Codifica per la compressione

Formati

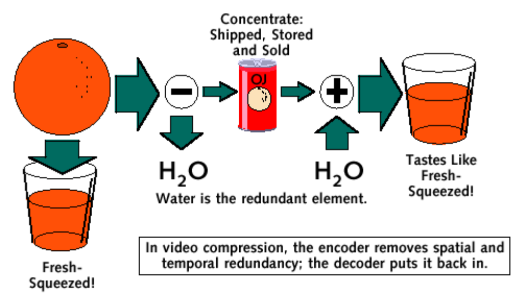
Elenco di alcuni formati tipici di immagini, e quanta memoria occupano:

PAL: 720 x 576	8 bit	24 bit
• QCIF: 160 x 120	19.200	57.600
• CIF: 320 x 240	76.800	230.400
• VGA: 640 x 480	307.200	921.600
• NTSC: 720 x 486	349.920	1.049.760
• Workstation: 1280 x 1024	1.310.720	3.932.160
• HDTV: 1920 x 1080	2.073.600	6.220.800
• 35mm slide: 3072 x 2048	6.291.456	18.874.368

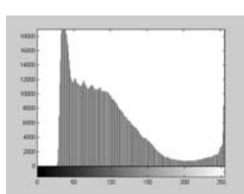
Occupazione su disco

PAL non compresso a colori:

1 fotogramma: 720 x 576 x 3 = 1.244.160   1.2Mb
(Hor x Vert x Rgb)
1 secondo: 1.244.160 x 25 (fps) = 31.104.000   31Mb
1 minuto: 31.104.000 x 60 (sec) = 1.866.240.000   1.9Gb
90 minuti: 1.866.240.000 x 90 (min) = 167.961.600.000
168Gb



Contenuto informativo di una immagine - 1

Una figura naturale con il suo istogramma

Contenuto informativo di una immagine - 2

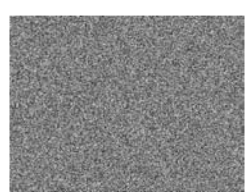
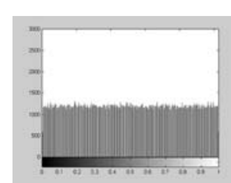
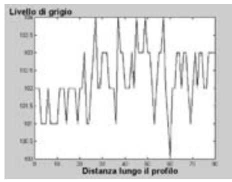
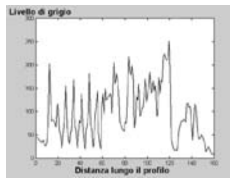



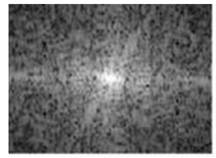
Immagine costituita da punti casuali e relativo istogramma

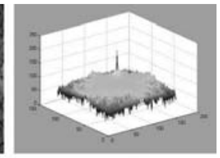
Contenuto informativo di una immagine - 3



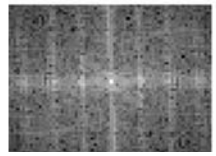


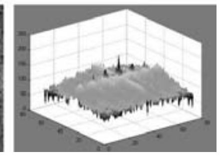
Contenuto informativo di una immagine - 4





cielo





loggiato

Misura del contenuto informativo

- Claude Elwood Shannon: *self-information*: l’evento *A* abbia una probabilità *p(A)* di accadere; la *auto informazione* associata all'accadere dell'evento *A* è:

$$i(A) = \log \frac{1}{p(A)} = -\log(p(A))$$

Misura del contenuto informativo -2

- l'evento è l'uscita di testa o croce nel lancio di una moneta, i due eventi hanno la stessa probabilità = 1/2, il contenuto informativo, dell'evento "testa" è pari a 1.
- Se un evento ha probabilità molto piccola 0,00005, il suo logaritmo è -19,93156857; la misura di informazione di questo evento è quasi pari a 20 >> 1.
- Ciò corrisponde alla nota espressione "un cane che morde un uomo non fa notizia, ma un uomo che morde un cane sì!".

Entropia informativa

- Il messaggio viene considerato come una serie di eventi (che possono essere correlati tra loro o totalmente indipendenti) generati da una sorgente di segnali, e la somma della misura di informazione associata a ciascun simbolo moltiplicata per la rispettiva probabilità permette di definire una quantità, chiamata *entropia*:

$$H = \sum_i p(s_i) i(s_i) = \sum_i p(s_i) \log \frac{1}{p(s_i)} = -\sum_i p(s_i) \log(p(s_i))$$

Codici efficienti

- Dal punto di vista pratico la difficoltà consiste nella precisione che si riesce ad ottenere nella stima dei valori di probabilità di ciascun simbolo presente nel messaggio stesso.

