



# 3. Fejezet : Adat formátumok

---

**The Architecture of Computer Hardware  
and Systems Software:  
An Information Technology Approach**

**3. kiadás, Irv Englander**

**John Wiley and Sons ©2003**

**Wilson Wong, Bentley College**

**Linda Senne, Bentley College**



# Adat formátumok

---

- Számítógépek
  - Minden adatformátumot bináris formában kezel és tárol
- Emberi kommunikáció
  - Beleértve a nyelveket, képeket és hangokat
- Adat formátumok:
  - Előírások az adatok számítógép által kezelhető formába konvertálásához
  - Különböző módok emberi adat ábrázolására a számítógép által feldolgozva és tárolva



# Adatok forrása

- Bináris bemenet
  - Egyedi inputként indul
  - Példa: billentyűzeti bemenet ***1+2=3 számolás***
  - Billentyűzet generál egy bináris számot minden gombhoz
- Analóg
  - Összefüggő adatok mint a hang vagy a képek
  - Szükséges egy hardver az adatok bináris számokká konvertálásához

**1+2=3**  
**számolás**



**Input**  
**eszköz**



**Számítógép**

**1101000101010101...**



# Általános adatábrázolás

Adat típus	Szabvány(ok)
Alfanumerikus	Unicode, ASCII, EDCDIC
Kép (bit térképes)	<ul style="list-style-type: none"><li>▪ GIF (graphical image format)</li><li>▪ TIF (tagged image file format)</li><li>▪ PNG (portable network graphics)</li></ul>
Kép (vektorgrafikus)	PostScript, SWF (Macromedia Flash), SVG
Grafikai megjelenés, ill. betűkészlet leírás	PostScript, TrueType
Hang	WAV, AVI, MP3, MIDI, WMA
Formátumozott dokumentum leírás	PDF (Adobe Portable Document Format), HTML, XML
Video	Quicktime, MPEG-2, RealVideo, WMV



# Belső adatábrázolás

- Kifejezi az
  - Input forrás összetettségét
  - A szükséges feldolgozás módját
- Optimális ábrázolás szempontjai
  - pontosság és felbontás
    - egyszerű fénykép szemben egy festmény-albummal
  - tömörség (tárolás és továbbítás)
    - több adat szükséges a jobb pontossághoz és felbontáshoz
    - *Tömörítés*: adatábrázolás tömörebb formában
    - *Metadata*: adat, ami leírja vagy meghatározza az adat értelmét
  - Módosítások egyszerűsödése:
    - Egyszerű audió szemben a Hi-Fi minőségű hanggal
  - Szabványosítás
    - *Szabadalmazott formátumok* adattároláshoz és feldolgozáshoz (WordPerfect vs. Word)
    - Tényleges szabványok: szabadalmazott szabványok az általánosan használt formákon alapulva (PostScript)



# Adattípusok: Számadat

---

- Matematikai számításokhoz használják
  - Összeadás, kivonás, szorzás, osztás
- Típusok
  - Integer (egész számok)
  - Real (tartalmaz egy tizedespontot)



# Adattípusok: Alfánumerikus

- Alfánumerikus:
  - Karakterek: *b T*
  - Számok: *7 9*
  - Írásjelek: *! ;*
  - Speciális karakterek: *\$ &*
- Számjegy karakterek és számok
  - Mindegyiket általános karakterként vesszük be
  - Számítógép konvertálja számokká a számításokhoz
    - Példák: változók deklarációja számokként a programozó által (int fut=0)
  - Karakterként kezeli, ha szöveggként kerül feldolgozásra
    - Példák: telefonszámok, irányítószámok



# Alfanumerikus kódok

---

- Tetszőleges bit-sorozatok választása a karakterek ábrázolására
  - Követelmény: input és output eszköznek ismernie kell a kódolást
  - Bináris szám értékének karakter ábrázolása összhangban van az abc-ben való elhelyezkedésével
    - Elősegíti a rendezést és keresést





