



1. GİRİŞ:

Ses, işareti, binlerce yıldır olduğu gibi bugün de insan-noğlunun en önemli iletişim aracıdır. Ses işaretini sayısal işleme çalışmaları son yirmi yılda ivmelenen bir biçimde yoğunluk kazanmıştır, önceleri, sayısal ses işleme için gerekli donanımından beklenen hızlı ve ekonomik olabilme, uygun bir hacime sığabilme gibi özelliklerin tümü aynı anda sağlanamadığından çalışmalar araştırma ağırlıklı kalmıştır, on yılı bile bulmayan bir süredir ise, *çok büyük ölçekli* tümdevre teknolojisindeki yenilikler uygulama alanında önemli gelişmelere neden olmuştur.

SAYISAL SES İLETİŞİMİ

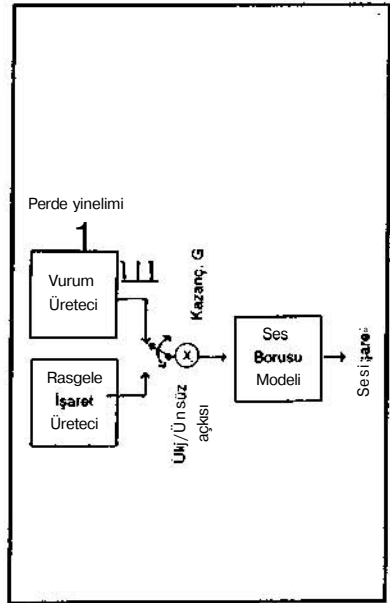
Uygulama açısından ses iletişimini, insanın insanla iletişimi ve insan makina iletişimi olarak ikiye ayırmak yararlı olacaktır, insanın insanla iletişimi, iki ya da daha çok insanın telefon ağı ya da başka bir veri iletim ortamı içindeki iletişimidir, buradaki temel sorun, kullanıcı sayısının ve kullanım hacminin sürekli artması nedeniyle varolan iletim ortamlarının daha verimli kullanılabilmesidir. Verimlilik, ses işaretinin iletimi için gerekli sıklık aralığının, işaretin niteliğinden olabildiğince az yitirerek daraltılması, bir başka deyişle iletim için gerekli ikil sayısının azaltılması ile gerçekleştirilebilir. Elektrik mühendisliğinin konu ile ilgili dalına ses sıkıştırma ya da ses kodlama adı verilmektedir.

insan makina iletişimi ise bir makinanın konuşmayı tanıması, konuşan kişiyi tanıması, kullanıcıya sesle yanıt vermesi, yazılı metnin konuşma-

ya, konuşmanın yazılı metine dönüştürülmesi gibi alt dalları içermektedir, insan makina iletişiminin, beden özürülülerin yaşamlarını kolaylaştırmaktan oyuncak endüstrisine, pilotlar gibi elleri ve gözleri sürekli meşgul insanlara ek olanaklar sağlamaktan güvenlik dizgelerine kadar çok geniş bir uygulama alanı vardır.

2. SES İŞARETİNİN ÜRETİMİ, ÖZELLİKLERİ VE ALGILANMASI

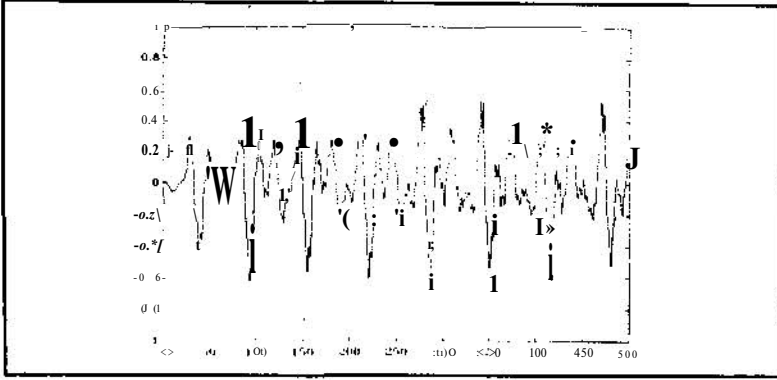
Ses işaretinin sıkıştırılabilmesi için taşıdığı özelliklerin bilinmesi ve bu özelliklerden yararlanmak gerekmektedir. Ses işareti, akciğerlerde sıkıştırılan havanın nefes borusu yoluyla, ses kırıları arasından yutak, gırtlak, ağız boşluğu ve burun boşluğuna üflenmesiyle oluşur^{*1}. Olayı kendi amaçlarımız doğrultusunda modellediğimiz zaman, ses kırılarından dudaklara kadar olan ve ses borusu olarak adlandırdığımız anatomik yapı, doğrusal ve zamanla değişen bir dizgeye soyutlanır. Öbek çizimi şekil-1 'de verilen bu dizgenin uyarımı, ses kırılerinin açılıp kapanmasıyla oluşan hava akımı, yanıtı ise dudaklar aracılığı ile boşluğa iletilen ses işaretidir. Ses kırıleri *ünlü* seslerin üretiminde yinelimli olarak açılıp kapanırlar. Bu



