



Εισαγωγή στην Πληροφορική και τον Προγραμματισμό Η/Υ

6^ο Μάθημα

Άλλοι τύποι δεδομένων. Κωδικοποίηση Χαρακτήρων ASCII/UNICODE.
Μετατροπή Αναλογικού Σήματος σε Ψηφιακό (ADC/DAC)

Λεωνίδας Αλεξόπουλος

Αν. Καθηγητής ΕΜΠ

E-mail: leo@mail.ntua.gr

Τηλ: 210 772-1666

Σύνοψη Μέχρι Σήμερα:

Στα προηγούμενα μαθήματα:

- Κωδικοποίηση – Επεξεργασία - Αποκωδικοποίηση
- Συστήματα Αρίθμησης – το Δυαδικό
- Bit/Byte/Word
- Κωδικοποίηση **ΑΡΙΘΜΩΝ** στο Δυαδικό
 - Παράσταση Ακέραιων Θετικών / Αρνητικών
 - Αναφέραμε + / - / x (όχι ÷) underflow/overflow ακεραίων
 - Fixed & Floating point (single/double –IEEE) Arithmetic
 - ΔΕΝ θα αναφέρουμε: +/ - / * / ÷ , underflow/overflow

ΣΗΜΕΡΑ:

- Κωδικοποίηση **ΑΛΛΩΝ** δεδομένων στο Δυαδικό
 - Χαρακτήρες, Εικόνα, Ήχος, Αναλογικό Σήμα
- Μετατροπή Αναλογικού Σήματος σε Ψηφιακό (ADC/DAC)

Προγραμματισμός σε Η/Υ

Εφαρμογές



Γλώσσες Προγραμματισμού

```

1 function rowTotals = rowsum
2 % Add the values in each row and
3 % store them in a new array.
4
5 x = ones(2,10);
6 [n, m] = size(x);
7 rowTotals = zeros(1,n);
8 for j = 1:n
9     Press Shift+Enter to rename 5 instances of 'j' to 'j'
10 end
11
12 function colsum = addToSum
13 colsum = 0;
14 thisrow = x(i,:);
15 for i = 1:m
16     colsum = colsum + thisrow(i);
17 end
18 end
19
20 end
    
```

Λειτουργικό Σύστημα

Γλώσσα Μηχανής

Assembly Language	Machine Code
add \$t1, \$t2, \$t3	04CB: 0000 0100 1100 1011
addi \$t2, \$t3, 60	16BC: 0001 0110 1011 1100
and \$t3, \$t1, \$t2	0299: 0000 0010 1001 1001
andi \$t3, \$t1, 5	22C5: 0010 0010 1100 0101
beq \$t1, \$t2, 4	3444: 0011 0100 0100 0100
bne \$t1, \$t2, 4	4444: 0100 0100 0100 0100
j 0x50	F032: 1111 0000 0011 0010
lw \$t1, 16(\$s1)	5A50: 0101 1010 0101 0000
nop	0005: 0000 0000 0000 0101
nor \$t3, \$t1, \$t2	029E: 0000 0010 1001 1110
or \$t3, \$t1, \$t2	029A: 0000 0010 1001 1010
ori \$t3, \$t1, 10	62CA: 0110 0010 1100 1010
ssl \$t2, \$t1, 2	0455: 0000 0100 0101 0101
srl \$t2, \$t1, 1	0457: 0000 0100 0101 0111
sw \$t1, 16(\$t0)	7050: 0111 0000 0101 0000
sub \$t2, \$t1, \$t0	0214: 0000 0010 0001 0100

Μικρολειτουργίες & Μικροπρογραμματισμός

Ψηφιακή Λογική



$x \setminus y$	0	1
0	0	0
1	0	1

Figure 1. Truth tables

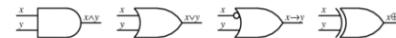
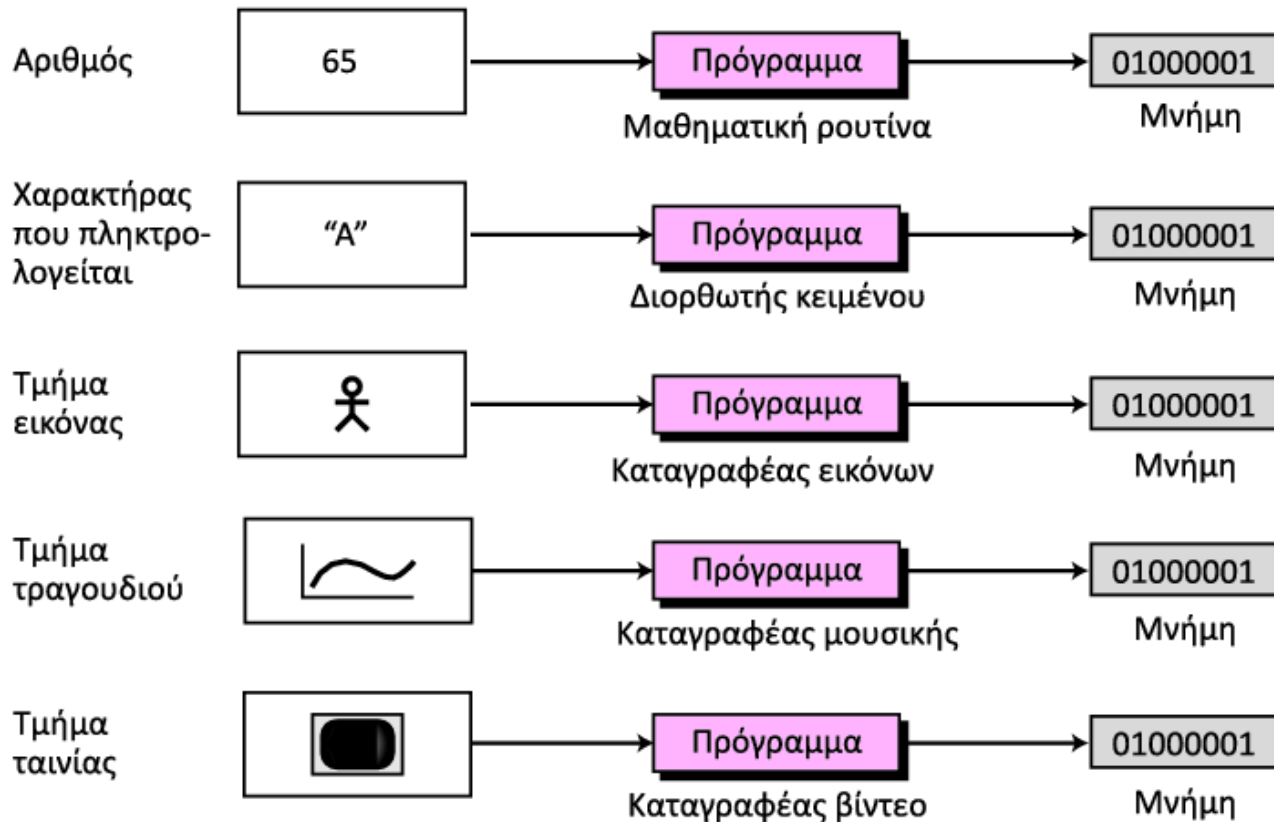
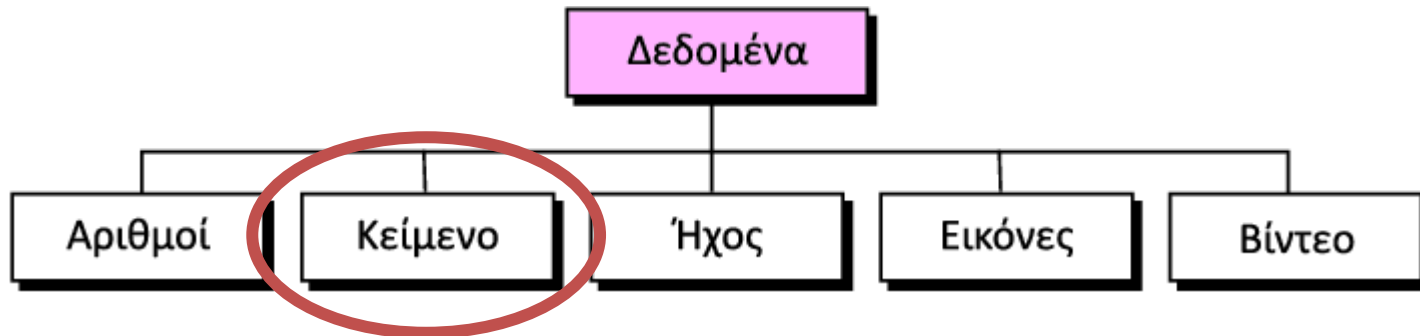


