

IEEE 754 standardında,

- Tek hassasiyetli sayılar için anlamlı rakam⁵ sayısı 6-9 arası (23 bit mantis)
- Çift hassasiyetli sayılar için anlamlı rakam sayısı 15-17 arasındır (52 bit mantis)

Yani float tipinde değerlerle işlem yaparken en az milyarda bir, en çok milyonda bir relatif hata ile karşılaşsınız. Buna göre tek hassasiyetli sayılarda 10 rakamdan oluşan 1 milyar sayısına, 15 eklediğinizde bu toplam kayan-nokta sayı olarak ifade edilemez ve en yakın kayan nokta sayıya, yani 1 milyara yuvarlanır.

Yine float tipinde bir değişkene, 11 rakamlı 12345678.901 değerini atarsanız, bu sayı bilgisayarda 1.2345679E7 (E7'nin anlamı Exponent (üst) değeri 7'dir) şeklinde 8 rakamlı bir sayı olarak depolanacaktır.

Eğer ortalama 7 basamak sizin için yeterliyse tek hassasiyetli, aksi halde çift hassasiyetli sayıları kullanmamız yerinde olur.

Metin Verisi

Yazılı bir metin içerisinde kullandığımız tüm harf, sayı, noktalama işareti gibi sembolleri/karakterleri sayısal olarak kodlayabiliriz. Mesela A harfini 65 (= 01000001₂) sayısı ile kodlayabiliriz. Siz klavyeden A tuşuna bastığınızda, klavyeden gelen bilgi, bilgisayarda 65 sayısı olarak anlaşılacaktır. Daha sonra bu harfi görüntülemek istediğinizde 65 sayısının karşılık geldiği harfin şekli, ekranınızda çizilecektir. Bu metotla yazılı bir metinde karşılaşabileceğimiz tüm karakterleri ifade edebiliriz. Fakat bu kodlamanın standart olması gerekir. Böylece bir bilgisayarda yazdığımız bir yazıyı, başka bir bilgisayarda aynen görüntüleyebiliriz. Hangi sayının hangi karaktere karşılık geldiğini söyleyen standartlar zaman içinde oluşturulmuştur. Bu standartlar, karakter ve karşılık gelen sayısal değeri gösterir ki bu eşlemenin oluşturduğu kümeye **karakter seti** (**character set**) diyoruz. Bunlardan en sık karşımıza çıkacak olanlar ASCII ve Unicode karakter setleridir.

ASCII (American Standard Code for Information Interchange), ilk oluşturulan karakter setlerinden biridir ve bu yüzden günümüzde en yaygın olarak kullanılan karakter setidir. İlk hali 7 bit sayılarla kodlanmıştı ve 128 karakterden oluşuyordu.

ASCII tablonun ilk 32 karakteri (tablonun ilk 2 satırı), ekranda görüntülenmeyen kontrol karakterlerini içerir. Mesela;

- Ekranda nerede olduğunuzu gösterir imleci (cursor) alt satıra taşıyan Satır besleme (Line Feed - **LF**) karakteri

⁵ Anlamlı rakam, en sağdaki veya en soldaki sıfır dizilerinin haricindeki rakamlardır. Çünkü bu sıfırlar üst şeklinde ifade edilebildikleri için sayı hassasiyetine etkileri olmaz.

- İmleci satırın en soluna taşıyan satır başı taşıyıcı (Carriage Return -**CR**) karakteri,
- Metin tipi veri transferinde kullanılan **STX** (Start-of-Text) ve **ETX** (End-of-Text) karakterleri
- Ekranda karakterleri geriye doğru silmekte kullandığımız geriye ok tuşu **BS** (BackSpace) karakteri

Hep bu gurubun üyeleridir. 32–127 arası karakterler ise noktalama işaretlerini, harfleri, rakamları ve bazı matematiksel işaretleri içerir.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

Şekil 3.7 7 bit ASCII karakter tablosu



ASCII Sembollerin Ekranda Basımı

ASCII karakterleri, DOS ve Windows içerisinde, sayısal değerlerini girerek ekrana bastırabilirsiniz. Bunun için sayısal tuş takımını açın (NumLock ışığı yanacak). ALT+sayı (ALT tuşunu basılı tutarken sayısal değeri sayısal tuş takımından tuşlayın) ile o sayıya karşılık gelen ASCII karakteri bulabilirsiniz. Mesela ALT+65, A harfini ekrana basacaktır.

Başlangıçta 8. bit 0 olarak atanıyordu. Daha sonra bu 8. bitin de yeni karakterler için kullanılması düşünüldü. Böylece yeni 128 karakter daha eklenerek oluşturulan bu yeni sete **uzatılmış (extended) ASCII** karakter seti denildi. Fakat bu eklenen 128 karakter standart olamadı. ASCII'nin değişik sürümleri ortaya çıktı. Geliştirilen ilk çözüm bu uzatılmış ASCII setlerini, bölgesel olarak gruplamak oldu. IBM tarafından bu yeni uzatılmış ASCII setlerine **kod sayfası (code page)** denilmiştir. Kod sayfalarının ilk 128 karakteri standart ASCII ile aynıdır. Son 128 karakter o bölge veya ülkenin karakterlerini içerir. Mesela kod sayfası 737 Yunan dili karakterlerini, kod sayfası 857 ise Türkçe karakterlerini içeriyordu.



