

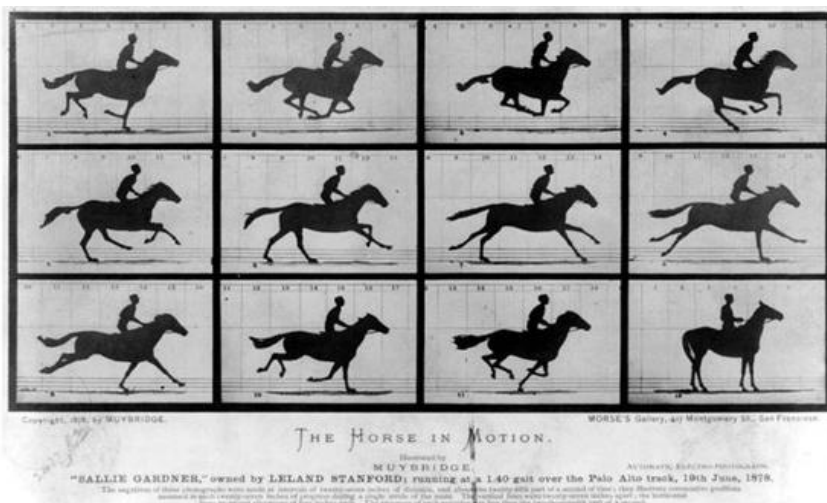
# Video Digitale

Prof. Sebastiano Battiato

Multimedia



## E. Muybridge: le prime sequenze

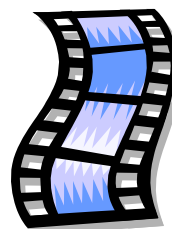


Multimedia



## Sequenze Video

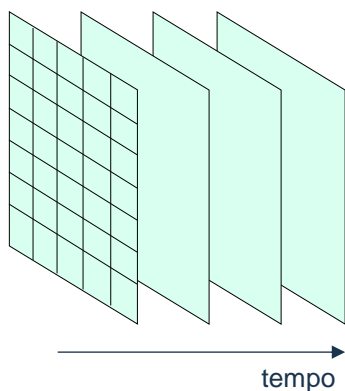
- Video come segnale discreto
  - ✓ Campionamento temporale della scena
  - ✓ Ad ogni istante la scena è “fotografata”
  
- Sequenza video
  - ✓ Successione di istantanee
  - ⇒ Fotogrammi (*frame*)



Multimedia



## Sequenze Video



- Sequenza video
  - ✓ Successione di frame
  
- Frame
  - ✓ Pellicola tradizionale (cinema)
  - ✓ Immagini digitali

Multimedia



## Codifica Video

Oltre che dalle necessità di contenere le dimensioni della rappresentazione delle immagini, a fine sia di memorizzazione sia di trasmissione, la **compressione di sequenze di immagini e video** trae origine dalla necessità di garantire la riproduzione delle sequenze con un adeguato *framerate*, quando la visualizzazione è on line rispetto alla trasmissione, come in alcune applicazioni killer quali la TV digitale o i sistemi di videoconferenza.

Multimedia



## Introduzione

Per esempio, lo standard TV PAL richiede la riproduzione, 25 volte al secondo, di frame di dimensioni pari a 720 (colonne) x 576 (righe). A 24 bpp questo darebbe luogo, anche trascurando altri aspetti non secondari della elaborazione del segnale TV, ad un bit rate minimo di  $720 \times 576 \times 25 \times 24 \approx 31$  Mbyte/sec, assolutamente improponibile per usi consumer.

L'estensione di metodi di compressione come quelli già visti per le immagini fisse, per esempio il JPEG, porta ad algoritmi di codifica che, sfruttando solo la ridondanza spaziale, non garantiscono qualità sufficiente ai bassi bit rate

Multimedia



## Standard di codifica

Per poter trasmettere o memorizzare dei file video è necessario definire degli standard riguardanti sia gli algoritmi di codifica/decodifica dei flussi multimediali, sia i protocolli necessari al loro trasferimento e al loro controllo sulla rete.

Gli standard cosiddetti “aperti” devono prevedere:

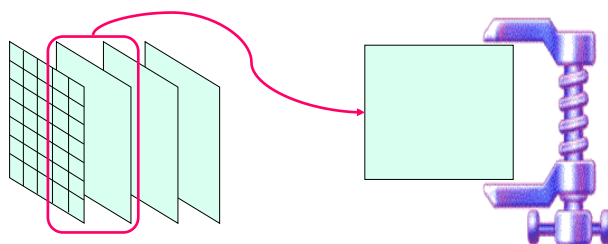
- ✓ La definizione e adozione di specifiche universalmente accettate;
- ✓ La definizione dei formati per dati multimediali e degli algoritmi utilizzati per codificarli.

Multimedia



## Codifica di Sequenze

- Sequenze come collezioni di immagini
  - ⇒ Frame codificati come immagini separate (JPEG, GIF, ...)



- Senza sfruttare la ridondanza temporale i risultati ottenuti sono scarsi.

Multimedia



## Codifica di Sequenze

- Il video non è solo una collezione di immagini
- In genere, la scena cambia lentamente e in parte
  - ⇒ Coerenza temporale tra fotogrammi vicini
  - ⇒ Ridondanza nel tempo (tra frame)

⇒ Tracciare il moto degli oggetti nella scena



Multimedia



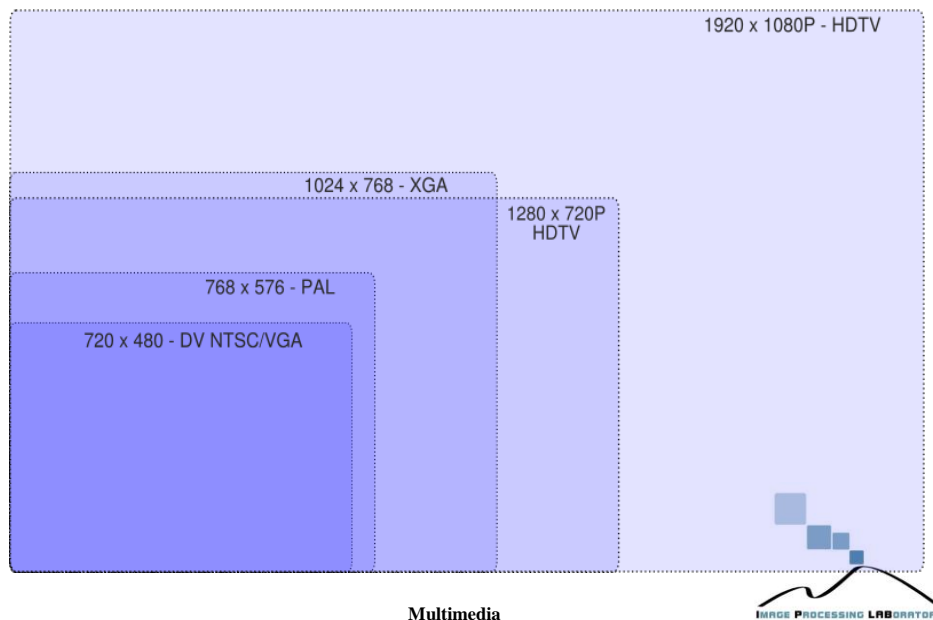
## HDTV

- High Definition TeleVision
- Aspect ratio 16:9
- Quattro formati
  - ✓ Half resolution HDTV, 960x540 pixel
  - ✓ 720p, 1280x720
  - ✓ 1080i, 1920x1080 interlacciato
  - ✓ Full HD, 1920x1080
- Sei volte più informazioni (per frame) rispetto al formato DVD