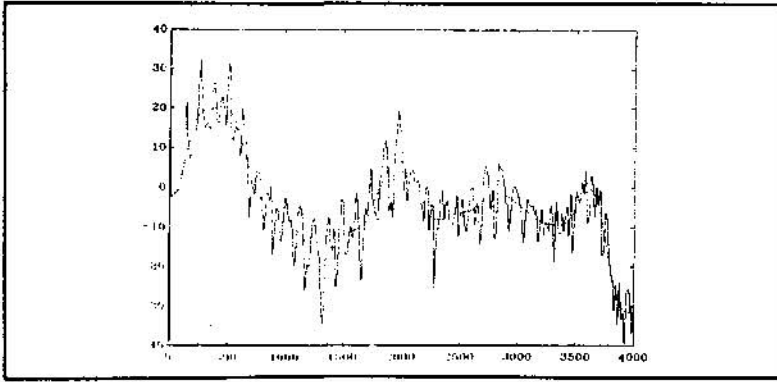
Şekil 1: SES İŞARETİ üretim modeli

Tuğrul MEMİŞOĞLU*

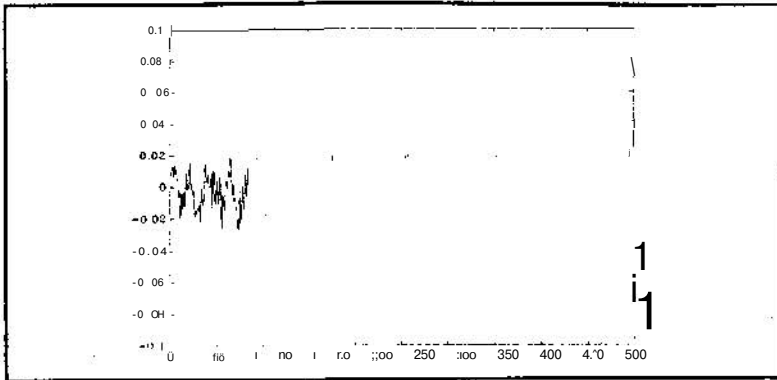
OO.D.T.Ü. Elektrik-Elektronik Mühendisliği Bölümü



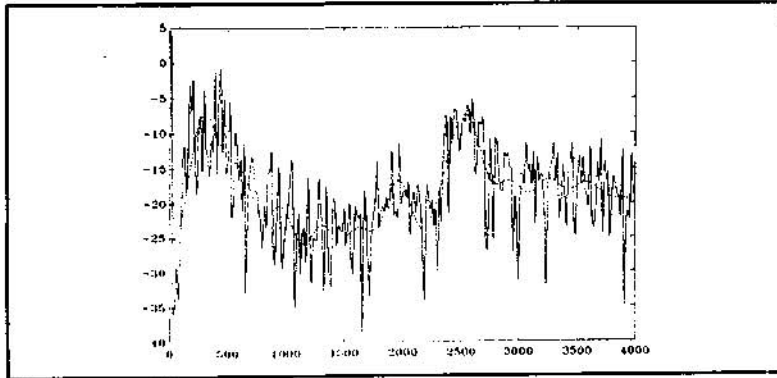
Şekil 2 a : ÜNLÜ BİR ses dilimi



Şekil 2 b: ŞEKİL 2 A VAKİ DİLİMİN genlik görünüşü



Şekil 2 c : ÜNSÜZ BİR ses dilimi (62.25 msec)



Şekil 2 d : ŞEKİL 2 c DEKİ DİLİMİN genlik görünüşü

nedenle ünlü sesler için uyarım, bir vurum dizisi olarak modellenebilir. Bu dizinin yinelimi *perde yinelimi* olarak adlandırılır. Ünsüz seslerde ise ses girişleri tümüyle açılır ve geniş görüneli bir rasgele süreç olarak modellediğimiz uyarımı yaratırlar.

Ses borusunun çınlanım sıklıklarına *formant* sıklıkları denir. Şekil-2.a'da ünlü bir ses dilimi, şekil-2.b'de bu dilimin genlik görünüşü, benzer biçimde şekil-2.c'de ünsüz bir ses dilimi, şekil 2.d'de ise bu dilimin genlik görünüşü verilmiştir. Şekil-2.b ve 2.d'deki kesikli çizgiler ses borusunun genlik tepkisini göstermektedir.

Ses işaretinin önemli özelliklerinden bazıları şöyledir:

- Ses örneklerinin uzun dönemdeki dağılımı yaklaşık Laplace dağılımıdır (Şekil-3). Daha basit bir anlatımla, işaret sıfır çevresinde yoğundur.
- Ses işareti çoğu kez örneklediği 8-10 kHz örnekleme hızına göre yavaş değişen bir işarettir, birbirini izleyen örnekler arasındaki genlik farkı görece düşüktür. Bir başka deyişle, her örnek kendi çevresindeki örneklerle ilişkin bilgi taşır. Bu nedenle, *kestirime* yönelik algoritmalar ses işleme konusunda çok başarılıdır.
- Ses işareti *durağan* bir işaret değildir. Ancak, 15-25 milisaniyelik zaman aralıkları için yerel anlamda durağan olduğundan söz edebiliriz. Bu özellik *uyarlamalı* ses kodlama yöntemlerinin daha başarılı olmasına neden olmaktadır.

insan kulağı bir Fourier görüneli çözümlenici gibi çalışmaktadır. Ancak kulak, her sıklık aralığına diğerleri ile eşit duyarlılıkta değildir. Alçak sıklıklarda yüksek enerjili işaretlerin varlığı, yüksek sıklıklardaki duyarlılığı azaltmaktadır. Bu olaya, *işitsel maskeleye* adı verilmektedir. Bundan başka kulak, ses işaretinin evresindeki değişimlere görece az duyarlıdır.

3. SES KODLAMA

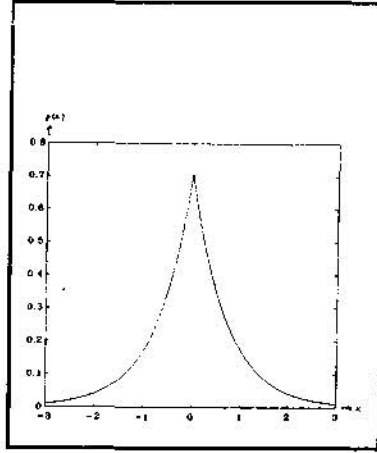
Ses kodlama yöntemlerini *dalga-biçimi kodlama* ve *kaynak kodlama*