# A 852-es kódlap

0→	20→¶	40→(	60→<	80→P	100→d	120→x	140→î	160→á	180→ı	200→Ł	220→■	240→-
1→@	21→§	41→)	61→=	81→Q	101→e	121→y	141→ž	161→í	181→Ā	201→Ź	221→ı	241→ˆ
2→0	22→_	42→*	62→>	82→R	102→f	122→z	142→Ā	162→ó	182→Ā	202→Ź	222→Ů	242→ˆ
3→♥	23→ı	43→+	63→?	83→S	103→g	123→{	143→Č	163→ú	183→Ě	203→ı	223→■	243→ˆ
4→♦	24→↑	44→,	64→@	84→T	104→h	124→	144→Ě	164→Ā	184→\$	204→ı	224→Ů	244→ˆ
5→♠	25→↓	45→-	65→A	85→U	105→i	125→}	145→Ĺ	165→a	185→ı	205→=	225→B	245→§
6→♣	26→→	46→.	66→B	86→V	106→j	126→~	146→Ī	166→ž	186→ı	206→ı	226→Ů	246→÷
7→·	27→←	47→/	67→C	87→W	107→k	127→û	147→ô	167→ž	187→ı	207→D	227→N	247→ˆ
8→■	28→_	48→0	68→D	88→X	108→l	128→Ç	148→ö	168→ĸ	188→ı	208→đ	228→ñ	248→ó
9→o	29→+	49→1	69→E	89→Y	109→m	129→ü	149→Ĺ	169→e	189→ž	209→D	229→ň	249→ˆ
10→o	30→▲	50→2	70→F	90→Z	110→n	130→é	150→Ī	170→	190→ž	210→Ď	230→Š	250→ˆ
11→♂	31→▼	51→3	71→G	91→[	111→o	131→â	151→Š	171→ž	191→ı	211→Ě	231→š	251→Ů
12→♀	32→	52→4	72→H	92→\	112→p	132→ä	152→š	172→Č	192→ı	212→đ	232→Ř	252→Ř
13→♪	33→!	53→5	73→I	93→]	113→q	133→û	153→ö	173→š	193→ı	213→N	233→Ů	253→ř
14→♯	34→"	54→6	74→J	94→^	114→r	134→ć	154→Ü	174→«	194→ı	214→Ī	234→ř	254→ı
15→✳	35→#	55→7	75→K	95→	115→s	135→ç	155→Ī	175→»	195→ı	215→Ī	235→Ů	255→
16→▶	36→\$	56→8	76→L	96→¯	116→t	136→ł	156→t	176→ı	196→-	216→ě	236→ý	
17→◀	37→%	57→9	77→M	97→a	117→u	137→ë	157→ł	177→ı	197→ı	217→ı	237→Ů	
18→↑	38→&	58→:	78→N	98→b	118→v	138→ő	158→x	178→ı	198→Ā	218→ı	238→t	
19→!!	39→'	59→;	79→0	99→c	119→w	139→ö	159→č	179→ı	199→ä	219→ı	239→ı	



# Karakterek ábrázolása

---

- ASCII – legszélesebb körben használt
- EBCDIC: IBM nagyszámítógép (hagyományos)
- Unicode: világméretű használatra kifejlesztve



# ASCII

---

- ANSI (American National Standards Institute) által kifejlesztett
- Ábrázolás
  - Latin betűk, arab számok, általános írásjelek
  - Plusz néhány ékezetes és más európai speciális karakterek
- ASCII
  - 7-bites kód: 128 karakter



# ASCII referenciacsík

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

74<sub>16</sub>  
111 0100



# EBCDIC

- Extended Binary Coded Decimal Interchange Code az IBM által kifejlesztve
  - Főként IBM vagy IBM kompatibilis nagyszámítógépek körében használatos
  - ASCII-ba/ból konvertáló szoftver hozzáférhető
  - Gyakori az archív adatoknál
  - Karakterkódok eltérnek az ASCII-től

	ASCII	EBCDIC
Space	20 <sub>16</sub>	40 <sub>16</sub>
A	41 <sub>16</sub>	C1 <sub>16</sub>
b	62 <sub>16</sub>	82 <sub>16</sub>



# Unicode

---

- Leggyakoribb, 16-bites forma, 65.536 karakter ábrázolására alkalmas
- ASCII Latin-I Unicode része
  - 0-tól 255-ig a Unicode táblában
- Soknyelvű: kódokat határoz meg
  - Majdnem az összes betű alapú karakterhez
  - Nagy része a kínai, japán és koreai „szóképeknek”
  - Vegyesen magánhangzók és szótag csoportok néhány nyelv igénye szerint
- Támogatja a szoftvermódosítást helyi nyelvekre



# Összehasonlító sorozat

- Alfabetikus rendezés ha a szoftver kezeli a vegyes kis- és nagybetű kódokat
- ASCII-ban a számokat helyezi előre; EBCDIC-ben utoljára
- ASCII rendezési sorrend karakterláncokra

Betűk							Szám karakterek					
Adam	A	d	a	m			1	011	0001			
Adamian	A	d	a	m	i	a	n	12	011	0001	011	0010
Adams	A	d	a	m	s			2	011	0010		