Figure 2. Logic gates

Άλλοι Τύποι Δεδομένων



Κωδικοποίηση Κειμένου

Κωδικοποίηση «**αλφαριθμητικών χαρακτήρων**» δηλ.

- **Γραμμάτων αλφαβήτου**: A,B,C,...,Y,Z,a,b,c,...,x,y,z
- **Αριθμητικών ψηφίων**: 1,2,3,...,9,0
- **Σημείων στίξης**: !,.,...
- **Άλλων συμβόλων**: +,-,{,%, ...
- **Χαρακτήρων ελέγχου**: DELETE, ENTER, κλπ

γίνεται σε ακολουθία δυαδικών ψηφίων **συγκεκριμένου** μήκους bit (πχ 8, 16 etc).

256 είσοδοι 8bit



Κώδικες ASCII και Unicode

- Αρχικά ο **κώδικας ASCII** άσκι -American Standard Code for Information Interchange κάνει χρήση ακολουθίας **7** δυαδικών ψηφίων, πράγμα που επιτρέπει την απεικόνιση, το πολύ, $2^7=128$ αλφαριθμητικών χαρακτήρων.

Παραδείγματα: **A** \leftrightarrow 1000001_2 **0** \leftrightarrow 0110000_2
 + \leftrightarrow 0101011_2 **DEL** \leftrightarrow 1111111_2

- Παρατηρούμε ότι **0** \leftrightarrow $0110000_2 = 48_{10}$!!! Παράλογο ?
Όχι, γιατί στον δυαδικό 0110000_2 αντικατοπτρίζεται ο **χαρακτήρας μηδέν** (που, π.χ. μπορεί να υπάρχει μέσα σε ένα κείμενο) και όχι η **αριθμητική τιμή** (δηλ. ο αριθμητικός ακέραιος) **μηδέν**.

Ο extended ASCII κώδικας είναι 8bit

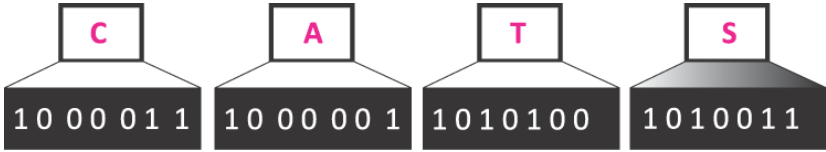
- Το πρότυπο **Unicode** παριστά τους χαρακτήρες με ακολουθίες **16** δυαδικών ψηφίων απεικονίζοντας έτσι $2^{16}=65536$ χαρακτήρες. Το πρότυπο Unicode είναι υπερσύνολο του ASCII και εκτός από την περιοχή του Extended ASCII (0_{10} - 255_{10}) έχει «περιοχές» που αντιστοιχούν στην ελληνική γλώσσα (880_{10}^6 975_{10}) Κινέζικη κλπ