olmak üzere iki ana başlıkta toplanabiliriz. Dalga-biçimi kodlama ve kaynak kodlama, zaman bölgesinde yapılabilecekleri gibi uygun bir dönüşüm bölgesinde de yapılabilirler. Dalga-biçimi kodlamada ses işaretinin dalga biçiminin korunarak yeniden üretimi amaçlanmaktadır. Kaynak kodlamada ise ses işaretinin gözlenerek, onu üreten dizgenin modellenmesi üzerinde durulmaktadır. Kaynak kodlayıcılar işaretin kendisini değil üreteç modelinin parametrelerini iletirler. Sesin yeniden üretimi ise modelin ilgili parametreleri uyarınca işletilerek *bireşimsel* ses elde edilmesi biçiminde olur. Bir başka anlatımla, kaynak kodlama tabanındaki bir iletişim dizgesi, çözümlenici—bireştirici ikilisinden oluşur. Vericide, çözümlenme yapılarak ses işareti bir dizi parametreye indirgenir. Alıcıda ise bu parametrelerden özgün işaretin benzeri elde edilir. Kaynak kodlama, dalga-biçimi kodlamadan ileri bir sıkıştırımı sağlar. Buna karşılık, kaynak kodlanmış sesin niteliği görece düşüktür ve kaynak kodlayıcılar görece karmaşıktır. Şekil-4'de dalga-biçimi ve kaynak kodlama yöntemleri ile yeniden üretilen sesin niteliğinin veri iletim hızı karşısındaki değişimi gösterilmektedir. Şekilde nitelik ölçüsü olarak *ortalama kanı*, bir başka deyişle çok sayıda dinleyicinin ortalama öznel değerlendirmesi kullanılmıştır. Değerlendirme özgün sese göre yapılır ve özgün ses 5 puan olarak değerlendirilir.

Ne tür bir kodlama yöntemi kullanılırsa kullanılsın sayısal ses işlemede ilk adım, ses işaretinin örneklenmesidir. Bunun için işaret önce 100-3,500 Hz arasını geçiren bir süzgeçle süzülür ve 8 kHz sıklıkta örneklerin (Şekil-5).

3.1. Dalga - Biçimi Kodlama Yöntemleri

Vurum Kod Kiplenme: Bu yöntemde her bir örnek diğerlerinden bağımsız olarak *nicemlenir*³¹ Nicemleme, işaretin değer aldığı kümeden, o kümenin sonlu bir alt-kümesine yapılan bir eşlemedir. Böylelikle, her bir örnek sonlu sayıda ikil ile ifade edilebilir. Şekil-6'de 8 düzeyli *düzensiz* bir nicemleyici



Şekil 3: LAPLACE olasılık yoğunluğu

gösterilmiştir. 8 düzeyi ifade edebilmek için 3 ikile gereksinim vardır.

Ses işaretinin sıfır çevresinde yoğun olma özelliği, sıfır çevresinde görece ince ve genlik arttıkça kabalaşan, düzensiz olmayan nicemleyicileri gündeme getirmiştir. Standartlaş-

mış düzensiz olmayan nicemleyiciler 64,000 ikil/saniyede yüksek nitelikli ses üretirler³¹.

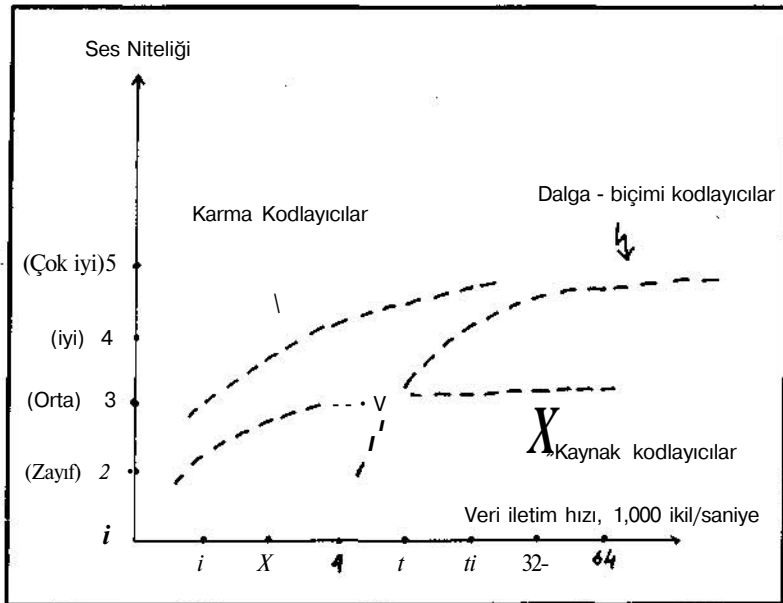
Nicemleyicinin önemli bir parametresi *adım boyu* 4'dür. Adım boyu sabit tutulabileceği gibi, işaretin yerel gücüne bağlı olarak zamanla değiştirilebilir. Adım boyu zamanla değişen vurum kod kiplenme dizgelerine uyarlamalı vurum kod kiplenme dizgeleri denir.

Ayrımsal Vurum Kod Kiplenme:

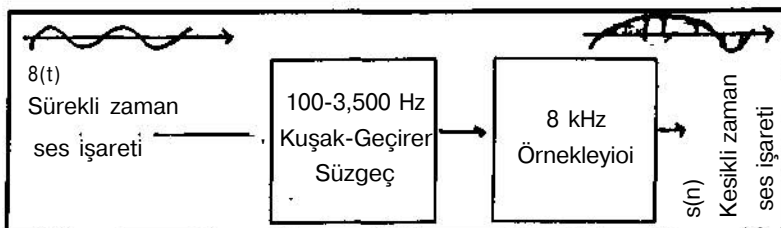
Bu yöntem doğrusal kestirim ilkesine dayanır³¹. $s(n)$ örneği $s(n-1)$, $s(n-2)$... $s(n-p)$ örneklerinin bir doğrusal bileşkesi olarak ifade edilir (Şekil 7).