Multimedia



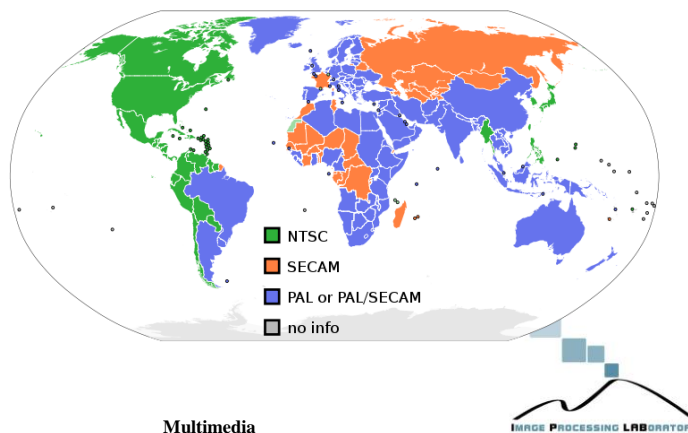
## Risoluzione dei formati



## Diverse risoluzioni

Almeno tre categorie di risoluzioni per i contenuti video:

- ✓ PAL o 576i;
- ✓ NTSC o 480i;
- ✓ SECAM;



## ASPECT RATIO

- L'aspect ratio è il rapporto tra larghezza/altezza dell'immagine.
  - ✓ **4:3** è il formato standard televisivo.  
Con la nascita dei Dvd e dei nuovi formati televisivi digitali è destinato alla scomparsa.
  - ✓ **16:9 Widescreen** è caratterizzato da dimensioni orizzontali più ampie del 4:3, con le proporzioni panoramiche tipiche dello schermo cinematografico. Ne esistono varianti più o meno allargate.
  - ✓ Il **Letterbox** permette di visualizzare il widescreen su schermi 4:3. L'immagine viene scalata fino a farlo rientrare nello schermo con l'aggiunta di due bande nere orizzontali sopra e sotto il video.
  - ✓ **Pan&Scan** è un'altra tecnica per visualizzare il widescreen su 4:3. La differenza col letterbox sta nel mostrare solo la parte centrale dell'immagine tagliandola sui lati.

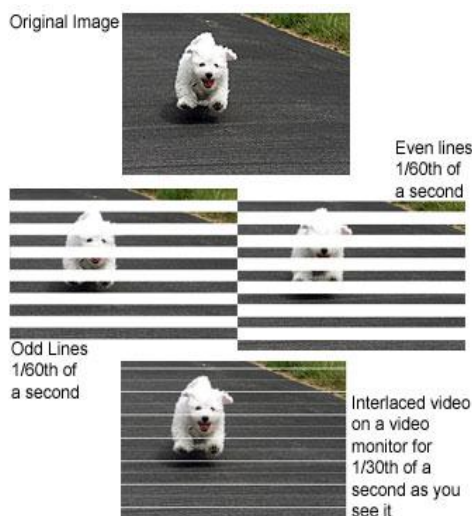


Multimedia



## Interlacing

- Immagine divisa in *field* pari o dispari
- Si alternano i *field* per metà tempo rispetto ai *fps* del filmato
- Su monitor si notano artefatti, sul televisore non accade



Multimedia



## Interlacing e deinterlacing

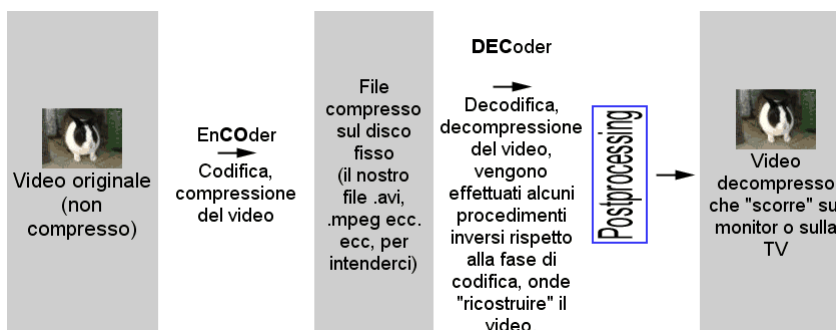


Multimedia



## La compressione video

Un *CoDec video* (*Co-Dec = Coder/Decoder*) è un software composto da due parti: l'**enCoder** che comprime la sequenza di immagini (video) archiviandola in un file ed un **Decoder** necessario per decomprimere la sequenza e poterla nuovamente visualizzare.



Multimedia





## Lossless/Lossy

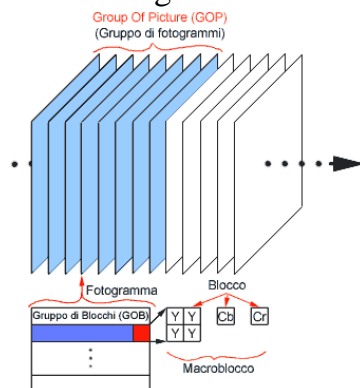
Anche le tecniche di compressione video possono essere suddivise in **tecniche *lossless***, dove la compressione è un processo perfettamente reversibile che avviene senza perdita di informazione e **tecniche *lossy*** dove la compressione non è reversibile, nelle quali video compresso e decompresso non sono più perfettamente identici in quanto al momento della compressione sono state volutamente eliminate alcune informazioni ritenute "sacrificabili".

Multimedia



## Video

Un video è costituito da una successione di immagini che si susseguono in rapida sequenza. Quindi quando si comprime un flusso video si stanno sostanzialmente comprimendo delle immagini.



Multimedia



## Come è possibile comprimere un video?

*Si utilizzano tecniche che sfruttano alcune caratteristiche intrinseche del video stesso, in combinazione con le caratteristiche del sistema visivo umano.*

In particolare è possibile comprimere un segnale video attaccando:

- la ridondanza spaziale e la ridondanza temporale
- le caratteristiche del sistema visivo umano

Multimedia



## Come è possibile comprimere un video?

In particolare è possibile comprimere un segnale video:

- *Rimuovendo la ridondanza (ripetitività) statistica contenuta in un video* e mantenendo solo le informazioni effettivamente utili; si cerca una rappresentazione "meno correlata" delle immagini, eliminando le "ripetizioni".

- ✓ Si può dimostrare che **pixels adiacenti**, vicini, all'interno di una stessa immagine, presentano caratteristiche molto simili per quel che riguarda il colore e la luminosità; la codifica intra-frames si occupa di rimuovere questa ripetitività altresì detta ridondanza spaziale all'interno dello stesso fotogramma.

Multimedia



## Come è possibile comprimere un video?

- ✓ Esiste inoltre una netta **correlazione non solo tra i pixel dello stesso fotogramma**, ma anche tra i **pixels di fotogrammi adiacenti**: un fotogramma ed i due vicini (il successivo ed il precedente) spesso risultano pressoché identici (fanno eccezione le situazioni in cui si hanno cambi di scena); questa **ridondanza temporale** tra fotogrammi vicini che ne sfrutta le loro minime differenze, viene trattata dalla **codifica inter-frames**.

Multimedia



## Come è possibile comprimere un video?

- ***Sfruttando alcune peculiarità del sistema visivo umano***: la **scarsa sensibilità dell'occhio alle alte frequenze video** soprattutto se si tratta di immagini in movimento. E' possibile "tagliare", buttar via, alcune informazioni soprattutto relativamente alle alte frequenze di un'immagine senza introdurre artefatti visibili. Il sistema visivo umano non è infatti in grado di percepire le variazioni nei dettagli di figure molto frastagliate; è molto difficile rendersi conto di una perdita di dettaglio nelle fronde di alcuni alberi in movimento; molto più semplice invece notare anche la più piccola variazione di colore o luminosità nell'azzurro di un cielo limpido e sereno sullo sfondo di un video.