Metin Yazıcı ve ASCII karakter seti

Çalıştığım işyerinde, UNIX sunucuya bağlı, saf metin (grafik içermeyen) çıktılar üreten nokta vuruşlu yazıcılar vardı. Bu yazıcılar sadece ASCII karakterlerden oluşan dosyaları basarlar. Oldukça hızlıdır ve aynı anda birkaç nüsha basabilirler. Genellikle ödeme makbuzu basımı için hala yaygın olarak kullanılmaktadırlar. Bu yazıcıların birinde, çıktı alt alta satırlar yerine merdiven basamağı şeklinde çıkıyordu. Yani şöyle;

```
1.satır
    2. satır
        3.satır
            . . . .
```

UNIX makine ve yazıcı, ASCII karakter setine göre konuşuyordu. Yukarıda bahsedilen ilk 32 karakter içerisinde LF karakteri yazıcıya alt satıra geçmesini söyler. CR ise, satırın sol en başına gitmeyi ifade eder. Görünen şu ki, Unix makine yalnızca LF karakterini her satır sonunda gönderiyor. Bu alt satıra geçmeyi sağlıyor. Fakat satır başına metni çeken CR karakteri gönderilmediğinden sağ alttan metin devam ediyordu.

Bizim yaptığımız yazıcı ayarlarına girip “her LF gördüğünde yazıcı otomatik olarak CR eklesin (LF=LF+CR)” seçeneğini aktif hale getirmek oldu ve sorun çözüldü.

IBM ve Microsoft gibi firmaların çözümleri yerine, gerçek bir standartlaşmaya gidilebilmesi için ISO (International Organization for Standardization) devreye girdi ve **ISO Latin-1** karakter setini geliştirdi. Bu karakter setinin ilk 128 karakteri standart ASCII ile aynıdır. Microsoft, Windows işletim sistemi için ISO Latin-1’in biraz değiştirilmiş halini **Windows Kod sayfası 1252** ismiyle kullanmaktadır.

ISO çalışmalarını daha da ileri götürerek daha çok Avrupa dillerinin karakterlerini içeren karakter setleri oluşturdu. Bu karakter setleri **ISO 8859 karakter seti grubu** adı altında toplandı. Bu set 16 gruptan oluşur. Türkçe için **ISO 8859-9** (buradaki 9, grup numarasını temsil ediyor) veya ISO Latin-5 diye bilinen karakter seti atanmıştır.



Windows Karakter Seti Uyumsuzluğu

Arkadaşlardan biri Windows içerisinde bir program aracılığıyla oluşturduğu metin dosyasının DOS penceresinde doğru görüntülenmediğinden ve metin yazıcıdan çıktı alındığında (ç,ş gibi) Türkçe karakterlerin bozuk olduğundan yakındı. Gerçekten de, Windows **Notepad** programıyla yazılan metindeki Türkçe karakterler, DOS penceresi içerisinde **TYPE** komutu ile bakıldığında bozuk görülüyordu.

Araştırmalarım sonucu Windows ve DOS komut satırı penceresinin farklı karakter setleri kullandığını öğrendim. Windows, karakter seti olarak, kayıt defteri (registry) içerisinde

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Nls\CodePage\

altında ACP hücreesindeki değeri (Benim makinemde CodePage 1254 –CP1254) kullanırken, DOS penceresi, **chcp** komutuyla görülebilecek (benim makinemde CodePage 857 –CP857) değeri kullanıyordu. Bu yüzden Windows ortamında basılan Ç harfi, CP1254'e göre bir sayısal değere çevrilip saklanırken, CP857'de aynı sayısal değer başka bir karaktere denk geldiğinden alakasız bir karakter ekrana basılıyordu.

Mevcut çok sayıda karakter setleri arası dönüşüm yapabilecek bir programı GNU (www.gnu.org) içerisinde buldum. Bu programın ismi **recode** ve 150 kadar karakter seti arası dönüşüm yapabiliyor. Windows için derlenmiş halini şu adreste bulabilirsiniz:

<http://www.iro.umontreal.ca/contrib/recode/HTML/index.html>

DOS penceresinde, recode programına dosyanın hangi karakter seti ile oluşturulduğunu ve hangi hedef karakter sete çevrileceğini söyleyince dosya doğru bir şekilde görüntülendi.

```
recode cp1254..cp857 < ilker.txt > lpt1
```

Yukarıdaki komut, ilker.txt dosyasını kaynak karakter set olan 1254'ten hedef karakter set olan 857'de tekrar oluşturuyor ve LPT1'e bağlı metin yazıcıya gönderiyor.