# A kódok kettő osztálya

- *Nyomtatható* karakterek
  - Megjelenik a képernyőn vagy a nyomtatón
- *Vezérlő* karakterek
  - Pozícionálja a kimenetet a képernyőn vagy a nyomtatón
    - ▣ VT: vertical tab (tabulátor)    ▣ LF: Line feed (újsor)
  - Tevékenységet vált ki
    - ▣ BEL: bell rings (csipogás)    ▣ DEL: aktuális karakter törlése
  - Állapotot közöl a számítógép és az I/O eszköz között
    - ▣ ESC: a soron következő karakter értelmezésének megváltoztatása





# Vezérlőkódok meghatározása

<b>NUL</b>	(Null) No character; used to fill space	<b>DLE</b>	(Data Link Escape) Similar to escape, but used to change meaning of data control characters; used to permit sending of data characters with any bit combination
<b>SOH</b>	(Start of Heading) Indicates start of a header used during transmission	<b>DC1, DC2, DC3, DC4</b>	(Device Controls) Used for the control of devices or special terminal features
<b>STX</b>	(Start of Text) Indicates start of text during transmission	<b>NAK</b>	(Negative Acknowledgment) Opposite of ACK
<b>ETX</b>	(End of Text) Similar to above	<b>SYN</b>	(Synchronous) Used to synchronize a synchronous transmission system
<b>EOT</b>	(End of Transmission)	<b>STB</b>	(End of Transmission Block) Indicates end of a block of transmitted data
<b>ENQ</b>	(Enquiry) A request for response from a remote station; the response is usually an identification	<b>CAN</b>	(Cancel) Cancel previous data
<b>ACK</b>	(Acknowledge) A character sent by a receiving device as an affirmative response to a query by a sender	<b>EM</b>	(End of Medium) Indicates the physical end of a medium such as tape
<b>BEL</b>	(Bell) Rings a bell	<b>SUB</b>	(Substitute) Substitute a character for one sent in error
<b>BS</b>	(Backspace)	<b>ESC</b>	(Escape) Provides extensions to the code by changing the meaning of a specified number of contiguous following characters
<b>HT</b>	(Horizontal Tab)	<b>FS, GS, RS, US</b>	(File, group, record, and united separators) Used in optional way by systems to provide separations within a data set
<b>LF</b>	(Line Feed)	<b>DEL</b>	(Delete) Delete current character
<b>VT</b>	(Vertical Tab)		
<b>FF</b>	(Form Feed) Moves cursor to the starting position of the next page, form, or screen		
<b>CR</b>	(Carriage return)		
<b>SO</b>	(Shift Out) Shift to an alternative character set until SI is encountered		
<b>SI</b>	(Shift In) see above		



# Billentyűzet bemenet

- *Scan kód*
  - Két különböző scan kód van a billentyűzeten
    - ▣ Egyik akkor keletkezik ha egy gombot lenyomunk, a másik, ha felengedünk egy gombot
  - A szoftver által terminálon vagy PC-n Unicode -dá, ASCII –vé vagy EBCDIC –ké konvertálva
- *Előny*
  - Könnyen alkalmazható különböző nyelveknél vagy billentyűzetkiosztásoknál
  - Különböző scan kód a gombok megnyomásához/ felengedéséhez és a billentyűzetkombinációkhoz
    - ▣ Példák: shift és control gombok



# Más alfanumerikus bemenet

- **OCR** (optical character reader/optikai jelolvasó)
  - Nyomtatott szöveget olvas be és szöveges adattá konvertálja
  - Különlegesen kódolt karakterek olvasására használják
  - Általános használat korlátozott a nagymértékű hibaarány miatt
    - mágneses felületek „optikai” felismerése (pl. csekkeken vízjel helyett)
- **Vonalkód olvasók**
  - Olyan alkalmazásokban használják, ami gyors, pontos, ismétlődő inputot igényel minimálisan képzett alkalmazottakkal
  - Példák: bevásárlóközpont pénztár- és leltár kezelése
  - Alfa numerikus adat a vonalkód olvasóban optikailag vonalakat használ
- **Mágneskártya olvasók:** alfanumerikus adat bankkártyákról
- **Hang**
  - Digitalizált audió rögzítése egyszerű, de alfanumerikus adattá konvertálása nehéz
    - Ismerni kell a hangmintákat egy nyelvben (*fonéma/beszédhang*) plusz a kiejtés-, nyelvtan- és mondattan szabályait



