

Compressione video

Occupazione su disco

PAL non compresso a colori:

1 fotogramma: $720 \times 576 \times 3 = 1.244.160$ 1.2Mb
(Hor x Vert x Rgb)

1 secondo: $1.244.160 \times 25$ (fps) = 31.104.000 31Mb

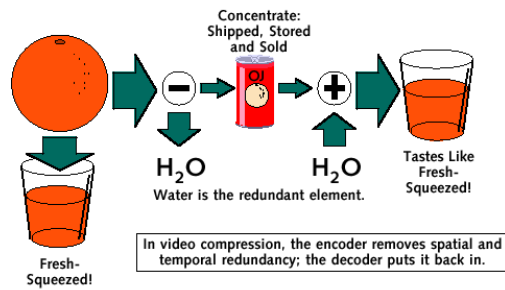
1 minuto: $31.104.000 \times 60$ (sec) = 1.866.240.000 1.9Gb

90 minuti: $1.866.240.000 \times 90$ (min) = 167.961.600.000

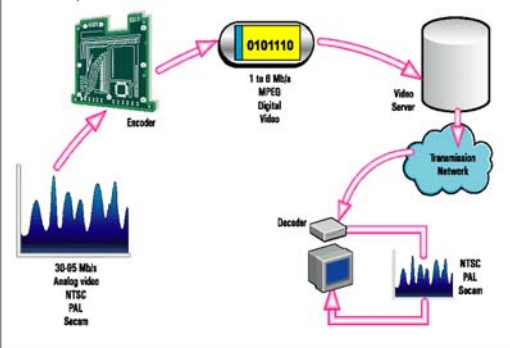
168Gb

Perchè comprimere il video

	Pixels/ Frame	Frame /sec	Bytes/ pel	Bit Rate
Video Telef.	~100k	10	1,5	12.2 Mbit/s
TV	~340k	30	2	170 Mbit/s
HDTV	~900k	60	2	885 Mbit/s



Video Compression



Come comprimere

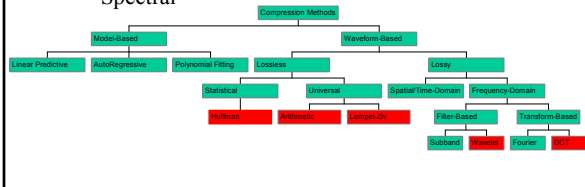
Eliminare ridondanza spaziale e temporale

- spaziale: intra-frame;
- temporale: inter-frame.

sfruttando proprietà del sistema visivo per rendere impercettibili le modifiche introdotte

What does compression do

- Reduces signal size by taking advantage of correlation
 - Spatial
 - Temporal
 - Spectral



Compression Issues

- Lossless compression
 - Coding Efficiency
 - Compression ratio
 - Coder Complexity
 - Memory needs
 - Power needs
 - Operations per second
 - Coding Delay

Compression Issues

- Additional Issues for Lossy Compression
 - Signal quality
 - Bit error probability
 - Signal/Noise ratio
 - Mean opinion score

Compression Method Selection

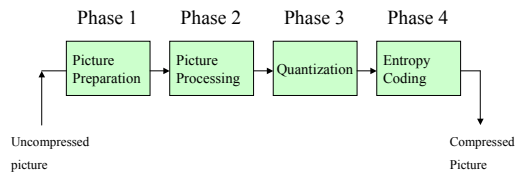
- Constrained output signal quality? – TV
- Retransmission allowed? – interactive sessions
- Degradation type at decoder acceptable?
- Multiple decoding levels? – browsing and retrieving
- Encoding and decoding in tandem? – video editing
- Single-sample-based or block-based?

