

FONDAMENTI DI INFORMATICA

Alessandro Renda

Dipartimento di Ingegneria e Architettura, Università degli Studi di Trieste

CENNI SULLA RAPPRESENTAZIONE DI ALTRE INFORMAZIONI

Anno Accademico 2024/2025

Rappresentazione di testi

Rappresentazione di informazione testuale

- C'è bisogno di uno *standard*, per agevolare lo scambio di informazioni testuali
- Codice tradizionale per la rappresentazione di caratteri alfanumerici:
 - ASCII: American Standard Code for Information Interchange
 - In italiano si pronuncia *aschi*
- 7 bit per ciascun carattere (originariamente). Alcuni esempi:

Carattere tastiera	Codice ASCII	Equivalente intero
(0101000	40
R	1010010	82
r	1110010	114
8	0111000	57
9	0111001	63

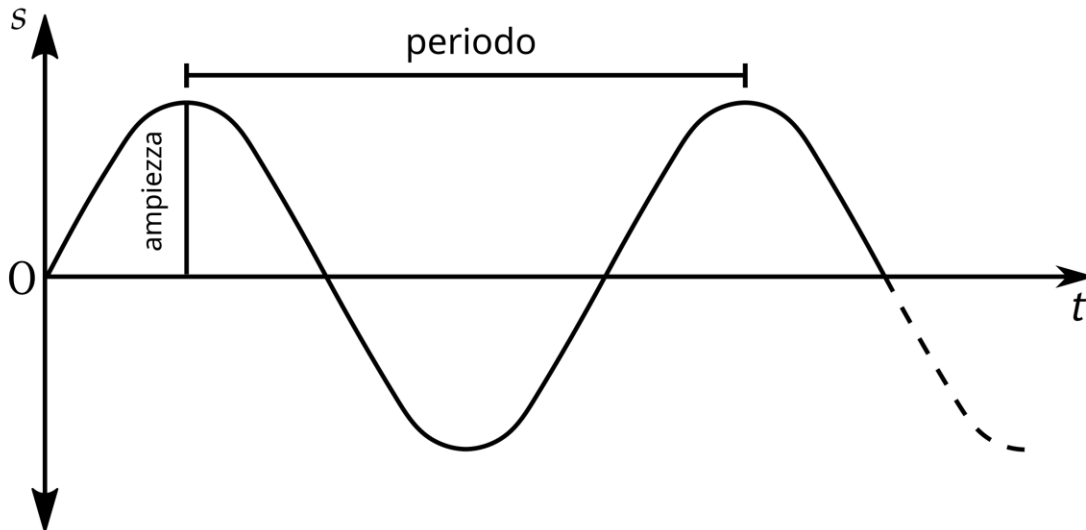
Rappresentazione di informazione testuale