ASCII TABLE

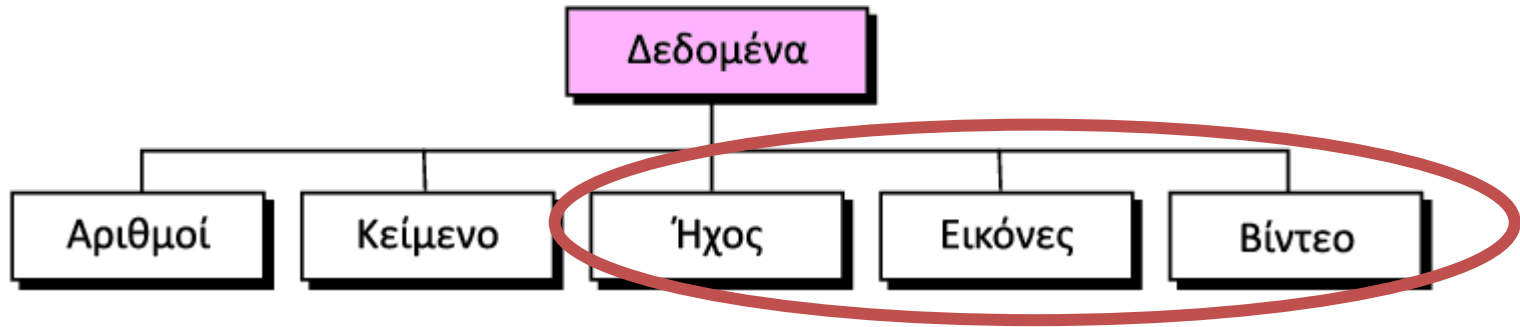
Decimal	Hexadecimal	Binary	Octal	Char	Decimal	Hexadecimal	Binary	Octal	Char	Decimal	Hexadecimal	Binary	Octal	Char
0	0	0	0	[NULL]	48	30	110000	60	0	96	60	1100000	140	`
1	1	1	1	[START OF HEADING]	49	31	110001	61	1	97	61	1100001	141	a
2	2	10	2	[START OF TEXT]	50	32	110010	62	2	98	62	1100010	142	b
3	3	11	3	[END OF TEXT]	51	33	110011	63	3	99	63	1100011	143	c
4	4	100	4	[END OF TRANSMISSION]	52	34	110100	64	4	100	64	1100100	144	d
5	5	101	5	[ENQUIRY]	53	35	110101	65	5	101	65	1100101	145	e
6	6	110	6	[ACKNOWLEDGE]	54	36	110110	66	6	102	66	1100110	146	f
7	7	111	7	[BELL]	55	37	110111	67	7	103	67	1100111	147	g
8	8	1000	10	[BACKSPACE]	56	38	111000	70	8	104	68	1101000	150	h
9	9	1001	11	[HORIZONTAL TAB]	57	39	111001	71	9	105	69	1101001	151	i
10	A	1010	12	[LINE FEED]	58	3A	111010	72	:	106	6A	1101010	152	j
11	B	1011	13	[VERTICAL TAB]	59	3B	111011	73	;	107	6B	1101011	153	k
12	C	1100	14	[FORM FEED]	60	3C	111100	74	<	108	6C	1101100	154	l
13	D	1101	15	[CARRIAGE RETURN]	61	3D	111101	75	=	109	6D	1101101	155	m
14	E	1110	16	[SHIFT OUT]	62	3E	111110	76	>	110	6E	1101110	156	n
15	F	1111	17	[SHIFT IN]	63	3F	111111	77	?	111	6F	1101111	157	o
16	10	10000	20	[DATA LINK ESCAPE]	64	40	1000000	100	@	112	70	1110000	160	p
17	11	10001	21	[DEVICE CONTROL 1]	65	41	1000001	101	A	113	71	1110001	161	q
18	12	10010	22	[DEVICE CONTROL 2]	66	42	1000010	102	B	114	72	1110010	162	r
19	13	10011	23	[DEVICE CONTROL 3]	67	43	1000011	103	C	115	73	1110011	163	s
20	14	10100	24	[DEVICE CONTROL 4]	68	44	1000100	104	D	116	74	1110100	164	t
21	15	10101	25	[NEGATIVE ACKNOWLEDGE]	69	45	1000101	105	E	117	75	1110101	165	u
22	16	10110	26	[SYNCHRONOUS IDLE]	70	46	1000110	106	F	118	76	1110110	166	v
23	17	10111	27	[ENG OF TRANS. BLOCK]	71	47	1000111	107	G	119	77	1110111	167	w
24	18	11000	30	[CANCEL]	72	48	1001000	110	H	120	78	1111000	170	x
25	19	11001	31	[END OF MEDIUM]	73	49	1001001	111	I	121	79	1111001	171	y
26	1A	11010	32	[SUBSTITUTE]	74	4A	1001010	112	J	122	7A	1111010	172	z
27	1B	11011	33	[ESCAPE]	75	4B	1001011	113	K	123	7B	1111011	173	{
28	1C	11100	34	[FILE SEPARATOR]	76	4C	1001100	114	L	124	7C	1111100	174	
29	1D	11101	35	[GROUP SEPARATOR]	77	4D	1001101	115	M	125	7D	1111101	175	}
30	1E	11110	36	[RECORD SEPARATOR]	78	4E	1001110	116	N	126	7E	1111110	176	~
31	1F	11111	37	[UNIT SEPARATOR]	79	4F	1001111	117	O	127	7F	1111111	177	[DEL]
32	20	100000	40	[SPACE]	80	50	1010000	120	P					
33	21	100001	41	!	81	51	1010001	121	Q					
34	22	100010	42	"	82	52	1010010	122	R					
35	23	100011	43	#	83	53	1010011	123	S					
36	24	100100	44	\$	84	54	1010100	124	T					
37	25	100101	45	%	85	55	1010101	125	U					
38	26	100110	46	&	86	56	1010110	126	V					
39	27	100111	47	'	87	57	1010111	127	W					
40	28	101000	50	(88	58	1011000	130	X					
41	29	101001	51)	89	59	1011001	131	Y					
42	2A	101010	52	*	90	5A	1011010	132	Z					
43	2B	101011	53	+	91	5B	1011011	133	[
44	2C	101100	54	,	92	5C	1011100	134	\					
45	2D	101101	55	-	93	5D	1011101	135]					
46	2E	101110	56	.	94	5E	1011110	136	^					
47	2F	101111	57	/	95	5F	1011111	137	_					

- 33 controls
- 95 glyphs
 - English alphabet
 - Uppercase
 - Lowercase
 - Digits
 - 31 punctuation marks and symbols

EXTENDED ASCII (8bit) περιλαμβάνει σημαντικά σύμβολα. Πχ Euro, ÷, ·, ≠, ≥, ≈, π



Άλλοι Τύποι Δεδομένων



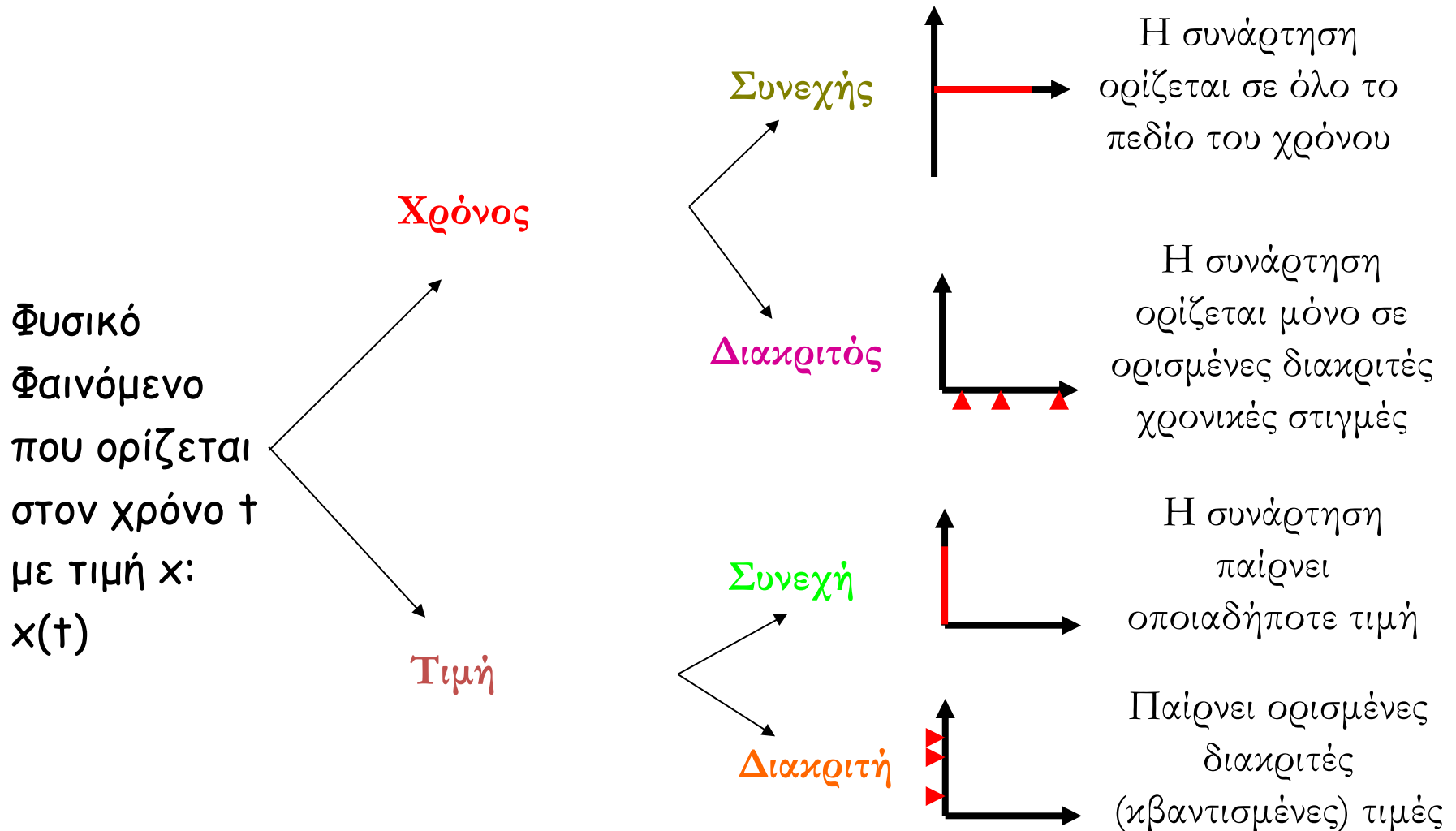
Πως ψηφιοποιούμε ήχο, εικόνα, βίντεο κ.α.?