ASCII karakter seti 2 noktada eksik kaldı. Birincisi, sadece 128 karakterin ifadesini mümkün kılmaktaydı. İkincisi, uluslararası ihtiyacın göz önünde bulundurulmamasıydı. Aynı şekilde ISO 8859 standart grubu da daha ziyade Avrupa dillerine hizmet veriyordu. Pek çok Asya ülkesinde 1 byte ile ifade edilebilecek 256 farklı karakter yetersiz kalmıştır. ISO 8859 grubuna ait ikinci eksik, farklı ISO dil setini kullanan kişiler arası mesajlaşmanın doğru gösterimi vermemesidir. Mesela Yunanistan'da yaşayan bir arkadaşınıza e-posta attığınızda, bu mesaj karşı bilgisayarda doğru olarak görüntülenemez. Çoğumuz yurtdışında yaşayan bir arkadaşımızla e-posta iletişimde Türkçe karakterlerin bozulduğuna şahit olmuşuzdur.

Uluslararası ihtiyaçları esas tutan ve dünyada tek bir karakter setinin kullanımını öngören yeni bir karakter seti oluşturulması düşünüldü. Bugün için, 96 binden fazla karakterden oluşan bu yeni karakter setinin adı **Unicode**. Unicode içerisinde, körlerin kullandıkları Braille alfabesinden, matematiksel sembollere, Uzakdoğu dilleri harflerinden Arap harflerine kadar oldukça geniş bir gösterim sağlanmıştır. Başlangıçta 16 bit kullanılarak 65536 farklı karakterin gösterimi amaçlanmıştı. Fakat daha sonra bu sayı yaklaşık 1.1 milyon karaktere çıkarılmıştır.

Semboller ve karşılık gelen sayıların kodlanması için ASCII'de olduğu gibi, sırayla artan sayılara sembolleri atamak mantığı kullanılamamıştır. Çünkü halen kullanılmakta olan ve Unicode'dan anlamayan sistemlerle uyumluluk söz konusudur. Mesela Unicode kullanılarak yazılan bir metnin eski sistemlerde okunabilmesi istenir. Bu amaçla **Unicode Dönüşüm Formatı (Unicode Transformation Format –UTF)** geliştirilmiştir. Bu format içerisinde UTF-8 formatı 8-bit uzunluğunda en çok 4 dizi, UTF-16 formatı 16 bit uzunluğunda en çok 2 dizi, UTF-32 ise tek bir 32 bit dizi içermektedir. ISO Latin-1 (veya standart uzatılmış ASCII) karakter seti, Unicode'un bir altseti olarak (Unicode'un ilk 256 karakteri) tanımlanmıştır. Böylece Unicode'a geçen uygulamalar uyumsuzluk problemi çekmeden çalışabilirler. Mesela UTF-8 formatı kullanıldığında, eğer karakter ISO Latin-1 setinin parçası ise sadece 1 Byte'lık bit dizisi gönderilmektedir. Windows, Linux, birçok Unix türevi ve Macintosh bilgisayarlarda kullanılan MacOS X işletim sistemi Unicode için destek vermektedir. Buna rağmen Unicode henüz dünya standardı olamamıştır.

Ψ	Ж	س	滄	Σ	ཨ
Yunanca	Rusça (Kiril)	Arapça	Çin/Japon/Kore Dilleri	Matematiksel	Tibet Dili

Şekil 3.8 Unicode Örnek Karakterler

 **E-Posta ve Web Sayfaları Karakter Kodlaması**

Özellikle Internet ortamında, web sayfalarının ve e-posta mesajlarının doğru olarak görüntülenebilmesi için, hangi karakter setinde kodlandığının bilinmesi gerektiği açıktır. Çünkü bilgisayarınıza ulaşan sadece bir sayı topluluğudur ve bu sayı topluluğunun hangi karakter setindeki karakterlerle ifade edileceği bilinemez. Bu amaçla e-posta mesajlarının başlık kısmında,

Content-Type: text/plain;charset="iso-8859-9" gibi bir bilgi;

HTML sayfalarının ise başlangıcında

<meta http-equiv="Content-Type" content="text/html; charset=utf-8"> şeklinde hangi

karakter setiyle oluşturulduğunu gösterir bilgiler bulunur. Bu sayede içerik doğru bir şekilde görüntülenebilir.

Eğer bu bilgiler eksikse veya karşı taraf bu karakter setini desteklemiyorsa içerik yanlış gösterilecektir. Internet Explorer içerisinden **Görünüm > Dil Kodlaması** kısmından kullanılacak karakter setini değiştirebilirsiniz.

Metin Dosyası (Text File) ve İkili Dosya (Binary File)

Dosyalar, metin ve ikili dosyalar olmak üzere 2 genel gruba ayrılabilir. İsimlerden de anlaşılacağı gibi metin dosyası insana yakın, yani insan tarafından anlaşılabilen bir dosyayı ifade eder. Bilgisayarın, ikili sayı sistemine tam olarak oturan 1 ve 0'larla çalıştığını daha önce görmüştük. Buradan ikili dosyanın, bilgisayara yakın, yani bilgisayar tarafından anlaşılabilen bir dosyayı ifade ettiğini anlayabiliriz.

