

Che cosa è, come funziona: Le origini del video digitale (La raccomandazione ITU-R BT.601)

ing. Marzio Barbero e
ing. Natasha Shpuza

1. Premessa

Alla base della rapida evoluzione dei sistemi di codifica digitale dell'informazione video sono le specifiche relative al campionamento e alla codifica individuate dalla Raccomandazione ITU-R BT.601. L'articolo [1] pubblicato nel 1982 da Elettronica e Telecomunicazioni forniva una completa e competente analisi della nuova raccomandazione: la prima versione risale infatti a tale anno. Questa scheda intende riproporre brevemente le caratteristiche principali di queste specifiche poiché esse spesso hanno implicazioni rilevanti nella definizione dei sistemi di compressione del segnale video e sulle tecniche di editing. La definizione della Rac. ITU-R BT.601 è l'evento alla base dello sviluppo e alla diffusione degli apparati di ripresa, di editing e di distribuzione e questa evoluzione è uno dei temi ricorrenti delle schede pubblicate in questa sezione "Che cosa è, come funziona" di Elettronica e Telecomunicazioni.

2. Cenni storici

Negli anni '70, le tecnologie numeriche si erano sufficientemente sviluppate da consentire all'industria di produrre i primi apparati digitali in grado di operare su segnali video di qualità adatta alla realizzazione di prodotti televisivi: nel 1981 l'UER e SMPTE riuscirono a definire i parametri essenziali per raggiungere un accordo a livello mondiale. In particolare i

ricercatori della Rai contribuirono, così come quelli dei maggiori laboratori e centri di ricerca degli enti televisivi, alla sperimentazione tecnica, nell'ambito dei gruppi di lavoro UER e del ITU, necessaria alla definizione di tali parametri.

La prima versione delle specifiche si limitava alla definizione dei parametri relativi al primo membro di una "famiglia estensibile di standard per la codifica digitale compatibile" basata sull'adozione di una frequenza di campionamento unica, pari a 13,5 MHz, per i formati di immagine 4:3 a 625 righe/50 Hz e 525 righe/60 Hz. La versione attuale, la quinta del 1995 [2], specifica anche i parametri per la codifica nel caso di formato d'immagine 16:9, utilizzando la frequenza di campionamento di 13,5 MHz oppure 18 MHz, quando è richiesto un proporzionale aumento della risoluzione orizzontale.

3. Motivazioni alla base dello standard

La decisione di passare da una codifica del segnale video di tipo analogico (NTSC, PAL, SECAM) ad una codifica di tipo numerico ha varie motivazioni:

- Consentire l'uso di reti digitali per trasportare il segnale: i collegamenti numerici si possono considerare "trasparenti" al tipo di informazione trasportata (ad esempio: dati, audio, video) e al formato di codifica

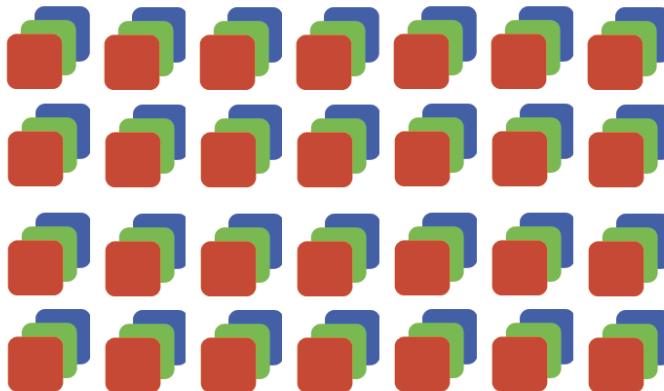


Fig. 1 - Nel caso di struttura di campionamento 4:4:4, le tre componenti sono campionate alla stessa frequenza e i tre campioni corrispondenti allo stesso elemento di immagine sono co-posizionati: ad esempio, nel caso di RGB i campioni rosso, verde e blu hanno lo stesso reticolo di campionamento.

dell'informazione (ad esempio, nel caso del video: formati a componenti o composti, a 625 o 525 righe)

- Utilizzare apparati in grado di elaborare l'informazione video in modo da consentire l'ideazione di programmi e l'utilizzazione di linguaggi tecnico-espressivi (ad esempio effetti speciali e studio virtuale) assolutamente non pensabili nel caso si fosse continuato ad operare nel dominio analogico.