Codici - 1

- Il problema della codifica di una immagine consiste nel trovare il codice di lunghezza minima
- Se un codice ha parole di codice c_1, c_2, .., c_M con lunghezza b_1, b_2, .., b_M allora il numero medio di bit richiesto dal codificatore è:

$$R = \sum_{k=1}^{M} b_k p_k$$

Codici - 2

- se R è vicino ad H (entropia della sorgente) il codice è ottimale;
- il problema della progettazione di codici è quindi quello della ricerca di codici ottimali, per cui cioé la lunghezza media delle parole di codice sia vicino all'entropia della sorgente del segnale.

Codici e ridondanza

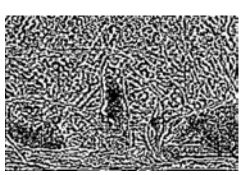
- I metodi di compressione basati su una stima accurata della probabilità dei simboli, e quindi che cercano di rendere minimo il valore dell'entropia, non sfruttano in modo adeguato la ridondanza dell'informazione.
- Sono stati ideati metodi che rinunciano a codificare in modo esatto il segnale, eliminando informazione ridondante e riducendo quindi a priori il contenuto informativo del messaggio e quindi l'entropia della sorgente.

Metodi *lossy* e *non lossy*

- Distinguiamo i metodi di compressione in due grandi categorie:
  - metodi *reversibili* o che non hanno alcuna perdita di informazione (chiamati in gergo *non lossy*),
  - metodi *non reversibili* o con perdita di informazione (chiamati in gergo *lossy*).
- Per i metodi *lossy* occorre un metodo per misurare la quantità di informazione che è stata persa.
- Si usa misurare la differenza tra l'immagine originale e quella ricostruita dopo il procedimento di compressione *lossy*.

Misure

- *i(x,y)*: immagine di input NxN pixel;
- *d(x,y)*: immagine compressa e decompressa, NxN pixel;
- errore: $e(x,y) = d(x,y) - i(x,y)$



Misura dell'errore - 1

La misura di errore più utilizzata è l'errore quadratico medio o la sua radice quadrata è il *root mean square error* RMS.

Errore quadratico medio

$$\overline{e^2} = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} e^2(x,y) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left[d(x,y) - i(x,y)\right]^2$$

RMS

$$e_{RMS} = \sqrt{\left(\overline{e^2}\right)}$$

Misura dell’errore - 2

L'errore - differenza - tra immagine originale e trasformata è *rumore.* Si può misurare l'errore come rapporto segnale disturbo SNR, che è il valore quadratico medio del rapporto segnale disturbo.

$$(SNR)_{ms} = \frac{\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} d^2(x,y)}{\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} e^2(x,y)}$$

Misura dell’errore - 3

rms del rapporto segnale disturbo

$$(SNR)_{rms} = \left[\frac{\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} d^2(x,y)}{\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} [d(x,y)-i(x,y)]^2} \right]^{1/2}$$

SNR sul valore di picco

$$PSNR = 20\ log_{10} \left\{ \frac{[picco]}{e_{rms}} \right\}$$

Misure di compressione

- *bit/pixel:* indica il numero medio di bit con cui ogni pixel dell'immagine è codificato; valori tipici sono compresi tra 0.2 e 1.5;
- *rapporto di compressione:* è il rapporto tra il numero di byte impiegati per codificare l'immagine di input e quelli impiegati per codificare l'immagine compressa.

Codificare senza perdita o con perdita di informazione

- Metodi non lossy
  - RLC;
  - Huffmann;
  - Aritmetico;
  - .....
- Metodi lossy
  - trasformate;
  - a dizionario.

Codifica - 1

L'immagine risulta così trasformata in una matrice di numeri interi.

Per arrivare a una forma compatibile con i sistemi di elaborazione, occorre rappresentare tali numeri in forma codificata acquisibile dai sistemi stessi.

Una forma di codifica può essere a lunghezza variabile, es: il codice Morse

Codifica - 2

Codice naturale binario, la lunghezza di tali parole dipende dal range di valori che assumono le uscite del quantizzatore (es. 8 bit per valori tra 0-255).

Codice di Gray, simile a quello naturale salvo che ogni parola differisce da quella che la precede nell'ordinamento naturale per un bit in una sola posizione.