İkili dosya, bilgisayar için anlamlı bilgiyi içerir. Genel olarak iki çeşittirler.

1. **Doğrudan çalıştırılabilir (executable) dosyalar:** Mesela Windows hesap makinesi programı calc.exe, doğrudan çalıştırılabilir ikili bir dosyadır. Bu dosyanın içeriği size bir şey ifade etmez fakat bilgisayar bunu işletip ekrana bir hesap makinesi programı getirebilir.
2. **Program veri dosyaları:** Saf metin düzenleyici uygulamalar (notepad gibi) haricinde bir programın yorumlamasıyla bir anlam ifade eden dosyalardır. Resim, ses ve video dosyaları örnek olarak verilebilir. Bu dosyalar sadece ilgili uygulama tarafından yorumlandığında bir anlam ifade ederler. Siz bunların içeriğine Notepad gibi bir metin düzenleyici vasıtasıyla bakarsanız bir sürü anlamsız karakter görürsünüz. Mesela PDF uzantılı Acrobat reader dosyaları böyledir. Acrobat Reader gibi PDF formatından anlayan bir program olmadan bu dosyalar kullanılamaz veya işlenemez.

Metin dosya ise, ASCII (veya Unicode) karakterlerden oluşur ve herhangi bir metin düzenleyici ile insan tarafından okunabilir ve işlenebilirler. Örnek olarak C dilinde yazılmış bir program kaynak kodunu (Source Code) veya Unix kullanıcı bilgilerinin tutulduğu passwd dosyasını veya Internet tarayıcı (Web browser) ile görüntülenen HTML dosyaları verebiliriz. Metin dosya satırlardan oluşur. Bunun için satır sonu karakteri her satır sonunda bulunur. Satır sonu karakteri ekrana basılmaz. Bunun yerine metin düzenleyici (Windows Notepad, DOS edit, UNIX vi uygulamaları) tarafından yorumlanarak, bir sonraki karakter satırbaşı yapıldıktan sonra ekrana basılır. Bunun dışında herhangi bir formatlama bilgisi içermezler. Bu yüzden Microsoft Word'de yazdığınız belgeler metin dosyası olmayacaktır. Çünkü Word dosyasının içerisinde ikili formda tutulan pek çok bilgi vardır (Resimler, fontlar, çizimler ve Word'ün kendine has kullandığı bazı bilgiler gibi). Bir Word belgesini Notepad ile açıp bakarsanız bunu daha iyi görebilirsiniz.

Hem metin hem de ikili dosyalar bilgisayarda 1 ve 0'lar şeklinde saklanırlar.

📌 FTP ve Satır Sonu Karakteri

Satır Sonu karakteri değişik işletim sistemleri tarafından değişik olarak tanımlanmıştır. Satır Besleme (LF) ve Satır Başı Taşıyıcı (CR) karakterleri esas alınan bu uygulamalar şöyledir:

DOS/Windows : CR + LF

UNIX : LF

Macintosh: CR

Bu yüzden bu sistemler arası metin dosya transferi yaparken, ftp gibi dosya transfer programlarına **metin aktarım modu (text transfer mode)** eklenmiştir. Böylece metin dosyalarının doğru olarak aktarılması mümkün olur. Mesela metin dosyasını Windows makinenizden Unix sisteme aktarırken, metin modunu seçerseniz, program Windows dosyasındaki satır sonlarını gösterir CR+LF karakterlerini sadece LF olarak karşı tarafa geçirecektir. Böylece dosyayı Unix üzerinde de herhangi bir metin düzenleyici ile düzgün olarak görüntüleyebilirsiniz. Eğer metin modunda aktarmazsanız Unix'te vi gibi bir metin düzenleyiciyle aynı dosyanın içeriğine bakarsanız satır sonlarında ^M karakterleri görürsünüz. Bunlar Unix tarafından gereksiz görünen CR karakterleridir.

Metin verisi txt, doc, rtf ve sxw gibi uzantılarla ifade edilebilir. **txt (text -metin)** uzantılı dosyalar genellikle sadece yazıyı oluşturan karakterleri içerir. **rtf (rich text format –zengin metin formatı)** yazıyı oluşturan karakterlerle birlikte font (büyüklük, renk, italik, kalın, altı çizili vs) bilgilerini de tutar. rtf uzantılı dosyalar, Windows'un yanı sıra Linux ve Macintosh gibi pek çok ortamda açılabilirler. Microsoft Wordpad ile rtf uzantılı dosyalar oluşturabilirsiniz. **doc (document)** formatı ise, yazı ile birlikte resim, video, şekil, ses gibi bilgileri de içerebilir. Bunların yanında formatlama, şekil çizimi, tablolar gibi konularda güçlü araçlar içerir. Microsoft Word ile bu formatta dosyalar oluşturabilirsiniz. rtf dosyalar içerisine resim koymadığınız müddetçe Microsoft Word ile hazırlanan doc uzantılı dosyalara oranla daha az yer kaplarlar. **sxw** ise Open Office paketindeki **writer** kelime işlem yazılımının (Microsoft Office içerisindeki Word yazılımının eşdeğeri) dosya formatıdır. Linux dünyasında sık olarak karşılaşılabiliyorsunuz. Bu arada Open Office paketinin Windows ortamında da çalıştığını ve ücretsiz olduğunu belirtelim