Decimal	Binary	Octal	Hex	ASCII	Decimal	Binary	Octal	Hex	ASCII	Decimal	Binary	Octal	Hex	ASCII	Decimal	Binary	Octal	Hex	ASCII
0	00000000	000	00	NUL	32	00100000	040	20	SP	64	01000000	100	40	@	96	01100000	140	60	`
1	00000001	001	01	SOH	33	00100001	041	21	!	65	01000001	101	41	A	97	01100001	141	61	a
2	00000010	002	02	STX	34	00100010	042	22	"	66	01000010	102	42	B	98	01100010	142	62	b
3	00000011	003	03	ETX	35	00100011	043	23	#	67	01000011	103	43	C	99	01100011	143	63	c
4	00000100	004	04	EOT	36	00100100	044	24	\$	68	01000100	104	44	D	100	01100100	144	64	d
5	00000101	005	05	ENQ	37	00100101	045	25	%	69	01000101	105	45	E	101	01100101	145	65	e
6	00000110	006	06	ACK	38	00100110	046	26	&	70	01000110	106	46	F	102	01100110	146	66	f
7	00000111	007	07	BEL	39	00100111	047	27	'	71	01000111	107	47	G	103	01100111	147	67	g
8	00001000	010	08	BS	40	00101000	050	28	(72	01001000	110	48	H	104	01101000	150	68	h
9	00001001	011	09	HT	41	00101001	051	29)	73	01001001	111	49	I	105	01101001	151	69	i
10	00001010	012	0A	LF	42	00101010	052	2A	*	74	01001010	112	4A	J	106	01101010	152	6A	j
11	00001011	013	0B	VT	43	00101011	053	2B	+	75	01001011	113	4B	K	107	01101011	153	6B	k
12	00001100	014	0C	FF	44	00101100	054	2C	,	76	01001100	114	4C	L	108	01101100	154	6C	l
13	00001101	015	0D	CR	45	00101101	055	2D	-	77	01001101	115	4D	M	109	01101101	155	6D	m
14	00001110	016	0E	SO	46	00101110	056	2E	.	78	01001110	116	4E	N	110	01101110	156	6E	n
15	00001111	017	0F	SI	47	00101111	057	2F	/	79	01001111	117	4F	O	111	01101111	157	6F	o
16	00010000	020	10	DLE	48	00110000	060	30	0	80	01010000	120	50	P	112	01110000	160	70	p
17	00010001	021	11	DC1	49	00110001	061	31	1	81	01010001	121	51	Q	113	01110001	161	71	q
18	00010010	022	12	DC2	50	00110010	062	32	2	82	01010010	122	52	R	114	01110010	162	72	r
19	00010011	023	13	DC3	51	00110011	063	33	3	83	01010011	123	53	S	115	01110011	163	73	s
20	00010100	024	14	DC4	52	00110100	064	34	4	84	01010100	124	54	T	116	01110100	164	74	t
21	00010101	025	15	NAK	53	00110101	065	35	5	85	01010101	125	55	U	117	01110101	165	75	u
22	00010110	026	16	SYN	54	00110110	066	36	6	86	01010110	126	56	V	118	01110110	166	76	v
23	00010111	027	17	ETB	55	00110111	067	37	7	87	01010111	127	57	W	119	01110111	167	77	w
24	00011000	030	18	CAN	56	00111000	070	38	8	88	01011000	130	58	X	120	01111000	170	78	x
25	00011001	031	19	EM	57	00111001	071	39	9	89	01011001	131	59	Y	121	01111001	171	79	y
26	00011010	032	1A	SUB	58	00111010	072	3A	:	90	01011010	132	5A	Z	122	01111010	172	7A	z
27	00011011	033	1B	ESC	59	00111011	073	3B	;	91	01011011	133	5B	[123	01111011	173	7B	{
28	00011100	034	1C	FS	60	00111100	074	3C	<	92	01011100	134	5C	\	124	01111100	174	7C	
29	00011101	035	1D	GS	61	00111101	075	3D	=	93	01011101	135	5D]	125	01111101	175	7D	}
30	00011110	036	1E	RS	62	00111110	076	3E	>	94	01011110	136	5E	^	126	01111110	176	7E	~
31	00011111	037	1F	US	63	00111111	077	3F	?	95	01011111	137	5F	_	127	01111111	177	7F	DEL

Rappresentazione di segnali audio

Rappresentazione di audio

- Il suono è un insieme di onde meccaniche che si propagano in un mezzo materiale
- Il suono è un'informazione di tipo *analogico*: assume valori in intervallo continuo
- Una nota, ad esempio, è una forma d'onda sinusoidale



https://it.wikipedia.org/wiki/Onda_sonora

Ampiezza: intervallo compreso tra massimo e minimo
Periodo: tempo per completare un ciclo
Frequenza: numero di cicli per unità di tempo