Καταγράφουμε σε **συγκεκριμένο (διακριτό) χρόνο**
(πχ 1/nsec) **διακριτοποιημένες τιμές**

Πως?

Με **Αναлого-Ψηφιακές Μετατροπείς**

Πεδία ορισμού & τιμών συναρτήσεων φυσικών μεγεθών



Πεδία ορισμού & τιμών συναρτήσεων

TIMEΣ	ΣΥΝΕΧΕΙΣ TIMEΣ	ΔΙΑΚΡΙΤΕΣ TIMEΣ	Συνήθης Φυσικός Κόσμος
ΧΡΟΝΟΣ			

Ο Η/Υ

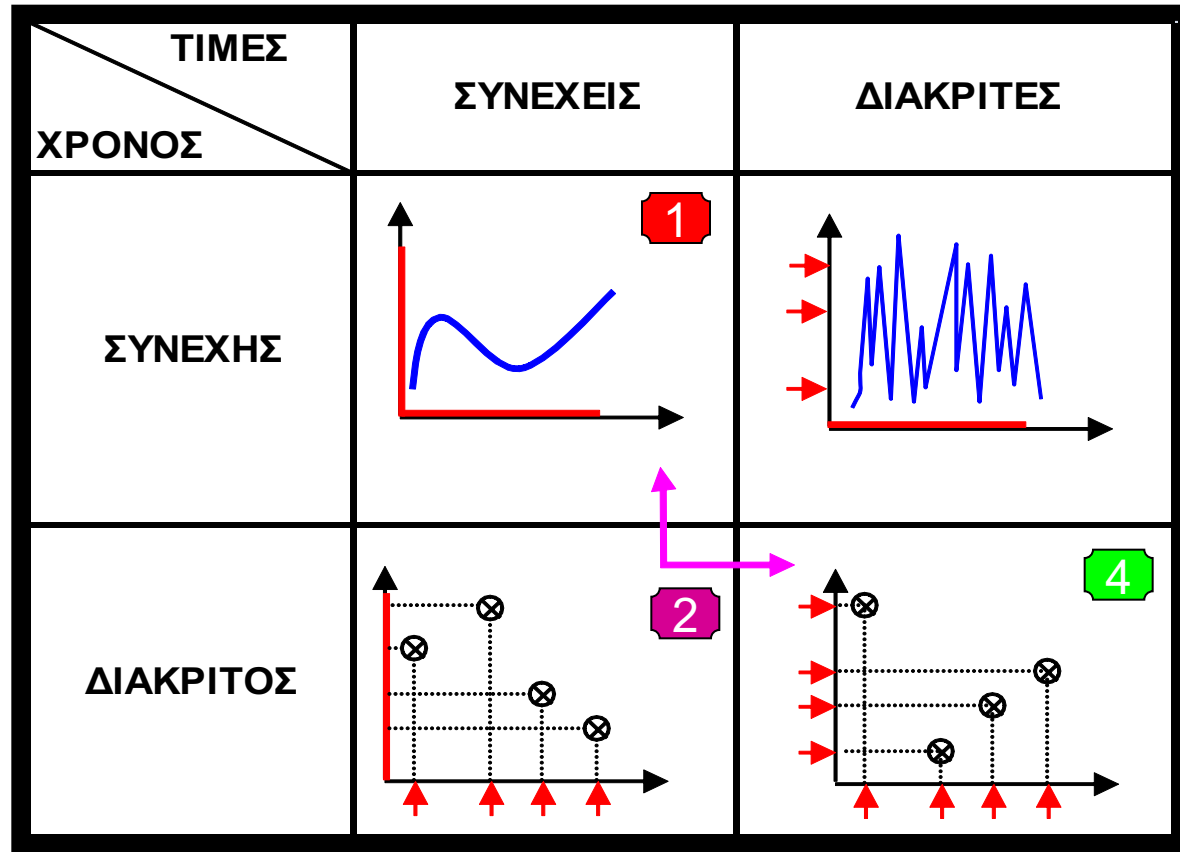
**Κωδικοποιεί, Επεξεργάζεται και Αποκωδικοποιεί
σε διακριτό χρόνο με διακριτές τιμές**

**(πχ 2 Mpixel, 17.6M αποχρώσεις χρωμάτων, CD:
44.1KHz/16 bit)**

Έτσι εκφράζονται μεγέθη ανά τακτά χρονικά διαστήματα (π.χ. πληθωρισμός, ισολογισμοί εταιριών κλπ)

ADC & DAC

- Η μετάβαση από το πεδίο **συνεχούς χρόνου – συνεχών τιμών (1)** που συναντούμε σε πλείστα φυσικά συστήματα στο πεδίο **διακριτού χρόνου – διακριτών τιμών (4)** που διαχειρίζεται ο Η/Υ, γίνεται μέσω μιας διάταξης **Αναлого-ψηφιακού Μετατροπέα (Analog to Digital Converter – ADC)**,
- Η αντίστροφη μετάβαση, γίνεται μέσω του **Ψηφιο-αναλογικού Μετατροπέα (Digital to Analog Converter – DAC)**
- Κατά τις δύο παραπάνω μετατροπές χρησιμοποιείται ως ενδιάμεσο στάδιο το πεδίο **(2)**. Κατά την μετατροπή **(1) → (2) → (4)** ένα μέρος της πληροφορίας χάνεται οριστικά και δεν μπορεί να ανακτηθεί κατά την αντίστροφη μετατροπή.



ADC & DAC

- **Διάστημα δειγματοληψίας (sampling interval) T_s** είναι το χρονικό διάστημα που μεσολαβεί ανάμεσα σε δύο διακριτές χρονικές στιγμές που «δύναται» το υπολογιστικό σύστημα να καταγράψει μια συνάρτηση $f(t)$. Σχετίζεται με τη **συχνότητα δειγματοληψίας (sampling rate) f_s** με την σχέση $f_s = 1/T_s$.
- Ο αριθμός bits που διατίθενται για την παράσταση της $f(t)$ εντός του Η/Υ κάνει φανερό ότι αυτή μπορεί να περιγραφεί προσεγγιστικά μόνο από πεπερασμένο αριθμός διακριτών επιπέδων. Αυτή η προσέγγιση της $f(t)$ παρίσταται ως $f_q(t)$. Η μεταβολή μεταξύ δύο διαδοχικών διακριτών επιπέδων ονομάζεται **βήμα κβαντισμού (quantization level) q** .