$$\hat{s}(n) = \sum_{k=1}^p \alpha_k \cdot s(n-k) \quad (1)$$

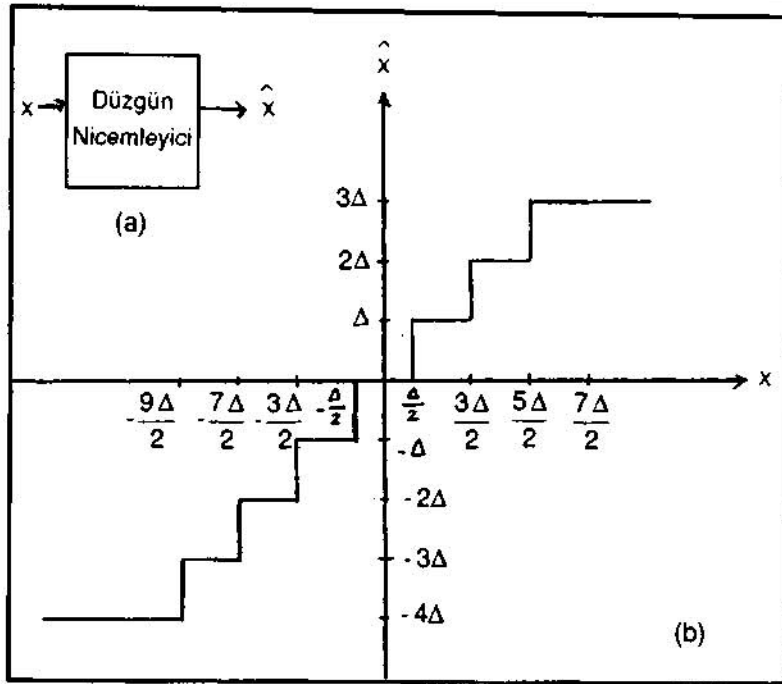
α_k sabitleri, $s(n)$ ile kestirimi $\hat{s}(n)$ arasındaki hatanın ortalama karesini en-azlayacak biçimde seçilirler.



Şekil 4: DEĞİŞİK TÜR KODLAYICILAR için veri iletim hızı-öznel ses niteliği ilişkisi

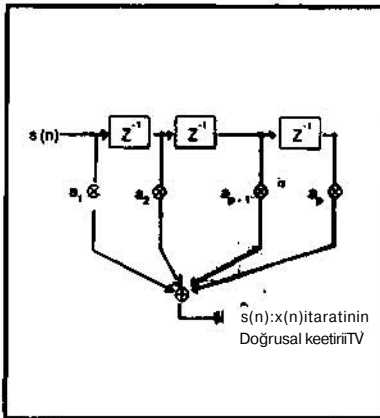


Şekil 5: KESİKLİ-ZAMAN ses işareti elde edilmesi



Şekil 6: a) NİCEMLEYİCİ öbek çizimi

b) 8 DÜZEYLİ DÜZGÜ bir nicemleyicinin girdi-çıkışı ilişkisi

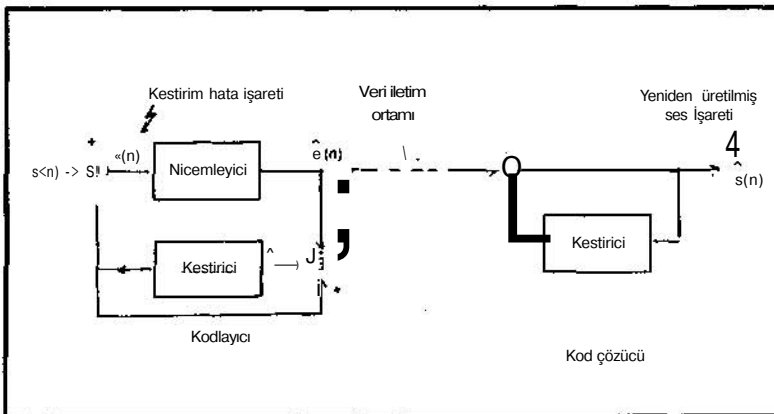


Şekil 7: DOĞRUSAL Kestirici öbek çizimi

Hata işaretinin gücü, $s(n)$ işaretinin örnekleri arasındaki istatistiksel ilişkinin yüksek olup olmamasına bağlıdır. Örnekler arasındaki istatistiksel ilişki arttıkça işaretin kestirilebilirliği artmakta, dolayısıyla hata işareti zayıflamaktadır.

Ayrımsal vuru kod kipleme kodlayıcı ve kod çözücü öbek çizimi Şekil-8'de verilmiştir. Bu dizgeler, kestirim hata işaretini iletirler. Hata işareti, ses işaretinden belirgin biçimde zayıf olduğu için daha az sayıda ikil ile nicemlenebilir.

En-uygun kestirici parametreleri ses işaretinin istatistiğine bağlıdır.



Şekil 8 : AYRIMSAL VURUM KOD kipleme dizgesi öbek çizimi

istatistik zamanla değiştiği için parametrelerin bu değişime uyarlanması gerekmektedir. Uyarlamalı kestirici kullanan ayrımsal vuru kod kipleme dizgelerine uyarlamalı ayrımsal vuru kod kipleme dizgeleri denir. Bu konudaki standart 32,000 ikil/saniyedir.

Delta Kipleme (DM): Delta kipleme³¹, ayrımsal vuru kod kiplemenin özel bir biçimidir. Ses işareti 64 kHz gibi yüksek bir hızda örneklendir. Örnekleme hızının yüksekliği nedeniyle ardarda gelen örnekler arasındaki ilişki iyice artar. Bir ikilik, yani iki düzeyli bir nicemleyici ve birinci dereceden bir kestirici kullanılır. Delta kiplemenin başarımlı, nicemleyici adım boyunun işaretin değişimine uyarlanması ile artırılmaktadır. Delta kipleme standardı 64,000 ikil/saniye, uyarlamalı delta kiplemenin standardı ise 16-32,000 ikil/saniyedir.

Delta kipleme dizgeleri, gerçekleştirilmesi görece kolay dizgelerdir. Bir ikilik nicemleyici kullanıldığı için bu dizgelerde dilim eş zamanlaması gibi bir sorun yoktur.

Alt-Kuşak Kodlama: Alt-kuşak kodlama³² bir sıklık bölgesi dalga-biçimi ses kodlama yöntemidir (Şekil-9). Ses işareti önce bir dizi kuşak-geçiren süzgeçten geçirilir. Süzgeç çıkışındaki işaretler özgün işarete göre dar kuşaklı olduklarından kesikli zamanda örneklenerek sıklık eksenine yayılırlar. Örnekleme çıkışında elde edilen işaretlere *alt-kuşak işaretleri* denir. Her bir alt-kuşak işareti diğerlerinden bağımsız olarak uygun bir zaman bölgesi dalga-biçimi kodlama yöntemiyle kodlanır. Ses işaretinin yeniden üretimi için ise bu işlemlerin tersi yapılır.