Rappresentazione di audio: digitalizzazione

- *Quantizzazione*: necessario discretizzare i valori continui

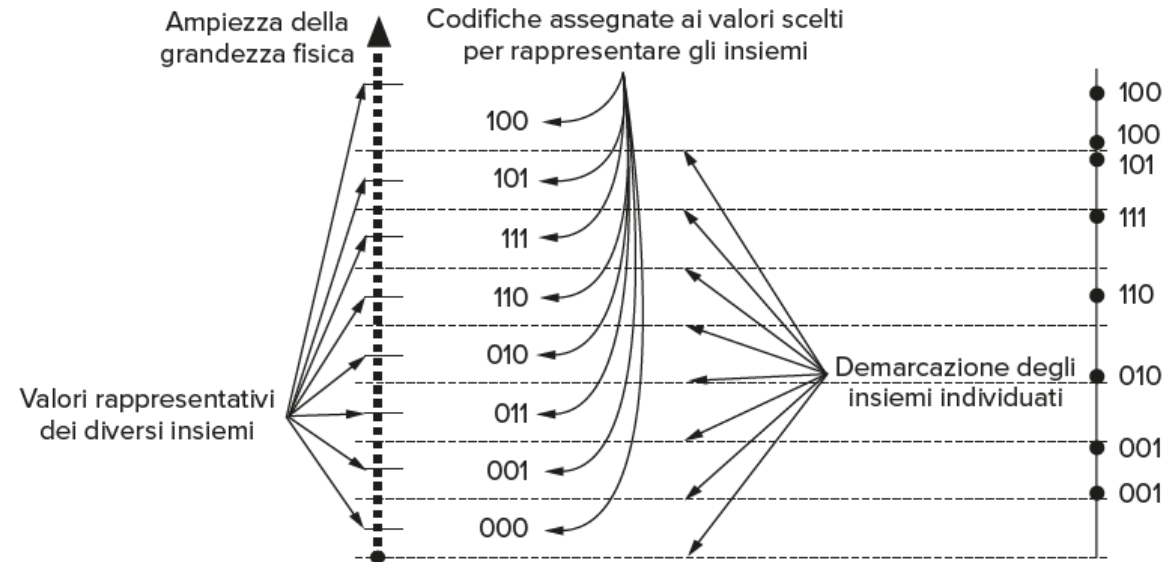


Figura 2.16 Esempio di quantizzazione dell'informazione associata a una grandezza fisica. L'insieme delle possibili entità di informazione viene suddiviso in un numero finito di sottoinsiemi, a ognuno dei quali viene associata una codifica digitale. In questo caso è prevista una quantizzazione su 8 livelli, dunque a 3 bit. Sulla destra è riportata la codifica di alcuni valori; si noti come valori diversi possano essere codificati nello stesso modo.

