

Addendum al materiale didattico del corso di
Elaborazione Numerica dei Segnali (*docente: G. Giunta*) su
compressione di immagini fisse e in movimento
con cenni agli standard JPEG e MPEG

N.B.: gli argomenti di base sulla DCT (tipo 2) sono trattati in:
A.V. Oppenheim, R.W. Shafer, J.R. Buck, "Discrete-Time Signal
Processing", Prentice-Hall, 1999, pag. 589-600.

Le applicazioni multimediali quali audio, immagini e video richiedono grosse quantità di dati per la loro rappresentazione. Ci sono tre motivi principali che hanno portato alla compressione dei dati nei sistemi multimediali:

- grande quantità di memoria richiesta;
- basse velocità di trasferimento dai supporti di memoria di massa che non permette l'esecuzione di operazioni in tempo reale sui dati (in special modo il video);
- larghezza di banda limitata dell'attuale rete che impedisce la trasmissione in tempo reale dei dati (audio e video).

1.1 Tecniche di compressione

Le moderne tecniche di compressione offrono una soluzione ai problemi sopra citati riducendo drasticamente la quantità dei dati che contengono l'informazione. I dati superflui possono venir rimossi senza perdere qualità e l'informazione ad essi associata presenta uno dei seguenti due tipi di ridondanza:

- **ridondanza statistica**, legata a fenomeni come la correlazione e la periodicità dei dati, si indica semplicemente con ridondanza e può essere a sua volta di due tipi:
 - ridondanza spaziale, quando si può predire il valore di un elemento dalla conoscenza dei valori degli elementi contigui allo stesso istante di tempo;

- ridondanza temporale, quando si può predire il valore di un elemento in base ai valori assunti dallo stesso in istanti di tempo differenti.

Questo genere di ridondanza può essere rimosso senza alcuna perdita di informazione.

- **ridondanza soggettiva**, legata alle caratteristiche psico-fisiche dei sistemi uditivo e visivo, si indica con il termine di irrilevanza. Essa è legata a quella parte di informazione che può essere rimossa senza conseguenze evidenti dal punto di vista percettivo. La rimozione dell'irrilevanza è irreversibile.

Le tecniche di compressione dei dati digitali si possono classificare in due grandi categorie a seconda che mirino alla riduzione della ridondanza o dell'irrilevanza presenti nel segnale. Si parla nel primo caso di tecniche di compressione senza perdite o *lossless*, nel secondo caso di tecniche con perdite o *lossy*.

1.1.1 Tecniche senza perdite

Queste tecniche permettono di riottenere una rappresentazione che è numericamente uguale all'originale. Vista l'elevata qualità dei dati ricostruiti, il fattore di compressione raggiungibile con tali metodologie non è molto elevato, in genere si ottengono rapporti di 2 : 1. Di seguito vengono accennate alcune tra le più comuni tecniche di compressione senza perdite.

Codifica entropica, sfrutta le statistiche relative ai dati. La più nota di queste codifiche è il codice di Huffman che mira a rappresentare i valori più probabili con simboli di lunghezza minima ed utilizzare simboli di lunghezza maggiore per valori meno frequenti.

Codifica run-length, mira a rappresentare con pochi bit sequenze formate dalla ripetizione di uno stesso valore. Quando si incontra un valore u con n occorrenze successive, si rappresenta la sequenza con la coppia (u, n) .

Codifica predittiva, basata sull'alta correlazione dei dati adiacenti temporalmente o spazialmente. Essa consiste nel predire il valore di un dato da codificare sulla base di dati già trasmessi. Se la predizione è fatta

bene l'entropia (informazione media) dell'errore di predizione è minore rispetto a quella dei dati originari e questo comporta un risparmio di bit nella fase di codifica.

Codifica a trasformate, nella quale anziché considerare i dati nel loro dominio originario, spaziale e/o temporale, si effettua una trasformazione lineare di blocchi di dati. Utilizzando tale trasformazione si ha una ridistribuzione dell'energia all'interno dei dati senza alterarli, cercando di ottenere quelle proprietà di correlazione e ripetizione, che permettono un miglior rendimento dei metodi precedentemente analizzati.