Alt-kuşak kodlamanın en önemli özelliği alt-kuşak işaretlerinin kaç ikile nicemleneceği konusunda ortaya konan esnekliktir. Kulağın görece duyarlı ve işaretin görece güçlü olduğu kuşaklara görece fazla ikil ayrılarak sesin nesnel ve öznel niteliği iyileştirilebilir. Alt-kuşak kodlama tipik olarak 16,000 - 32,000 ikil/saniye veri iletim hızı bölgesinde kullanılmaktadır.

Dönüşüm Kodlama: Bu yöntemde³, ardarda gelen belli sayıdaki ses örneğinden oluşan bir \bar{x} vektörüne doğrusal bir A dönüşümü uygulanır.

$$\bar{\Theta} = A \cdot \bar{x} \quad (2)$$

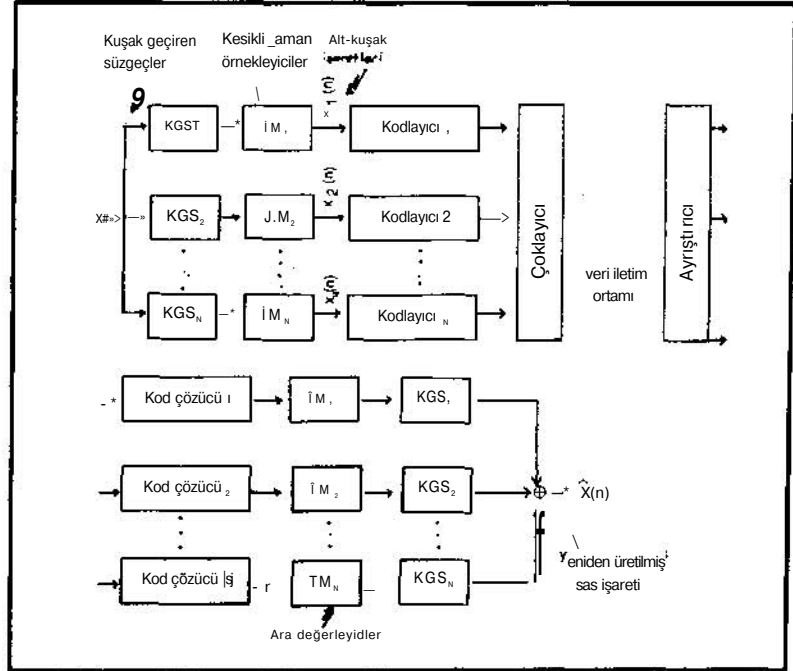
Ardından, *dönüşüm katsayıları* olarak adlandırdığımız $\bar{\Theta}$ vektörünün elemanları, güçlerine bağlı olarak belirlenen sayıda ikile nicemlenir.

Dönüşümden amaçlanan, aralarında istatistiksel ilişki bulunmayan *dönüşüm katsayıları* elde etmektir. Böylece her bir katsayı diğerlerinin taşımadığı bir bilgiyi taşıyacak, gereksiz veri iletimi engellenmiş olacaktır. Bu amaca kesin olarak ulaşmak için *Karhunen-Loève* dönüşümü kullanmak gerekmektedir, ne varki *Karhunen-Loève* dönüşümü için gerekli işlem sayısı bu dönüşümün bugünkü donanım teknolojisi ile gerçekleştirilebilmesini olanaksız kılar, başarımlı açısından en iyi olmayan, fakat gerçekleştirilmesi kolay kesikli Fourier dönüşümü, kesikli kosinüs dönüşümü gibi dönüşümler kullanılır.

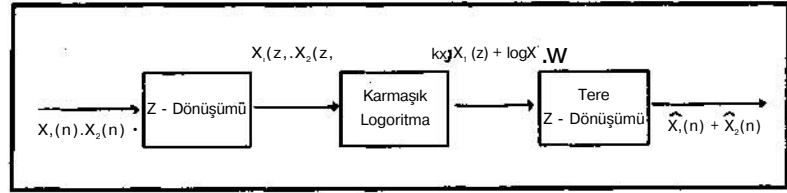
Ses işaretinin zamanla değişir olması, dönüşüm kodlama sonucu elde edilen katsayıların ağırlıklarının da zamanla değişir olmasına neden olmaktadır. Bu yüzden katsayıları ayrılan ikil sayılarının, katsayıların ağırlıklarına uyarlanması gerekmektedir.

3.2. Kaynak Kodlama Yöntemleri
Kaynak kodlama yöntemlerinin hemen tümü, ses borusu ve onu uyaran işaretin ayrı ayrı modellenmesi temeline dayanır.

Kanal Kaynak Kodlayıcı: Bu tür kodlayıcılarda ses borusunun, 15-25 milisaniyelik bir gözlem dilimi için yerel sıklık yanıtı, gözlenen ses diliminin ünlü ya da ünsüz olduğu, ünlü ise perde periyodu çözümüleme sonucu elde edilir. Ses borusunun sıklık yanıtını çıkarsamak için ses işareti önce bir dizi kuşak-geçiren süzgeçten geçirilir. Ardından her bir süzgeç çıktısı bir zarf seziciden geçirilerek işaretin o kuşaktaki yerel gücü bulunur²². Uyarım işaretinin



Şekil 9: ALT-KUŞAK kodlama dizgesi öbek çizimi



Şekil 10: EVRİŞİM İŞLEMİNİN toplama işlemine dönüştürülmesi

üretimi için ünlü/ünsüz kararı ve perde yineli bilgisi ayrı bir çözümleme ile elde edilir.

Homomorfik Kaynak Kodlama:

Ses işaretini $u(n)$ işareti ile uyarılan ses borusu vurum yanıtı $h_v(n)$ olarak ifade edelim:

$$s(n) = u(n) \cdot h_v(n) \quad (3)$$

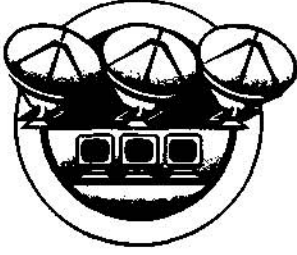
Eşitlikten de de görüleceği gibi ses işareti, uyarım işareti ile ses borusu vurum yanıtının evrişiminden oluşmuştur. Bu nedenle, uyarım işareti ve ses borusu vurum yanıtı, ses işaretinden doğrusal bir süzgeçleme ile elde edilemez. Homomorfik işaret işleme, bu ve benzeri sorunların çözümünü amaçlar. Birbiri ile evrişmiş $x_1(n)$ ve $x_2(n)$ gibi iki işaretten Şekil-10'da öbek çizimi verilen dizge ile $x_1(n) \vee x_2(n)$ işaretlerinin toplamı elde edilir. $\hat{x}_1(n)$ ile

$x_2(n)$ ile $\hat{x}_2(n)$ ile $x_2(n)$ arasındaki ilişki bire birdir, yani bu işaretlerin birinden diğeri üretilebilir.

