

Codifica video:

gli standard di compressione ISO/IEC MPEG – ITU-T

Paola Sunna

Rai - Centro Ricerche e Innovazione Tecnologica

1. INTRODUZIONE

Nella teoria delle comunicazioni, le tecniche per la riduzione del bit-rate hanno come scopo quello di eliminare nel modo più efficiente possibile tutte le ridondanze presenti nel segnale audio-video originale, in particolare le ridondanze spaziali a livello di singolo quadro e le ridondanze temporali tra quadri successivi di una sequenza televisiva.

La trasmissione completa dell'informazione contenuta nel segnale televisivo numerico, descritto nelle raccomandazioni *ITU-R BT.601* [1], *ITU-R BT.709* [2] e *ITU-R BT.2020* [3], rispettivamente per la *standard*, la *high* e la *ultra-high definition*, richiederebbe la generazione di un flusso binario caratterizzato da un bit-rate estremamente elevato (da 166 Mbps fino a 12Gbps) che eccederebbe la capacità di un canale di trasmissione tradizionale; per ovviare a questo problema è necessario adottare sistemi in grado di comprimere l'informazione originaria al fine di ridurre la banda occupata.

Questo articolo fa riferimento ad una serie di standard di compressione definiti congiuntamente dagli organismi di standardizzazione *ISO/IEC MPEG* e *ITU-T*, standard che sono riusciti ad ottenere un consenso a carattere universale sia per quanto riguarda la loro adozione da parte di diversi settori merceologici (informatica, telecomunicazioni, *broadcasting* e *consumer*) sia a livello di copertura geografica.

Scopo di questo articolo è quello di fornire una panoramica delle principali differenze tra i due sistemi di compressione del segnale video AVC (Advanced Video Coding)/H.264 e HEVC (High Efficiency Video Coding)/H.265; in particolare HEVC è il successore di AVC ed è stato rilasciato ufficialmente come standard nel 2013 dal gruppo di lavoro congiunto ISO/IEC MPEG e ITU-T denominato JCT-VC (Joint Collaborative Team - Video Coding). HEVC, come AVC, si basa sulla codifica a blocchi dell'immagine e sullo sfruttamento della ridondanza spaziale e temporale che caratterizza il segnale video, ma rispetto ad AVC, i singoli tool di codifica sono stati ottimizzati, al fine di ottenere la stessa qualità soggettiva (qualità visiva percepita) con una riduzione del bit-rate dell'ordine di circa il 50%.

2. ISO/IEC MPEG E ITU-T

L'**ISO/IEC MPEG** (*Moving Picture Experts Group*), formalmente ISO/IEC JTC 1/SC 29/WG 11, è un comitato tecnico congiunto delle organizzazioni internazionali ISO (International Standardization Organization) e IEC (International Electrotechnical Commission) incaricato di definire standard per la rappresentazione e compressione dei segnali audio, video ed altre tipologie di contenuti multimediali in formato digitale. L'**ISO** è una *federazione non governativa*, costituita nel 1947, che abbraccia oltre 130 enti normatori di altrettante nazioni a livello mondiale. L'ISO promuove lo sviluppo e l'unificazione normativa per consentire e facilitare lo scambio dei beni e dei servizi. Coordina l'ambiente scientifico, tecnologico ed economico e fissa riferimenti vincolanti per una vastità di settori quali informatica, meccanica, elettrotecnica, ecc. I lavori dell'ISO sono il risultato di lunghi accordi internazionali e danno luogo a Norme Internazionali (*International Standard*). I Paesi aderenti all'accordo, tramite i singoli comitati di standardizzazione nazionali, s'impegnano a introdurre gli *International Standard* nelle corrispondenti norme nazionali. L'**IEC** è, invece, un'organizzazione internazionale che si occupa della definizione di standard in materia di elettricità, elettronica e tecnologie correlate. La IEC è stata fondata nel 1906 e ad essa attualmente partecipano più di 60 paesi.

L'**ITU-T** (*International Telecommunication Union - Telecommunication Standardization Sector*) è l'agenzia delle Nazioni Unite specializzata nelle tecnologie per l'informazione e le comunicazioni. **ITU-T** riunisce esperti da tutto il mondo allo scopo di sviluppare standard internazionali, noti come *Raccomandazioni ITU-T*, che hanno come oggetto le tecnologie dell'informazione e della comunicazione (*Information and Communication Technology, ICT*).