Multimedia



## Compressione statica

- Conversione dello spazio di colore: da RGB a YUV 4:1:1
- **Blocking**: l'immagine viene suddivisa in macroblocchi di 16x16 pixels
- DCT, Discrete Cosine Transform
- Quantizzazione
- Zig-Zag Scanning (Scansione a Zig-Zag)

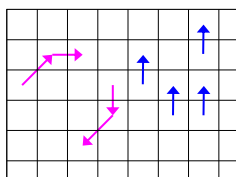
**Sostanzialmente si esegue una compressione in stile JPEG!!!!**

Multimedia



## Block Matching

- Esplicita l'idea di rendere uniforme il moto di pixel vicini
  - ✓ Frame partizionato in blocchi non sovrapposti
  - ✓ Un vettore di moto per ogni blocco

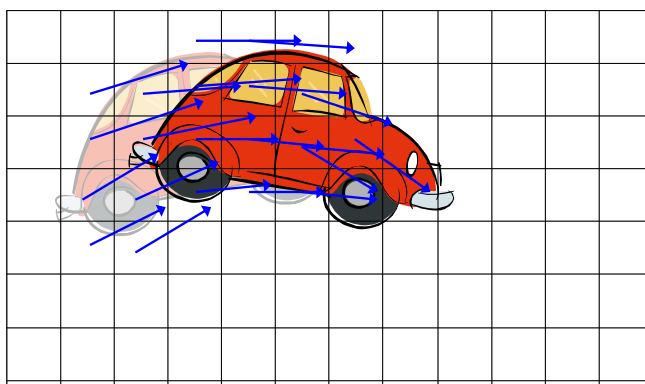


Multimedia



## Esempio

- Vettori di moto

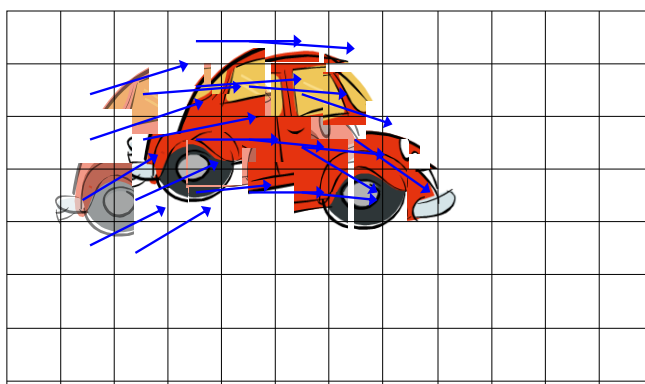


Multimedia



## Esempio

- Compensazione del moto (per blocchi)

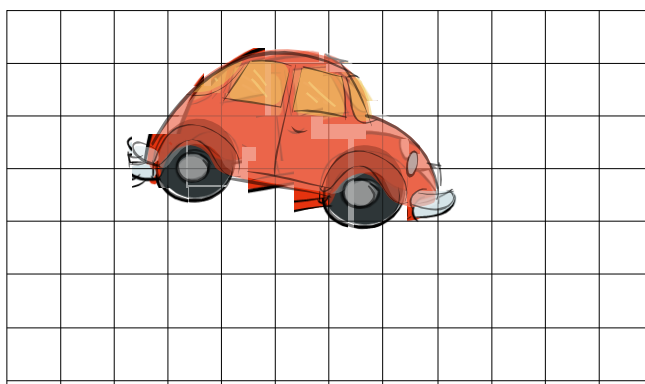


Multimedia



## Esempio

- Differenza tra *previsione* e *frame reale*

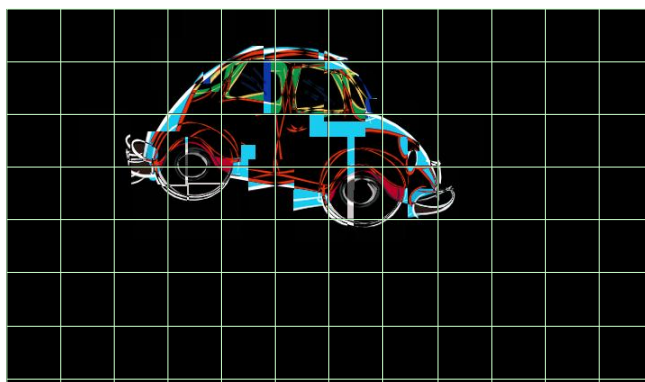


Multimedia



## Esempio

- Errore di compensazione

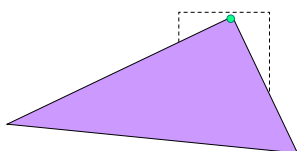


Multimedia



## Stima per Blocco

- Ricerca del blocco corrispondente (*block matching*)
  - ✓ Ricerca esaustiva in una finestra intorno al pixel centrale

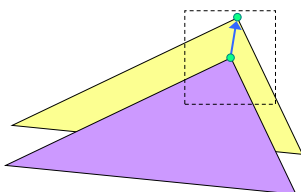


Multimedia



## Stima per Blocco

- Ricerca del blocco corrispondente (*block matching*)
  - ✓ Ricerca **esaustiva** in una finestra intorno al pixel centrale
- Ecco il vettore di moto!



Multimedia



## Stima per Blocco

Si noti che la ricerca per il blocco di riferimento viene generalmente effettuata in un **intorno limitato** del blocco di partenza in quanto se la ricerca fosse effettuata su tutto il fotogramma, i tempi per la codifica potrebbero allungarsi in modo non accettabile; l'efficienza del metodo è assicurata dal principio della **ridondanza spaziale**: *la probabilità di trovare un blocco simile man mano che ci si allontana dal blocco di partenza, diminuisce in modo esponenziale con l'aumentare della distanza.*

➤ Esistono algoritmi di matching più veloci

- ✓ 2D-log
- ✓ Ricerca in tre passi

Multimedia



## 2D-Log

- Si parte dalla posizione corrispondente a moto nullo
- Si verificano 5 punti a stella
- Si ripete centrando su quello che da errore minore
- La finestra di ricerca è dimezzata se il punto migliore è
  - ✓ quello centrale
  - ✓ sul bordo della massima finestra di ricerca