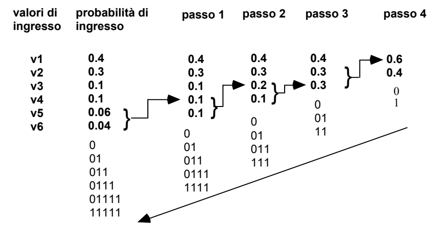
Codici univocamente decodificabili sono quelli per cui una sequenza di parole di codice può essere decodificata in un solo modo. ad es. il codice: c1 = 0; c2 = 1; c3 = 01; c4 = 10 non è univocamente decodificabile, poiché la sequenza 0011 può essere decodificata come c1c1c2c2 oppure come c1c3c2.

Gray code su 3 bits

000 001 011 010 110 111 101 100

Codice compatto: Huffman

Un codice compatto soddisfa la proprietà che la lunghezza media della parola di codice è minore o uguale di quella di ogni altro codice univocamente decodificabile.
Il codice di Huffmann è un codice compatto.
Associa parole di codice di minore lunghezza a stringhe di maggiore frequenza relativa.



Nella prima fase si costruisce la tabella di frequenze, nella seconda fase alle frequenze si associano i bit via via 0 o 1.

Huffmann: esempio - 1

	Simbolo	Probabilità	codice
a	α	0,4	*cod(a)*
b	β	0,2	*cod(b)*
c	c	0,2	*cod(c)*
d	d	0,1	*cod(d)*
e	e	0,1	*cod(e)*

I due simboli *d, e* hanno la probabilità minima e ad essi assegniamo la parola di codice 0 e 1 rispettivamente, che chiamiamo per brevità *cod*:

	Simbolo	Probabilità	codice
a	α	0,4	*cod(a)*
b	β	0,2	*cod(b)*
c	c	0,2	*cod(c)*
d'	d e	0,2	*cod* : 0, 1

Huffmann: esempio - 2

Ripetiamo l'operazione precedente e assegniamo a *c* e *d'* il codice: *cod'*+0 e *cod'*+1 rispettivamente, ma il codice per *d'* già comprendeva la parte *cod,* quindi ora è diventato: *cod'*+10 e *cod'*+11.

	Simbolo	Probabilità	codice
a	α	0,4	*cod(a)*
c'	χ δ ε	0,4	*cod'*+0 *cod'*+10 *cod'*+11
b	b	0,2	*cod(b)*

Assegniamo ora a *c' code"*+0 e a *b code"*+1 e otteniamo la nuova tabella per un alfabeto di due simboli

	Simbolo	Probabilità	codice
b'	β χ δ ε	0,6	*cod"*+1 *cod"*+00 *cod"*+010 *cod"*+011
a	α	0,4	*cod(a)*

Huffmann: esempio - 3

L'ultimo passaggio consiste nell'assegnare il codice ai simboli *b'* e *a*, che essendo quelli a frequenza massima saranno composti da un solo simbolo rispettivamente 0 che assegniamo a *cod"* e 1 che assegniamo a *cod(a)*. Il risultato finale, operando tutte le sostituzioni è

	Simbolo	Probabilità	codice
a	α	0,4	1
b	β	0,2	01
c	c	0,2	000
d	d	0,1	0010
e	e	0,1	0011

Compressione reversibile o senza errore

Un processo di compressione può non essere reversibile, in particolare le trasformazioni di mapping e di quantizzazione sono responsabili di garantire la reversibilità. In generale, il requisito di decodificabilità univoca rende reversibile invece la trasformazione di codifica.

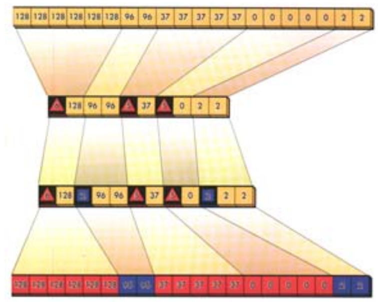
I metodi di compressione reversibili sono anche detti **non lossy**, mentre quelli non reversibili sono detti **lossy**, ovvero senza o con perdita di informazione.

RLC - 1

Il metodo *RLC* di mapping (*Run Lenght Coding*) può essere applicato in modo monodimensionale o bidimensionale.

E’ molto semplice e sfrutta la probabilità che pixel vicini abbiano lo stesso valore per raggrupparli in tratti in cui si registrerà la lunghezza e il valore di colore o di grigio corrispondente.

RLC - 2



Compressione non reversibile

Si possono adottare metodi **lossy** nel ricostruire un’immagine quando l’errore generato dal metodo non influenza in modo negativo il funzionamento dell’applicazione che la usa.