$x_1(n)$ ile $x_2(n)$ dizileri birbirinden ayrışık ise basit bir pencereleme ve ardından $x_1(n) + x_2(n)$ 'i elde etmek için yapılan işlemlerin tersi yapılarak amaçlanan süzgeçleme gerçekleştirilmiş olur.

Yukarıda özetlediğimiz işaret işleme yönteminden yararlanarak ses borusunun genlik yanıtı başarıyla kestirilebilmekte, uyarım işaretinin parametreleri de elde edilebilmektedir²⁴.

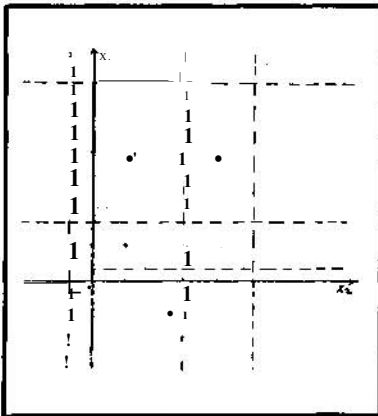
Doğrusal Kestirimli Kaynak Kodlama: Bu kodlama yönteminde ses işareti örneği $s(n)$, geçmiş ses işareti örnekleri ve uyarım işareti örneğini $u(n)$ 'in bir doğrusal bileşkesi biçiminde ifade edilir⁽²⁴⁾:



Konuşma tonuna konusunu sözcük tanıma ve sürekli konuşma tonuna olarak ikiye ayırabiliriz. Sözcük tanuna, birbirinden yalıtılmış ve yalıtılmamış sözcükleri tanıma olarak tekrar ikiye ayrılabilir.

$$s(n) = \sum_{k=1}^p x_k \cdot s(n-k) + G \cdot u(n) \quad (4)$$

Doğrusal kestirim çözümlemesi ile, α^k katsayıları ve kazanç katsayısı G elde edilir. Bu işlem her 15-25 milisaniyelik dilim için yinelenir. Uyarım işaretini üretebilmek için ise dilimin ünlü/ünsüz oluşu, ünlü ise peroe yinelimi gibi bilgilere gereksinim vardır. Doğrusal kestirimli kaynak kodlama standardı 2,400 ikil/saniyedir.



Şekil 11-a) İKİ BOYUTLU BİR vektörün elemanlarının sayıl nitelemesine bir örnek

3.3. Ses Kodlamada Güncel Konular: Karma Kodlama

Günümüzde ses kodlama araştırmaları, 4,000-8,000 ikil/saniye ve daha düşük veri iletim hızlarında yüksek nitelikli ses üretme konusunda yoğunlaşmıştır. Kaynak kodlayıcılarla üretilen sesin niteliği, şekil-4'den de görülebileceği gibi belli bir düzeyin üzerine çıkamamaktadır. Bunun en önemli nedeni, şekil-1'de gösterilen uyarım modelinin fiziksel gerçeklerden gereğinden fazla uzak, fazla basit olmasıdır. Sorunun çözümü için, kaynak kodlamada en-uygun uyarım işa, atının üretimi için dalga-biçimi kodlama ilkelerini kullanan yeni kodlama yöntemleri önerilmiştir. Bu yöntemler, kaynak ve dalga-biçimi kodlama yöntemlerini kaynaştırdığı için *karma kodlama* yöntemleri olarak adlandırılmaktadır^{5*}.

8,000 ikil/saniyede ses iletimi için çok *vurumlu doğrusal kestirimli* karma kodlama yöntemi⁶⁾ güçlü bir adaydır. Bu yöntemde, bir ses dilimi için uyarım işareti uygun sayıda vurdan oluşan bir dizi olarak alınır. Örneğin, 10 milisaniyelik bir dilim için 10 vuruş yeterli olabilmektedir. Vurumların genlikleri ve zaman boyutundaki yerleri teker teker, bir kapalı döngü arama yordamı içinde en-iyilenmektedir. Çok vurumlu doğrusal kestirimli kodlamada, uyarım bilgisinin iletimi için saniyede iletilen 8,000 ikilin yarıdan çoğu kullanılmaktadır. Bu nedenle ses borusu modelinin parametrelerini iletmek zorlaşmaktadır. Bu sorun, bi-

razdan değineceğimiz vektörel nicemleme ile aşılmaktadır.

4,000 ikil/saniye veri iletim hızında ise *kod uyanımlı doğrusal kestirimli kodlama*^{P1} ve *rasgele uyanımlı doğrusal kestirimli kodlama*[^] umut veren iki yöntemdir. Bu yöntemlerde en uygun uyarım dizisi, önceden depolanmış aday uyarım kümesi içinden bir kapalı-döngü arama yordamıyla belirlenir.

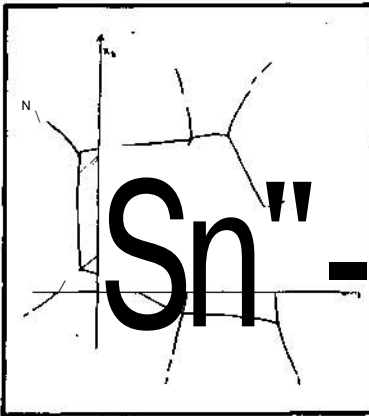
4. VEKTÖREL NİCEMLEME

Vektörel niemleme, sayısal işaret işleme alanının son on yıldaki en popüler konusudur⁹⁾. Kesikli bir bilgi kaynağı çıktılarını, birbirinden istatistiksel anlamda bağımsız bile olsalar, öbek öbek nicemlemenin, *sa-yıl* yani tek tek nicemlemeye göre az bozunum yaratacağı bilgikurumsal gerçeğine dayanır.