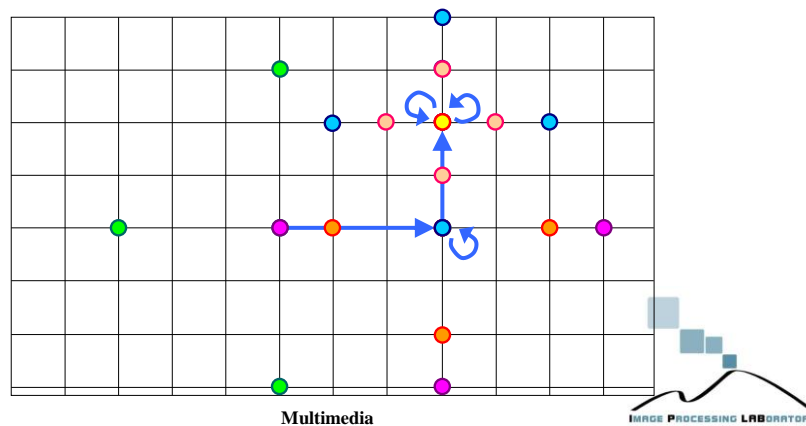
Multimedia





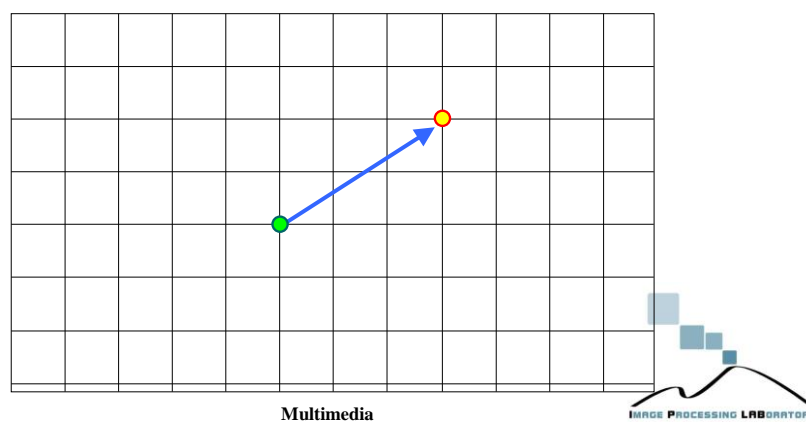
## 2D-Log - Esempio

- Ricerca della corrispondenza migliore



## 2D-Log - Esempio

- Vettore di moto



## Ricerca in Tre Passi

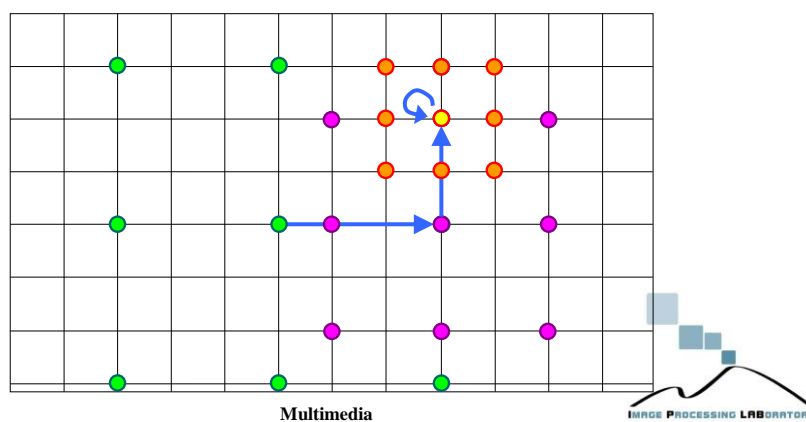
- Si parte dalla posizione corrispondente a moto nullo
- Si verificano 9 punti a stella e sul bordo
- Si ripete centrando su quello che da errore minore
- La finestra di ricerca è dimezzata ad ogni passo

Multimedia



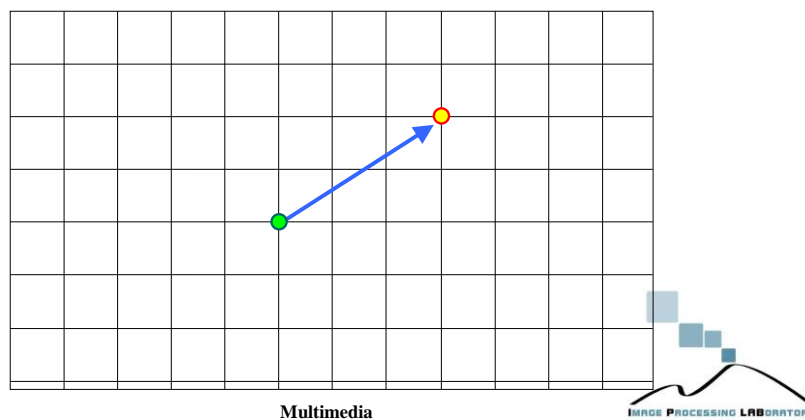
## Ricerca in Tre Passi: Esempio

- Ricerca della corrispondenza migliore



## Ricerca in Tre Passi: Esempio

- Vettore di moto



## Stima per Blocco

- Modella
  - ✓ Moto 2D
  - ✓ Traslazione
- Non modella
  - ✓ Rotazione
  - ✓ Traslazione lungo l'asse ottico
  - ✓ Movimenti di camera (pan, zoom, tilt, ...)

Multimedia



## Compressione dinamica

- *Motion Estimation* (ME) e *Motion Compensation* (MC)
- *Blocking*: l'immagine viene suddivisa in macroblocchi di 16x16 pixels
- DCT, Discrete Cosine Transform
- Quantizzazione
- Zig-Zag Scanning (Scansione a Zig-Zag)

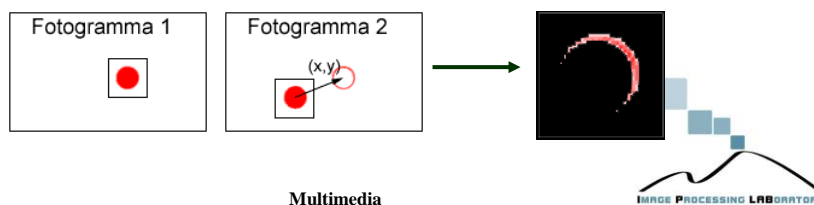
Multimedia



## Motion Estimation (ME)

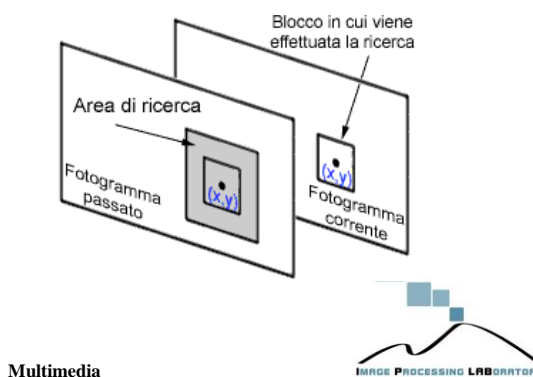
L'encoder individua tra i fotogrammi adiacenti (nel solo fotogramma precedente, o nel precedente e nel successivo a seconda dei casi) **il blocco più simile** (se non uguale).

Dopodichè viene associato al blocco su cui è stata effettuata l'analisi, un **vettore di moto**, cioè una coppia di numeri tipo  $(x,y) = (-1,4)$  che individuano sul piano ipotetico rappresentato dal fotogramma, il vettore di spostamento, che indica verso e entità dello spostamento del blocco passando dal fotogramma 1 al fotogramma 2.



## Motion Compensation (MC)

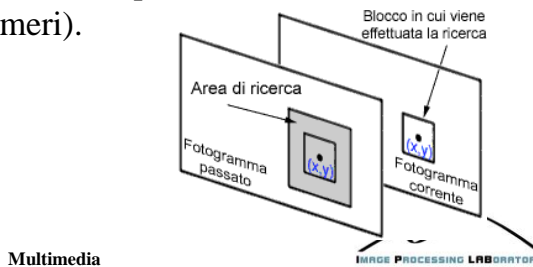
- Viene creato il **blocco differenza**, in pratica viene sostituito il blocco originale su cui è stata effettuata la ricerca, con il risultato che si ottiene sottraendo dal blocco molto simile trovato del fotogramma 1 (reference frame) il blocco in questione nel fotogramma 2.



## Motion Compensation (MC)

Se tutto ha funzionato a dovere, ovvero l'encoder non ha commesso errori nella ricerca del blocco di riferimento (nel fotogramma 1), si ottiene un netto vantaggio consistente nel fatto che il nuovo blocco "differenza" sarà costituito da un numero decisamente inferiore di dati.

Saranno infatti presenti molti zeri, l'entropia sarà minore, la codifica più efficiente. Le uniche informazioni aggiuntive da considerare saranno quelle relative al vettore di moto (una coppia di numeri).