La trasformata più utilizzata è sicuramente la Trasformata Discreta Coseno (DCT), che presenta ottime proprietà di compressione dell'energia del blocco, ma ne esistono altre [3] come la Trasformata Discreta di Fourier (DFT) e la Trasformata di Karhunen-Loeve.

1.1.2 Tecniche con perdite

L'utilizzo di queste tecniche è possibile solo quando si accetta una certa approssimazione dei dati originari. La perdita di qualità che ne deriva è ripagata dagli alti fattori di compressione da 20 – 200 : 1 per arrivare fino a 1000 : 1. Per questo motivo, sempre più spesso, queste tecniche sono preferite a quelle senza perdite. Fra queste la più comune è la quantizzazione.

Quantizzazione, può essere vista come una funzione surriettiva, ma non iniettiva, che viene applicata ai dati. Lo scopo di questa applicazione è quello di ridurre il numero dei valori possibili. In generale tale funzione si realizza dividendo il dominio in un certo numero di sottoinsiemi ed assegnando ad ognuno un valore, detto di quantizzazione. A seconda della dimensione del dominio si parla di quantizzazione scalare (dimensione unitaria) o vettoriale.

1.2 Gli standard

Viste le numerose metodologie di compressione per applicazioni multimediali sono stati stabiliti dall'ISO (*International Organization for Standardization*) degli standard mondiali. La standardizzazione di queste tecniche è di fondamentale importanza per ridurre il costo dei codificatori e decodificatori permettendo anche l'integrazione tra apparecchiature di produttori differenti. Limiteremo la trattazione a due tra gli standard

di compressione di immagini e video, citati di seguito, perchè a questi è rivolto il lavoro di tesi:

- JPEG utilizzato per la compressione di immagini;
- MPEG utilizzato per la compressione di video ed audio associato in applicazioni *motion-intensive*.

1.2.1 Lo standard JPEG

Questo standard [4] è nato per applicazioni sia su immagini bianco e nero che a colori, con un coefficiente di compressione medio pari a 15 : 1. Il nome JPEG deriva dal gruppo di studio che lo ha sviluppato (Joint Photographic Experts Group). L'immagine a colori da trattare si considera nel formato YUV, con Y componente di luminanza ed U e V componenti di cromaticanza. L'algoritmo verrà descritto per la sola luminanza Y, considerando quindi l'immagine come fosse in bianco e nero. Per la compressione di immagini a colori si può pensare di applicare il seguente metodo alle tre componenti, l'operazione può avvenire, o interamente su una componente alla volta in modo separato, oppure intercalando in modo alternato i campioni di ognuna delle tre componenti.

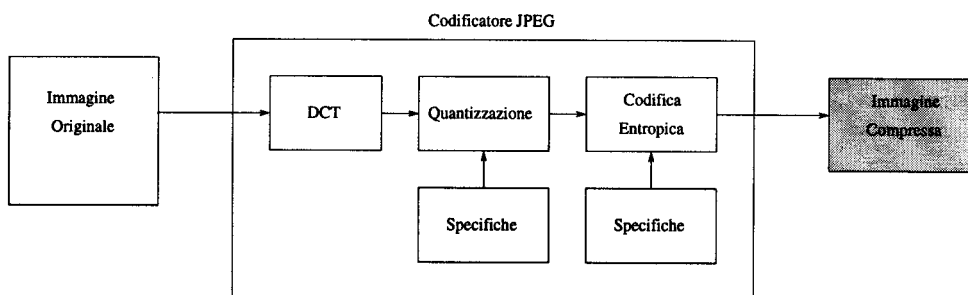


Figura 1.1: Schema del codificatore JPEG.

Il codificatore JPEG, come si può vedere nello schema di figura 1.1, è composto da tre sezioni principali che svolgono la funzione di eliminare la ridondanza e l'irrelevanza spaziale. I tre blocchi sono:

- Trasformata Discreta Coseno (DCT);
- Quantizzazione;

- Codifica entropica.