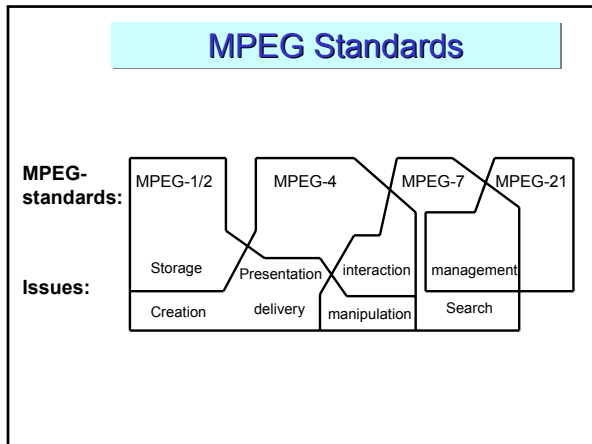
Video Taxonomy

Compressed		Uncompressed	
Tape	Streaming	Digital	Analog
Digital Betacam	MPEG-1	D-1 (CCIR 601)	Betacam
DVCPPro (D-7)	MPEG-2	D-2	VHS
DV	MPEG-4	D-5	S-Video
DVCAM	M-JPEG		Video-8
Digital-8	H.261		Hi-8
D-9	H.263+		Betamax
DVCPPro50	Real		
	Sorenson		
	Cinepak		
	Indeo		
	Video for Windows		

Image Compression System

- Basic Units of an Encoder

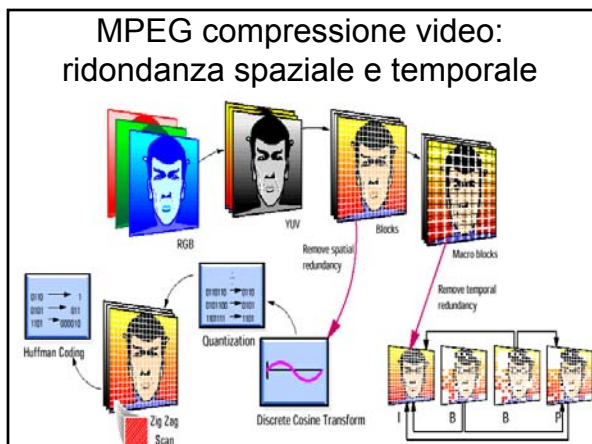




- ### MPEG Family
- MPEG – 1
 - Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mb/s
 - MPEG – 2
 - Generic coding of moving pictures and associated audio
 - For broadcasting & studio work
 - MPEG – 3
 - no longer exists (has been merged into MPEG-2)
 - MPEG – 4
 - Very low bit rate audio-visual (integrated multimedia) coding

- ### MPEG Family
- MPEG – 7
 - Multimedia content description interface
 - MPEG – 21
 - Vision statement
 - To enable transparent & augmented use of multimedia resources across a wide range of networks and devices
 - Objectives
 - To understand how the elements fit together
 - To identify new standards which are required if gaps in the infrastructure exist
 - To accomplish the integration of different standards

- ### MPEG1: Video Encoding
- The MPEG standards
 - do not define an encoding process
 - define syntax of the coded stream
 - define a decoding process



- ### Video Compression
- Digital video compression algorithms operate on a sequence of bit-mapped images
 - *Spatial compression (intra-frame)*
 - Compress each individual image in isolation
 - *Temporal compression (inter-frame)*
 - Store the differences between sub-sequences of frames

Spatial Compression

- Compress method is similar to image compression
 - Lossless
 - No information loss
 - Compression ratios is lower
 - Lossy
 - Some information loss
 - Compression ratios is higher
- Why recompressing video is unavoidable
 - The compressor used for capture are not suitable for multimedia delivery
 - For post-production

Temporal Compression

- *Key frames*
 - Certain frames in a sequence are designated as key frames
- *Difference frame*
 - Each of the frames between the key frames is replaced by a difference frame
 - Records only the differences between the frames

Codifica a blocchi con trasformate

- ogni trasformata concentra l'informazione in una regione limitata, dove si hanno maggiori frequenze spaziali;
- trasformata teoricamente più efficiente è KLT, ma le funzioni base dipendono dall'immagine;
- trasformata del coseno è meno efficiente ma più economica.