3. STANDARD DI COMPRESSIONE

Nel corso degli ultimi vent'anni, **ISO/IEC MPEG** e **ITU-T** hanno pubblicato congiuntamente tre standard, che si sono susseguiti nel tempo ottenendo

sempre una diffusione universale per quanto riguarda sia l'aspetto geografico (standard adottati ovunque nel mondo) sia l'aspetto merceologico (standard adottati in più campi applicativi). Si tratta degli standard:

- **H.262/MPEG-2**, datato 1995 [4]
- **H.264/AVC** (*Advanced Video Coding*), datato 2003-2004 [5]
- **H.265/HEVC** (*High Efficiency Video Coding*), datato 2013 [6]

I tre standard sopra menzionati, **MPEG-2**, **AVC** e **HEVC**, condividono lo stesso approccio basato su:

- a) compressione senza perdita di informazione basata sullo sfruttamento della *ridondanza spaziale* (correlazione tra pixel adiacenti nel quadro, scomposizione dell'immagine a blocchi), della *ridondanza temporale* (correlazione tra quadri/semiquadri nel tempo) e sulla *codifica entropica*;
- b) compressione con eliminazione dell'*irrelevanza*, ossia di quell'informazione non più ricostruibile dal decodificatore, ma non percepibile dal sistema visivo umano (codifica psico-visiva);
- c) compressione con perdita di informazione legata al processo di *quantizzazione*;

ma si differenziano per i *tools* che sono stati integrati in ciascuno di essi al fine di migliorare le prestazioni, in termini di efficienza di compressione, rispetto al predecessore.

Per i dettagli su **MPEG-2** si rimanda allo standard [4] mentre nel seguito si descrive il funzionamento dei più recenti standard **AVC** e di **HEVC**.

4. AVC (ADVANCED VIDEO CODING)

Nel 2001, gli Organismi di standardizzazione **ISO/IEC MPEG** e **ITU-T**, e in particolare per quest'ultimo il **VCEG** (*Video Coding Expert Group*), costituiscono il **JVT** (*Joint Video Team*) per la definizione di un sistema avanzato di codifica, denominato **H.264/AVC** con

prestazioni superiori a quelle di *MPEG-2*.

Nel corso del 2003 *AVC* viene integrato come **Parte 10** dello standard *MPEG-4 ISO/IEC 14496* e, con il nome di **H.264**, in ITU [5] e l'approvazione finale congiunta da parte di **ISO** e **ITU** avviene nell'ottobre del 2004.

Il sistema *AVC* specifica la codifica del video, **VCL** (*Video Coding Layer*), ed il formato con cui organizzare i dati video per il trasporto e la memorizzazione, **NAL** (*Network Abstraction Layer*).

Lo standard *AVC*, così come avviene nel caso di *MPEG-2*, definisce la sintassi del flusso dati ed il metodo di decodifica e, costituendo un nuovo sistema di codifica, non è compatibile con i ricevitori *MPEG-2*.

La versione attuale include 21 profili ciascuno dei quali ottimizzato per una applicazione specifica. Di seguito si riportano i profili rivolti principalmente alle applicazioni in ambito televisivo:

- a) *Baseline Profile*, destinato ad applicazioni a basso ritardo *end-to-end*, applicazioni mobili, videotelefonia;

- b) *eXtended Profile*, per applicazioni mobili e per streaming;
- c) *Main Profile*, rivolto ad applicazioni diffusive SDTV (*Standard Definition TeleVision*). Utilizza 8 bit per la rappresentazione digitale dei campioni e il formato di cromaticità 4:2:0;
- d) *High profile*, rivolto ad applicazioni diffusive HDTV (*High Definition Television*) e memorizzazioni su supporti ottici. Utilizza 8 bit per la rappresentazione digitale dei campioni e il formato di cromaticità 4:2:0;
- e) *High 10 profile*, supporta il formato di cromaticità 4:2:0 e 10 bit per la rappresentazione digitale dei campioni;
- f) *High 4:2:2 profile*, supporta il formato di cromaticità 4:2:2 e fino a 10 bit per la rappresentazione digitale dei campioni;
- g) *High 4:4:4 profile*, supporta fino al formato di cromaticità 4:4:4 e rappresentazioni digitali dei campioni fino a 12 bit.

I profili *High 10 profile*, *High 4:2:2 profile* e *High Profile 4:4:4* vengono generalmente utilizzati per applicazioni professionali di contribuzione, produzione/editing in studio.

Lo schema a blocchi di riferimento per il codificatore *AVC* è riportato nella figura 1.