I campioni dell'immagine in ingresso, che presentano valori nell'intervallo $[0, 2^p - 1]$, con p numero di bit, vengono traslati della quantità -2^{p-1} . Nella prima sezione (DCT) i nuovi dati così ottenuti, presi a blocchi di dimensioni 8×8 , sono trasformati nel dominio delle frequenze attraverso la Trasformata Discreta Coseno in due dimensioni, definita dalle seguenti espressioni:

$$F(u, v) = \frac{C(u)}{2} \frac{C(v)}{2} \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \frac{(2x+1)v\pi}{16} \cos \frac{(2y+1)u\pi}{16}$$

dove

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{per } k = 0 \\ 1 & \text{per } k > 0 \end{cases}$$

e $f(x, y)$ rappresenta il campione dell'immagine di posizione (x, y) .

Intuitivamente la DCT può essere vista come un'analizzatore armonico. Il segnale in ingresso viene decomposto in 64 segnali base ortogonali, ognuno dei quali è rappresentato da una coppia di frequenze spaziali (u, v) . I valori $F(x, y)$ così calcolati sono chiamati in generale coefficienti DCT; il termine $F(0, 0)$ è detto DC perché rappresenta una media dei campioni del blocco, mentre i restanti 63 sono detti AC.

Solitamente i valori dei campioni dell'immagine variano lentamente lungo le due direzioni, orizzontale e verticale, portando ad uno spettro in frequenza concentrato attorno le basse frequenze. Questo fenomeno sta alla base della tecnica di compressione basata sulla DCT e normalmente per un'immagine molti dei coefficienti spettrali sono nulli o vicini allo zero quindi non necessitano di essere codificati.

Il secondo passo è rappresentato dalla quantizzazione dei coefficienti DCT. Tale operazione permette di rappresentare con precisione non troppo elevata e differente i diversi coefficienti, incrementando così il fattore di compressione. Risulta vantaggioso adottare quantizzazioni più grossolane in corrispondenza delle alte frequenze e viceversa per le basse, sfruttando il fatto che il sistema visivo umano è un filtro passa basso e quindi la percezione dell'errore di quantizzazione è più sensibile alle basse frequenze. La quantizzazione come operazione è definita dalla divisione di ciascun coefficiente per il corrispettivo passo di quantizzazione, seguita dall'arrotondamento all'intero più vicino:

$$F_Q(u, v) = \text{integer round} \left(\frac{F(u, v)}{\Delta(u, v)} \right).$$

Il passo di quantizzazione $Q(u, v)$ si ottiene, per ogni coefficiente DCT, moltiplicando i valori di una apposita tabella di quantizzazione per un fattore che esprime la qualità desiderata dell'immagine compressa. Queste tabelle di partenza sono state ottenute attraverso esperimenti psico-visivi per ridurre il più possibile la visibilità degli artefatti introdotti, un esempio è riportato in figura 1.2.

8	6	5	8	12	20	26	31
6	6	7	10	13	29	30	28
7	7	8	12	20	29	35	28
7	9	11	15	26	44	40	31
9	11	19	28	34	55	52	39
12	18	28	32	41	52	57	46
25	32	39	44	52	61	60	51
36	46	48	49	56	50	52	50

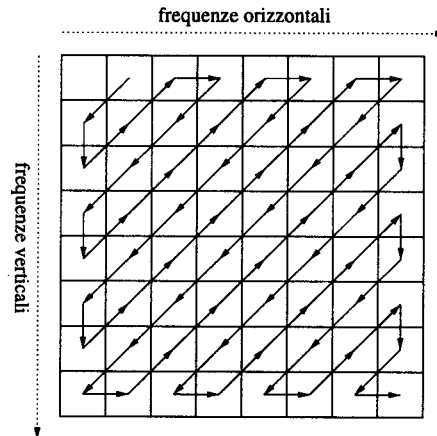
Figura 1.2: Tabella di quantizzazione JPEG.