DCT (Discrete Cosine Transform)

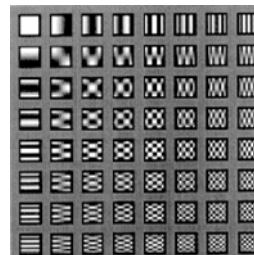
- DCT è una trasformata con funzioni ortogonali, l'inversa è ancora una DCT;
- la trasformata è separabile: componenti orizzontale e verticale sono indipendenti;
- la DCT trasforma ciascun blocco 8x8 in una serie di 64 forme d'onda associando ad esse un peso;
- la prima componente di frequenza è chiamata DC o componente continua;
- la base di funzioni è indipendente dall'immagine;
- le componenti di frequenza sono direttamente correlate alle caratteristiche della visione umana;
- esiste un algoritmo di calcolo veloce.

DCT diretta e inversa

$$y_{kl} = \frac{c(k)c(l)}{4} \sum_{i=0}^7 \sum_{j=0}^7 x_{ij} \cos\left(\frac{(2i+1)k\pi}{16}\right) \cos\left(\frac{(2j+1)l\pi}{16}\right)$$

$$x_{ij} = \sum_{k=0}^7 \sum_{l=0}^7 y_{kl} \frac{c(k)c(l)}{4} \cos\left(\frac{(2i+1)k\pi}{16}\right) \cos\left(\frac{(2j+1)l\pi}{16}\right),$$

Le 64 componenti di una DCT



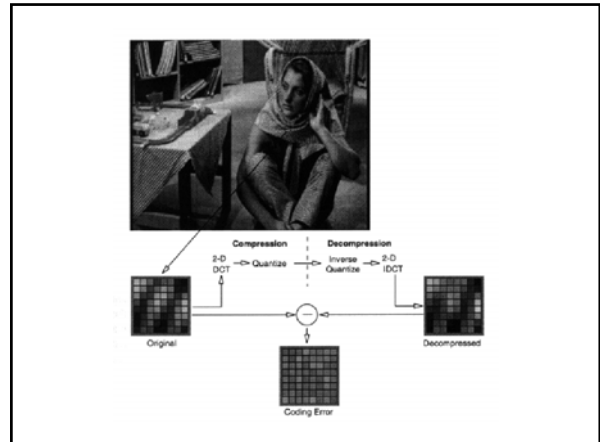
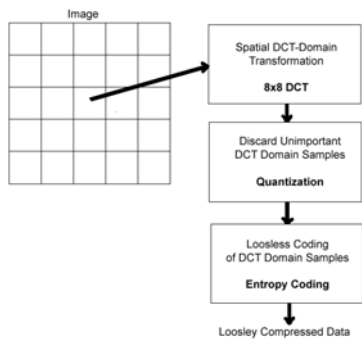
Trasformata Discreta del Coseno

- Porta dal dominio del tempo al dominio delle frequenze. Divide la forma d'onda nelle sue componenti frequenziali, espresse solo in forma di cos();

- La DCT in sé non effettua alcuna compressione (anzi). Semplicemente trasforma i pixel sorgente in una forma nella quale è possibile identificare la ridondanza;

- Non tutte le frequenze spaziali sono simultaneamente presenti, quindi all'uscita della DCT avremo alcuni coefficienti con un valore rilevante, ma la maggior parte sarà molto prossima a zero.

Schema di compressione DCT

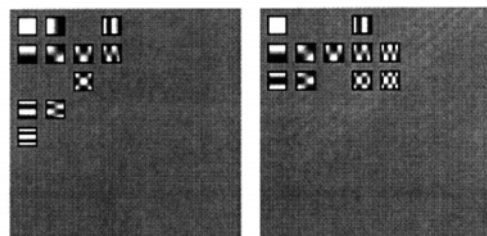


Quantizzazione

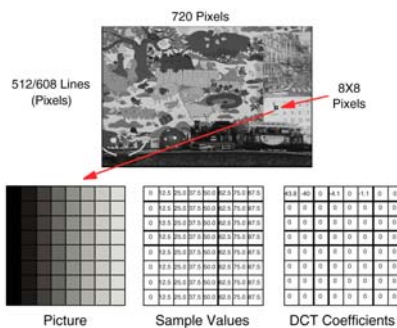
- Questo è il processo che introduce perdita di qualità del segnale video;
- I coefficienti della DCT si dividono per una matrice di quantizzazione:
 - Valori sempre più alti man mano che cresce la frequenza
 - Fattore di scala
- Molti coefficienti a 0;
- Si organizzano a zig-zag;

