 2, çok kademeli dalgacık ayrıştırma ve tekrar oluşturma işlemini göstermektedir.



Şekil 2. Çok Kademeli Ayrık Dalgacık Dönüşümü Ayrıştırma ve Tekrar Oluşturma İşlemleri

Ses işleme uygulamalarında DWT sinyal gösterimi kullanılmasının nedeni DWT'nin verilen bir sinyalle eşleştirilebilecek oldukça esnek bir sinyal gösterim ailesi olması ve ses verisi sıkıştırma görevine iyi uygulanabilir olmasıdır [5]. Dalgacık dönüşümü kullanılarak ses sıkıştırmanın anahtar noktalarından biri sıfır olmayan dalgacık katsayılarının sayısını azaltmak için gerekli olan en uygun taban aramadır. Ancak en uygun taban arama büyük bir programlama tasarımı gerektirir. Bunu telafi etmek için bu uygulamada başka bir sıkıştırma tekniği kullanılmaktadır. Bu teknik en uygun genel eşik kullanan ayrık dalgacık ayrıştırmasına dayanmaktadır. Bu teknik ses sıkıştırmada başarıyla kullanılmıştır.

4. İŞİTSEL MODEL

İşitsel model insan algılaması üzerine yapılan çalışmalara dayanmaktadır. Bu çalışmalarda insanların bütün frekansları aynı şekilde duymadığı ortaya konmuştur. Ortamda bulunan farklı seslerin etkileri ve insan algılamasındaki kısıtlamalar dolayısıyla bir ses sinyalindeki algılanamayan kısımların çıkartılabileceği fikri ortaya atılmıştır. İşitsel modeli meydana getiren iki temel unsur mutlak duyma eşiği ve işitsel maskeleme modelleridir.

4.1 Mutlak duyma eşiği

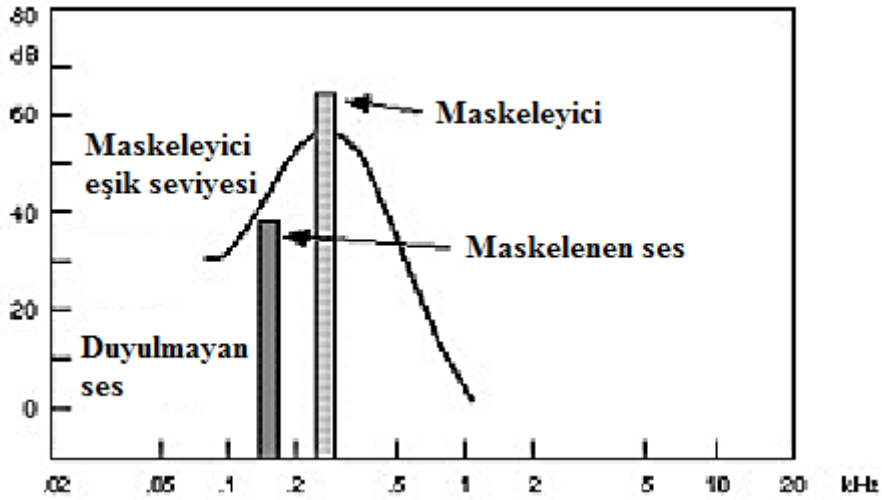
İnsanlar 20 Hz – 20.000 Hz arasındaki frekansları duyabilirler. Ancak bu bütün frekansların aynı şekilde duyulduğu anlamına gelmez. Konuşmayı meydana getiren frekanslar daha kolay duyulur. Alt ve üst frekanslara yakın sesler ise daha zor duyulur. İnsanlar konuşmayı meydana getiren düşük frekanslardaki sesleri 20 kHz civarındaki yüksek frekanslardaki seslerden daha iyi duyabildiği için düşük frekanslardaki ton değişimlerini yüksek frekanslardakinden daha kolay ayırt ederler.

Örneğin 500 Hz - 600 Hz arasındaki fark, 17.000 Hz – 18.000 Hz arasındaki farka göre daha çabuk anlaşılır. Duyma konusunda yapılan birçok çalışmanın neticesinde, 20 Hz – 20.000 Hz arasındaki frekansların birbirine eşit olmayan, doğrusal artış göstermeyen ve duyulan sese bağlı olan kritik bant genişliklerine bölünebileceği ortaya konmuştur. Kritik bant genişliği içinde bulunan sesler insanlar tarafından zor ayırt edilir.