## ITU



L' *International Telecommunication Union* è l'agenzia delle Nazioni Unite specializzata nel campo delle telecomunicazioni.

L' *ITU-Telecommunication Standardization Sector (ITU-T)* è l'organo responsabile degli studi tecnici, dei problemi tariffari ed emette "*Recommendations*" per standardizzare le telecomunicazioni in tutto il mondo.

Multimedia



## ITU-T SG16 Q.6

Lo *Study Group 16*, conosciuto come *Video Coding Expert Group (VCEG)*, è attivo nella standardizzazione di ogni aspetto del multimedia (architetture, protocolli, tecnologie mobili, qualità del servizio...).

Ha focalizzato i suoi sforzi su IPTV, sistemi di videoconferenza, codifica di audio e video.

Ha prodotto la "*V-series modem Recommendation*" che ha standardizzato i modem prima dell'avvento dell'ISDN.

Ha prodotto standard anche per la tecnologia fax e il protocollo VoIP.

Multimedia



## ISO



L'*International Organization for Standardization* è una rete degli istituti nazionali di standardizzazione di 163 Paesi composta da membri governativi e privati.

E' impegnata nella produzione di standard qualitativi in svariati settori.

Opera in collaborazione con la *International Electrotechnical Commission (IEC)*

Multimedia



## ISO/IEC JTC 1/SC 29/WG 11

Il gruppo di lavoro 11 del subcomitato 29 del Joint Technical Group 1 di ISO/IEC è anche chiamato *Moving Picture Expert Group*, in breve **MPEG**.

Si occupa di definire standard per la rappresentazione in forma digitale di audio, video e altre tipologie di contenuti multimediali in un'ampia varietà di applicazioni.

Multimedia



## Video Standard

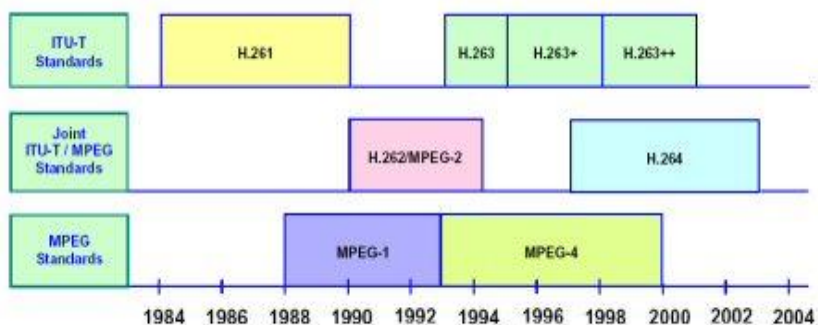


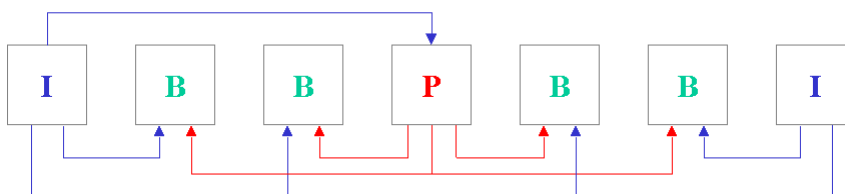
Figure 1. Progression of the ITU-T Recommendations and MPEG standards.

Multimedia



## Frames I/P/B

- Gli standard MPEG prevedono la classificazione dei frame in tre tipi: I, B, P .
- *I frame* è un frame video completamente indipendente.
- *P frame (predictive frame)* si basa su un precedente I frame.
- *B frame (bi-directional frame)* è costituito da informazioni ricavate sia da I frame che P frame (anche successivi) attraverso interpolazione .



Multimedia

IMAGE PROCESSING LABORATORY



## Intra-Frames

I fotogrammi di tipo **I**, chiamati anche **Intra-Frames** o **Key-Frames** (fotogrammi chiave), sono fotogrammi che vengono codificati utilizzando le informazioni contenute nel fotogramma stesso e non contengono nessun riferimento od informazione sui fotogrammi adiacenti; in pratica sono compressi alla stregua di un'immagine singola, allo stesso modo di quando un'immagine viene salvata in formato JPEG. Nessun tipo di compressione temporale (ovvero compressione che tiene conto anche dei fotogrammi successivi e/o precedenti) viene applicata a questi fotogrammi.

Multimedia



## Intra-Frames

In genere i fotogrammi chiave vengono inseriti dal codec ogni qualvolta vi sia un **repentino cambiamento** tra due immagini successive. Se inoltre viene specificato un **intervallo massimo** tra un fotogramma chiave ed il successivo il codec dovrà necessariamente inserire un fotogramma chiave anche se non strettamente necessario.

Multimedia



## P-Frames

Il fotogramma P, (**Predicted frames**) viene codificato utilizzando informazioni acquisite in base al fotogramma che lo precede, sia questo di tipo I o di tipo P. Ogni macroblocco di 16x16 pixels di un P-Frame **può essere codificato** in modo indipendente (come nel caso dell'I-Frame) oppure può essere **compensato**, cioè bilanciato utilizzando informazioni del fotogramma precedente.

Utilizzando le somiglianze tra fotogrammi successivi i fotogrammi P risultano essere più piccoli dei corrispondenti I-Frames.

Multimedia



## P-Frames

Un fotogramma di tipo P **contiene le informazioni della posizione** (X',Y') nel fotogramma corrente in cui si è spostato un blocco che aveva coordinate (X,Y) in quello precedente (Motion Estimation/Compensation).

Lo **svantaggio** dell'utilizzo di questo tipo di fotogrammi si ha in fase di **decodifica**; è infatti necessario "ricostruire" ciascun fotogramma P prima di poterlo visualizzare, e per far questo si deve sempre partire dal fotogramma P seguente all'ultimo fotogramma chiave.

Multimedia



## B-frames / Bi-directional encoding

Per i fotogrammi di tipo **B** la ricerca del moto (motion estimation/ compensation) è effettuata non solo sul fotogramma **precedente** (come nel caso di P-Frames) ma anche sul fotogramma **successivo**. La codifica ed anche la decodifica risultano quindi decisamente più complesse.

Sostanzialmente i fotogrammi B sono di tipo "**Bidirezionale**", nel senso che fanno riferimento sia a ciò che li precede, sia a quello che segue. Inserire in un fotogramma informazioni che si riferiscono ad un fotogramma successivo è possibile solo alterando l'ordine in cui i fotogrammi vengono archiviati all'interno del file video compresso.

Multimedia



## Esempio

Supponiamo di avere 4 fotogrammi da comprimere.

Il primo di questi sarà necessariamente un fotogramma chiave (I-Frame), mentre vogliamo che i successivi due siano B-Frames (che generalmente hanno una dimensione di 1/4 del P-Frame corrispondente). L'ultimo deve essere necessariamente un P-frame, in quanto i fotogrammi B necessitano dopo di loro di qualcosa da cui essere derivati. In sequenza avremmo:

n° fotogramma	1	2	3	4
tipo di fotogramma:	I	B	B	P

I fotogrammi verranno archiviati all'interno del filmato in questo modo:

n° fotogramma	1	4	2	3
tipo di fotogramma:	I	P	B	B

Multimedia



## Esempio

Dopo aver codificato l'I-Frame, l'encoder salta avanti di due fotogrammi e codifica quello che è destinato ad essere il fotogramma P (ovvero il quarto) e lo codifica come se seguisse immediatamente l'I-frame:

n° fotogramma	1	4
tipo di fotogramma:	I	P

Questo processo genererà un P-frame di dimensioni superiori a quello che si avrebbe codificando come P-frame il 2° fotogramma, in quanto generalmente vi saranno più cambiamenti (ovvero differenze) tra il 1° fotogramma ed il 4° che non tra il 1° ed il 2°. Tuttavia, l'utilizzo dei due B-frame porterà complessivamente ad una riduzione del numero di informazioni (dimensioni) necessarie alla codifica (come ho già detto un B-frame occupa 1/4 delle dimensioni di un P-frame).