📌 Ücretsiz PDF Belge Oluşturucu Yazılım

Kitap şeklinde belgeler için PDF formatı Internet'te oldukça yaygındır. Acrobat Reader gibi ücretsiz programlarla PDF formatındaki dosyaları görüntüleyebileceğimizi biliyoruz. Eğer PDF belge oluşturmak istersek, ücretli PDF dosya oluşturma araçları satın almalıyız. Fakat Internet'te bu iş için kullanabileceğiniz **PDFCreator** isimli ücretsiz bir yazılım var. Yazılım tıpkı bir yazıcı gibi kullanılıyor. Fakat yazıcı çıktısı yerine PDF dosya üretiyor. Web adresi şöyle;

https://sourceforge.net/project/showfiles.php?group_id=57796

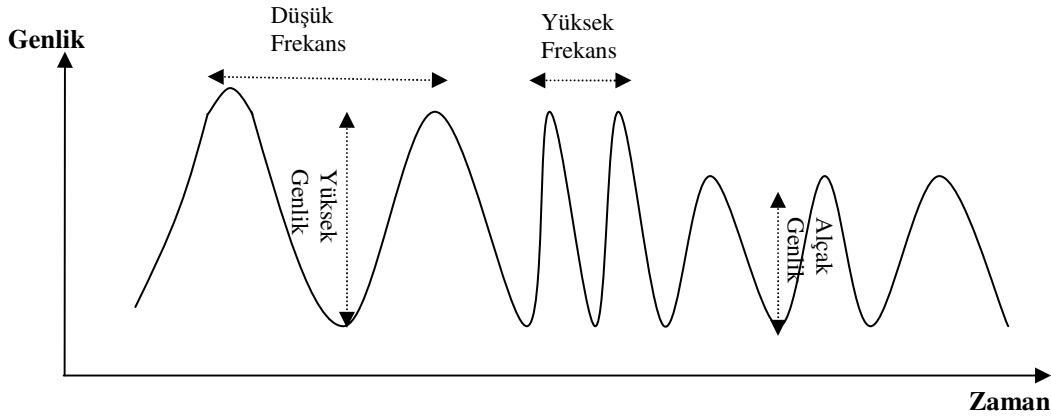
Ses Verisi

Sesin bilgisayar tarafından nasıl saklandığına bakmadan önce, Allah'ın harika bir sanatı olan işitme yeteneğimize yakından bakalım. Çünkü oradan kopya çekeceğiz.

Ses üreten her cisim, etrafını saran havada (veya sıvı veya katı ortamda) titreşimler meydana getirir. Cismi saran ilk hava tabakası bunu sonraki hava moleküllerine iletir. Bu şekilde ses kulağımıza kadar gelir. Kulağımızda bulunan kulak zarı, hava moleküllerinin çarpmasıyla titreşir ve bu titreşimden oluşan sinyaller beyne ulaşır.

Hava ortamında oluşan titreşimler çok seri itme ve çekme (veya ileri-geri) hareketi şeklindedir. Bunu denizin dalga hareketine benzetebiliriz. Bu yüzden molekülden moleküle iletilen bu titreşim, **ses dalgası** olarak adlandırılır. Hava moleküllerinin *ne kadar hızlı* ileri-geri hareketi yaptığı bilgisine, ses dalgasının **frekansı** denir. Frekans birimi **Hertz**'dir (Hz) ve hava moleküllerinin saniyede kaç defa titreştiğini ifade eder. Mesela alkış sesinin frekansı 100 Hz'dir. Yani bu ortamda hava molekülleri saniyede 100 defa titreşir. İnsan kulağı 16Hz – 20KHz arasındaki ses dalgalarını duyabilir. Eğer sesin frekansı düşük ise **kalin (bas)**, yüksek ise **ince (tiz)** ses duyarız. Müzikle uğraşanlar için frekansın duyulan ses üzerinde oluşturduğu etki **ses perdesi (pitch)** olarak bilinir.

Ses bir dalga olduğu için frekansı ile birlikte **genlik (Amplitude)** değeri de vardır. Bu değer hava moleküllerinin *ne kadar sert* itme ve çekme hareketi yaptığına bağlıdır. Bu özellik sesin yüksek (sesli) veya alçak (sessiz) olmasını belirler. Ses şiddeti birim **Desibel (dB)** dir. Mesela bir piyano tuşuna hafifçe dokunursak bağlı olan piyano teli hafifçe titreşecektir. Eğer tuşa sert basarsak tel de etrafındaki havayı sertçe itip-çekeceği için yüksek sesle duyarız. Bir insan konuşmasının ses şiddeti değeri 60 dB'dir. 130 dB ses, insana acı verecek kadar şiddetlidir.