Rappresentazione di audio: digitalizzazione

- *Campionamento*: necessario codificare grandezze che variano nel tempo

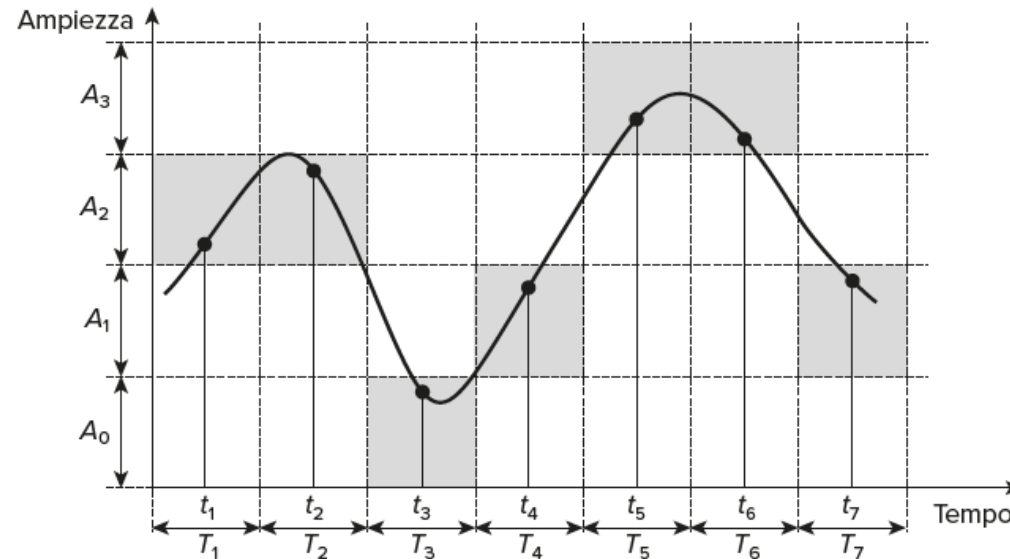
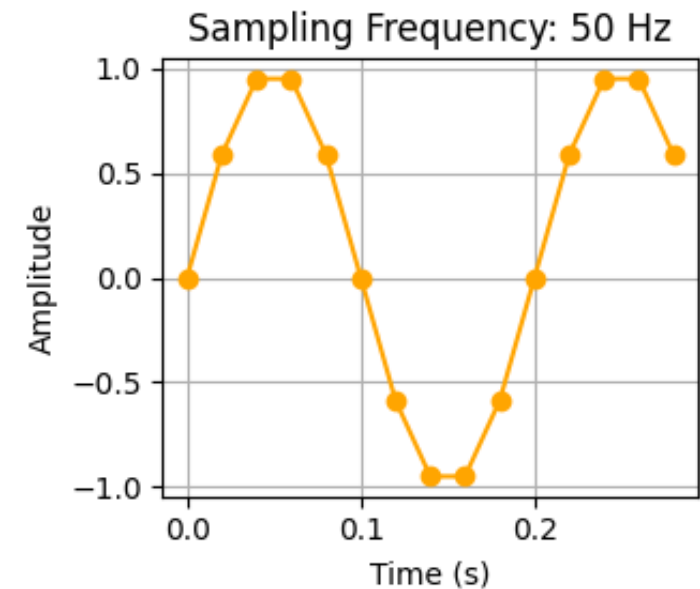
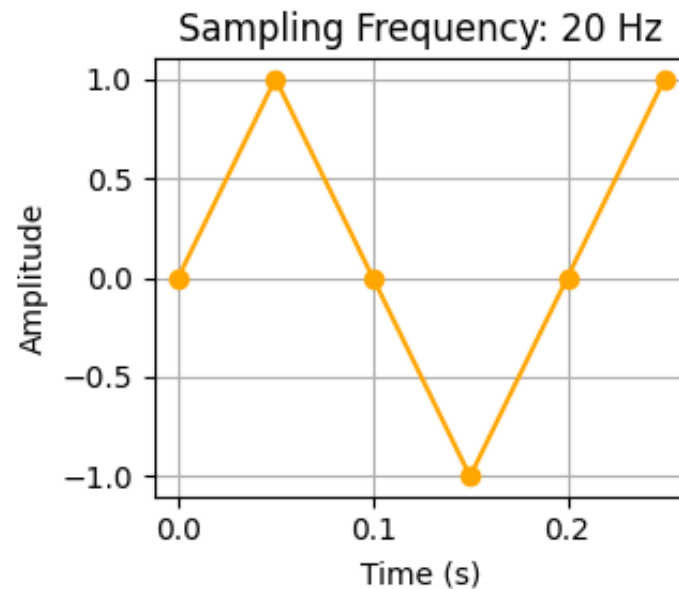
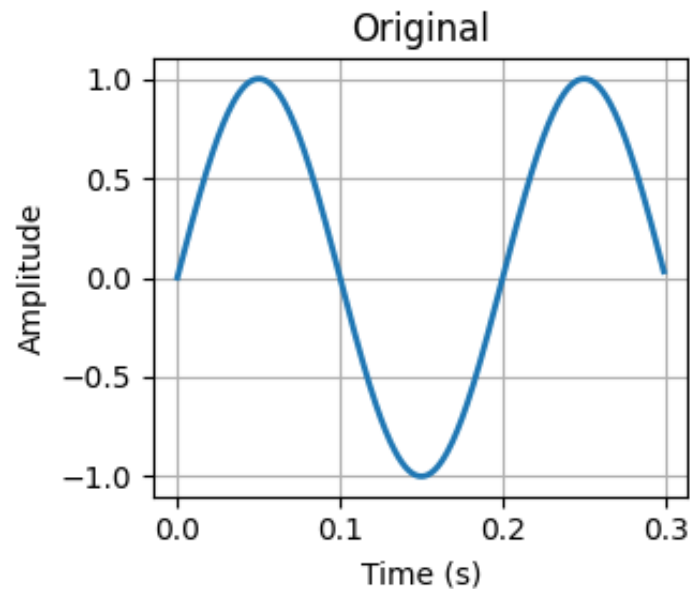