Multimedia



## Lo standard MPEG

Nel **1988** l'organismo internazionale di standardizzazione ISO-IEC (*International Organization for Standardization and International Electrotechnical Commission*) si è assunto il compito di sviluppare uno standard per la compressione e la rappresentazione del video digitale e dell'audio ad esso associato che fosse adatto alla memorizzazione su dispositivi di memoria di massa (dischi ottici, DAT) e alla trasmissione su canale di telecomunicazione (ISDN, LAN, TV).

Il *Moving Picture Expert Group* (MPEG) è il comitato internazionale nato in seno all'ISO per raggiungere tale obiettivo. Formalmente MPEG è il gruppo di lavoro 11 del subcomitato 29 del Joint Technical Group 1 dell'ISO-IEC

Multimedia



## Il formato MPEG-1

Questo formato è nato per rispondere all'esigenza di memorizzare filmati su compact disc. E' stato definito come standard nel 1992 ed utilizza un bitrate costante di 1,15 Mbit/sec per il video e dai 384 ai 198 Kbit/sec per l'audio.

Da un punto di vista qualitativo, l'obiettivo prefissato era il raggiungimento della qualità VHS. In pratica, per ottenerla, il video è codificato a 352x288 pixel (288 linee orizzontali da 352 punti ciascuna) per quanto riguarda la luminosità, mentre per quanto riguarda il colore l'immagine è ulteriormente divisa per due ed è pertanto codificata a 176x144

Multimedia



## Il formato MPEG-1

Per ottenere il video con la qualità SIF(352x240), il codec MPEG-1 effettua una serie di operazioni di compressione delle immagini che sfruttano non solo l'algoritmo *DCT*, ma anche le differenze tra un fotogramma e l'altro.

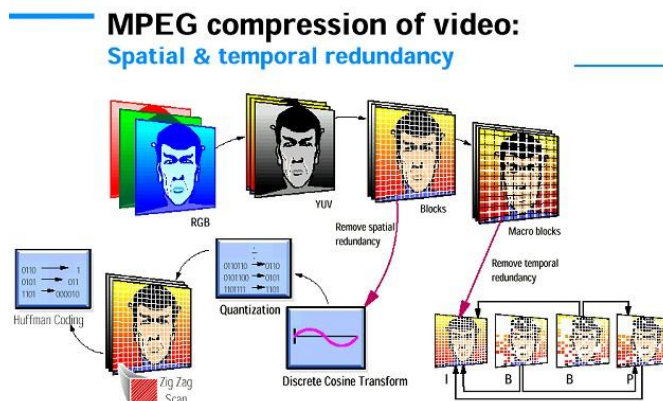
Anzichè memorizzare tutti i fotogrammi per intero, se ne memorizzano soltanto alcuni come tali (ad intervalli prefissati e regolari), e tra di essi ci si limita a memorizzare una serie di frames incompleti nei quali vengono "scritte" solo le informazioni che subiscono delle variazioni rispetto alle immagini precedenti.

Multimedia



## Il formato Mpeg-1

Riassumendo, lo schema di codifica MPEG-1 è:



MPEG1 gestisce solo frame (progressive scan). L'ingresso di MPEG1 è il SIF (352x240), composto solo da 1 field

Multimedia



## Il formato MPEG-2

- Questo standard è stato sviluppato partendo dall'MPEG-1 ed esiste dal 1994. L'obiettivo dell'MPEG-2 era quello di creare un formato flessibile ed adatto a varie applicazioni, in grado anche di codificare in digitale le immagini con una qualità equivalente a quella analogica definita come broadcast (corrispondente alla qualità delle trasmissioni televisive), e l'audio con quella cinematografica, utilizzando flussi di dati fino a 60 Mbit/sec.
- La caratteristica principale dell'MPEG-2 è la sua scalabilità, ovvero la possibilità di creare soluzioni di codifica e decodifica più o meno complesse in base al tipo di prodotto da realizzare, aggiungendo poi altre caratteristiche quali la possibilità di trasmettere il flusso multimediale su reti a larga banda, assicurando una buona robustezza nei confronti degli errori della rete, il trasporto parallelo di molteplici canali audio, le funzioni di protezione e di controllo di accesso al flusso, per citarne i principali.

Multimedia



## Il formato MPEG-2

Per consentire all'industria di procedere gradualmente con l'implementazione dello standard, il comitato di lavoro dell'MPEG ha definito una serie di livelli e di profili in base ai quali ogni soluzione tecnica può essere sviluppata e verificata. Non tutte le combinazioni portano ad un sottoinsieme di specifiche valide, per questo i cinque profili ed i quattro livelli si combinano solo in 11 soluzioni (e non già venti).

I profili sono: **Simple (SP)**, **Main (MP)**, **SNR**, **Spatial Scalable**, **High**.

Mentre i livelli sono **Low (LL)**, **Main (ML)**, **High 1440 (H-14)**, **High (HL)**

Profilo	Livello	Pixel Orizzontali	Pixel Verticali	Frame Rate Max	Bitrate Max (Mbit/sec)
Simple	main	720	576	30	15
Main	low	352	288	30	4
Main	main	720	576	30	15
Main	high 1440	1440	1152	60	60
Main	high	1920	1152	60	80
SNR Scalable	low	352	288	30	3(4)
SNR Scalable	main	720	576	30	10(15)
Spatially Scal.	high 1440	720 (1440)	576 (1152)	30 (60)	15(40 o 60)
High	main	352 (720)	288 (576)	30 (30)	4 (15 o 20)
High	high 1440	720 (1440)	576 (1152)	30 (60)	20 (60 o 80)
High	high	960 (1920)	576 (1152)	30 (60)	25 (80 o 100)

Multimedia



## II MPEG-2

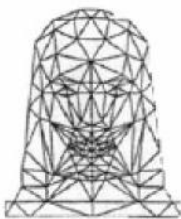
- MPEG-2 introduce i concetti di *frame picture* e di *field picture*, associando ad essi i metodi di codifica basati sulla predizione sul frame e predizione sul field.

Multimedia



## MPEG4

- L' MPEG4 usa fundamentalmente lo stesso algoritmo di compressione di MPEG1 e MPEG2, ma in modo molto più efficiente.
- La differenza sostanziale è che il sistema riesce a distinguere i **vari livelli di un immagine**, lo sfondo e i primi piani. Se lo sfondo rimane uguale nei fotogrammi successivi non verranno memorizzati, risparmiando così prezioso spazio.
- Inoltre è possibile elaborare queste immagini più semplicemente, **estrapolando gli attori o gli oggetti dallo sfondo** con grande facilità.



Multimedia

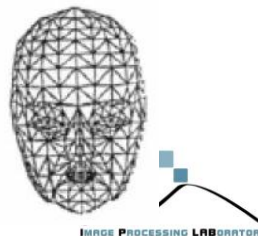


IMAGE PROCESSING LABORATORY

## H.264

- Codec di nuova generazione
  - ✓ aka: MPEG-4 part 10 o AVC (Advanced Video Coding)
  - ✓ standard congiunto ITU-T Video Coding Experts Group – ISO Moving Picture Experts Group
- A parità di bitrate, qualità percepita e risoluzione raddoppiano rispetto a MPEG2
- Incremento della complessità di codifica e decodifica
- I supporti HD-DVD e Blue Ray adottano H.264

Multimedia





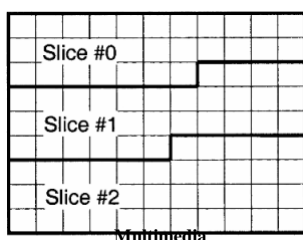


## Slice

Una **slice** è un insieme di macroblocchi che vengono processati in sequenza.