Şekil 3.9 Ses dalgasının karakteristiği

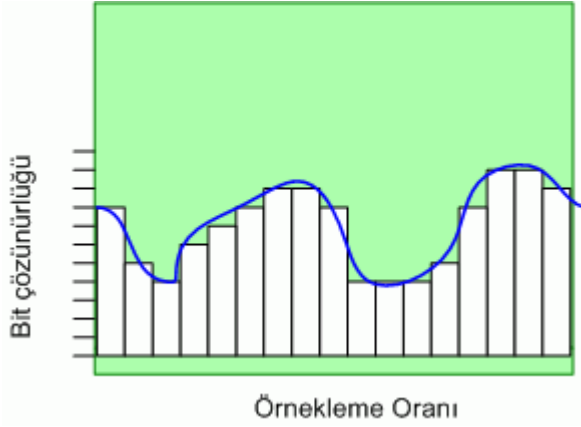
Bu bilgilerden ses dalgasının karakteristiğini, bir şekilde bit dizileri halinde kodlayabilirsek, o sesi bilgisayar ortamında saklayabileceğimizi anlarız. Ses verisi analog formdadır. Bu yüzden her andaki ses bilgisini kaydetmek yerine (ki bu sonsuz büyüklükte yer ihtiyacını doğurur), belli anlardaki ses bilgisini kaydedebiliriz. Analog ses verisini, belli zaman aralıklarına bölüp o andaki ses verisini kaydetme işlemine **örnekleme (sampling)** ve bu işi yapan elektronik devreye de **analog-dijital çevirici (Analog to Digital Converter -ADC)** denilir. Örnekleme işleminin ne kadar sık yapıldığı bilgisi **Örnekleme oranı (sampling rate)** ile belirtilir. Eğer örnekleme anları çok aralıklı olursa, ses kalitesi düşük olur. Tabii örneklenmeyen anlar kaydedilemeyecektir. Fakat başta da dediğimiz gibi duyularımızı tatmin edecek kadarı bize yeter.

Herhangi bir frekanstaki dalganın hem tepe hem de çukur değerlerinin yakalanabilmesi için en az 2 katı bir hızla örneklenmesi gerekir. İnsan kulağı en çok 20KHz frekansındaki sesleri işitebilir demiştik. Buna göre gerçek dünya seslerinin doğru bir şekilde kaydedilebilmesi için en azından 40 KHz örnekleme oranıyla kaydedilmesi gerekir. Yani ortamdaki ses, saniyede 40,000 defa örneklenir.

Bizim algılayabileceğimiz ses frekansının en çok 20KHz olması, etrafımızdaki bütün seslerin de en çok 20KHz frekansta olabileceği anlamına gelmez. 20kHz'den yüksek frekansta sesler etrafımızda bulunur ve bunlar örnekleyici tarafından yanlış olarak, duyulabilen frekansta sesler olarak kaydedilebilirler. Bunlar mevcut sesin ahengine uymayacakları için bizim tarafımızdan sesin kalitesini düşüren gürültü olarak algılanırlar. Bunu önlemek için yüksek frekanslı seslerin kaydını önleyen filtre devreleri kullanılır. Bu filtrelerin çalışabilmesi için örnekleme hızı biraz daha artırılır ve 44.1 KHz veya 48KHz hızlarında örnekleme yapılır. Böylece ses kalitesi çok daha iyi olur.

Bilgisayarda ses verisinin kaydedilmesi, **mikrofonlar** aracılığıyla olur. Dış dünyadan gelen hava moleküllerinin titreşimleri, mikrofonun diyaframında (bu kulak zarı gibi, titreşimleri algılayabilecek bir yapıdır) elektriksel sinyale çevrilir. Bu sinyal, ses kartımızdaki Analog-dijital çevirici tarafından örneklenir. Örnekleme sırasında, elektriksel sinyal bit dizisine çevrilir. Her örnekleme işlemi için kullanılacak bit dizisi uzunluğu **bit çözünürlüğü (bit resolution)** olarak ifade edilir. Bit dizisi ne kadar büyük olursa o kadar hassas olarak ses verisini ifade etmemiz mümkün olur. Eğer 4 bit kullanırsak bununla 16 değişik durumu ifade edebiliriz. 16 bit ile 65536 değişik seviye elde ederiz. Bit çözünürlüğü sesin kalitesi yanında sesin şiddeti için de bir ölçüttür. Kaydedilen sesin alabileceği ses şiddeti aralığı 8 bit için 48dB, 16 bit için 96 dB'dir. Demek oluyor ki, ortamdaki seslerin seviyelerini doğru olarak kaydetmek istiyorsak yüksek bit çözünürlüğünde kayıt yapmalıyız. Böylece fısıltıyla konuşan ve bağırarak konuşan iki insana ait ses seviyeleri doğru olarak kaydedilir. Bunun yanında, 8 bitlik kayıtlarda genellikle "arka plan tıslaması" denilen ve sesin kalitesini bozan bir unsur da sesle birlikte duyulacaktır. 16 bitlik bit çözünürlüğü ise bundan çok etkilenmez ve duyularımızı tatmin için oldukça iyi bir değerdir.