I 63 coefficienti AC di ogni blocco, ora quantizzati, vengono ordinati secondo una scansione detta a *zig-zag*, come mostrato in figura 1.3.

Questo ordinamento rende più redditizia l'ultima fase, la codifica. Infatti posizionando i coefficienti quantizzati delle basse frequenze, che sono in prevalenza diversi da zero, prima di quelli delle alte frequenze, si aumenta la possibilità di ottenere lunghe sequenze di zeri.

I coefficienti DC sono trattati diversamente in quanto esiste una forte correlazione tra coefficienti DC di blocchi adiacenti. La tecnica utilizzata per questi è una codifica differenziale tra il coefficiente DC analizzato e quello che lo precede spazialmente.

Per ottenere infine una forma più compatta, che rappresenti l'immagine compressa, si effettua una codifica entropica. Tale processo si può dividere in due passi, da prima viene sfruttata una codifica *run-length* che dà in uscita un codice intermedio, convertito poi nella sequenza binaria finale da una codifica di Huffman ottenuta grazie a specifiche tabelle. Quanto visto rappresenta il processo di codifica, per quello di decodifica si eseguono esattamente le stesse procedure solo nell'ordine inverso.

Figura 1.3: Scansione *zig-zag* dei coefficienti AC.

1.2.2 Lo standard MPEG

MPEG è stato realizzato per la compressione di sequenze video a colori di tipo *motion-intensive* e dell'audio ad esse associato. Il nome MPEG (Moving Picture Expert Group) deriva dalla commissione che ne ha iniziato lo studio e la successiva realizzazione. Il primo standard prodotto, conosciuto come MPEG-1 [5], è nato esclusivamente per applicazioni video con risoluzione abbastanza ridotta, 320×240 punti, con un *bitrate* di 1 – 1.5 Mbit/s e solo per sequenze di tipo progressivo, come ad esempio nel caso della memorizzazione su CD-ROM. MPEG-2 [6], l'evoluzione di MPEG-1, è più versatile ed oltre a garantire la compatibilità con MPEG-1 può trattare formati con risoluzioni differenti, come il CCIR 601 con 720×486 punti ed anche HDTV con 1920×1080 , con un *bitrate* molto più elevato di 2 – 40 Mbit/s. MPEG-2 supporta sia sequenze di tipo progressivo che interlacciato ed attualmente viene utilizzato nella TV digitale via satellite (DBS), nella TV via cavo e per applicazioni come la video-comunicazione su rete ISDN. MPEG-2 è per ora l'ultima versione dello standard, risalente al 1993 ed è quella utilizzata nel lavoro di tesi. Per gennaio del 1999 è prevista la nascita del nuovo standard MPEG-4 [7], dedicato alla videotelefonìa con bassissimi *bitrate* minori di 64 Kbit/s, che utilizzerà una codifica ad oggetti.

Prima di passare alla descrizione dell'algoritmo, che realizza MPEG-2, risulta utile analizzare brevemente le caratteristiche ritenute importanti per questo standard.

Accesso casuale, fondamentale per la riproduzione delle sequenze a partire da un qualunque punto. Per soddisfare a questa caratteristica si deve permettere che il *bitstream*, relativo alla sequenza, sia accessibile in punti intermedi codificati solo con riferimento a se stessi.

Ricerca avanti/indietro veloce, si deve poter scandire l'intero *bitstream* visualizzando le immagini selezionate per ottenere l'effetto di avanzamento o riavvolgimento veloce, usando i punti intermedi di accesso.

Riproduzione inversa, è una caratteristica richiesta solo da particolari applicazioni, tuttavia la sua introduzione non dovrebbe apportare un costo troppo elevato in termini di memoria.

Sincronizzazione audio-video, il segnale video deve essere sincronizzato perfettamente con quello audio anche quando i due segnali provengono da sincronismi leggermente sfalsati.