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Componenti selezionate

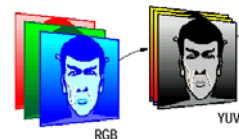


Esempio

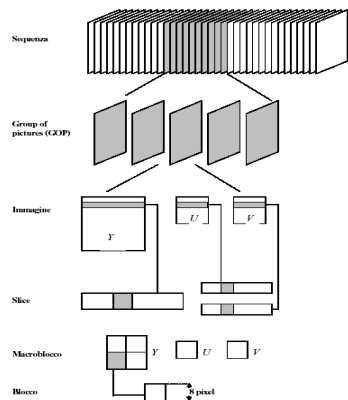


Correlazione cromatica

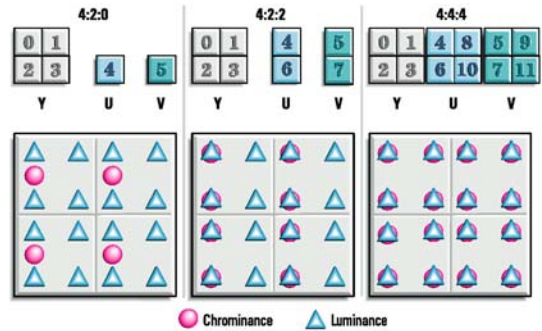
- Il sistema visivo umano è più sensibile a variazioni di luminosità che a variazioni cromatiche;
- conversione da RGB a YUV;
- sottocampionamento spaziale di U e V: 4:2:2 o 4:1:1



Suddivisione in blocchi - 1



Suddivisione in blocchi - 2

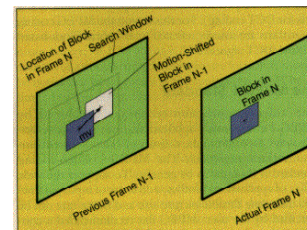


Campionamenti tipici

		Luminance	Chrominance
NTSC	50 Interlaced Fields per Second	720 x 243	360 x 243
	30 Progressive Frames per Second	720 x 486	360 x 486
PAL, SECAM	50 Interlaced Fields per Second	720 x 288	360 x 288
	25 Progressive Frames per Second	720 x 576	360 x 576
SIF NTSC		352 x 240	176 x 120
SIF PAL, SECAM		352 x 288	176 x 144

Compressione Inter-frame

- Sottocampionamento temporale: si saltano immagini; in fase di decodifica: ricostruzione delle immagini mancanti per interpolazione;
- Motion-Compensated prediction, errore di predizione e vettore di moto di un blocco 16x16 vengono codificati.



Standard di video coding

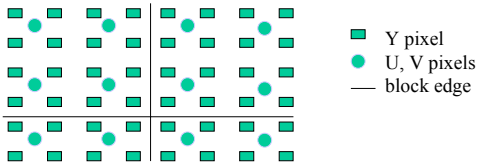
Organizz.	Standard	Bit Rate	Applicaz.
ITU-T	H.261	px64 kbit/s p=1..30	ISDN Video Conf
ISO	ISO 11172-2 MPEG-1 Video	1.2 Mbit/s	CD-ROM
ISO	ISO 13818-2 MPEG-2 Video	4-80 Mbit/s	SDTV, HDTV
ITU-T	H.262 MPEG-2	4-80 Mbit/s	SDTV, HDTV
ITU-T	H.263	64 kbit/s	Video tel.
ISO	CD 14496-2 MPEG-4 Video	24-1024 kbit/s	
ITU-T	H.263 Ver. 2	< 64kbit/s	Video tel.
ITU-T	H.263L	< 64kbit/s	Video tel.