Burada kafa karıştırabilecek diğer bir birim de, **bit oranıdır (bit rate)** ve birim zamandaki ses verisi için gereken depolama alanını gösterir. Mesela 128kbps (kilo bits per second) bit oranı, sesin bir saniyesi için 128 kbit (veya 16 KByte) depolama alanının kullanılacağını gösterir.



Şekil 3.10 Analog ses dalgası ve örnekleme yoluyla elde edilen dijital hali

Ses kaydıyla ilgili bilmemiz gereken son konu, sesin bir veya daha çok kanal üzerinden kaydedilebileceğidir. **Mono** olarak bilinen 1 kanalın kullanıldığı durumda, kayıt esnasında sadece 1 mikrofon kullanılır. **Stereo** iki (veya daha çok) kanallıdır ve kayıt esnasında iki (veya daha çok) mikrofon kullanılır. Böylece ses, iki kulağımıza göre daha iyi bir şekilde oluşturulabilir. Özellikle ortamda oluşan seslerin konumları da bu yöntemle fark edilebilir. Mesela stereo kaydedilen bir orkestrada, sağ taraftaki müzik aleti, sağ hoparlör tarafından daha canlı çalınacaktır. Mono bir sesin, iki farklı hoparlörden çalınmasının aynı etkiyi vermeyeceğine dikkat edin. Günümüzde daha yeni olarak **kuşatma (surround)** ses sistemleri çıkmıştır. Bu teknik sanki sesin olduğu ortamın içinde bulunuyorsunuz hissinin verilebilmesi için tasarlanmıştır. **Dolby stereo** namı altında sinema salonlarında sıkça karşılaşılabilirsiniz.

CD kalitesi denilen ses kalitesini yakalamak için, analog ses, saniyede 44,100 defa örneklenir. Her örneklemede 16 bit uzunluğunda stereo veri oluşturulur. Buradan CD kalitesinde bir sesin her saniyesi için yaklaşık 172 KB bellek alanı gerektiği anlaşılır.

Bu bit dizilerinden ses oluşturulmak istenirse bu sefer ses kartı, gelen bit dizilerini, **Dijital-Analog çevirici (Digital-to-Analog Converter -DAC)** yardımıyla, elektrik sinyaline çevirir. Bu sinyal **hoparlör (speaker)** üzerinde bulunan diyaframı hareket ettirir. Bu diyafram etrafındaki hava moleküllerini titreştirerek ses tekrar oluşturulur.

Ses bilgisi wav veya mp3 gibi formatlarda olabilir. **WAV (sound WAVE)** dış dünyadaki herhangi bir sesi kodlayabilir. **MP3** (MPEG audio Layer 3) ise bu gün baskın ses formatıdır. Kodlama ile birlikte veri üzerinde sıkıştırma da yapar. Bu yüzden wav formatına nispeten çok daha az yer

kaplar. Dış dünyadaki bir sesi wav yerine MP3 formatında kaydetmek yaklaşık 10 kat bellek alanı tasarrufu sağlar. Yani bir ses verisi 10 KB wav dosyası oluyorsa, sesin aynı kalitede (örnekleme ve bit çözünürlük değerleri aynı) fakat MP3 formatında saklanmasıyla 1 KB'lık dosya oluşacaktır.



WAV – MP3 çevrimi

Arkadaşlardan birisi, wav formatında kaydedilen ses verisini, VCD player'da dinleyebilmek için MP3 formatına çevirmeye çalıştığından bahsetti. Bu çevirme işleminden sonra 16 MB olan wav dosyasının boyutu 65 MB'a kadar büyüyordu. Halbuki MP3 aynı kalitede sesi çok daha küçük alanlara sığdırabilme konusunda wav formatından daha yeteneklidir. Arkadaş birçok farklı ürünü internette indirdiğini fakat sonucun pek değişmediğinden yakındı.

Wav – MP3 çevirici yazılım içerisinden dosyaları incelemek için baktığımda, wav dosyasının 8 KHz örnekleme hızında ve 4 bit ile Mono formatında kaydedildiğini gördüm. Yani ses verisinin her bir saniyesi için 4KB (8000 x 0.5 Byte) alan harcanıyordu. Buna karşın çıktı olarak üretilmekte olan MP3 dosyasının 44.1 KHz ve 128kbps ile Stereo formatında oluşturulduğunu gördüm. Yani ses verisinin her bir saniyesi için 16KB (128kbps) alan harcanıyordu. Çıktı dosyasının bit çözünürlüğü ve örnekleme oranları gibi özellikleriyle oynadıktan sonra, dinlenebilir bir ses kalitesi için, 16 KHz ve 24 kbps (3 KB/s) ile mono formatında ses dosyası oluşturuldu. Bu sayede oluşan MP3 dosyasının boyutu 12 MB'ın altına kadar düştü.