Bir ses dalgasının frekansının işitme üzerindeki etkilerini ortaya koymak için yapılan çalışmalarda her bir frekansta belirli bir minimum güç seviyesinin altındaki seslerin duyulmadığı ortaya konmuş ve insanların duyabildiği frekans aralığı için bir eşik seviyesi belirlenmiştir. Buradan hareketle ses sıkıştırma işlemlerinde eğer bir sinyalin frekans bileşenlerinin güç seviyeleri mutlak duyma eşiğinin altına düşerse bu frekansların insanlar tarafından duyulamayacağı varsayılır ve bu frekans bileşenleri yok kabul edilir.

4.2 İşitsel maskeleme

İnsanlar frekanslardaki çok küçük değişimleri algılayamazlar. Örneğin 1001 Hz'lik bir sinyalden 1000 Hz'lik bir sinyali ayırt edemezler. Bu durum eğer iki sinyal aynı anda alınırsa daha da zorlaşır. Bunun yanı sıra, 1000 Hz'lik bir sinyal 1010 Hz, 1100 Hz, 990 Hz gibi yakın frekanslardaki sinyallerin duyulmasını da etkiler. Bu kavram maskeleme olarak adlandırılır. Eğer 1000 Hz'lik sinyal güçlüyse yakın frekanslardaki sinyalleri maskeler ve maskelenen bu frekanslar dinleyici tarafından duyulamaz. Maskelenen bir sinyalin duyulabilmesi için bu sinyalin gücünün maskeleyici tonun frekansı ve gücü tarafından belirlenen eşik seviyesinden daha yükseğe getirilmesi gerekir. Gürültü de maskeleyici olabilir. Eğer gürültü çok güçlüyse bir ses tonunu maskeleyebilir.



Şekil 3. Maskeleyici ve Maskelenen Ses

Dolayısıyla sıkıştırma algoritmalarında ton maskeleyiciler, gürültü maskeleyiciler ve maskeleyicilerden kaynaklanan maskeleyici etkileri belirlenmelidir. Bu maskeleyicilerin maskeleyici eşik seviyelerinin altına düşen bütün frekans bileşenleri yok kabul edilir. Bir frekans bileşeninin belirli bir zaman dilimi için sabit olup olmayacağı ve bir frekans spektrumundaki keskin bir tepe noktası olup olmadığı bilinmesi gereken önemli unsurlardır.

5. SIKIŞTIRMA İŞLEMİ, UYGULAMA VE TESTLER

Uygulama VOIP servislerinde kullanılacağı için uygulamada basitlik ve hız ön planda tutulmuştur. Ses sıkıştırma işleminde kullanılan DWT tabanlı kodlayıcının önemli bileşenlerinden birisi maskeleme modelidir [9,10]. Maskeleme modelinde amaçlanan her bir alt bantın işitsel eşik seviyesini dalgacık kısıtına dönüştürmektir. İşitsel maskeleme hem maskeleme sinyalinin hem de maskelenen sinyalin zaman ve frekansına bağlıdır.

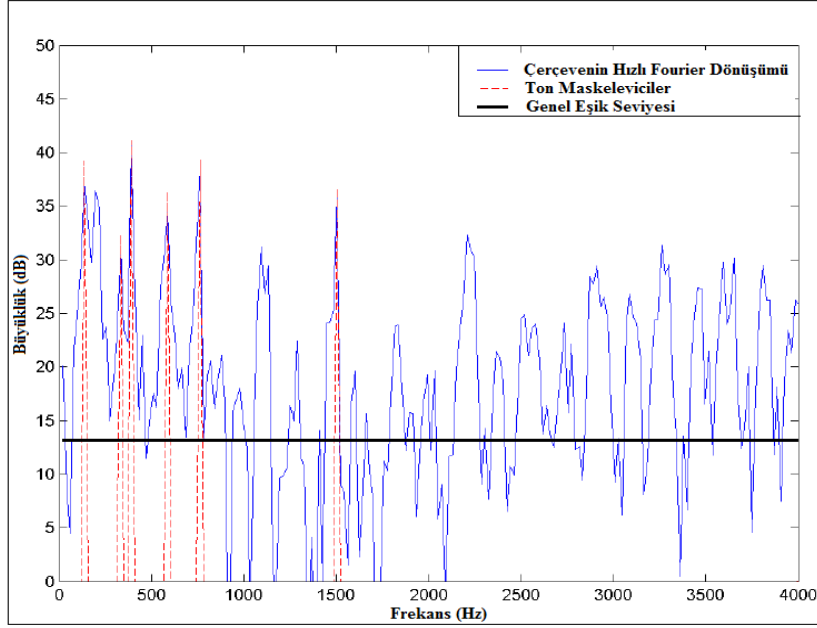
Yapılan çalışmalarda [9,10] her bir frekanstaki toplam maskelenen güç her bir frekanstaki sinyal bileşenlerinin maskelenmiş güçlerinin toplanmasıyla tahmin edilmiştir. Böylece her bir bantta bulunan minimum maskelenmiş güç elde edilir. Dinleyen bir kişi her bir frekanstaki güç spektrumunun maskelenmiş olan gürültü gücünden az olması koşuluyla tekrar üretilmiş ses sinyalindeki ilave gürültüye tolerans gösterecektir [10]. Maskelemenin bazı dezavantajları olsa bile birçok kodlama tekniğinde başarıyla kullanılmıştır [9,10].