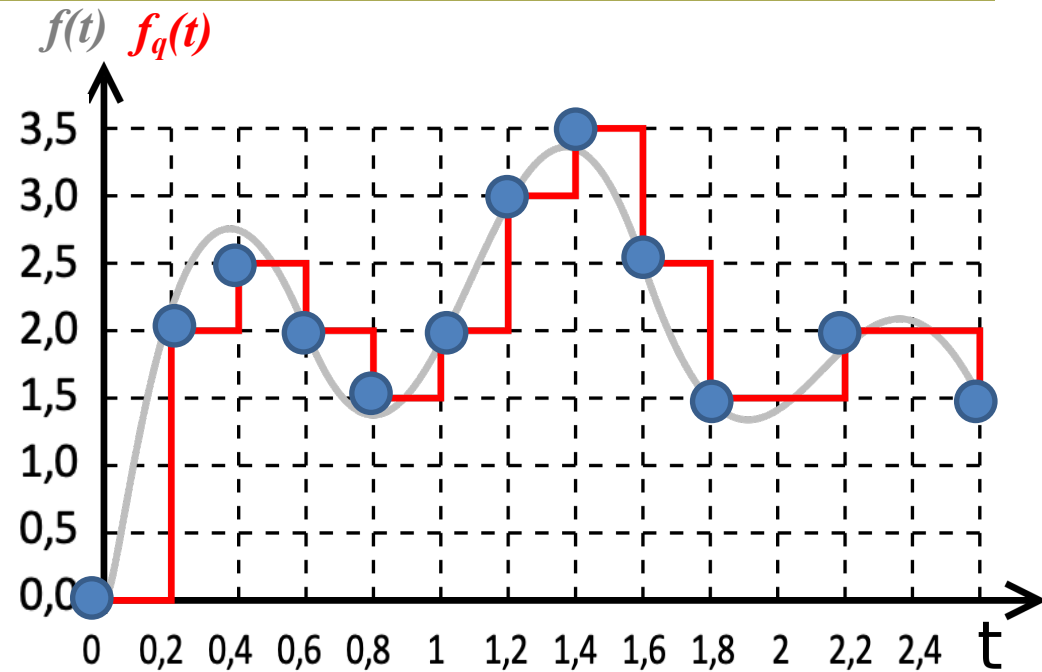
Από τον **συνεχή χρόνο & συνεχείς τιμές** στο **διακριτό χρόνο & διακριτές τιμές**

- $f(t)$: μορφή (1).
- $f_q(t)$: μορφή (3).

Βήμα Κβαντισμού

$q=0.5$

Χρόνοι Δειγματοληψίας	$f(t)$	$f_q(t)$
0.0	0	0.0
0.2	2.10	2.0
0.4	2.70	2.5
0.6	2.05	2.0
0.8	1.37	1.5
1.0	1.97	2.0
1.2	1.70	3.0
...



Διάστημα Δειγματοληψίας $T_s=0.2 \text{ sec}$
 Συχνότητα $f_s=1/T_s = 5\text{Hz}$

Από τον **συνεχή χρόνο & συνεχείς τιμές** στο **διακριτό χρόνο & διακριτές τιμές**

Έστω 3 bit ADC, 5Hz

2δικό: 000 → 111

Min/Max?

→ Min:0 Max:7(=2ⁿ-1)

Πόσα Διαστήματα?

→ 7 (=2ⁿ-1)

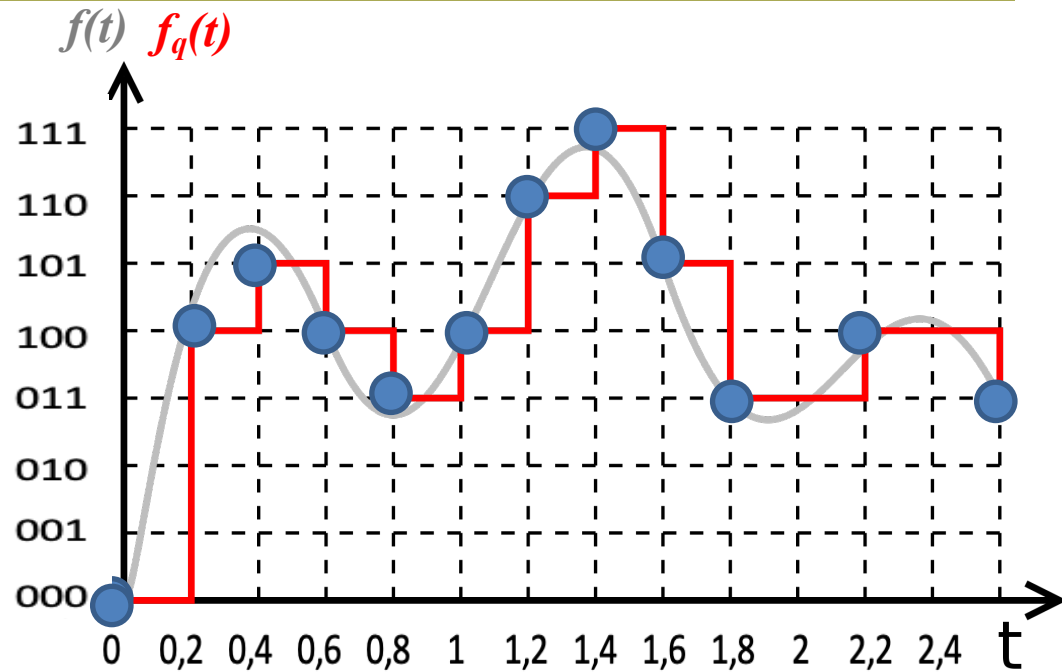
Φυσικό μέγεθος:

- Min: 0 σε Max:3,5. Μέγεθος: 3,5

Βήμα Κβαντισμού?