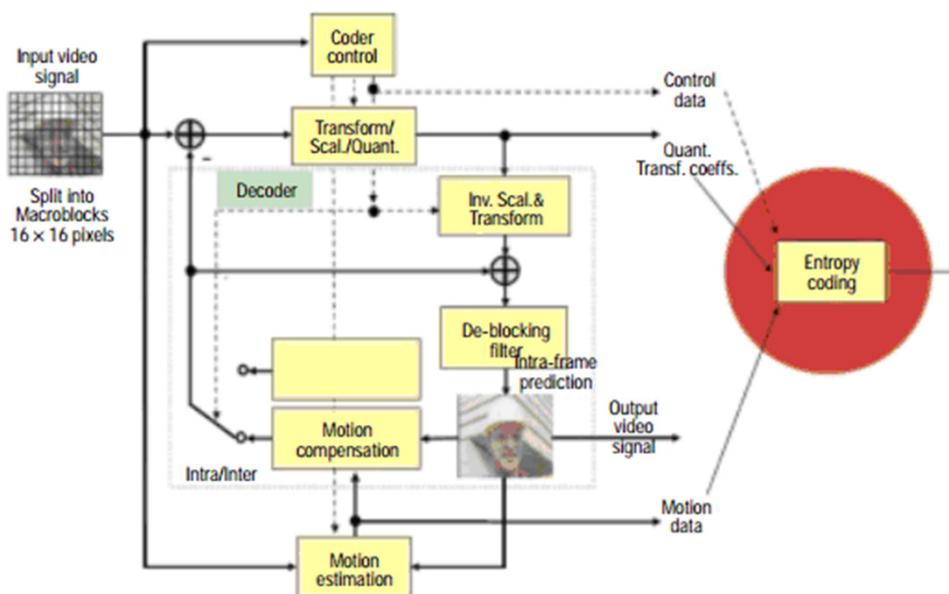


Fig. 1 – Codificatore AVC

L'efficienza di codifica di un codificatore **AVC** realizzato secondo la norma H.264 è superiore di circa il 50% rispetto a quella di MPEG-2 per le differenze riportate qui di seguito:

- a) H.264 utilizza blocchi di dimensione e forma variabile rispetto al blocco di dimensione fissa 16×16 di MPEG-2 realizzando, in questo modo, un risparmio di bit-rate che può arrivare fino al 15%;
- b) la precisione nella stima dei vettori movimenti in H.264 è più precisa che in MPEG-2: (1/4 di pixel in H.264 contro 1/2 in MPEG-2) e consente di ridurre il *bit-rate* necessario alla codifica fino al 20%;
- c) H.264 utilizza fino a un massimo di cinque quadri per la stima del movimento contro i due di MPEG-2 con un guadagno di *bit-rate* compreso tra il 5 e il 10%;
- d) riduzione della ridondanza spaziale: H.264 utilizza una **trasformata intera** invece della **DCT** (*Discrete Cosine Transform*) allo scopo di ridurre la perdita di precisione in seguito alla trasformata inversa;
- e) quantizzazione: H.264 utilizza un maggiore numero di livelli di quantizzazione, 52 contro i 31 di MPEG2;
- f) codifica entropica: H.264 utilizza tecniche più complesse quali **CAVLC** (*Context-Adaptive*

Variable Length Code) e **CABAC** (*Context-based Adaptive Binary Arithmetic Coding*), ma più efficienti rispetto all'uso di tabelle **VLC** (*Variable Length Code*) statiche di MPEG-2;

- g) uso di un *deblocking filter*: lo standard H.264, a differenza di quello MPEG-2, utilizza per il *deblocking* un filtro adattativo che consente di migliorare le aberrazioni visive (presenza di una struttura a blocchi che permea tutta l'immagine) dovute alla perdita di dati per effetto di un processo di compressione molto spinto, quale quello effettuato in questo caso sul segnale video. H.264, a differenza di MPEG-2, utilizza questo filtro adattativo allo scopo di ridurre l'effetto della disposizione in blocchi sulla sequenza decodificata, effetto che pregiudica drasticamente la qualità percepita dall'utente finale.

La maggiore efficienza (definibile come la riduzione di *bit-rate* ottenibile a parità di qualità soggettiva) di H.264 rispetto a MPEG-2 si paga però in termini di aumento della complessità sia del codificatore sia del decodificatore, come risulta evidente dalle indicazioni riportate nella tabella 1.

Nel 2009 è stata rilasciata l'estensione di AVC che copre la codifica del segnale 3D, stereoscopico e multi viste.