İşitsel maskeleme modelinin yapısının karmaşık olması dolayısıyla bu uygulamada maskeleme modeli basitleştirilmiştir. Kullanılan model bir sinyalde sadece maskeleyici tonları tespit etmekte ve bütün frekanslar için genel bir eşik seviyesi vermektedir. Önceki çalışmalarda [11] kullanılan model gürültüyü de işlemekte ve her bir alt bant için bir eşik seviyesi vermektedir. Bu uygulamada kullanılan model şekil 4'de gösterilmektedir.

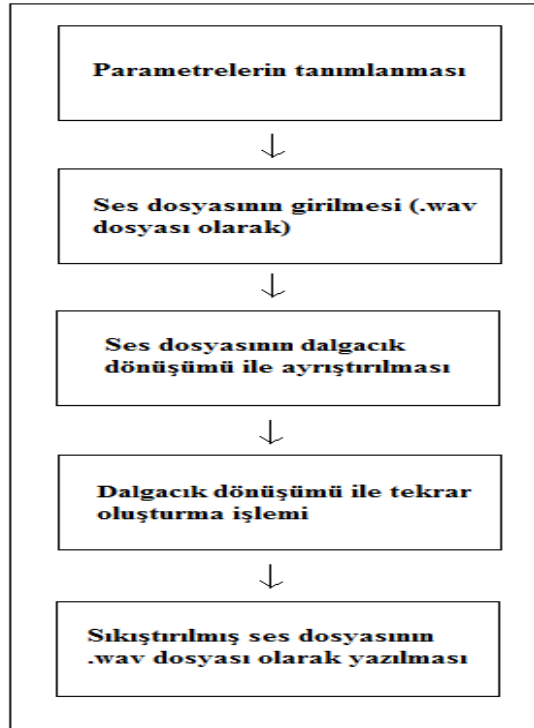
Uygulamada sıfır olmayan dalgacık katsayılarını azaltmak için en uygun taban arama gerçekleştirilmemiştir. Bunun yerine en uygun genel eşik kullanan ayrık dalgacık ayrıştırması ile sıkıştırma gerçekleştirilmiştir. Yapılan çalışmalarda [9] en iyi sonuçların Daubechies dalgacık tekniği ile edildiği görülmüştür. Yapılan çalışmada [9], dalgacık doğruluk düzeyi (polinomun yaklaşım düzeyi) 10 ve 5 kademeli dalgacık ayrıştırma kullanılmıştır. Bu seçimler en uygun taban arama zorunluluğunun üstesinden gelmeyi sağlayacaktır.

Uygulama kodunun akış diyagramı şekil 5'de gösterilmektedir. Uygulama parametrelerin tanımlanması ile başlamaktadır. Dalgacık dönüşümü parametreleri, ayrıştırma ve tekrar oluşturma işlemlerinde kullanılan parametreler burada tanımlanmıştır. İkinci aşamayı ses dosyasının girilmesi oluşturmaktadır. Benzetim çalışmaları önceden kaydedilmiş ses dosyaları üzerinde yapılmıştır. Üçüncü aşamada ilk aşamada girilen çerçeve boyutu ve dalgacık dönüşümü parametreleri kullanılarak ses dosyası ayrıştırılmaktadır. Dördüncü aşamada ayrık dalgacık dönüşümünün ikinci aşaması gerçekleştirilmekte ve ses dosyası tekrar oluşturulmaktadır. Uygulamanın son aşamasında

dalgacık dönüşümü ile tekrar oluşturulmuş olan ses
.wav uzantılı dosyaya yazılmaktadır



Şekil 4. Bir Çerçevdeki Ton Maskeleyici Tespiti



Şekil 5. Uygulama Kodu Akış Diyagramı

CD Ses kalitesinde 1411 Kbps bit oranına sahip 10 farklı wav dosyası ile yapılan testlerin FLAC (Free Lossless Audio Codec) kayıpsız sıkıştırma yöntemi

ve MP3 sıkıştırma yöntemi ile yapılan karşılaştırmaları tablo 1'de gösterilmektedir. FLAC kayıpsız bir sıkıştırma olup, FLAC dönüştürme

işlemleri dBpoweramp Music Converter programı ile gerçekleştirilmiştir. MP3 kayıplı bir sıkıştırma olup, MP3 dönüştürme işlemleri Nero 7 Ultra Edition programı ile gerçekleştirilmiştir.

