

Compressione delle immagini

1 Concetti di base

I *dati* sono il mezzo con cui è rappresentata l'*informazione*. Diverse quantità di dati possono essere usate per rappresentare la stessa informazione. Le tecniche di compressione si basano sull'idea di eliminare la ridondanza nei dati che codificano un'immagine.

Dal punto di vista matematico, la compressione trasforma un'immagine (che è una matrice di pixel) in un insieme non correlato di elementi (che potrebbe non avere un formato visuale).

Queste tecniche sono importanti per l'archiviazione e (soprattutto) per la trasmissione di immagini e video.

2 Ridondanza relativa

La **ridondanza relativa** tra due sistemi di codifica è la quantità

$$R = 1 - \frac{1}{C}$$

dove:

- $C = \frac{n_1}{n_2}$ è il **rapporto di compressione**;
- n_1 e n_2 sono i numeri di unità di rappresentazione necessari per l'immagine, rispettivamente nel sistema di codifica di partenza e in quello di arrivo.

La ridondanza relativa permette di confrontare in modo quantitativo due sistemi di codifica:

- con $n_2 \gg n_1$, si hanno $C \rightarrow 0$ e $R \rightarrow -\infty$, che indicano un'espansione (passando al sistema di codifica di arrivo, vengono aggiunti dei dati ridondanti);
- con $n_2 \ll n_1$, si hanno $C \rightarrow \infty$ e $R \rightarrow 1$, che corrispondono a una compressione (rimozione di dati ridondanti);
- se $n_2 = n_1$, allora $C = 1$ e $R = 0$, cioè non c'è differenza di ridondanza tra i due sistemi di codifica.

Ad esempio, con un rapporto di compressione di 10 : 1 ($C = 10$), si ha $R = 0.9$, cioè il 90 % dei dati nel sistema di codifica di partenza è ridondante.

3 Tipi di ridondanza

Esistono diversi tipi di ridondanza:

- ridondanza **di codifica**: la codifica non è ottimizzata in base al numero di occorrenze dei vari livelli d'intensità presenti nell'immagine;
- ridondanza **interpixel** (o **spaziale**): la rappresentazione dell'immagine non tiene conto di determinati pattern che si ripetono;
- ridondanza **psicovisuale**: la rappresentazione dell'immagine contiene informazioni che sono ignorate dal nostro sistema visivo, e potrebbero quindi essere rimosse senza degradare la qualità percepita.

La compressione consiste nella riduzione/eliminazione di una o più di queste ridondanze.

4 Ridondanza di codifica

Le tecniche per la riduzione della ridondanza di codifica si basano su una concezione probabilistica del dato immagine: si trattano i valori di grigio di un'immagine come valori assunti da una variabile casuale.

Lo strumento di base è allora l'*istogramma normalizzato*, che, per ogni livello di grigio r_k , indica la probabilità $p(r_k)$ con cui esso occorre nell'immagine.

Sia $l(r_k)$ il numero di bit richiesti dal sistema di codifica adottato per rappresentare il valore di grigio r_k (ad esempio, in un'immagine a 8 bit si ha $l(r_k) = 8 \forall k$). Allora, il numero di bit richiesti, in media, per rappresentare il valore di grigio di un pixel (supponendo che l'immagine abbia livelli di grigio nell'intervallo $[0, L - 1]$) è:

$$L_{\text{avg}} = \sum_{k=0}^{L-1} l(r_k) p(r_k)$$

Rimuovere la ridondanza di codifica significa minimizzare tale quantità. Per farlo, si usano codifiche a lunghezza variabile (**variable length coding**), che assegnano rappresentazioni di meno bit ai valori più frequenti.

5 Ridondanza interpixel

Se nell'immagine sono presenti strutture regolari, il valore di un pixel diventa prevedibile in base ai valori di quelli adiacenti, cioè si ha una correlazione spaziale tra i valori dei pixel.

Tale correlazione può essere sfruttata per comprimere l'immagine, passando dalla rappresentazione visuale (cioè come matrice di pixel) a una più efficiente, ma non visuale. Questa trasformazione è chiamata **mapping**.

Se un'immagine presenta sia ridondanza di codifica che ridondanza interpixel, è possibile applicare una combinazione di variable length coding e mapping.

5.1 Run length coding

Una semplice tecnica di mapping è il **run length coding**, che sfrutta la presenza di sequenze (run) di valori costanti nell'immagine: ciascuna riga di pixel

$$f(x, 0), f(x, 1), \dots, f(x, N - 1)$$

viene trasformata in una sequenza di run

$$(g_1, w_1), (g_2, w_2), \dots$$

dove

- g_i rappresenta l' i -esimo valore di grigio incontrato sulla riga;
- w_i è il numero di pixel consecutivi aventi tale valore.