Tab. 1 – Complessità di AVC rispetto a MPEG-2

Profilo	Stima preliminare dell'efficienza rispetto a MPEG2	Aumento della complessità stimata per il decodificatore H264*
Baseline	Circa 1,5	Circa 2,5 volte
Extended	Circa 1,75	Circa 3,5 volte
Main	Circa 2	Circa 4 volte
*Il codificatore è circa 8 volte più complesso		

5. HEVC (HIGH EFFICIENCY VIDEO CODING)

Dopo la finalizzazione di **AVC High Profile** nel 2004, **ITU-T VCEG** e **ISO/IEC MPEG** hanno fatto nuovamente partire un'attività esplorativa al fine di capire se ulteriori miglioramenti in termini di efficienza di compressione fossero possibili rispetto ad **AVC**. La risposta è stata affermativa e nel 2010 è stato creato un gruppo di lavoro congiunto **JCT-VC (Joint Collaborative Team-Video Coding)** che ha emesso una "call for proposals" tramite la quale sono state raccolte 27 proposte. I *tools* migliori sono stati verificati, selezionati e raccolti per dare vita al nuovo sistema di compressione **H.265/HEVC**. La versione *draft* dello standard è stata rilasciata a Gennaio 2013 mentre **ITU-T** ha rilasciato quella ufficiale ad Aprile 2013 [6] e **ISO/MPEG** a Novembre 2013.

Come i suoi predecessori **MPEG-2** e **AVC**, **HEVC** continua ad essere un sistema di compressione basato sulla codifica a blocchi e sullo sfruttamento della ridondanza spaziale e temporale che caratterizza il segnale video, ma i diversi *tools* di codifica sono stati ottimizzati allo scopo di ottenere la stessa qualità soggettiva di **AVC** utilizzando però soltanto

metà del *bit-rate*. **HEVC** inoltre è stato concepito per essere efficiente anche per risoluzioni video, **UHD-1** e **UHD-2**^{Nota 1}, superiori all'alta definizione e per sfruttare architetture capaci di *processing* parallelo.

Lo schema a blocchi del codificatore HEVC è riportato in figura 2 e ricorda quello di **AVC** tranne per le differenze sintetizzate di seguito:

- la dimensione del blocco di codifica è passata da 16x16 pixel ad un massimo di 64x64 *pixel* al fine di sfruttare meglio le ridondanze tra *pixel* adiacenti, particolarmente evidenti al crescere della risoluzione del segnale video;
- l'accuratezza delle predizioni *intra* (nell'ambito dello stesso quadro) e *inter* (tra quadri diversi) è stata migliorata aggiungendo più modi di codifica e filtri interpolatori più precisi;
- tecniche sofisticate sono state adottate per compattare ulteriormente l'occupazione dei vettori movimento;
- le prestazioni della codifica entropica sono state ottimizzate;

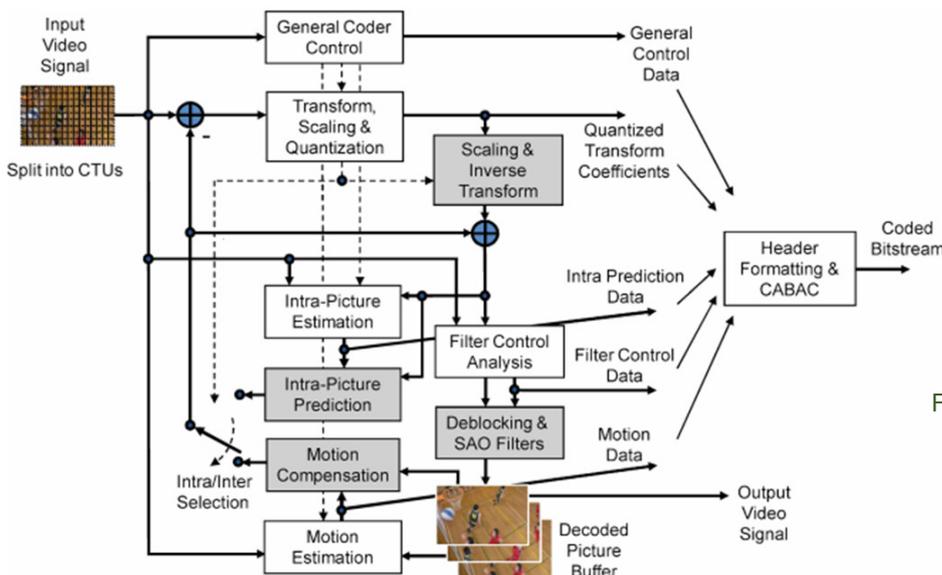


Fig. 2 – Codificatore HEVC (tratto da [7])

Nota 1 - **UHD (Ultra High Definition)**: **UHD-1** si riferisce ad una risoluzione di 3840x2160 pixel, quattro volte superiore a quella dell'alta definizione; **UHD-2** si riferisce ad una risoluzione di 7680x4320 pixels, otto volte quella dell'alta definizione.