Robustezza nei confronti dell'errore, visto che il *bitstream* sarà memorizzato o trasmesso, comunque ci si deve aspettare l'introduzione di possibili errori. In questi casi la codifica di sorgente utilizzata (metodo di compressione) deve essere in grado di evitare malfunzionamenti del sistema.

Ritardo di codifica/decodifica, dipende sostanzialmente dalla particolare applicazione considerata, per la videotelefonìa, ad esempio, non può superare i 150 ms per non far perdere la natura colloquiale di tale applicazione. In altri campi sono possibili anche ritardi maggiori, comunque in generale, bisognerà trovare un compromesso tra qualità e ritardo accettabile, che quindi dovrà essere considerato come un parametro di sistema.

Editabilità, nel senso di poter accedere ed elaborare le informazioni relative a ciascun quadro direttamente nella forma codificata. D'altro canto per avere elevati rapporti di compressione le immagini non possono essere codificate a se stanti, quindi per soddisfare entrambe queste esigenze si può prevedere l'esistenza di un'unità minima di immagini, detta GOP, codificata in modo indipendente rispetto al resto della sequenza. In tal modo è possibile modificare i dati del *bitstream* che riguardano una

singola unità in maniera autonoma.

Flessibilità, prevedere una larga flessibilità del formato delle sequenze da trattare, in termini di dimensioni per quadro e di frequenza di quadro.

L'algoritmo adottato per la realizzazione di MPEG-2 prevede due metodi diversi per la codifica delle immagini. Il primo, sfruttando la ridondanza temporale contenuta nella sequenza, codifica le immagini con riferimento a quelle precedenti e/o future (codifica *interframe*), il secondo invece considera l'immagine come se fosse statica e la codifica in modo indipendente dalle altre (codifica *intraframe*). L'utilizzo contemporaneo di queste due tecniche garantisce un elevato rapporto di compressione e la possibilità di accesso casuale al *bitstream*. In base all'applicazione di queste tecniche le immagini vengono classificate in tre tipi come visibile in figura 1.4:

1. Intra (I), codificate come immagini statiche;
2. Predette (P), codificate facendo riferimento ad un quadro precedente;
3. Interpolate (B), codificate facendo riferimento ad un quadro precedente e/o ad uno successivo.

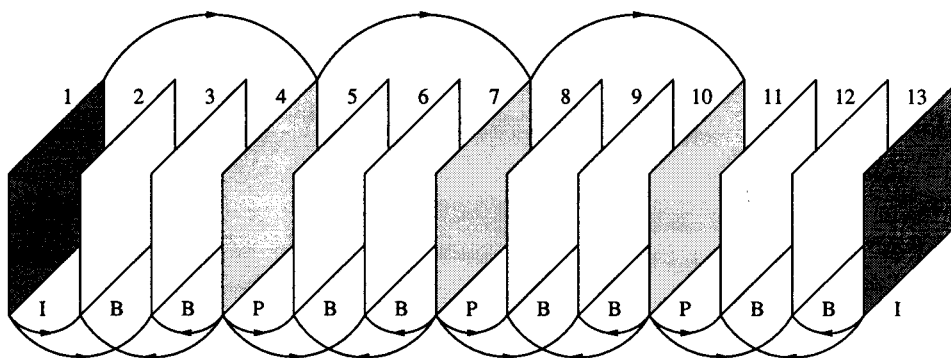


Figura 1.4: Esempio di GOP.

Le immagini di tipo intra (I) sono codificate con riferimento solo a se stesse usando una tecnica basata sulla DCT simile al caso del JPEG. Queste immagini forniscono i punti di accesso casuale ed hanno anche il compito di mantenere l'allineamento tra il trasmettitore ed il ricevitore.

infatti trasmettendo solo immagini di tipo predette o interpolate si ha un fenomeno di deriva della qualità dovuto all'accumolo degli errori in ricezione. L'effetto appena descritto si elimina introducendo un'immagine intra ogni tanti quadri, che fornisce così un nuovo riferimento per le immagini successive. Il rapporto di compressione che si ottiene è abbastanza basso.