$q = \text{Μέγεθος} / \text{Διαστήματα}$

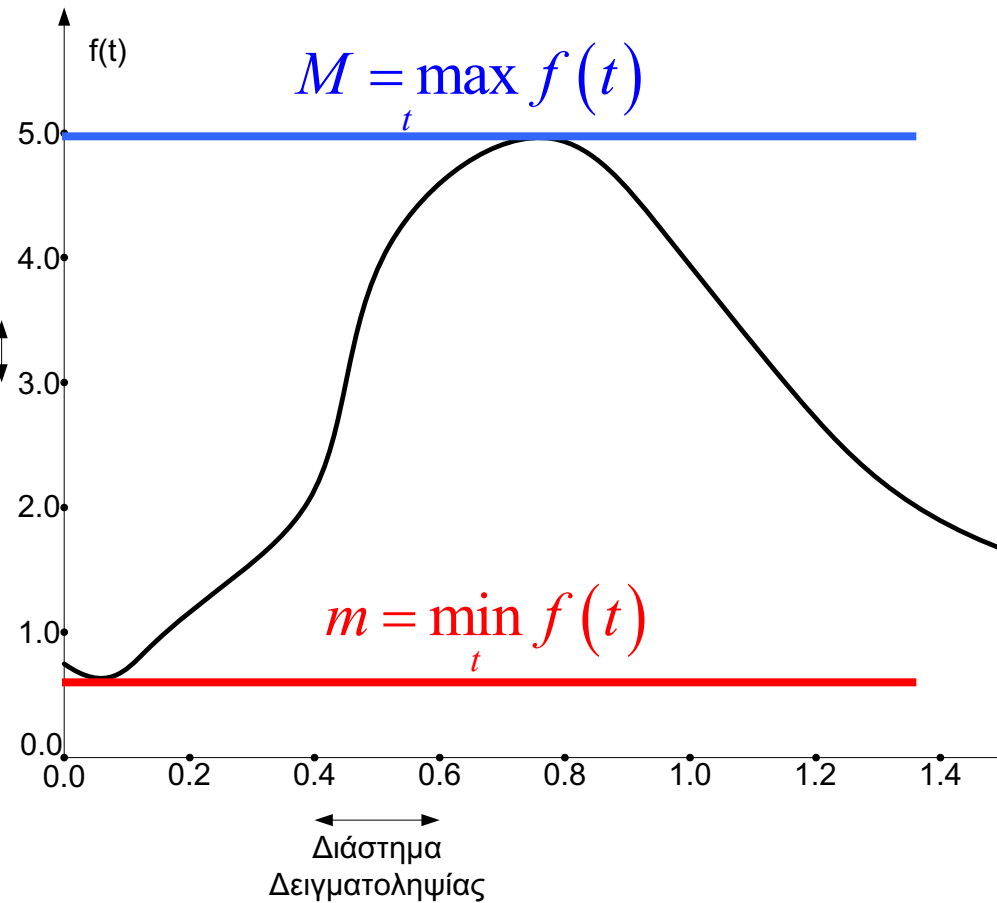
$3,5 / 7 = 0.5$



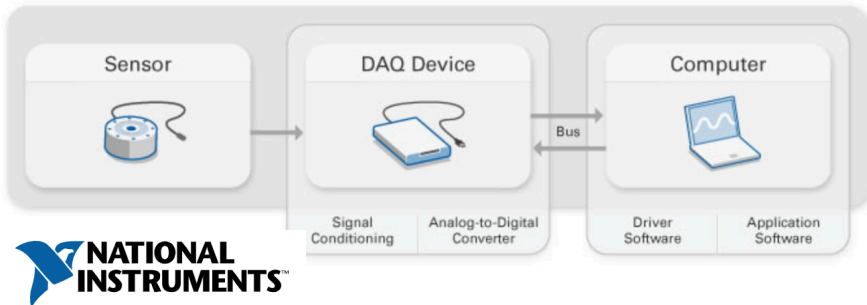
Ποιο θα είναι το βήμα κβαντισμού αν n=4 bit?
 $q = 3,5 / 15 = 0.23333$ (σχεδόν το μισό)

Από τον συνεχή χρόνο & συνεχείς τιμές στο διακριτό χρόνο & διακριτές τιμές

- Αναλογική Συνάρτηση $f(t)$
- Σε έναν ADC n -bit :
 - Αντιπροσωπεύονται οι 2^n αριθμοί:
 - 0.....0 (n φορές το “0”)
 - 0...01
 - ...
 - 1...10
 - 1.....1 (n φορές το “1”)
 - Δηλαδή $2^n - 1$ διαστήματα
 - Για να έχουμε καλή κωδικοποίηση αντιστοιχούμε το m στο 0...0 και το M στο 1...1. Άρα το βήμα κβαντισμού είναι: $q = \frac{M - m}{2^n - 1}$

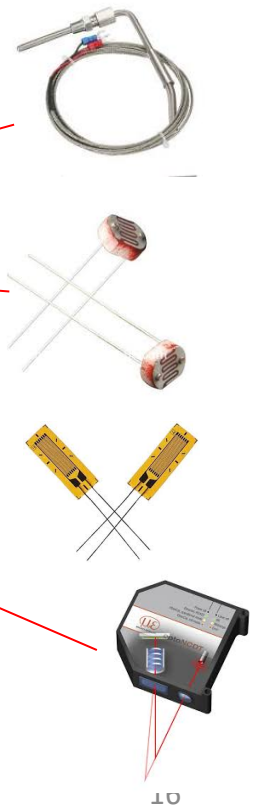


Αναлого-ψηφιακή Μετατροπή Αναλογικού Σήματος: Αισθητήρας → ADC → Η/Υ



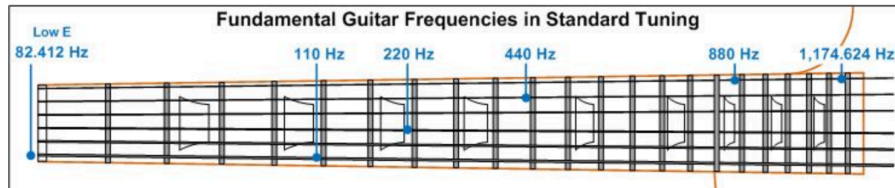
Αισθητήρες: μετατρέπουν τα φυσικά μεγέθη σε αναλογικά σήματα (δηλ. συνεχείς τάσεις μικρής ισχύος)

ΑΙΣΘΗΤΗΡΑΣ	ΦΑΙΝΟΜΕΝΟ
Thermocouple	Temperature
Photo Sensor	Light
Microphone	Sound
Strain Gage, Piezoelectric Transducer	Force and Pressure
Potentiometer, LVDT, Optical Encoder	Position and Displacement
Accelerometer	Acceleration
pH Electrode	pH
CMOS / CCD	Image

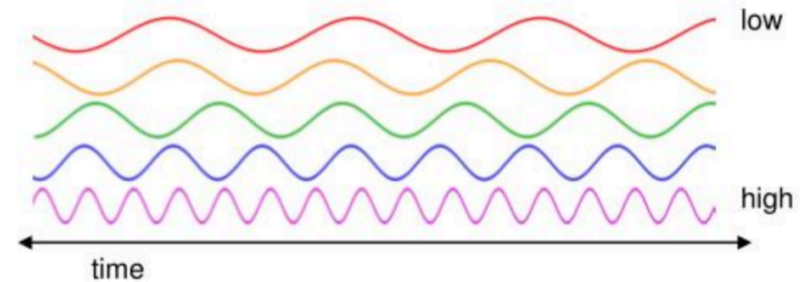
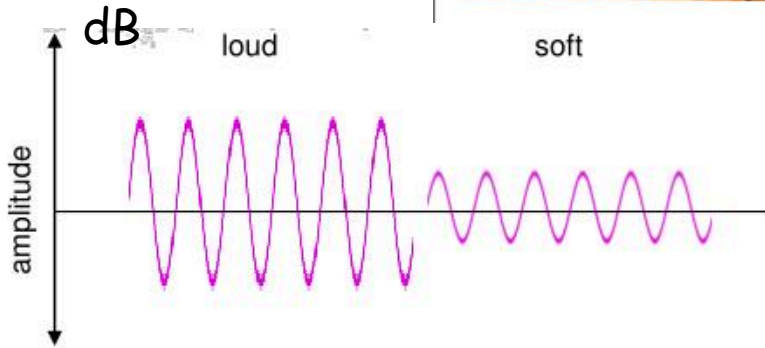