Codec

Codec	Manufacturer	Description	S.	P.	Version	Size
c:\windows\system32\mpeg4c12.dll	Microsoft Corporation	Microsoft MPEG-4 V...	OK	C	3.01.0.2094	248,31 KB (...)
c:\windows\system32\mpeg4v2.dll	Microsoft Corporation	Microsoft MPEG-4 V...	OK	C	5.3.0000005.0	16,50 KB (...)
c:\windows\system32\mpeg4v3.dll	Intel Corporation	Intel® Video S.10	OK	C	8.5.10.15.2.05	725,50 KB (...)
c:\windows\system32\mpeg4v32.dll	Microsoft Corporation	Microsoft MPEG-4 V...	OK	C	5.00.2134.1	27,27 KB (...)
c:\windows\system32\mpeg4v32.dll	Microsoft Corporation	Microsoft MPEG-4 V...	OK	C	4.4.3385	163,77 KB (...)
c:\windows\system32\mpeg4v32.dll	Microsoft Corporation	Microsoft MPEG-4 V...	OK	C	4.4.3385	262,37 KB (...)
c:\windows\system32\mpeg4v32.dll	Microsoft Corporation	Microsoft MPEG-4 V...	OK	C	5.00.2195.6612	35,77 KB (...)
c:\windows\system32\mpeg4v32.dll	Intel® Corporation	Intel® Video S.10	OK	C	Not Available	194,50 KB (...)
c:\windows\system32\mpeg4v32.dll	Radix Inc.	Radix Inc.	OK	C	1.10.0.0	106,00 KB (...)

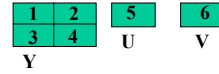
H.261

- basso bit rate per videotelefono e videoconferenza;
- basato su motion compensation e codifica per trasformata;
- formato immagine:
 - QCIF: 176 x 144 pixels;
 - CIF (~CGA): 352 x 288 pixels;
- color sampling: 4:2:0

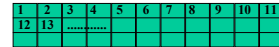


H.261 - Blocchi, Macro blocchi e gruppi di blocchi

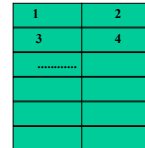
MacroBlocco:
6 blocchi 8x8,
4 per Y
2 per U e V



GOB in un video:
33 MB



GOB in una picture: 16 MB

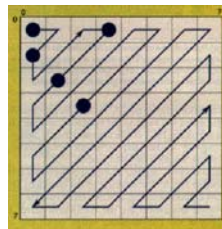


H.261 - DCT e Quantizzazione

I blocchi vengono trasformati DCT.

I coefficienti della trasformata vengono quantizzati in modo uniforme.

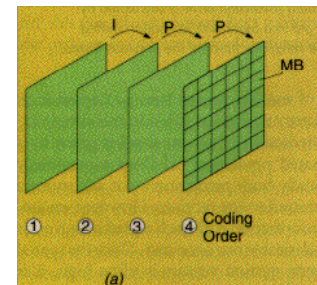
ZigZag scan: l'array dei coefficienti viene attraversato a "zigzag" creando una successione 1D; si creano lunghe serie di 0 che permettono di risparmiare spazio.



Compressione Intra-Frame

La sequenza di immagini è organizzata in:

- I-picture: codificate DCT
- P-picture: interframe prediction



Picture Type

- I (intra) pictures
 - Code without reference to other pictures
 - Low compression rate
- P (predicted) pictures
 - Code using motion compensated prediction from a past I or P picture
 - Higher compression rate than I picture
- B (bidirectional-predicted) pictures
 - Code bidirectional interpolation between the I or P picture which preceded & followed them
 - Highest compression rate