Figura 2.17 Andamento di una grandezza fisica nel tempo, campionamento e successiva quantizzazione. L'intervallo di tempo considerato è suddiviso nei sottointervalli di campionamento T_i e per ciascuno di essi viene scelto l'istante di campionamento t_i in cui il valore della grandezza viene rilevato. In questo caso la quantizzazione è su quattro livelli, quindi il risultato del campionamento, espresso dalla successione dei livelli associati ai campioni, è $A_2A_2A_0A_1A_3A_3A_1$ che corrispondono ai riquadri in grigio nella figura (se i livelli fossero codificati in binario, la successione potrebbe essere, per esempio, 10 10 00 01 11 11 01).

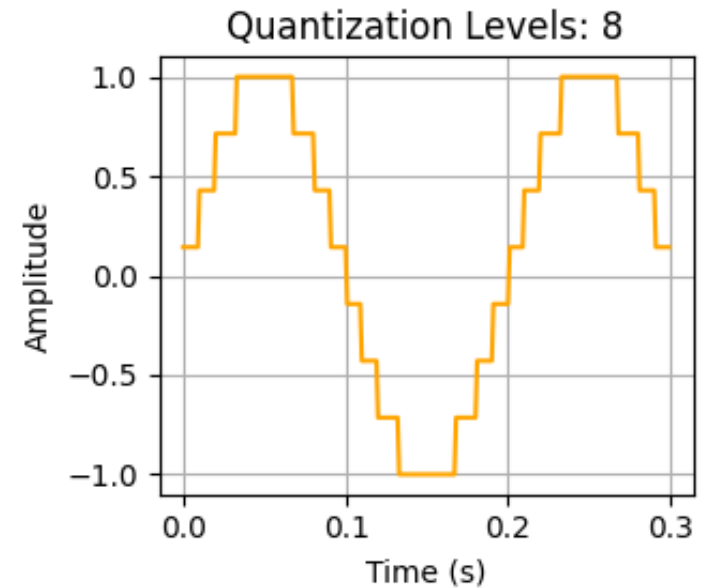
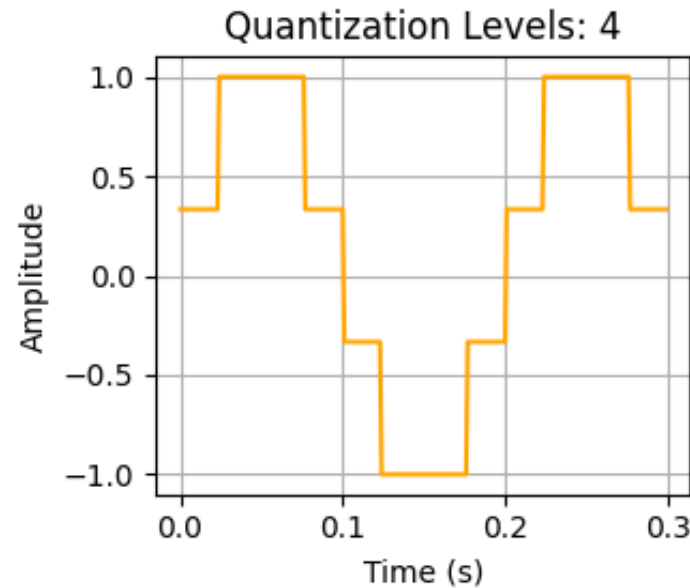
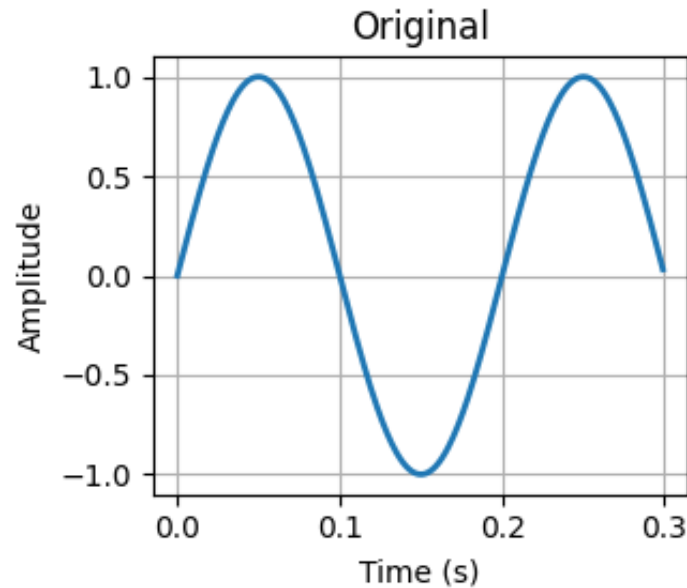
Rappresentazione di audio