Le immagini di tipo predette (P) sono codificate usando una tecnica predittiva con riferimento ad una sola immagine passata di tipo intra o predetta. Il rapporto di compressione è significativamente maggiore di quello ottenuto per le immagini intra .

Le immagini di tipo interpolate (B) sono codificate usando due immagini di riferimento, di tipo intra o predette, una nel passato e l'altra nel futuro. Questo tipo di immagine non viene mai usata come riferimento e presenta i più alti rapporti di compressione tra le tre.

Nella figura 1.4 si può osservare un esempio di GOP (*Group of Pictures*), che come già detto in precedenza, rappresenta l'unità minima codificabile in modo indipendente. Il GOP viene definito dall'utente, in base alle esigenze specifiche dell'applicazione, in fase di codifica specificando il tipo e numero di quadri in esso contenuti.

Sempre in figura 1.4 è rappresentata la soluzione tipicamente adottata per il formato CCIR 601 con un GOP formato da 12 quadri, di cui uno I, tre P, ed otto B. I quadri sono ordinati secondo la sequenza di visualizzazione mentre quella di codifica, che si può ricavare osservando le frecce indicanti i quadri di riferimento, è:

I1 P4 B2 B3 P7 B5 B6 P10 ...

Per la descrizione del codificatore si fa riferimento allo schema di figura 1.5, nel quale x_n rappresenta l'immagine originale in ingresso al tempo n , \tilde{x}_n indica invece la sua stima elaborata attraverso il predittore P . In quest'ultimo la stima \tilde{x}_n viene ottenuta dall'immagine precedente ricostruita \hat{x}_{n-1} , salvata nella memoria di quadro Z^{-1} . L'errore di predizione, indicato con d_n , è ottenuto come differenza tra l'immagine originale e la sua stima, c_n è il suo trasformato in frequenza attraverso la DCT e \hat{c}_n indica lo stesso dopo l'operazione di quantizzazione. Infine \hat{d}_n e \hat{x}_n rappresentano rispettivamente l'errore e l'immagine originale ricostruiti. I dati necessari alla predizione ed i valori quantizzati vengono codificati con un codice a lunghezza variabile *VLC* e producono un *bitstream* che viene memorizzato su un *buffer* per poter essere poi trasmesso su un mezzo trasmissivo a velocità costante. Lo stato di riempimento del *buffer* viene impiegato come segnale di ricezione per controllare, attraverso il passo di

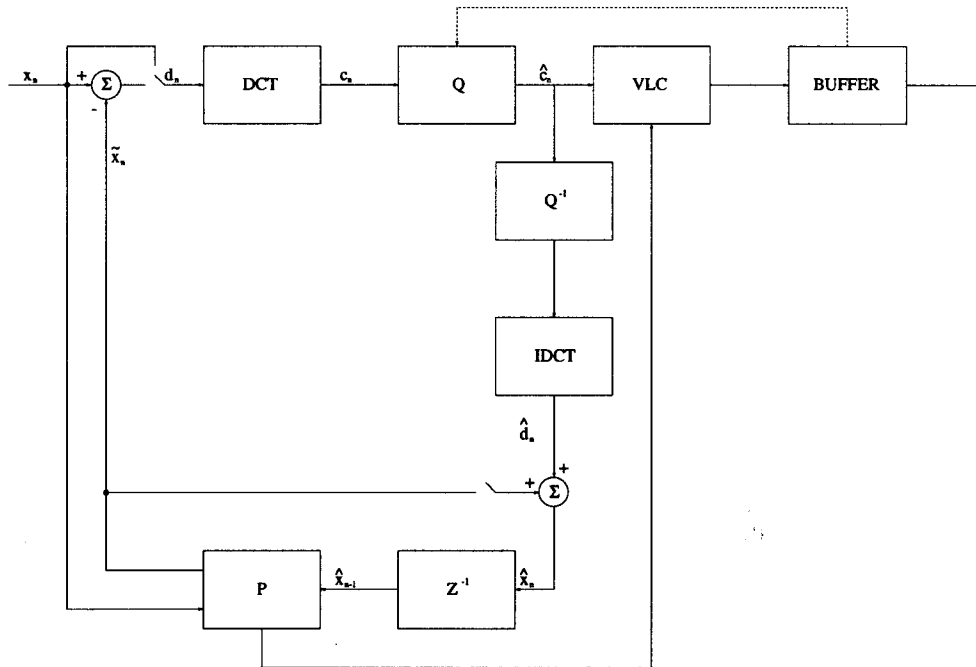


Figura 1.5: Codificatore MPEG.

quantizzazione, il flusso di dati da trasmettere e quindi il *bitrate*.