# Kép adat

- Fényképek, ábrák, ikonok, rajzok, táblázatok és grafikonok
- Kétféle megközelítés:
  - *Bit-térképes* vagy *raszteres (rácso) képek*
    - Képek és festmények összefüggő váltakozása
  - *Objektumos* vagy *vektorgrafikus*
    - *grafikai objektumokból* áll össze, mint a vonalak és görbék, geometriailag meghatározva
- Különbségek:
  - Kép minősége
  - Tárhely igény
  - Továbbadás időtartama
  - Módosítás könnyedsége



# Bit-térképes képek

- Valóságű, folytonosan változó árnyékú, színű, élességű és szerkezetű képeknél használják
  - Példák:
    - ▣ Scanner-elt képek
    - ▣ Képtár egy *rajz*-program által létrehozva
- Főként nagymértékben részletezett és meglehetősen egyszerű feldolgozás igényű képeknél használják
- Input eszközök:
  - Scanner-ek
  - Digitális kamerák és videó felvevő eszközök
  - Grafikus input eszközök, mint az egér és a toll
- *Képszerkesztő*- vagy *rajz program* kezeli
  - Szerkesztő eszközök, amelyek könnyebbé teszik a fárasztó bit-ről bit-re való feldolgozást






# Bit-térkép megjelenítés

- Monochrome: fekete vagy fehér
  - 1 bit pixelenként
- Szürke: fekete, fehér illetve 254 árnyalata a szürkének
  - 1 byte pixelenként
- Színes képek: 16 szín, 256 szín, vagy 24 bit-es true color (16,7 millió szín)
  - 4, 8, és 24 bit-es (3 x 8 bit-es) képek



# Bit-térképes képek tárolása

- Gyakran nagy fájlok
  - Példa: 800×600 pixel, mindhárom szín (RGB) 1 byte-on tárolva  ~1.5MB –os file-t ad (1.37 MB)
- A file méret függ a
  - *Felbontástól* (pixelek száma inch-enként)
    - Kihatással van a kép élességére és részletességére
  - *Mélységtől*: szürke vagy összetett színek megjelenítéséhez használt bit-ek száma
    - *Paletta (színválaszték)*: színtábla, ami minden pixelhez kódot használ az aktuális szín értéke helyett
  - Adattömörítéstől





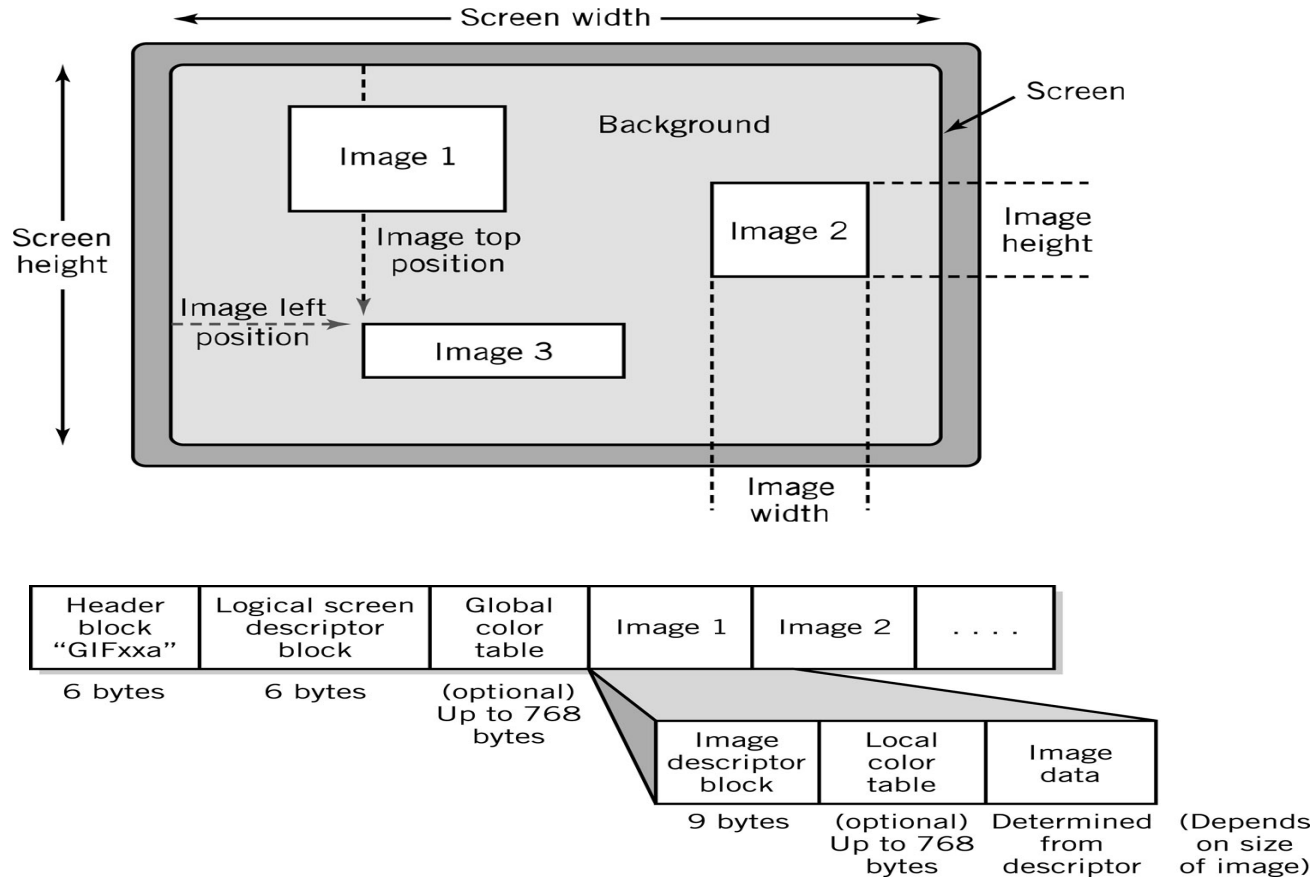
# GIF (Graphics Interchange Format)