Questa codifica risulta tanto più efficiente quanto più l'immagine è composta da zone costanti. Se, invece, non ci sono abbastanza valori ripetuti, la dimensione dell'immagine aumenta: nel caso peggiore, ogni singolo pixel forma un run di lunghezza 1, e la codifica di un run richiede più bit rispetto a quella di un semplice valore di grigio (perché deve essere memorizzata anche la lunghezza).

6 Ridondanza psicovisuale

La compressione mediante l'eliminazione della ridondanza psicovisuale si basa sulle nostre capacità percettive. Ad esempio, la nostra attenzione si focalizza sui contorni degli oggetti, piuttosto che sugli specifici livelli di grigio presenti all'interno di tali oggetti.

In base a questo principio, un'immagine può essere compressa riducendo la quantizzazione, cioè il numero di livelli di grigio: così, servono meno bit per la codifica di ciascun pixel.

La riduzione della quantizzazione tende però a introdurre dei falsi contorni, che comportano un notevole degrado della qualità percepita, ma esso può essere attenuato mediante apposite tecniche. Un esempio è l'**IGS (Improved Gray-scale Quantisation)**, che smussa i falsi contorni aggiungendo nelle loro vicinanze dei valori (pseudo)casuali, calcolati in base a una somma:

1. Inizialmente, il valore della somma è 0.
2. Successivamente, per ogni pixel, essa viene aggiornata, sommando al valore (a 8 bit) di tale pixel i 4 bit meno significativi della somma precedente.¹ Poi, i 4 bit più significativi della nuova somma vengono usati come codifica del pixel corrente nell'immagine compressa.

Esempio di applicazione della tecnica IGS

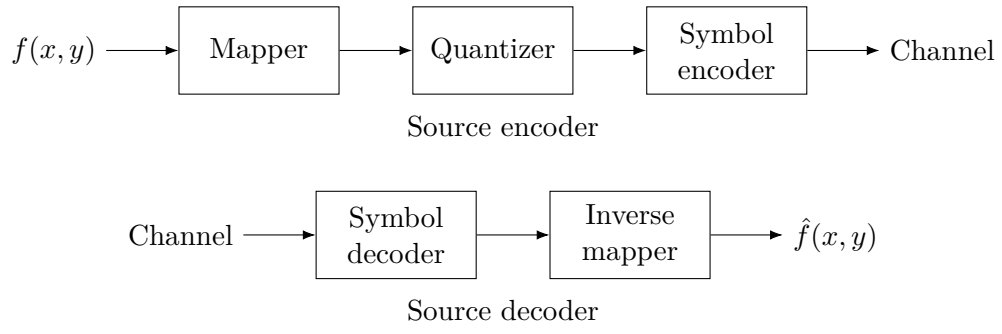
| Pixel | Valore | Somma | Codifica IGS |
|---------|-----------|----------------------------------|--------------|
| | | 0000 0000 | |
| i | 0110 1100 | $0110\ 1100 + 0000 = 0110\ 1100$ | 0110 |
| $i + 1$ | 1000 1011 | $1000\ 1011 + 1100 = 1001\ 0111$ | 1001 |
| $i + 2$ | 1000 0111 | $1000\ 0111 + 0111 = 1000\ 1110$ | 1000 |

7 Struttura di un sistema di compressione

Un sistema di compressione è composto da vari moduli, finalizzati all'eliminazione di diversi tipi di ridondanza. I principali sono:

- *Mapper*: agisce sulla ridondanza interpixel.
- *Quantizer*: opera sulla ridondanza psicovisuale. Esso è presente solo nei sistemi di compressione **lossy**, perché, riducendo la quantizzazione, provoca un degrado non reversibile dell'immagine (a differenza delle tecniche di codifica utilizzate per gli altri tipi di ridondanza, che possono essere invertite per ricostruire esattamente l'immagine originale, e si dicono quindi **lossless**).
- *Symbol encoder*: agisce sulla ridondanza di codifica.

¹Come eccezione a questa regola, se il valore del pixel corrente è **1111 xxxx** (cioè ≥ 11110000), la somma diventa semplicemente pari a tale valore (senza quindi aggiungere i 4 bit meno significativi della somma precedente). Ciò serve a evitare un overflow, che altrimenti potrebbe trasformare un pixel bianco (**1111**) in uno nero (**0000**), quando invece l'effetto desiderato è solo una piccola variazione del livello di grigio.



In generale, le tecniche lossy permettono di ottenere rapporti di compressione maggiori rispetto alle tecniche lossless.