Παραδείγματα Ψηφιακής Επεξεργασίας Σήματος: Ηχος

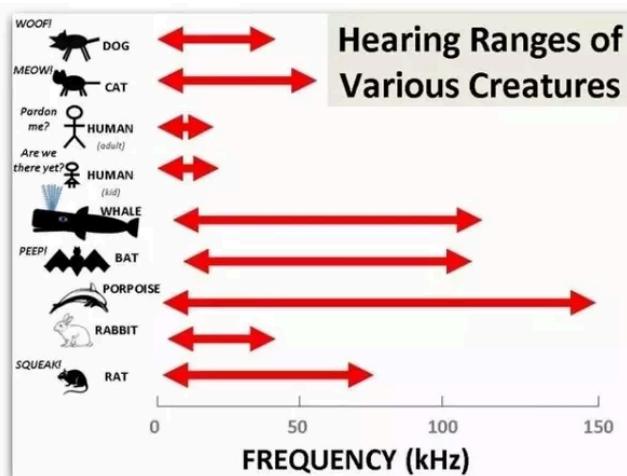
ένταση ήχου



Συχνότητα

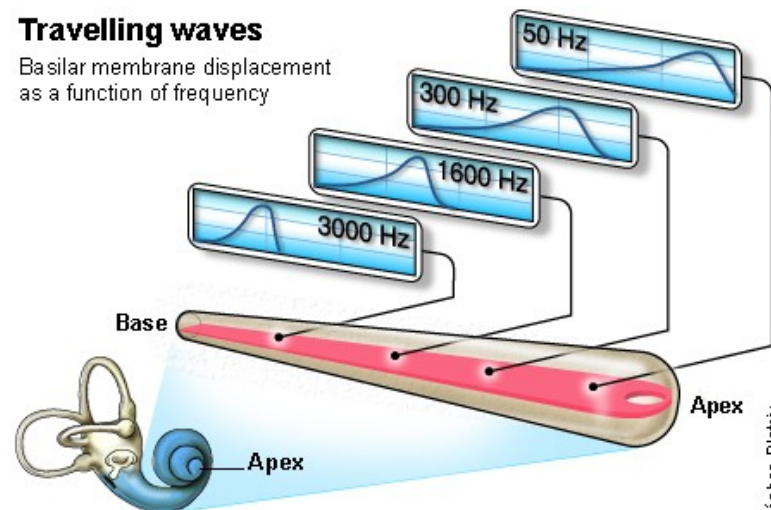


Ανθρώπινο αυτί: 0-->20kHz



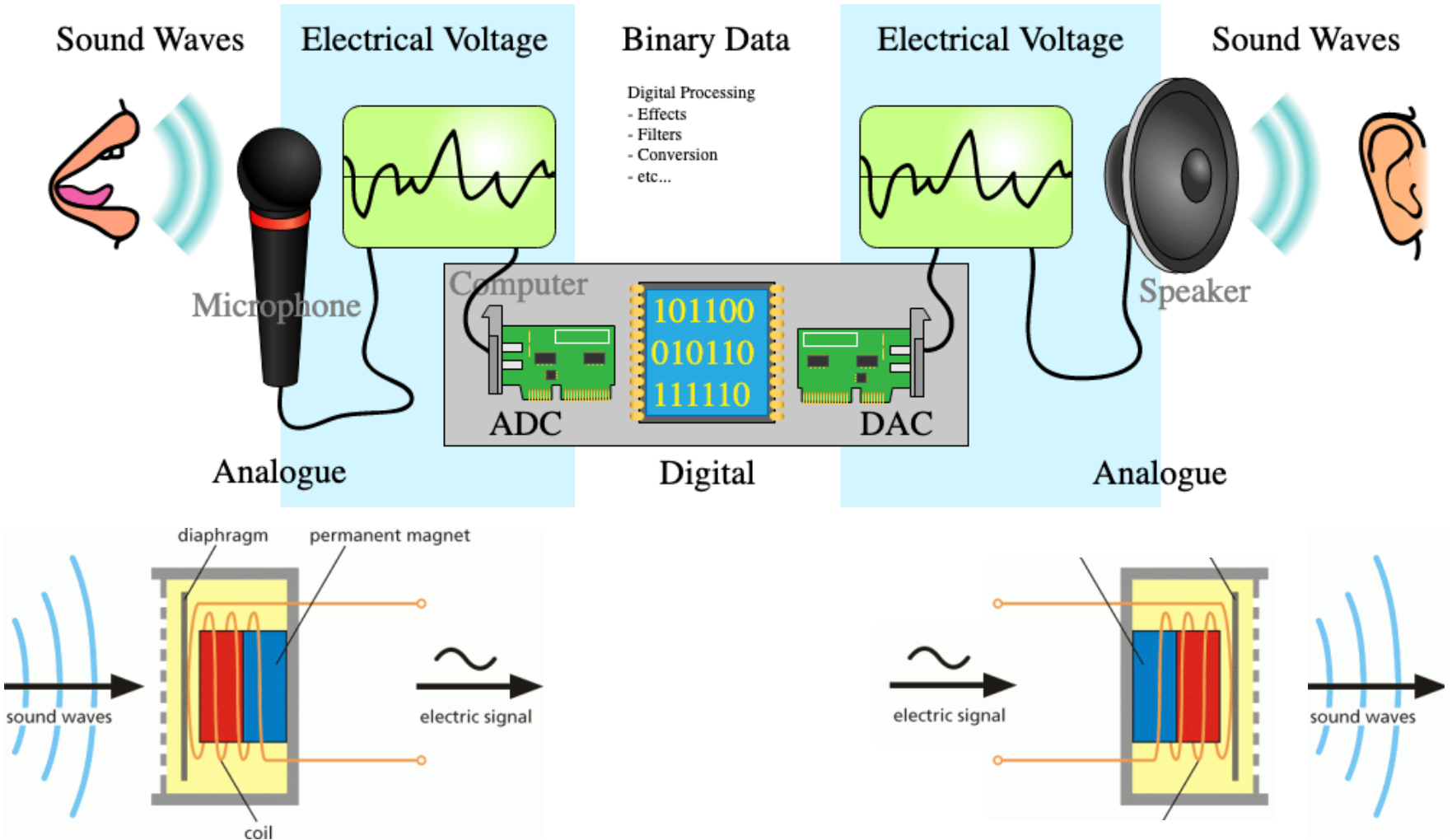
Travelling waves

Basilar membrane displacement as a function of frequency



Παραδείγματα Ψηφιακής Επεξεργασίας Σήματος: Ηχος

<https://www.bbc.co.uk/bitesize/guides/z7vc7ty/revision/1>



Παραδείγματα Ψηφιακής Επεξεργασίας Σήματος: Ηχος

<https://www.bbc.co.uk/bitesize/guides/z7vc7ty/revision/1>

Παράδειγμα: CD sound Audio Resolution (uncompressed)

Sampling frequency (sampling rate): 44.1 kHz (εύρος ζώνης των ήχων που γίνονται αντιληπτοί από το ανθρώπινο αυτί είναι από τα 0 KHz έως και τα ~20 KHz)

Bit depth (quantization): 16 bit per sample (0 to 65,535) vs DVD 24 bits 0 to 16,777,216 (quantization of sound pressure level: decibel)