All are compressed using the MPEG version of JPEG compression

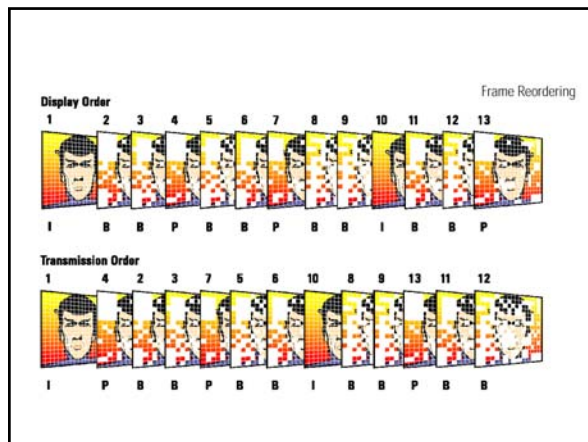
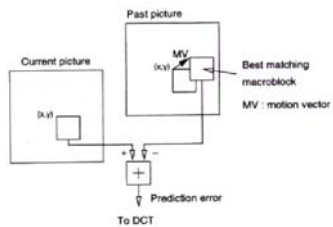
Compression performance of MPEG 1

Type	Size	Compression
I	18 KB	7:1
P	6 KB	20:1
B	2.5 KB	50:1
Avg	4.8 KB	27:1

H.261 - Motion Vector coding

Viene effettuata codifica per differenza.

La differenza di motion vector è codificata VLC.



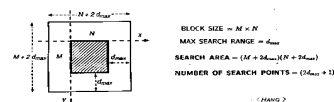
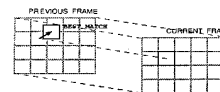
An object moving between frames



Area of potential change

Motion Compensation

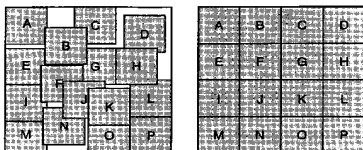
Matching Algorithm for Motion Estimation



mmc6.06

Motion Compensation

Model for Block-Based Motion Compensation



mmc6.04

Motion compensation -1

Ricerca pixel corrispondenti in due immagini consecutive.

Si calcola $e(x,y,t) = I(x, y, t) - I(x-u, y-v, t-1)$

$I(x-u, y-v, t-1)$ è la predizione di $I(x, y, t)$

Motion compensation - 2

Lo standard impone la ricerca solo per traslazioni.

Criterio di ricerca: non su tutto il frame ma solo in un intervallo $[-p, p]$;

$p = 15$ per *talking heads*

$p = 63$ per azioni di sport

Piccoli blocchi $N \times M$ soddisfano "smoothness", ma spendono tempo;

Blocchi grandi sono più efficienti.

Lo standard propone $N=M=16$ pixel

Motion Compensation - 3

Criterio di matching: Maximum Absolute Error (MAE):

$$MAE(i, j) = \frac{1}{M \times N} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} |c(x+k, y+l) - r(x+i+k, y+i+l)|$$

Dove c frame corrente, r frame di riferimento

Quando MAE è minimo, r è il blocco di best match e la coppia (i, j) individua il motion vector

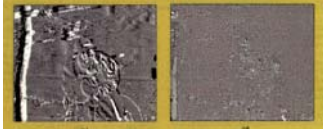
H.261 - Esempio

frame N-1



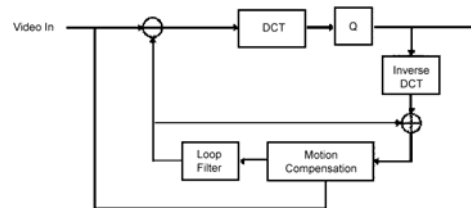
frame N
con motion
vector

prediction
error senza
motion
compensation
($N - (N-1)$)



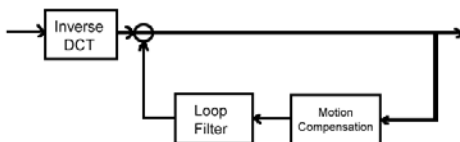
con motion
compensation

H.261 - Encoder



Loop filter è semplice smoothing

H.261 - Decoder



FORMATI DI REGISTRAZIONE A CONFRONTO

	Video Cd	X-Video	SuperVideo Cd	X-Video	Dvd
Formato di compressione	Mpeg1	Mpeg1	Mpeg2	Mpeg2	Mpeg2
Formato immagini in pixel	352 x 288	352 x 288 a 720 x 576	480 x 576	720 x 576	720 x 576
Bit rate video in kbps	1.150	fino a 3.500	fino a 2.600	fino a 9.800	fino a 9.800
Bit rate audio in kbps	224	224	32-384	32-384	variable
Durata video su Cd da 650 Mb in minuti	74	24-74	35-70	15-70	fino a 20

ISO/IEC 11172-2 : MPEG 1



Requisiti:

- video generico (VHS) a 1~1.5 Mbit/sec;
- adatto a CD-ROM;
- accesso diretto a un frame in un tempo limitato;
- avanti/indietro veloce;
- audio e video sincronizzati;
- decoder pratico e facile da implementare.