Dosya boyutunun çok küçülmemesinin bir sebebi, dış ortamdaki sesin kaydedilmesi yerine, wav formatında kaydedilen ses verisinin MP3 formatına çevrilmesidir. Diğer sebebi Wav formatında kaydedilmiş verinin zaten çok düşük örnekleme oranı ve bit çözünürlüğüyle kaydedilmiş olmasıdır.



Windows içinden WAV - MP3 çevrimi

WAV dosyaların MP3 formatına dönüştürme işlemi Windows ses kaydedicisi ile yapılabilir. Windows ses kaydedicisi Donatılar -> eğlence altındadır. Veya doğrudan Başlat -> Çalıştır menü seçeneğinden sndrec32.exe yazarak uygulamayı çalıştırabiliriz. Wav dosyasını açtıktan sonra Dosya -> Özellikler seçeneğini seçeriz. Açılan pencerede şimdi dönüştür düğmesine basarız. Açılan pencerede Biçim kısmında MPEG Layer-3 seçeneğini, alttaki kutuda ise bu biçime ait seçeneklerden uygun birisini seçmelisiniz. Uygun seçimi bulurken birçok deneme yanılma yapmanız gerekebilir.

Bilgisayarın dış dünyadaki sesleri kaydedip tekrar çalması, ses üretmek için kullandığı yöntemlerden birisidir. Buna **sayısallaştırılmış ses** diyoruz. Çünkü analog formdaki ses bilgisi dijital/sayısal hale çevrilir. Bilgisayarın ses üretmek için kullandığı diğer yöntem **sentezlemedir**. Bu yöntem, notaların bir çalgı aletinin (gitar, piyano vb.) nasıl çalınacağını göstermesi gibi, ses donanımına gidecek komutların kullanılması esasına dayanır.

Sentezleme insan sesi gibi karmaşık seslerin üretilmesi konusunda oldukça yetersizdir. Fakat çalgı aleti seslerinin bu yolla üretilmesi oldukça kolaydır. Sentezleme tekniğinin en yaygın örneği **MIDI (Musical Instrument Digital Interface)** dosyalarıdır. Bir MIDI dosyasının içerisinde temel olarak, hangi çalgı aletinin, hangi notayı ne kadar süre ile çalacağı bilgisinin tutulduğunu düşünebilirsiniz.

Bu dosyalar sesin kendisi yerine üretilme biçimini sakladıkları için, sayısallaştırma tekniğiyle üretilen dosyalara oranla çok daha az yer kaplarlar.

Günümüz ses kartları, gerçekçi müzik üretimi için **Dalga tablosu sentezi (Wave Table Synthesis)** tekniğini kullanırlar. Bu teknikte, müzik aletlerine ait gerçek sesler ses kartının üzerindeki bir bellek bölgesinde kaydedilir. Tekrar üretim sırasında bu sesler kullanılarak gerçeğe çok yakın bir müzik parçası elde edilebilir.

Resim Verisi

Yine önce gerçek dünya... Etrafımızdaki nesnelere çarpan foton parçacıkları, gözümüzün retina denilen arka kısmına ulaşır. Burada ışığa duyarlı 2 tip hücre yaratılmıştır; Çubuk hücreler ve Koni hücreler. Çubuk hücreler az ışık ortamlarında görme işinde çalışırken, koni hücreler ayrıntı ve renk algılamasında çalışırlar. Koni hücreler başlıca üç rengi algılayabilecek yapıdadır: Kırmızı, Yeşil ve Mavi (Red, Green, Blue veya kısaltma olarak RGB). Bu demek oluyor ki, görebildiğimiz tüm renkleri bu üç rengin değişik kombinasyonları ile ifade edebiliriz. Mesela sarı rengi, aynı oranda kırmızı ve yeşil renklerini karıştırarak elde edebiliriz. Bu bilgiyi bilgisayarlarda görüntü ifadesinde kullanacağız...

Gerçek dünya görüntülerinin renk çeşitliliği ve aynı rengin değişik tonlamaları çok fazladır. Bu renkler cismin tüm yüzeyi boyunca süreklilik arzederler. Bunun yanında, bu görüntüler ayrıntı noktasında kusursuzdur. Ne kadar yaklaşırsak o kadar ayrıntı görürüz. Bu kadar renk ve ayrıntıyı, bilgisayar ortamına taşırken şu modeli izleyeceğiz;

1. Resmi noktalara böleceğiz. Böylece resmi analog formdan dijital forma çevireceğiz.
2. Her noktaya Kırmızı, Yeşil ve Mavi renklerinin kombinasyonundan oluşan bir renk atayacağız. Böylece o noktaya ait renk bilgisini kodlayabileceğiz.