Vektörel nicemlemeyi basit bir örnekle anlamaya çalışalım:

İki boyutlu bir vektör uzayı düşünelim ve bu uzaydaki bir $(x_1, x_2)^T$ vektörünün elemanlarının sayıl nicemlendiği durumu inceleyelim. x_1 ve x_2 'nin nicemlenmesi, ait oldukları tek boyutlu uzayda en-küçük bozunum yaratacak biçimde yapılmış olsun. Olayı şekil 11-a'daki gibi iki boyutta ele aldığımız zaman sayıl nicemlemenin getirdiği kısıtlamaları görebiliriz. Vektör uzayı yalnızca dikdörtgenler biçiminde bölüntülenebilmekte ve nicemleyici çıktıları yalnızca bir kafes üzerindeki noktalar olabilmektedir. x_1 ve x_2 'yi bir vektör olarak nicemlediğimizde ise bu kısıtlamalar ortadan kalkmaktadır. Gerek bölüntü kümeleri ve gerekse bu kümeleri temsil edecek noktaların seçimi uzayın tümü üzerindeki bozunumu en-küçültme koşulunu sağlayacak biçimde özgürce yapılabilmektedir. (Şekil-11.b).

Vektörel nicemleme önceleri, daha çok kaynak kodlamaya uygulanmış, 1,000 ikil/saniyenin altında veri iletim hızlarına inebilme çalışmaları yapılmıştır, örnek olarak, 800 ikil/saniye veri iletim hızı olan doğrusal kestirimli bir kaynak kodlayıcı verilebilir⁽¹⁰⁾. Benzer biçimde, sesin *anlaşılabilirliğinin* korunarak 200 ikil/saniyeye inildiği söylenmektedir.



Şekil 11-b) İKİ BOYUTLU BİR vektörün nicemlenmesine bir örnek

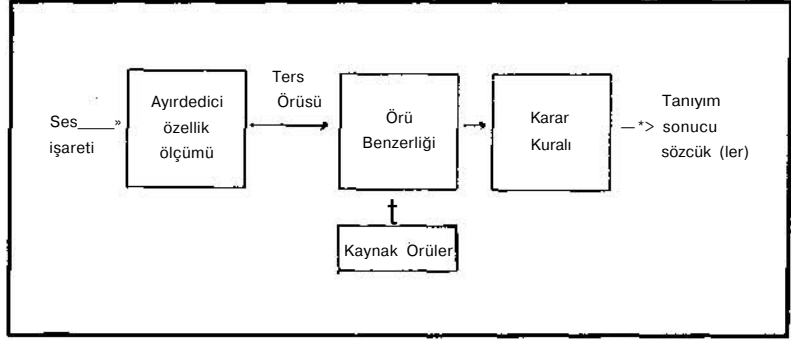
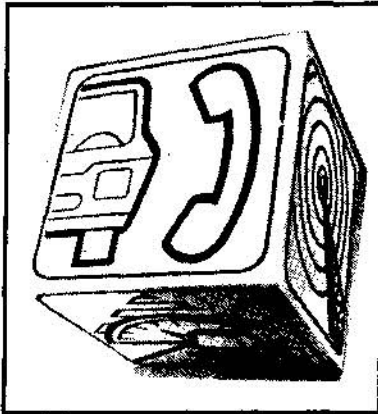
Son yıllarda ise ağırlık dalga-biçimi ve özellikle de karma kodlamaya kaymış durumdadır.

5. KONUŞMA TANIMA

Konuşma tanıma konusunu^(1,3) *sözcük tanıma* ve *sürekli konuşma tanıma* olarak ikiye ayırabiliriz. Sözcük tanıma, birbirinden yalıtılmış ve yalıtılmamış sözcükleri tanıma olarak tekrar ikiye ayrılabilir. Sürekli konuşma tanıma ise belli bir konu ya da amaçla kısıtlı konuşmayı tanıma ve kısıtsız konuşma tanıma biçiminde sınıflandırılabilir.

Sözcük tanıma dizgeleri, sonlu sayıda sözcükten oluşan bir dağarcık tabanında kurulur. Bu tür tanıma dizgeleri konuşan kişiye bağımlı ya da bağımsız olabilir.

Yalıtılmış sözcük tanımadaki sözcüklerin birbirinden yalıtılması, konuşan kişinin her sözcükten sonra yeterince duraksaması ile gerçekleştirilir. Şekil 12'de bir yalıtılmış sözcük tanıma dizgesinin öbek çizimi verilmiştir. Tanıma işlemi için önce, her bir sözcüğün diğerlerinden ayırabileceği düzeyde özelliklerinin belirlendiği bir çözümleme yapılır. Anılan çözümleme gerçekte bir veri sıkıştırımından başka bir şey değildir. Doğrusal kestirim çözümlemesi ve kuşakgeçiren süzgeçlerle yapılan görüğe çözümlemesi çokça kullanılan ayırıcı özellik belirleme yöntemleridir. Çözümleme sonucu elde edilen parametreler dizisine, o sözcüğe ilişkin *test örüsü* denir. Test örüsü, dağarcıktaki tüm sözcükler için, evvelce bir *belletme* evresi sonucu hazırlanıp derlenmiş kaynak örüleri ile karşılaştırılır. Han-



Şekil/7 12 : YALITILMIŞ SÖZCÜK tanıma dizgesi öbek çizimi

gisine en yakın ise, o sözcüğün söylenmiş olduğu kararına varılır.

Sürekli konuşmayı tanımak için sesi yalnızca bir dalga-biçimi olarak ele almak yeterli değildir, ses işareti kullanılan dilin kuralları içinde bir mesaj taşımakta, konuşan kişinin o anki ruh durumu gibi birçok etkeni de yansıtmaktadır. Bu nedenle, dilbilgi, dilbilim, anlambilim gibi bilim dallarından yararlanmak gerekmektedir. Sürekli konuşmayı tanımak için ses işareti, kullanılan dilin özelliklerine göre hece, yarı hece, temel ses (phone-me) gibi birimlere ayrılarak çözümlenmekte, bu birimler arasındaki ilişkiler modellenmeye çalışılmaktadır.