- La precisione con cui il suono può essere riprodotto dipende da due parametri
 - Frequenza di campionamento



Rappresentazione di audio

- La precisione con cui il suono può essere riprodotto dipende da due parametri
 - Risoluzione in bit



Rappresentazione di audio

- Le piattaforme streaming audio usano
 - Frequenza di campionamento: 44100 Hz
 - https://it.wikipedia.org/wiki/Frequenza_di_campionamento#Audio
 - Risoluzione: 16 bit (65536 livelli) o 24 bit (16M livelli)
 - https://it.wikipedia.org/wiki/Profondità_di_bit

Rappresentazione di immagini e video

Rappresentazione di immagini

- Anche le immagini sono informazioni di tipo analogico
- È necessaria una digitalizzazione
- L'unità minima convenzionale di un'immagine digitale è detta *pixel*, deriva dall'inglese *picture elements*

Rappresentazione di immagini



Quante informazioni vengono memorizzate per ciascun pixel?
Dipende.

Rappresentazione di immagini

- Ad esempio:

Binary Image



1 bit

Grayscale Image



8 bit = 1 byte





Original Image



24 bit = 3 byte

Rappresentazione di immagini

- Grafica *raster* (o *bitmap*): rappresentazione per cui ciascun pixel viene codificato come valore binario senza segno, associato ad uno specifico colore
- Per le immagini a colori, uno schema di codifica comune è **RGB**
 - dall'inglese *Red, Green, Blue*
 - solitamente, 8 bit per ciascun colore

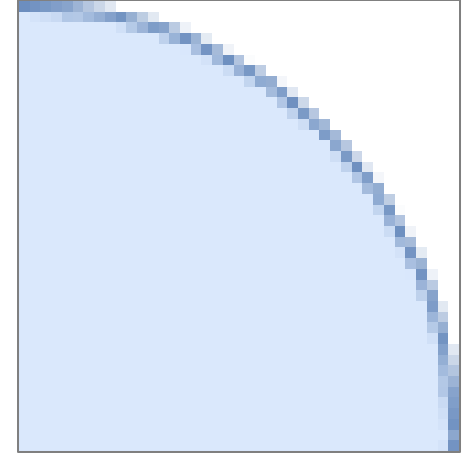
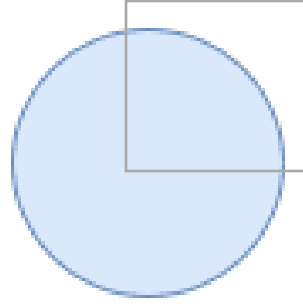
R	G	B	Hex	Colore	
255	0	0	0 × <i>FF0000</i>	Rosso	
0	255	0	0 × <i>00FF00</i>	Verde	
0	0	255	0 × <i>0000FF</i>	Blu	
255	255	0	0 × <i>FFFF00</i>	Giallo	