Il codificatore usa per la compressione, come già accennato, due diverse tecniche:

- riduzione della ridondanza temporale;
- riduzione della ridondanza spaziale.

Riduzione della ridondanza temporale, è prevista solo per codifiche *interframe* quindi con immagini di tipo P e B. Si ottiene attraverso l'impiego della predizione moto compensata. L'idea fondamentale di tale metodo è quella di considerare localmente l'immagine da codificare come ottenuta dalla traslazione di un'immagine precedente e/o successiva nel tempo. Il termine localmente sta ad indicare che il moto non deve essere necessariamente uguale per tutta l'immagine, ma può variare da zona a zona. L'ipotesi restrittiva che rende questa tecnica facilmente implementabile ma che contemporaneamente ne limita la possibilità di impiego, è il fatto di considerare il moto solo di tipo traslatorio.

Per capire meglio come funziona la compensazione del moto consideriamo

esclusivamente i quadri di tipo P in cui la predizione viene fatta solo con riferimento al passato. Nella descrizione del metodo si continua ad utilizzare la terminologia dello schema di figura 1.5.

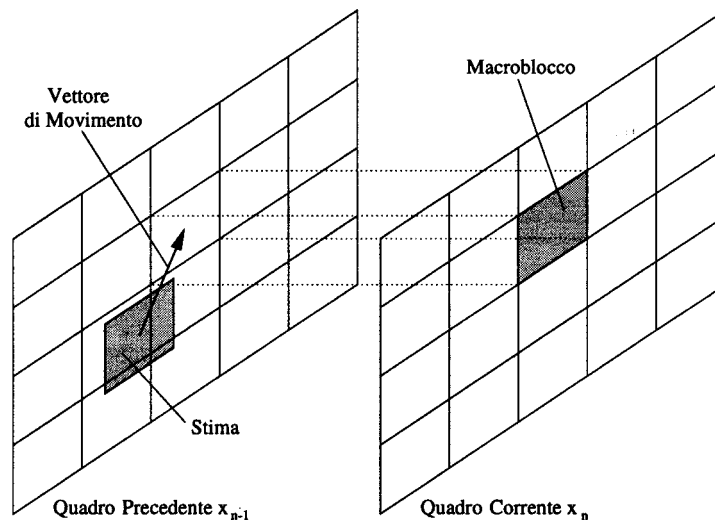


Figura 1.6: Compensazione del movimento.

L'immagine corrente x_n , come illustrato in figura 1.6, viene suddivisa in blocchi di ridotte dimensioni ed ogni blocco si ricerca nell'immagine precedente x_{n-1} in modo da trovare un vettore di movimento rappresentativo dello spostamento del blocco nel passaggio dal quadro al tempo $n - 1$ a quello corrente.

La stima \tilde{x}_n dell'intera immagine corrente si ottiene da quella precedente traslando i blocchi delle quantità espresse dai rispettivi vettori di movimento. L'immagine x_n originale viene ricostruita quindi da quella precedente x_{n-1} , disponendo dei vettori di movimento e dell'errore di predizione dato da $x_n - \tilde{x}_n$.

In realtà come immagine di riferimento non si può considerare l'immagine originale al tempo $n - 1$, ma si utilizza quella ricostruita, poiché solo di questa si dispone in ricezione.