---

- 1987-ben fejlesztette ki a CompuServe
- GIF89a animált képek lehetősége
  - Lehetővé teszi a képek egymást követő megjelenítését rögzített időközönként
- Színkorlát: 256
- LZW (Lempel-Zif-Welch) algoritmussal tömörítve
- Főként vonalrajzoláshoz, művészi és nagyméretű azonos színű területeket tartalmazó képekhez
- *Veszteségmentes tömörítés*



# GIF (Graphics Interchange Format)





# JPEG

## (Joint Photographers Expert Group)

- Több mint 16 millió színre képes
- Megfelelő nagy részletességű fényképekhez és festményekhez
- *Veszteséges tömörítési* algoritmust alkalmaz, amely
  - Adatokat eldobva csökkenti a file méretét és az átviteli sebességet
  - Csökkentheti a kép felbontását, elmoshatja az éles vonalakat
  - Minden egyes mentés "rontja" a kép minőségét



# Más bit-térképes formátumok

- TIFF (Tagged Image File Format/címkézett kép file formátum): *.tif*
  - Kiváló minőségű képek feldolgozása, különösképp kiadásnál
- BMP (BitMaPped/Bittérképes): *.bmp*
  - Eszköz-független formátum Microsoft Windows környezethez: pixel színek a kimeneti eszközöktől függetlenül tárolva
- PCX: *.pcx*
  - Windows Paintbrush formátuma
- PNG: (Portable Network Graphics): *.png*
  - Internetes alkalmazásoknál a GIF és JPEG lecserélésére tervezték
  - Szabadalom mentes
  - Továbbfejlesztett *vesztésmentes* tömörítés
  - Nincs animáció támogatás



# Vektorgrafikus képek

- Csomagok *rajzolásával* vagy ábratáblázattal készítik
- Különböző színű vonalakból és alakzatokból áll össze
- A számítógép lefordítja a geometriai formulákat a kép megalkotásához
- Tárhely mérete függ a kép összetettségétől
  - Utasítások száma egy vonal rajzolásához, alakzatok, minták mennyisége
- Filmek, így a *Shrek* és *Toy Story* megalkotásához is vektorgrafikát használtak



# Vektorgrafikus képek

---

- Matematikai képleteken alapul
  - Könnyű mozgatni, méretezni és elforgatni az élesség elvesztése nélkül, a bit-térképesek képekkel szemben
- Kevesebb helyre van szüksége, mint a bit-térképes képeknek
- Nem tud képeket vagy festményeket ábrázolni
- Nem lehet közvetlenül megjeleníteni vagy nyomtatni
  - Bit-térképesé kell konvertálni, mivel a rajzgép bit-térképes



# Ismertebb vektorgrafikus programok

- A legtöbb vektorgrafikus formátum szabadalmazott
  - File kiterjesztéseket is beleértve .wmf, .dxf, .mgx, és .cgm
- Macromedia Flash: alacsony sáv szélességű animáció
- Micrographx Designer: tervrajzkészítés, a termékek illusztrálásához
- CorelDraw: vektoros illusztráló, elrendező, bit-térkép készítő, képszerkesztő, festő és animáló program
- Autodesk AutoCAD: építészeknek, mérnököknek, tervezőknek, és tervfüggő szakembereknek
- W3C SVG (Scalable Vector Graphics) XML Web leíró nyelven alapul
  - Nem szabadalmazott