Rappresentazione di immagini

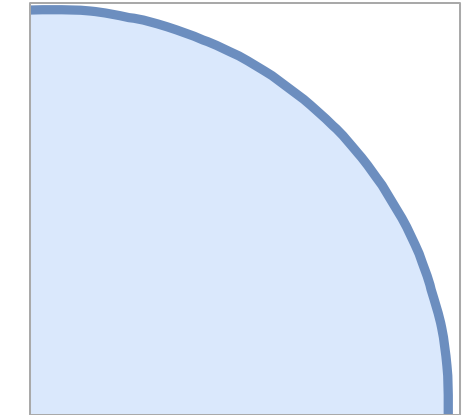
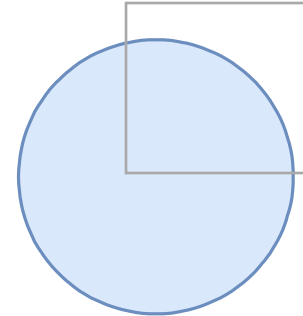
- *Grafica vettoriale*: rappresentazione per cui un'immagine è un insieme di primitive geometriche (punti, linee, curve parametriche)
 - Le forme vengono modellate tramite calcoli matematici
 - Possono essere ingrandite o ridotte senza perdita di informazione
 - Occupa generalmente meno spazio su disco rispetto alla grafica raster
- Formati di file comuni:
 - PDF, SVG, EPS

Rappresentazione di immagini

- Grafica raster



- Grafica vettoriale



Rappresentazione di video

- Sequenza di immagini, dette «frame»
- Frequenza di campionamento espressa in *frame per second* (FPS) o *frame rate*
 - PAL: 25 fps
 - NTSC (US): 29,97 fps
- In ogni caso sufficiente a far percepire la scena con la stessa sensazione di movimento che l'essere umano percepisce guardando la realtà

Compressione dei dati

Compressione dei dati

- Numero in *single precision*: 32 bit
- Numero in *double precision*: 64 bit
- Carattere in formato *ASCII*: 8 bit
- Un campione di una traccia audio: 8 bit
- Un pixel di un immagine codificata RGB: 24 bit

Compressione dei dati

- Un romanzo di 300 pagine (ipotizzando codifica ASCII)
 - In media, circa 300 parole per pagina: $\sim 100\,000$ parole in tutto
 - In media, circa 5 caratteri per parola: $\sim 500\,000$ caratteri in tutto
 - 8 bit per carattere: $\sim 4\,000\,000$ bit ~ 0.5 MB
- Un minuto di registrazione (ipotizzando $f_s = 44100$ Hz, $D_{\text{bit}} = 16$)
 - $\sim 42\,000\,000$ bit ~ 5 MB
- Una fotografia scattata con uno smartphone (ipotizzando $\sim 4000\text{px} \cdot 3000\text{px}$ e 24 bit per pixel)
 - $\sim 288\,000\,000$ bit ~ 36 MB

Compressione dei dati

- È importante considerare anche il *costo* e l'*efficienza* della codifica

⚠ Spazio di archiviazione quasi esaurito Se esaurisci lo spazio, non puoi creare, modificare e caricare file.

- *Compressione dei dati*: riduzione del numero di bit richiesti per rappresentare un'informazione
- Classificazione generale delle tecniche di compressione
 - *Lossless*: senza perdita di informazione (ad esempio, ZIP, RAR)
 - *Lossy*: con perdita di informazione (ad esempio JPEG)
- Rapporto di compressione: $\frac{\text{dimensione } file \text{ originale}}{\text{dimensione } file \text{ compresso}}$

Compressione dei dati

- In alcune applicazioni, posso accettare di *perdere informazione (lossy compression)* in relazione all'uso che ne devo fare
 - Immagini: JPEG PNG GIF
 - Nell'esempio: 600KB vs 50KB
 - Audio: MP3 (MPEG-I, layer3)
 - Video: MPEG-2, MPEG-4
- Alcune buone ragioni:
 - Limiti occhio/orecchio umano
 - Canali di trasmissione di capacità limitata (ad esempio, servizi di streaming)
 - Limiti del terminale (ad esempio, smartphone)