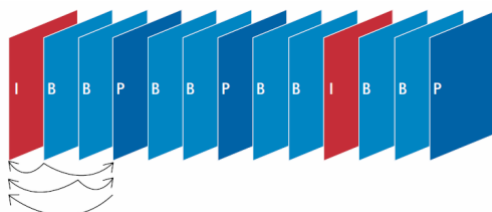
Un frame può essere suddiviso in una o più slice.

Una slice è “self-contained”, cioè non sono necessarie informazioni di altre slice per decodificarla, questo permette di decodificarle in qualsiasi ordine. Le informazioni di altre slice potrebbero essere necessarie per la fase di deblocking.



## Frames

Come negli standard precedenti, anche H.264 divide lo stream in GOP (Group Of Picture). Un GOP è composto da I-frame, P-frame e B-frame (inizia sempre con un I-frame).



Multimedia

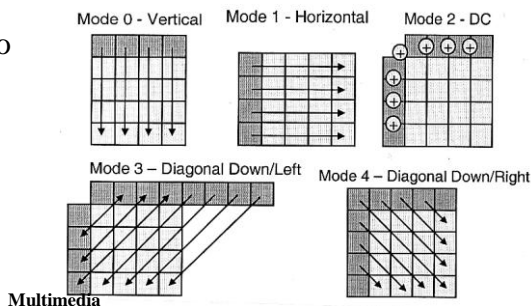


## Intra-frame

H.264 introduce un **intra-prediction scheme** per ridurre la dimensione degli I-frame.

Per codificare un blocco 4x4 all'interno di un macroblocco, si verifica se i pixel del bordo fanno match con i pixel dei blocchi adiacenti all'interno dello stesso macroblocco in 9 modi diversi, in tal caso vengono copiati (tranne per DC in cui si calcola la media).

I modi non rappresentati sono le altre diagonali.



## Intra-frame



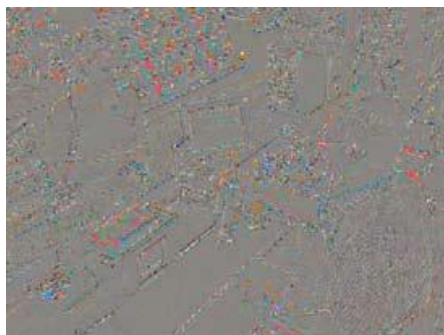
•(A) Immagine originale



• (B) Immagine intra-predicted

•(A) viene codificata per similitudine tra blocchi adiacenti ottenendo (B), infatti si può notare dell'aliasing lungo gli edge.

## Intra-frame



(C) Immagine residuo



(D) Immagine di output

(C) È usato per ricreare da (B) l'immagine di output (D).

Si nota che (C) contiene le alte frequenze.

Nello stream un I-frame è codificato come la coppia [(B),(C)].

Multimedia



## Inter-Frame Prediction nei P-frame

I frame Predicted sono codificati per compensazione del moto dagli I-frame.

Il moto è compensato per partizioni di macroblocchi; per la luma le dimensioni permesse sono 16x16, 16x8, 8x16 e 8x8, mentre per la chroma resta solo 8x8. Nel caso di partizioni 8x8, si può effettuare un ulteriore partizionamento come in figura.

	16x16	16x8	8x16	8x8
M Types	0	0 1	0 1	0 1 2 3
	8x8	8x4	4x8	4x4
8x8 Types	0	0 1	0 1	0 1 2 3

Multimedia



## Inter-Frame Prediction nei P-frame

In decodifica, la compensazione avviene cercando l'area corrispondente nel frame di riferimento indicato da un indice e calcolando il vettore di moto.

L'accuratezza del vettore di moto è al **quarto** di pixel. Nel caso il vettore non sia intero, i pixel delle posizioni "non intere" sono ottenuti per interpolazione applicando un filtro specifico nel caso della luma, per interpolazione bilineare la chroma.

La possibilità di compensare il moto ad un quarto di pixel, a metà o all'intero è il maggior miglioramento dello standard rispetto ai precedenti, introduce maggiore precisione e maggiore flessibilità.

I vettori di moto possono puntare all'esterno del frame.

Multimedia

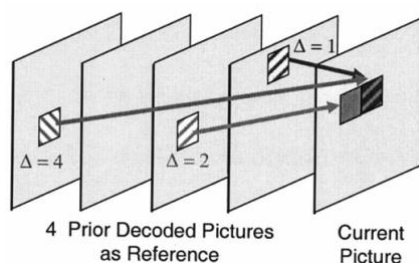


## Inter-Frame Prediction nei P-frame

La finestra di ricerca per i macroblocchi è grande 2048x2048 pixel, inoltre si tiene conto di **resizing** e di **rotazioni**.

Come in H.263++, non si ha un solo frame di riferimento, ma può esserlo qualsiasi frame precedente (che si trova già nel buffer dell'encoder o del decoder).

Dove non si ha alcun match, si usano blocchi intra-coded.



Multimedia



## Inter-Frame Prediction nei B-frame

Nella codifica di un blocco, i frame Bi-direzionali possono usare un valore medio ottenuto dalla compensazione da due frame di riferimento differenti.

Anche i macroblocchi dei B-frame sono partizionati con lo schema visto in precedenza. Il vettore di moto è calcolato come per i P-frame

Multimedia



## Quantizzazione

H.264 utilizza la DCT con alcune modifiche per codificare un video e farne una quantizzazione.

La DCT è calcolata su blocchi 4x4 per ridurre al minimo l'effetto di ringing.

Inoltre i blocchi delle basi possono essere ingranditi in due modi:

- ✓ con una trasformata gerarchica che estende la dimensione dei blocchi usati per codificare la chroma a basse frequenze a 8x8;
- ✓ Permettendo all'encoder di estendere la dimensione dei blocchi usati per le basse frequenze della luminanza a 16x16;

Multimedia



## Deblocking filter

La codifica per blocchi e macroblocchi introduce degli artefatti lungo i margini di essi. H.264 introduce un deblocking filter all'interno del ciclo di stima del moto per ridurre gli artefatti.

Questo filtro effettua uno smoothing dei margini dei blocchi con un algoritmo adattivo: se il valore assoluto della differenza tra i valori dei pixel è molto elevata rispetto a quello che si possa ottenere dalla quantizzazione, si ipotizza di essere in presenza di un bordo dell'immagine, quindi non si applica il filtro (protegge le alte frequenze);

Multimedia



## Deblocking filter



Multimedia



## Profili

H.264 contiene un insieme di video coding tools differenti tra loro. In base all'applicazione per cui bisogna codificare il video, si sceglie un sottoinsieme di essi chiamato "profilo".

In questo modo si può realizzare un encoder snello e ottimizzato per il determinato scopo includendo solo gli strumenti necessari, allo stesso modo si limita la complessità del decoder.

Per esempio un video per cellulari può sfruttare la bassa risoluzione dei dispositivi, un video tratto da una telecamera di sicurezza deve avere una buona risoluzione per identificare i volti.

Multimedia



## Baseline Profile

Questo profilo è stato ideato per applicazioni in sistemi dove la potenza di calcolo è limitata.