ISO/IEC 11172-2 : aspetti nuovi

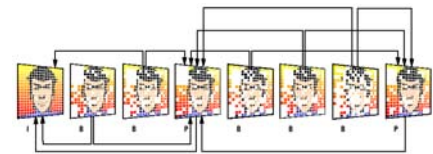
- motion compensation bidirezionale
- gruppo di figure (GOP)
- motion compensation a livello di mezzo pixel (half pel compensation)
- quantizzazione pesata sulla visione umana

Motion Estimation - 1

- In realtà la grossa riduzione del BitRate si ha sfruttando la ridondanza temporale:
 - Una volta codificata un'immagine di riferimento come spiegato prima (I-Frame) cerco similarità con i frames adiacenti;
 - Innanzitutto controllo che il macroblocco in esame non sia rimasto identico (o quasi) nel frame successivo, nel qual caso non lo invio;
 - Altrimenti, cerco nelle vicinanze per identificare se si è mosso. Se lo trovo, invio un vettore di movimento;
 - Codifico secondo lo schema visto in precedenza solo se non ho trovato il macroblocco da nessuna parte nel frame;
 - Se sono tra una *I* e una *P*-Picture, interpolo i vettori di movimento;
 - Non è la panacea, però. Se commetto un piccolo errore e calcolo i frames successivi solo partendo dal precedente, l'errore si propaga rapidamente.

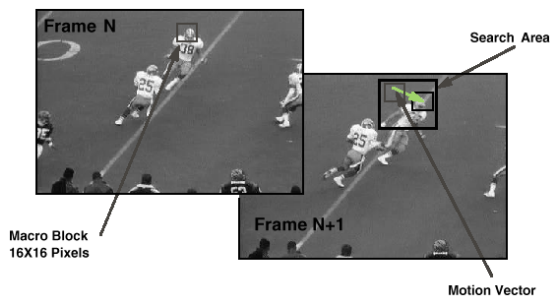
Motion Estimation - 2

Per risolvere il problema si creano GOP (Group Of Pictures) con diversi tipi di Picture all'interno.



- Intra-coded Picture
- Predictive coded Picture
- Bi-directionally coded Picture
- I-frames: contain full picture information
- P-frames: predicted from past I or P frames
- B-frames: use past and future I or P frames
- Transmit I frames every 12 frames or so.

Motion Estimation - 3



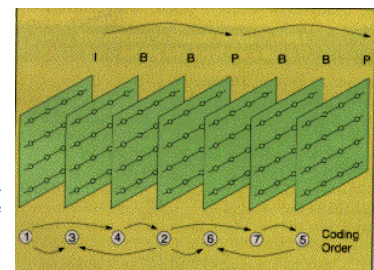
Ricerca dei blocchi corrispondenti per la compensazione del moto

ISO/IEC 11172-2 : Parametri nuovi

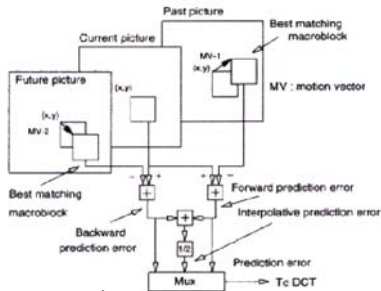
dimensione figura:
fino a 4096x4096

frequenza immagini:
23.97 ~ 60

GOP contiene oltre a
frames *I* e *P* anche
frames *B*



ISO/IEC 11172-2 : B-frames



B-frames sono ottenute per motion compensation predittivo da P-frames passati o futuri

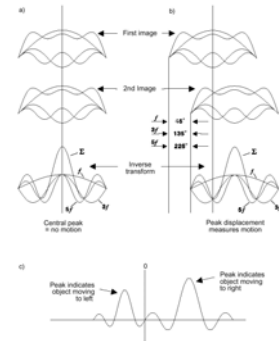
Come avviene la ricerca del Macroblocco ?