Audio bitrate (απαιτούμενη ταχύτητα δεδομένων για αποστολή, αποθήκευση κτλ)
Frequency × bit depth × channels = bit rate

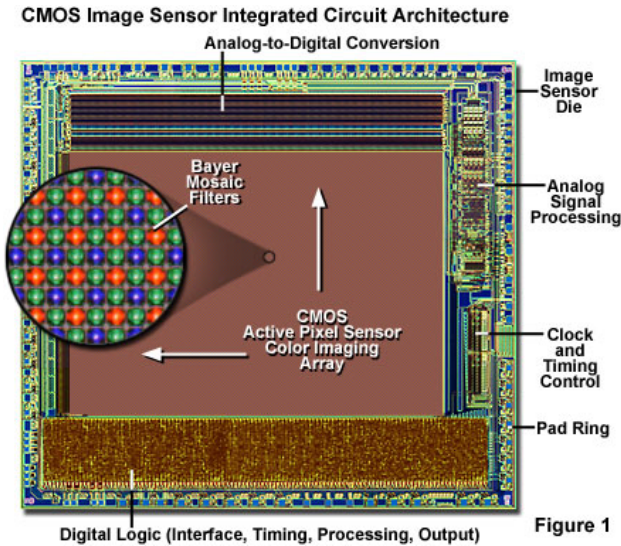
CD bitrate: 44,100 samples per second × 16 bits per sample × 2 channels (stereo) = 1,411,200 bits per second (or 1.2 Mbps) Mega bits / sec

Άρα ένα 4λεπτο (240 second) τραγούδι σε αυτό το bit rate θα δημιουργήσει αρχείο μεγέθους:
 $14,411,200 \times 240 = 3,458,688,000 \text{ bits} = 3,458,688,000 / 8 \text{ bytes} = 428,586,000 \text{ bytes}$
Or $428,586,000 / 2^{20} = 40.37 \text{ megabytes}$

Άρα κάποια μορφή συμπίεσης (.i.e. MP3, MPEG) είναι απαραίτητη.

Παραδείγματα Ψηφιακής Επεξεργασίας Σήματος: Εικόνα

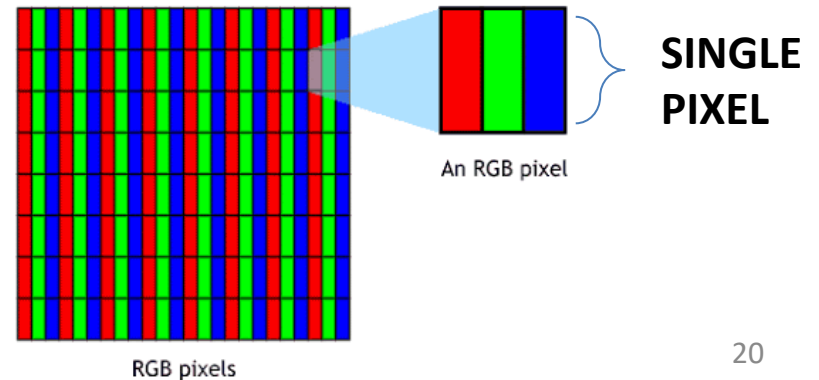
ΕΙΣΟΔΟΣ (πχ κάμερα, CMOS)



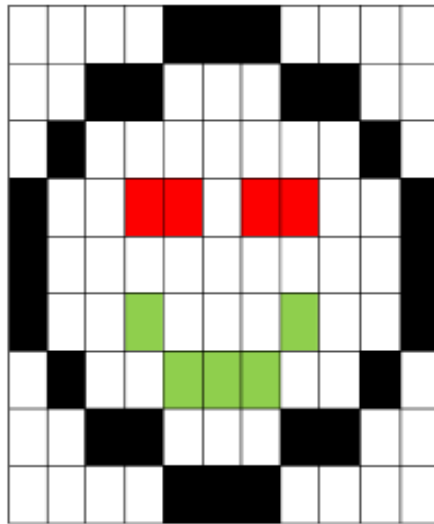
Bayer Color Filter Mosaic Array



ΕΞΟΔΟΣ (πχ LED οθόνη)



Παραδείγματα Ψηφιακής Επεξεργασίας Σήματος: Εικόνα RGB



(α)

	1	2	3	4	5	6	7	8	9	10	11
1	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255
2	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255
3	255, 255, 255	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0	255, 255, 255
4	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0
5	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0
6	0, 0, 0	255, 255, 255	255, 255, 255	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0	255, 255, 255	255, 255, 255	0, 0, 0
7	255, 255, 255	0, 0, 0	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255	0, 0, 0	255, 255, 255
8	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255
9	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255	0, 0, 0	0, 0, 0	0, 0, 0	255, 255, 255	255, 255, 255	255, 255, 255	255, 255, 255

(β)

Σχήμα 2.16. (α) Διαμερισμός pixels με βάση το χρωματικό μοντέλο RGB εικόνας 9x11 – (β) αναπαράσταση με την ψηφιογραφική μέθοδο – μέγεθος 2376 bits.

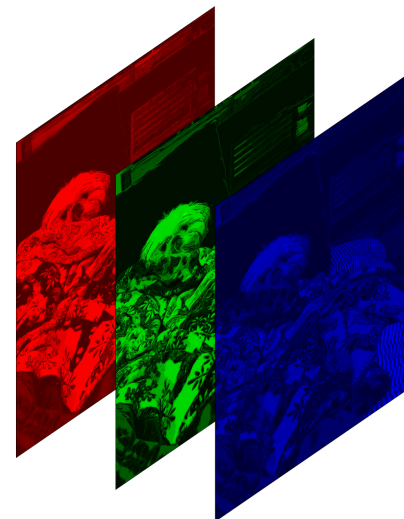
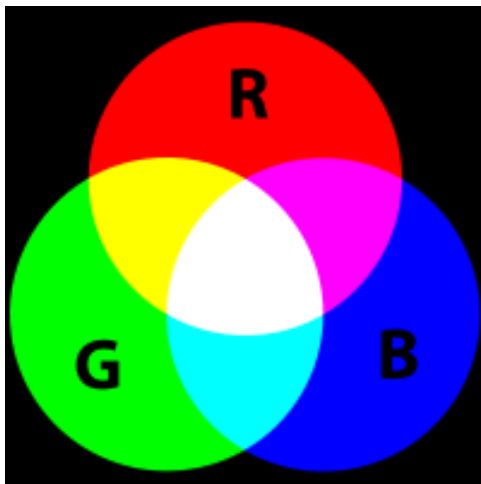
Παραδείγματα Ψηφιακής Επεξεργασίας Σήματος: Εικόνα RGB

Χρώμα	Κόκκινο	Πράσινο	Μπλε	Χρώμα	Κόκκινο	Πράσινο	Μπλε
Μαύρο	0	0	0	Κίτρινο	255	255	0
Κόκκινο	255	0	0	Κυανό	0	255	255
Πράσινο	0	255	0	Ματζέντα	255	0	255
Μπλε	0	0	255	Λευκό	255	255	255



256 (8bit) αποχρώσεις RED
256 (8bit) αποχρώσεις GREEN
256 (8bit) αποχρώσεις BLUE

8+8+8= 24-bit image:
16,7M colors



END