Il BP è il più flessibile, da buone performance in un real-time encoder e può essere integrato in dispositivi come telecamere di rete.

Il profilo offre bassa latenza, il che è importante in applicazioni real-time e rende disponibile il controllo pan/tilt/zoom (PTZ) nelle PTZ network cameras.

Questo profilo non supporta: B-frame, CABAC, field coding, PAFF e MBAFF, slice data partitioning.





## High Profile

Incluso nella specifica del BD-ROM Video della Blue-ray Disc Association, nella specifica di HD-DVD del DVD Forum e negli standards DVB (digital video broadcast) delle trasmissioni televisive europee.

Si pensa che in breve HP supererà MP nell'interesse delle implementazioni a breve termine dell'industria in quanto HP incrementa la compressione senza complicare troppo l'implementazione rispetto a MP.

Supporta video a 8-bit con campionamento 4:2:0.

Multimedia



## High 10 profile

Supporta il campionamento 4:2:0 con una rappresentazione a 10 bit.

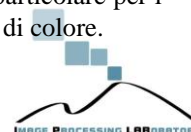
## High 4:2:2 profile

Supporta il campionamento 4:2:2 per la crominanza, con una rappresentazione a 10 bit per campione.

## High 4:4:4 profile

Supporta il campionamento 4:4:4 per la crominanza, con una rappresentazione a 12 bit. Supporta codifica lossless di regioni e una trasformata particolare per i video dello spazio RGB che evita i problemi di cambio di spazio di colore.

Multimedia



## Livelli

H.264 ha 11 livelli o gradi di abilità per limitare richieste di performance, di bandwidth e di memoria definendo delle costanti.

Ogni livello definisce il bit-rate e il tasso di codifica in macroblocchi per secondo al fine di definire diverse risoluzioni da QCIF a HDTV e oltre. Definisce inoltre la dimensione del multipicture buffer, il video bit-rate in un range da 64kbps a 240Mbps.

Più alta è la risoluzione, più alto è il livello richiesto.

Multimedia



## Livelli

Levels in H.264/AVC

Level Number	Typical Picture Size	Typical frame rate	Maximum compressed bit rate (for VCL) in Non-FRExt profiles	Maximum number of reference frames for typical picture size
1	QCIF	15	64 kbps	4
1b	QCIF	15	128 kbps	4
1.1	CIF or QCIF	7.5 (CIF) / 30 (QCIF)	192 kbps	2 (CIF) / 9 (QCIF)
1.2	CIF	15	384 kbps	6
1.3	CIF	30	768 kbps	6
2	CIF	30	2 Mbps	6
2.1	HHR (480i or 576i)	30 / 25	4 Mbps	6
2.2	SD	15	4 Mbps	5
3	SD	30 / 25	10 Mbps	5
3.1	1280x720p	30	14 Mbps	5
3.2	1280x720p	60	20 Mbps	4
4	HD Formats (720p or 1080i)	60p / 30i	20 Mbps	4
4.1	HD Formats (720p or 1080i)	60p / 30i	50 Mbps	4
4.2	1920x1080p	60p	50 Mbps	4
5	2kx1k	72	135 Mbps	5
5.1	2kx1k or 4kx2k	120 / 30	240 Mbps	5

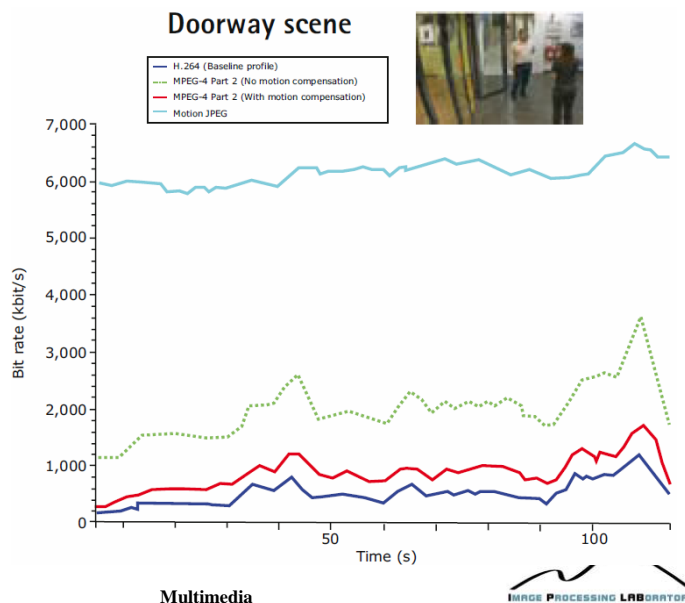
Multimedia



## Statistiche

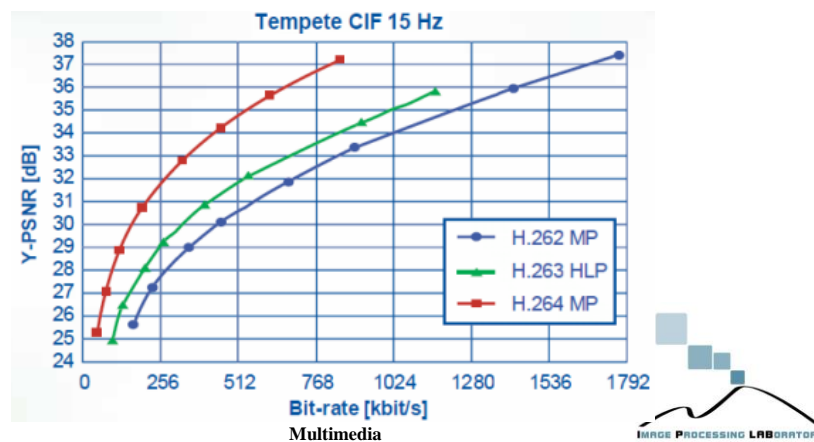
Test effettuato con una telecamera di video sorveglianza interna.

H.264 offre la migliore compressione



## Statistiche

H.264 ha migliorato il rapporto rate-distortion rispetto agli standard precedenti.



## Conclusioni

A parità di bit-rate con gli standard precedenti, H.264 offre una migliore qualità del video; a parità di qualità invece offre una migliore compressione riducendo il bit-rate. Se usato nelle trasmissioni satellitari, può raddoppiare il numero di programmi trasmessi rispetto all'uso di H.262/MPEG-2 .

E' supportato da diversi software tra cui QuickTime, Flash, dalla Playstation3, dagli I-Pod, da servizi di streaming quali YouTube.

Si pensa che H.264 rimpiazzerà tutti gli standard precedenti, anche se VCEG e MPEG stanno lavorando come Joint Collaborative Team on Video Coding (JVT-VC) per produrre una nuova generazione di codifiche per video digitali.

Multimedia



## Riferimenti

Articoli di alcuni ideatori di H.264:

[http://ip.hhi.de/imagecom\\_G1/assets/pdfs/csvt\\_overview\\_0305.pdf](http://ip.hhi.de/imagecom_G1/assets/pdfs/csvt_overview_0305.pdf)

<http://www.fastvdo.com/spie04/spie04-h264OverviewPaper.pdf>

Flyers

[www.itu.int/dms\\_pub/itu-t/oth/1D/0C/T1D0C0000010002PDFE.pdf](http://www.itu.int/dms_pub/itu-t/oth/1D/0C/T1D0C0000010002PDFE.pdf)

Siti ufficiali:

<http://www.itu.int>

<http://www.iso.org>

Alla pagina seguente, tutti gli standard ufficiali e i link alle librerie:

<http://www.itu.int/itu-t/recommendations/index.aspx?ser=H>

Multimedia