• Block Matching

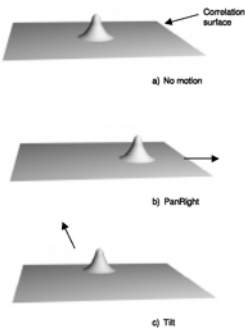
- Enorme quantità di calcolo;
- Fatta a diversi livelli di dettaglio.

• Phase correlation

- Trasformata di Fourier sui blocchi da comparare;
- Sottraggo le componenti frequenziali;
- Antitrasformata di Fourier;
- Ottengo dei picchi che rivelano la direzione e l'entità dei movimenti (ma non la loro posizione);
- Block Matching ma solo in direzioni e distanze predeterminate.



Phase Correlation

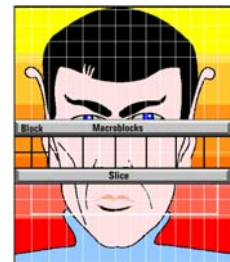


In realtà il processo avviene in 2 dimensioni, per cui si ottiene una superficie di correlazione con tanti picchi quanti sono i movimenti nell'area che si sta esaminando.

È importante capire che con la correlazione di fase si ottiene informazione sull'entità e la direzione del movimento, ma non dove nell'immagine questo avvenga, quindi si dovrà cercarlo, ma l'efficienza della ricerca è incredibilmente superiore a prima.

Slices

- Non sono necessarie per la compressione, ma sono molto utili nel caso di errori nel canale.
- Risincronizzano il flusso video
- Permettono di definire aree di interesse



ISO/IEC 13818-2 e ITU-T H.262 : MPEG-2



Requisiti:

- 4 ~ 15 Mbit/sec;
- supporta video interlacciato;
- TV broadcast digitale, TV cavo/satellite, HDTV, servizi video in rete (es. ATM);
- ricerca e playback avanti-indietro e veloce;
- video codifica scalabile per diversi livelli di qualità;
- sincronizzazione audio/video per molteplici stream;
- strutturazione dei pacchetti per trasmissione digitale;
- audio esteso per effetti surround (Dolby) e multilingua.

Profili di qualità di MPEG-2

Profile	Target Bit-Rate	Quality	Application
Low	4 Mb/s	352 x 240 x 30	VHS equivalent
Main	15 Mb/s	720 x 480 x 30	Broadcast quality
High - 1440	60 Mb/s	1440 x 1152 x 30	HDTV
High	80 Mb/s	1920 x 1080 x 30	Film production

Livelli di qualità e applicazioni

Level	Application
Simple	Same as main but without B-frames, primarily intended for software decoders
Main	Low-cost single chip implementation for cable TV and satellite uplink compression
Spatial	Main with spatial scalability, e.g. HDTV
SNR	Spatial with SNR scalability
High	SNR with 4:4:4 chrominance in the macroblocks

Profili e Livelli

HIGH		4:2:0 1920x1152 80 Mb/s I,P+B			4:2:2 + 4:2:0 1920x1152 100 Mb/s I,P+B
		4:2:0 1440x1152 60 Mb/s I,P+B		4:2:0 1440x1152 60 Mb/s I,P+B	4:2:2 + 4:2:0 1440x1152 80 Mb/s I,P+B
MAIN	4:3:0 720x576 15 Mb/s I+P	4:2:0 720x576 15 Mb/s I,P+B	4:2:0 720x576 15 Mb/s I,P+B		4:2:2 + 4:2:0 720x576 15 Mb/s I,P+B
		4:2:0 352x288 4 Mb/s I,P+B	4:2:0 352x288 4 Mb/s I,P+B		
LEVEL PROFILE	SIMPLE	MAIN	SNR	SPATIAL	HIGH