- e) dopo il filtro di *de-blocking* è stato aggiunto un altro filtro che migliora ulteriormente la qualità del segnale ricostruito;
- f) sono stati introdotti *tools* per il *processing* parallelo. Questi *tools* però non sono obbligatori e il loro utilizzo o meno è a discrezione del costruttore.

La tabella 2 riporta in dettaglio le differenze tra **AVC** e **HEVC**.

La complessità dell'encoder è di circa 10 volte superiore a quella di un codificatore **AVC** mentre la complessità del decoder **HEVC** è di circa 2-3 volte superiore a quella di **AVC**.

I profili attualmente definiti nella specifica sono:

- 1) **Main Profile**: utilizza il formato di crominan-

za 4:2:0 e una rappresentazione digitale dei campioni su 8 bit;

- 2) **Main 10 Profile**: utilizza il formato di crominanza 4:2:0 e una rappresentazione digitale dei campioni a 8 e 10 bit;

- 3) **Main Still Profile**: utilizza il formato di crominanza 4:2:0 e una rappresentazione digitale dei campioni su 8 bit come nel caso *Main Profile*, ma si utilizza per la codifica di singole immagini (*single still picture*).

Ulteriori estensioni di **HEVC** verranno rilasciate nei prossimi due anni; in particolare:

- estensioni per il supporto di formati di crominanza 4:2:2 e 4:4:4 fino a 14 bit per la rappresentazione digitale dei campioni;
- scalabilità spaziale;
- codifica 3D stereoscopica e multi vista.

Tab. 2 – Confronto tra AVC e HEVC (realizzazione delle funzionalità)

Funzionalità	AVC High Profile	HEVC Main Profile
Coding Unit Structure	16x16 Macroblocks	Coding Tree Block : 64x64, 32x32, 16x16 Coding Unit : 64x64 to 8x8 in quad-square-tree structure
Prediction Structure	Partitions down 4x4	Prediction Unit : 64x64 down to 4x4; square, symmetric and asymmetric
Transform Structure	8x8 and 4x4 transforms	Transform Units : 32x32, 16x16, 8x8, 4x4
Core Transform	<i>Integer DCT</i>	<i>Integer DCT</i> type for 32x32, 16x16, 8x8, 4x4 <i>Integer DST</i> type for 4x4
Intra predictions	Intra predictions (9 directions)	Intra prediction (33 angular modes with planar and DC modes)
Luma interpolation	6-tap filtering to ½ pel followed by bilinear interpolation to ¼ pel	8-tap filter to ½ pel and 7-tap filter to ¼ pel
Chroma interpolation	Bilinear interpolation	4-tap to 1/8 pel
Motion Vector	Motion Vector Prediction	Advanced Motion Vector Prediction (spatial and temporal)
Entropy Coding	CABAC, CAVLC	CABAC
De-blocking filter	In-loop deblocking filter	In-loop deblocking filter
Sample Adaptive Offset	no	yes
Dedicated tools for parallel processing	Slices	Slices, dependent slices, tiles and wavefronts

BIBLIOGRAFIA

- [1] [Recommendation ITU-R BT.601](#), *Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios*, 1983 e versioni successive
- [2] [Recommendation ITU-R BT.709](#), *Parameter values for the HDTV standards for production and international programme exchange*, Novembre 1993 e versioni successive
- [3] [Recommendation ITU-R BT.2020](#), *Parameter values for ultra-high definition television systems for production and international programme exchange*, Agosto 2012 e versioni successive
- [4] [Recommendation ITU-T H.262](#), *Information technology – Generic coding of moving pictures and associated audio information – Video*, 1994 e versioni successive
- [5] [Recommendation ITU-T H.264](#), *Advanced video coding for generic audiovisual services*, Maggio 2003 e versioni successive
- [6] [Recommendation ITU-T H.265](#), *High efficiency video coding*, Aprile 2013 e versioni successive
- [7] G. J. Sullivan, J-R. Ohm, W-J. Han, T. Wiegand, [Overview of the High Efficiency Video Coding \(HEVC\) Standard](#), in "IEEE Transactions on Circuits and Systems for Video Technology", Volume 22, Numero 12, Dicembre 2012, pag 1649-1668