Tablo 1. Önerilen Yöntem ile FLAC Kayıpsız Sıkıştırmanın ve MP3 Kayıplı Sıkıştırmanın Karşılaştırılması

Sıkıştırma Tekniği	Bant Genişliği	Sıkıştırma Oranı
Önerilen Yöntem	282 Kbit/s	5:1
dBpoweramp Music Converter	705 Kbit/s	2:1
MP3	128 Kbit/s	12:1

64 Kbps bit oranına sahip wav dosyaları ile yapılan testlerin sonuçları tablo 2’de gösterilmektedir. 64 Kbps PSTN sabit telefon hatlarında erişilebilecek maksimum bant genişliğidir.

Tablo 2. Önerilen Yöntemin Sabit Telefon Hatlarında Kullanımı

Sıkıştırma Tekniği	Bant Genişliği	Sıkıştırma Oranı
Önerilen Yöntem	12.8 Kbit/s	5:1

Sıkıştırma işlemi sonucunda elde edilen değerlerin VOIP uygulamalarında kullanılan sıkıştırma yöntemleriyle karşılaştırılması tablo 4’de verilmiştir. 20 adet 30 sn’lik 64 Kbps bit oranlarında kaydedilmiş olan denek sesleri ile hesaplamalar yapılarak ortalaması alınmıştır. Karşılaştırma işlemi ITU tarafından standartları belirlenmiş G711, G726, G728, G729 ve G723

kodlayıcı aileleriyle yapılmıştır. Tablo 4’de sadece ses sıkıştırma işleminde farklı yöntemlerde harcanan süreler gösterilmektedir. Ancak ITU standartlarına göre 150 ms altında olması gereken uçtan uca taşıma süresi sadece ses sıkıştırma için harcanan süreye bağlı değildir. Sesin alıcı ve gönderici arasındaki hatta taşınması için harcanan süre de dikkate alınmalıdır.

Elde edilen ses kalitesini belirleyen MOS skoru hesaplanırken 10 erkek ve 10 kadın olmak üzere toplam 20 denekten alınan veriler kullanılmıştır. MOS (Mean Opinion Score) skorunun hesaplanmasında farklı yöntemler bulunmasına rağmen pratik bir çözüm olduğu için genelde Sübjektif Değerlendirme metodu kullanılır. Sübjektif Değerlendirme metodunda bir grup erkek ve kadın denek duydukları seslere tablo 3’de ki veriler ışığında not verirler. Tüm notların ortalaması MOS skorunu verir.

Tablo 3. MOS Skorunun Derecelendirilmesi

MOS	Kalite	Ses Kalitesindeki Bozulma
5	Mükemmel	Farkedilemez
4	İyi	Farkedilebilir ancak rahatsız edici değil
3	Orta	Biraz rahatsız edici
2	Düşük	Rahatsız edici
1	Kötü	Oldukça rahatsız edici

Sıkıştırma işlemi sonucunda elde edilen ses kalitesi telefon hatlarında elde edilen sese göre çok daha iyidir. Ancak mp3 ve wma ses dosyalarına göre daha düşüktür. Uygulama testlerinde elde edilen ortalama sıkıştırma oranı 5:1’dir. Testlerde örnek başına ortalama 5 bit kullanılmıştır. Uygulamanın diğer sıkıştırma yöntemlerine göre avantajı sıkıştırma işlemi diğer yöntemlere göre daha kısa sürede gerçekleştirmesidir. Elde edilen ses kalitesi VOIP uygulamalarının gereksinimlerini fazlasıyla karşılamaktadır.