# PostScript

---

- *Oldalleíró nyelv*: műveletek és tulajdonságok listája, amely jellemzi az összes objektumot a nyomtatandó oldalon
  - ASCII-ben vagy Unicode-ban tárolva
  - Az értelmező program a számítógépen vagy a kimeneti eszközön olvassa a PostScript-et a kép konvertálásához
- Méretezhető betű támogatás
  - Többi objektumhoz hasonló meghatározás





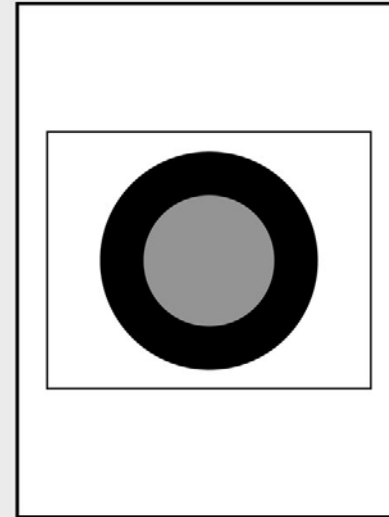
# PostScript program

```
288 396 translate % move origin to center of page
0 0 144 0 360 arc % define 2" radius black circle
fill

0.5 setgray % define 1" radius gray circle
0 0 72 0 360 arc
fill

0 setgray % reset color to black
-216 -180 moveto % start at lower left corner
0 360 rmoveto % and define rectangle
432 0 rmoveto % ...one line at a time
0 -360 rmoveto
closepath % completes rectangle
stroke % draw outline instead of fill

showpage % produce the image
```





# Karakterek ábrázolása

---

- Karakterek tárolása Unicode-hoz vagy ASCII-hez hasonló formátumokban
  - A szöveg elsősorban tartalom szerint kerül feldolgozásra és tárolásra
- Megjelenítési feltételek, mint a betűkészlet tárolása karakterrel
  - Szövegmegjelenés az elsődleges tényező
    - Példa: betűkészlet a Windows-ban
- *Glyph-ek*: Macintosh kódolási minta, amely tartalmazza a karakterek azonosítási és megjelenési feltételeit



# Bit-térképes és vektorgrafikus képek összehasonlítása

Bit-térképes	Vektorgrafikus
Pixel térkép	Geometriailag meghatározott formák
Fénykép minőség	Összetett rajzok
Festő program	Rajz program
Nagyobb tárhely igény	Nagyobb számítási igény
Kép méretének növelése életlenséget okoz	Objektumok méretezése egyenletes
A kimenet felbontását a kép felbontása korlátozza	A kimenet felbontását a kimeneti eszköz korlátozza



# Videó képek

- Nagymennyiségű adatra van szükség
  - A videokamera 720x576/704x480 (PAL/NTSC) felbontású 24 bit-es képet rögzít 25/30 kép/másodperc sebességgel.
  - Ez 248 Mbit/sec a CCD-n, a szalagon átlagosan 28.8 Mbit/sec.
  - 1 perces film 1.7 Gbit és egy 1 órás film közel 13 Gbyte!
- Lehetőségek a file méret csökkentésére: csökkenti a kép méretét, korlátozza a színek számát, csökkenti a képkocka-arányt
- A videohoz való hozzájutási módtól függően
  - *Videó folyam*: a videó úgy jelenik meg, ahogy letöltöttük a Web szerverről
    - Példa: videó konferencia
  - Helyi adat (file a DVD-re vagy a rendszerre letöltve) a jobb minőségért
    - MPEG-2: kódolási példa: IBBPBBPBBI
      - Intra, Bidirectional, Predictive, film minőségű képek, magas tömörítéssel, tekintélyes feldolgozási képességet igényel,
      - $3 \div 10$  Mbit/sec tömörítési arányra optimalizálva, ez a DVD szabványa is. A file méret = 95% videó-adat + 5% tömörítetlen audió-adat.