6. SES İŞLEMEDE KULLANILAN DONANIM

Çok büyük ölçekli tümdevre teknolojisinde yakın dönemde olan gelişmeler sayısal ses işleme konusunun gerek araştırma ve gerekse uygulama yönünden çok zenginleşmesine

neden olmuştur. Çok büyük ölçekli tüm devre teknolojisinin sayısal işaret işlemeye en önemli katkılarından biri gerçek zamandaki uygulamalarda neredeyse bir devrim yaratan *sayısal işaret işleyicilik* (Q_{a-y} sayısal işaret işleyicilerin komut kümeleri genel amaçlı mikro işlemcilerden büyük ölçüde farklıdır ve tümüyle sayısal işaret işlemeye yöneliktir. Günümüzde TMS320C30 ve DSP32C gibi 32 ikil sözcük uzunlukta ve kayan-noktalı sayısal işaret işleyiciler yaygın olarak kullanılmaktadır. Bu işlemcilerin ulaştığı işlem hızı saniyede 12-16 milyon komut düzeyindedir.

VLSI teknolojisinin bir diğer önemli katkısı *iş-istasyonu* olarak adlandırığımız bilgisayarlardaki gelişmedir, iş istasyonları, işlem hızı, bellek ve disk sığası açısından birkaç yıl öncesinin büyük boy bilgisayarlarını geride bırakmaktadır. Ederleri ise anılan büyük boy bilgisayarların çok altındadır.

KAYNAKLAR

- (1) Flanagan, J.L., **Speech Analysis, Synthesis and Perception**, 2nd Ed., Springer-Verlag, New York, 1972.
- (2) Rabiner, L.R., Schafer, R.W., **Digital Processing of Speech Signals**, Englewood Cliffs, NJ: Prentice-Hall, 1978.
- (3) Jayant, N.S., Noll, P., **Digital Coding Of Waveforms**, Englewood Cliffs, NJ: Prentice-Hall, 1984.
- (4) Markel, J.D., Gray, A.H., **Linear Prediction of Speech**, Springer-Verlag, New York, 1976.
- (5) Jayant, N.S., **"Coding speech at low bit rates"** IEEE Spectrum, pp. 58-63, August 1986.
- (6) Atal, B.S., Remde, J.R., "A new model of LPC excitation for producing natural-sounding speech at low bit rates" Proc. 1982 IEEE Int. Conf on Acoustics, Speech and Signal Processing, Paris, France, pp. 614-617, 1982.
- (7) Schroeder, M.R. and Atal, B.S. "Code-excited linear prediction

(CELP): High quality speech at low bit rates,"Proc. 1985 IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 937-940, 1985.

(8) Atal, B.S., Schroeder, M.R. "Stochastic coding of speech at very low bit rates," Proc. Int. Conf. Commun., ICC84, part 2, pp. 1610-1613, May 1984.

(9) Abut, H. (Editör), *Vector Quantization*, New York, NY: IEEE Press, 1990.

(10) Wong, D.Y., Juang, B-H., Gray, A.H.Jr., "An 800 bit/s vector quantization LPC

vocoder," IEEE Trans on Acoustics Speech and signal Processing, vol. 30, no. 5, pp. 770-780, October 1982.

(11) Lea, W.A. (Editör), *Trends in Speech Recognition*, Englewood Cliffs, NJ: Prentice-hall, 1980.

(12) Reddy, FL, "Speech recognition by machine: a review," Proceedings of the IEEE, vol. 64, no. 4, pp. 501-531, April 1976.

(13) Rabiner, L.R., Levinson, S.E., "Isolated and connected word

recognition-theory and selected applications," IEEE Transactions on Communications, vol. 29, no.5, pp.621-659, May 1981.

(14) IEEE Micro Magazine, *Special Issue on Digital Signal Processing*, December 1986.

(15) IEEE Micro Magazine, *Special Issue on Digital Signal Processing*, December 1988.

(16) IEEE Micro Magazine, *Special Issue on Digital Signal Processing*, October 1990.

MATAMATİK



Hazırlayan : M. Serhat ÖZYAR

Değerli Matematikçiler,

3 79. sayımızdaki sorulara ilişkin elimize şu ana kadar dört mektup ulaştı. Sayın Mustafa Küçük (İzmir) 41 ve 42. sorularımıza, Sayın İsmail SEVİNÇ (Bursa) 41. sorumuza, Sayın Musa TÜRKEL yine 41. sorumuza ve 41. sorumuzu köşemiz için hazırlama inceliğinde bulunan Sayın A. NECAH da 42. sorumuza yanıt göndermişler. Sn. A. NECAH ayrıca 37. sorumuzla ilgili yayınladığımız çözümlerdeki uyarımızı dikkate alarak bize yeni bir çözüm göndermiş. Bu arada 378. sayımızda doğru çözüm gönderenler arasında ismini duyurduğumuz, ancak ilini yanlışlıkla Adana olarak belirttiğimiz Sn. Musa TÜRKEL'den özür diliyoruz.

Çoğunluğumuzun tatillerini tamamlamakta olduğu şu günlerde mektup göndererek köşemize katkıda bulunanlara teşekkür ediyor, tüm matematikçilere esenlik dolu günler diliyoruz.

SORU : 45. EMO İLE SÖZMETİK

(A. NECAH)

Bir aritmetik işlemden rakamların belli harflerle değiştirilmesi ile düzenlenen türdeki bulmacalar, daha önceki MATEMATİK sayfalarında yer almıştı. Bu tür bulmacalara, SÖZCÜK'ün ilk hecesi ile ARİTMETİK'in son iki hecesinin birleşmesinden oluşan "SÖZMETİK" deyimini yakıştırıyorum.

"EMO"daki harflerin harmanlanması sonucu oluşan üç ayrı SÖZMETİK bulmaca aşağıda verilmiştir. Birincisinin bir çözümü vardır. Diğer ikisi ise ondalık sistemde çözümsüzdür. Ama ondalık sayı sistemi dışındaki (en çok 16 tabanlı "hexadecimar sisteme kadar olan başka) sayı sistemlerinde çözümleri vardır.

SÖZMETİK 1'in ondalık sistemdeki çözümünü, SÖZMETİK 2 ve SÖZMETİK 3'ün hangi sayı sisteminde çözümünü ve çözümü veren rakamları bulunuz.