Più la predizione è accurata minori saranno i dati da dover trasmettere, ottenendo così elevati rapporti di compressione. Come si può intuire è di fondamentale importanza la scelta della dimensione dei blocchi:

- con blocchi piccoli si incrementa l'accuratezza della stima diminuendo quindi l'errore di predizione, ma aumentando il numero di

vettori di movimento ed il costo computazionale;

- con blocchi grandi, viceversa, aumenta l'errore di predizione ma diminuisce il numero di vettori dei movimento ed il costo computazionale.

La dimensione scelta deve essere un compromesso tra i due casi sopra citati, in generale MPEG utilizza dei blocchi di 16×16 *pixel*, definiti macroblocchi in quanto ognuno contiene esattamente 4 blocchi 8×8 usati per il calcolo della DCT. In particolare MPEG-2 ha la possibilità, solo per sequenze di tipo interlacciato, di considerare dei macroblocchi di dimensione ridotta 16×8 *pixel*.

Facendo sempre riferimento ad un quadro di tipo P, non tutti i macroblocchi dell'immagine devono obbligatoriamente essere codificati in modo *inter*. L'algoritmo trova sempre per ogni macroblocco una stima, poi calcola la varianza dell'errore di predizione e solo se questa risulta minore di quella dell'immagine originale lo codifica come *inter*, altrimenti il macroblocco viene codificato come *intra*.

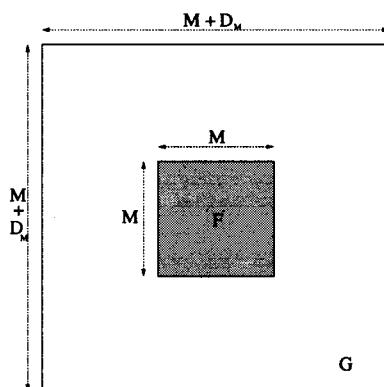


Figura 1.7: Area di ricerca nella tecnica *block-matching*: F rappresenta un macroblocco di dimensioni $M \times M$ nel quadro corrente, G l'area di ricerca in quello precedente, D_M è lo spostamento massimo nelle due direzioni.

Per quanto riguarda la stima del movimento, lo standard MPEG non specifica la procedura che deve essere adottata. Le tecniche comunemente utilizzate sono del tipo *block-matching*, in cui il macroblocco da stimare viene paragonato a blocchi della stessa dimensione interni ad un'area di ricerca nel quadro precedente, come illustrato in figura 1.7.

La stima viene fatta prendendo il vettore di movimento \overline{mv}^* che minimizza una opportuna funzione di costo dell'errore di predizione:

$$\overline{mv}^* = \arg \min_{\overline{mv} \in G} \left\{ \sum_{\bar{p} \in F} C(F(\bar{p}) - G(\bar{p} + \overline{mv})) \right\}.$$

La scelta della funzione di costo C e la tecnica di ricerca del minimo sono lasciate libere all'implementazione.

Quanto visto per i quadri P vale anche per le immagini di tipo B, dove in aggiunta si stima lo spostamento dall'immagine successiva. La predizione dal futuro permette di codificare in modo *inter* anche quelle aree, cosiddette scoperte, che non presentano nessun riferimento nel passato, come ad esempio un nuovo oggetto che compare nella scena. Le scelte per la codifica di ogni macroblocco dell'immagine di tipo B sono molteplici:

- con la sola predizione dal passato, come per i quadri P;
- con quella solo dal futuro;
- nel modo interpolato, cioè utilizzando entrambe le predizioni futura e passata con un'operazione di media;
- in modo *intra*.

L'utilizzo dei macroblocchi interpolati è una particolarità dello standard MPEG-2 ed è grazie a questa che si ottengono elevati fattori di compressione a parità di qualità dell'immagine.

Riduzione della ridondanza spaziale, viene applicata, con caratteristiche diverse, sia ai macroblocchi dell'immagine codificati *intra*, che a quelli dell'errore di predizione quando si usa una codifica *inter*. La procedura è sostanzialmente analoga a quella trattata per lo standard JPEG, con la Trasformata Coseno Discreta, la quantizzazione e la codifica entropica.