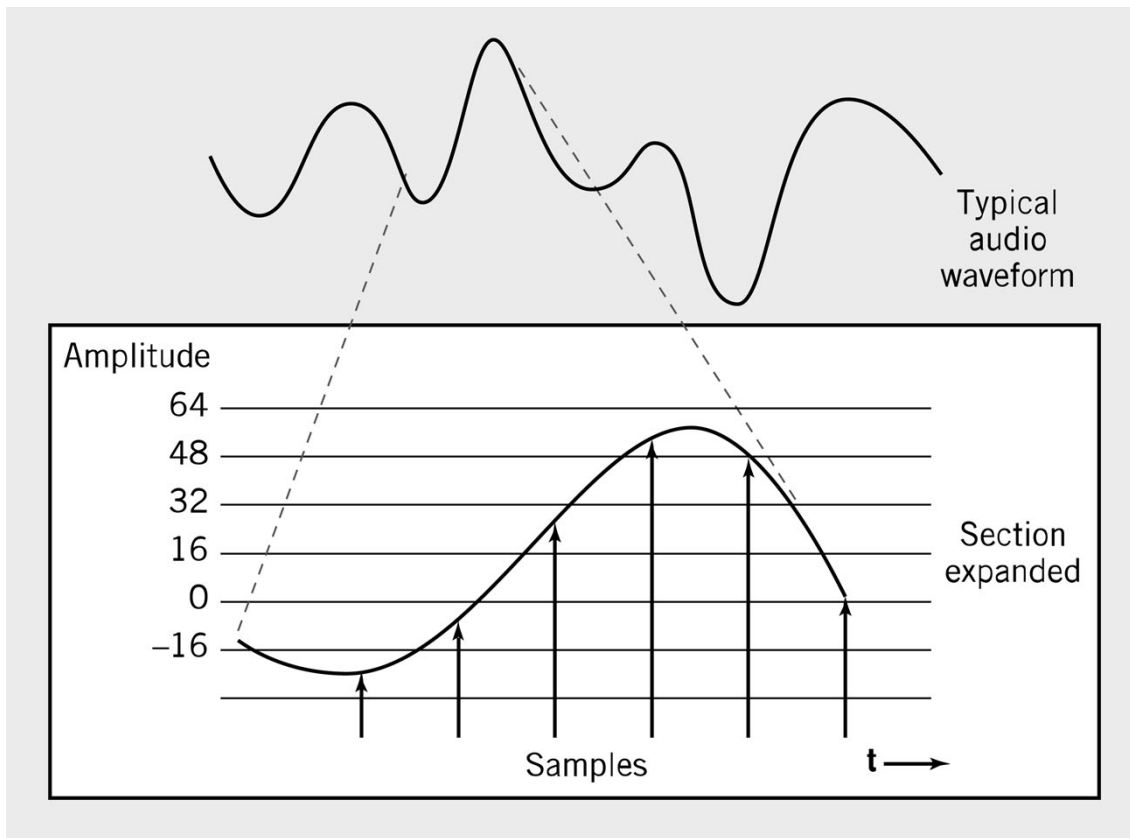
# Audió adat

---

- Átviteli és feldolgozási igényei kisebbek, mint a videoé
- *Waveform audio*: a hang digitális ábrázolása
- *MIDI* (Musical Instrument Digital Interface): utasítások a hangok előállításához
- Az *A-D (analóg-digitális) konverter* analóg hangot konvertál a digitális értékekre



# Waveform Audio



Mintavételezési  
arány  
rendszerint  
50kHz



# Mintavételezési arány

- A hangból vett mintavételek száma másodpercenként a rögzítési folyamat alatt
  - 1000 mintavétel másodpercenként = 1 kHz (kilohertz), azaz ezred másodpercenként 1.
  - Példa: Audió CD mintavételezési aránya = 44.1 kHz, azaz 226 tízmilliomod másodpercenként 1.
- Különböző minőségű rögzítések:
  - 8 bit-es kvantáltság a rádió minőségű rögzítéshez,
  - 16 bit-esnél 44.1 kHz CD, 48 kHz a Hi-Fi minőségű rögzítéshez, 2 sztereó csatornán,
  - 12 bit-es kvantáltság 32 kHz-en 2 x 2 sztereó csatornán rögzíti a hangot,
  - 24 bit-es kvantáltság 192 kHz-en 5.1 hangkép audió DVD-n.