Tablo 4. Önerilen Yöntemin VOIP Sıkıştırma Yöntemleri ile Karşılaştırılması

Sıkıştırma Yöntemi	Bit Oranı (kbps)	MOS Skoru	Sıkıştırma Gecikmesi (ms)
G.711 PCM	64	4.1	0.75
G.726 ADPCM	32	3.85	1
G.728 LD-CELP	16	3.61	3 to 5
G.729 CS-ACELP	8	3.92	10
G.729 x 2 Encodings	8	3.27	10
G.729 x 3 Encodings	8	2.68	10
G.729a CS-ACELP	8	3.7	10
G.723.1 MP-MLQ	6.3	3.9	30
G.723.1 ACELP	5.3	3.65	30
Önerilen Yöntem	12	3.87	3.5

6. SONUÇLAR

Bu çalışmada VOIP uygulamaları için geliştirilmiş olan bir ses sıkıştırma yöntemi ele alınmıştır. VOIP uygulamaları gecikmeye duyarlı olup, uçtan uca ses taşıma süresinin kaliteli bir servis için ITU standartlarına göre 150 ms altında bir değerde olması gerekmektedir [2]. VOIP uygulamalarında ses sıkıştırılmakta ve daha sonra taşınmaktadır. Ses sıkıştırma sırasında bu işlemi gerçekleştiren yönlendiricinin veya VOIP geçidinin işlemci ve bellek gibi kaynakları kullanılmaktadır. Yoğun ses taşınan bilgisayar ağlarında sistem kaynaklarının yüklü olması durumunda ses sıkıştırma süresi uzamakta veya değişkenlik gösterebilmektedir. Ses sıkıştırma süresinin uzaması uçtan uca olan taşıma süresinin de uzamasına neden olabilmektedir. Bunun sonucunda servis kalitesinde düşmeler meydana gelebilmektedir.

Bu çalışmada kullanılan ses sıkıştırma işlemi dalgacık dönüşümü teknikleri ile gerçekleştirilmiştir. Yöntem basit ve hızlı bir sıkıştırma işlemi gerçekleştirmekte olup VOIP uygulamalarının gereksinimlerini karşılamaktadır. Yöntem fazla işlemci gücü ve bellek gerektirmemekte ve sıkıştırma işlemini çok kısa bir süre de tamamlamaktadır. Alınan sonuçlara göre bu sıkıştırma yöntemi VOIP uygulamalarına uygun gözükmektedir. Bu yöntem özellikle düşük kapasiteli yönlendirici ve VOIP geçitler için uygundur.

Çalışma da kullanılan algoritmanın benzetimleri MATLAB altında gerçekleştirilmiş ve Fedora Linux işletim sisteminde C++ dili ile yazılan örnek bir uygulama ile algoritma test edilmiştir. Çeşitli paket boyları ile gerçekleştirilen testlere göre yöntem diğer sıkıştırma yöntemlerine göre çok daha kısa sürede sıkıştırma işlemini gerçekleştirmektedir. Elde edilen sıkıştırma oranları, ses kalitesi ve

sıkıştırma için harcanan süreler diğer yöntemlerle karşılaştırılmalı olarak tablo 4’ de gösterilmiştir.

KAYNAKLAR

1. Zeadally, S., Siddiqui, F., Kubher, P., (2004). Voice over IP in intranet and Internet environments, *Communications, IEE Proceedings*, **151** (3), 263 – 269.
2. International Telecommunication Union (ITU), (1996). One-Way Transmission Time (Recommendation G. 114).
3. Shenquan, W., Mai, Z., Dong, X., Wei Z., (2006). Design and implementation of QoS-provisioning system for voice over IP, *IEEE Transactions on Parallel and Distributed Systems*, **17**(3), 276 – 288.
4. Hillenbrand, M., Gotze, J., Muller, P., (2005). Voice over IP – Considerations for a Next Generation Architecture, *Proceedings of 31st EUROMICRO Conference on Software Engineering and Advanced Applications*, 386 – 393.
5. Tewfik, A.H., Sinha, D., Jorgensen, P.,(1992). On the optimal choice of a wavelet for signal representation, *IEEE Transactions on Information Theory*, **38**(2), 747 – 765.
6. Weeks, M., (2006), *Digital Signal Processing Using MATLAB and Wavelets*, Jones & Bartlett Publishers, Boston.
7. Walnut, D. F., (2004), *An Introduction to Wavelet Analysis*, Birkhäuser Boston, Boston.
8. Fleet, P. V., (2008), *Discrete Wavelet Transformations: An Elementary Approach with Applications*, Wiley-Interscience, New York.
9. Sinha, D., Tewfik, A., (1993). Low Bit Rate Transparent Audio Compression using Adapted Wavelets, *IEEE Trans. ASSP*, **41** (12).
10. Agbinya, J.I., (1996). Discrete Wavelet Transform Techniques in Speech Processing, *IEEE Tencon Digital Signal Processing Applications Proceedings*, 514-519.
11. Srinivasan, P., Jamieson, L. H., (1998). High Quality Audio Compression Using an Adaptive Wavelet Packet Decomposition and Psychoacoustic Modeling, *IEEE Transactions on Signal Processing*, **46** (4).