# MIDI

---

- Zene jelölési rendszer, amely lehetővé teszi a számítógépek kommunikációját a szintetizátorokkal
- Utasítások, melyeket a MIDI műszerek és a MIDI hangkártyák használnak a hang visszajátszásához.
  - Nem alkalmas beszéd- vagy énekhang tárolására vagy visszajátszására
  - Sokkal tömörebb, mint a waveform
  - 3 perc = 10 kB





# Audió formátumok

---

- *MP3*

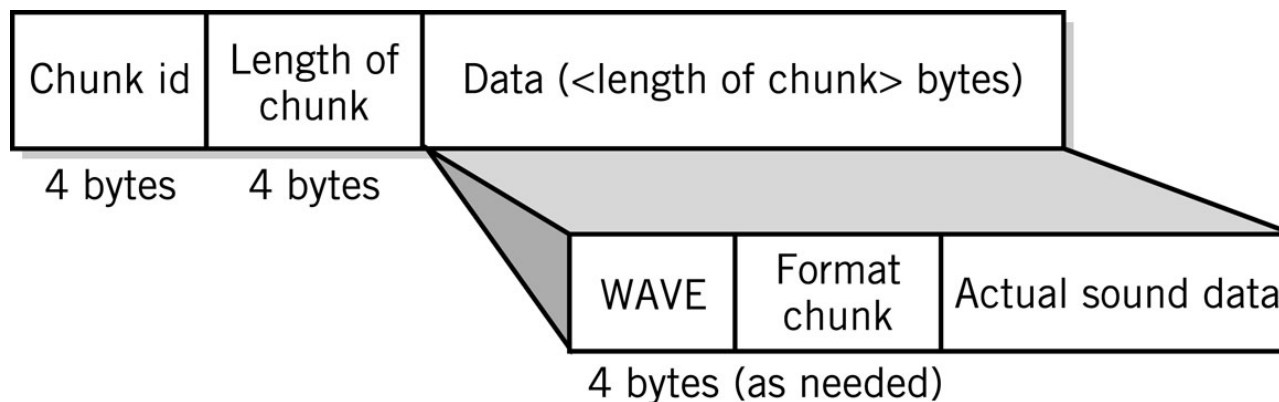
- Az MPEG-2 leszármazottja (ISO *M*oving *P*icture *E*xperts *G*roup)
- *Psychoacoustic*–kus tömörítési technikát alkalmaz a tárhelyigény csökkentéséhez
- Eldobja az ember számára hallhatatlan hangokat: veszteséges tömörítés

- *WAV*

- Microsoft fejlesztette ki a multimédiás környezetének a részeként
- Általános felhasználású veszteségmentes tömörítés



# .WAV hang formátum





# Adattömörítés

- *Tömörítés*: adatrögzítés, amelynek kevesebb tárhelyre van szüksége
- *Tömörítési arány*: méretes file kicsinyítése
- *Veszteség mentes*: a fordított algoritmus visszaállítja az adatot eredeti formájába
  - Példák: GIF, PCX, TIFF, WAV
- *Veszteséges*: adatvesztés lép fel a kisebb méret és a gyorsabb letöltés kedvéért
  - Sokkal nagyobb tömörítési arányok, általában 1-től 10-ig
  - Példa: JPEG
  - Gyakori a multimédiában
- MPEG-2: mindkettőt használja 100:1-es arányhoz



# Tömörítő algoritmusok

- Ismétlődés

- 0 5 8 7 0 0 0 0 3 4 0 0 0 → 0 1 5 8 7 0 4 3 4 0 3

- Példa: nagy területek ugyanazzal a színnel

- Minta helyettesítés

- Mintákat keres az adatban

- Behelyettesíti az új mintákat, szótárat hoz létre

- Példa: 45-ből 30 byte plusz szótár

⌘	Pe	*	pi	❖	ed
☀	er	◎	ck	⊕	pe
☆	Pi				

- Peter Piper picked a peck of pickled peppers.

- ⌘t☀☆p☀\*⊕◎❖ a ⊕◎ of \*◎❖ ⊕pps.



# Belső számítógép adatformátum

---

- Minden adat bináris számként van rögzítve
- Értelmezés alapjai
  - A számítógép képes a műveletek végrehajtására
  - Programozási nyelvek adattípus támogatása az alkalmazások készítéséhez



# Öt egyszerű adattípus

- *Boolean*: kettő értékű változó vagy konstans, amelynek értéke igaz vagy hamis
- *Char*: változó vagy konstans, amelynek értéke alfanumerikus karakter
- *Enumerated* (felsorolt):
  - Felhasználó által meghatározott adattípusok, a definícióban megadott értékeket veheti fel
    - Type DayOfWeek = Mon, Tues, Wed, Thurs, Fri, Sat, Sun
- *Integer*: pozitív vagy negatív egész számok
- *Real*:
  - Számok tizedesponntal
  - Számok, amelyek értéke olyan nagy vagy kicsi, hogy meghaladja a számítógép integer tárolási határát