La raccomandazione è stata definita con lo scopo di avere la maggior parte dei parametri in comune nel caso dei formati a 625 e 525 righe al fine di consentire economie di scala nella realizzazione degli apparati e di facilitare lo scambio internazionale dei programmi. In particolare si è scelto di avere frequenze di campionamento uguali (per formati a 525 e 625 righe) e una codifica basata su tre componenti.

4. La famiglia estensibile di standard compatibili per la codifica digitale

4.1 4:4:4 e 4:2:2

I segnali possono essere sotto forma RGB, cioè corrispondenti ai tre colori primari (rosso, verde e blu, ovvero RGB *red-green-blue*). In

questo caso tutte le tre componenti sono campionate alla stessa frequenza: i membri della famiglia che rispondono a queste caratteristiche sono denominati 4:4:4. In questo caso la struttura di campionamento prevede che i campioni relativi alle tre componenti siano co-posizionati (figura 1).

A partire dai segnali elettrici precorretti di gamma E'_R , E'_G e E'_B all'uscita dei sensori della telecamera è possibile ottenere come loro combinazione lineare altri tre segnali digitali denominati Y (luminanza), C_R e C_B (differenza-colore), anche in questo caso la struttura è 4:4:4, cioè le tre componenti sono caratterizzate dalla stessa frequenza di campionamento e sono co-posizionati.

A partire dalle componenti Y, C_R e C_B nel formato 4:4:4 è possibile ottenere il formato 4:2:2, caratterizzato dal fatto che le due componenti differenza-colore sono limitate in banda e campionate a metà della frequenza con cui è campionata la luminanza.

Ovviamente la limitazione di banda implica che i formati 4:2:2 siano meno adatti ad applicazioni in cui si richieda una elevata qualità anche in presenza di numerose e complesse elaborazioni del segnale video.

D'altro canto la scelta di prevedere il formato 4:2:2 è essenzialmente economica: si riduce del 25% il numero di campioni da memorizzare o trasmettere. Inoltre sia i sistemi per la diffusione, sia quelli analogici (PAL, SECAM, NTSC) che quelli digitali (DVB, DVD) sono caratterizzati da forte limitazione di banda delle componenti di colore e quindi non vi sono degradamenti nel passaggio dal formato professionale a quello di diffusione.

Quando vennero definiti i parametri di codifica, molta attenzione fu fatta sulla scelta della struttura di campionamento e sulla banda necessaria per le componenti, per assicurare buone prestazioni anche nel caso di operazioni di postproduzione complesse, quali il chroma-key o intarsio [1]. Particolare cura è posta nelle specifiche per i filtri a cui sottoporre i segnali Y o R,G e B (figura 2) e i segnali differenza-colore sia nel caso siano ottenuti direttamente

a partire dai segnali analogici (figura 3), sia a partire dai segnali digitali mediante filtri numerici (figura 4).

Nel caso di struttura di campionamento 4:2:2 ciascuna coppia relativa ai due segnali differenze-colore deve essere co-posizionata spazialmente con i campioni di posizione dispari della luminanza (figura 5).

4.2 La quantizzazione

Le parole digitali che rappresentano i valori dei campioni possono essere a 8 bit o a 10 bit, gli otto bit più significativi sono la parte intera, mentre i restanti due bit, se presenti, sono da considerare parte frazionaria e se non sono indicati sono da supporre uguali alle cifre binarie 00

Con 8 bit sono disponibili 256 livelli di quantizzazione equipaziati (da 0000 0000 a 1111 1111 in notazione binaria, o da 0 a 255 in notazione decimale). Le configurazioni 0 e 255 sono riservate per la sincronizzazione, mentre quelle da 1 a 254 sono destinate al video.

Per consentire opportuni margini operativi (ad esempio per evitare che calcoli per realizzare filtri digitali generino valori video non rappresentabili correttamente) l'informazione di luminanza occupa solo 220 dei livelli disponibili: il nero corrisponde al livello 16 e il bianco al livello 235. Analogamente i segnali differenze di colore devono occupare solo 225 livelli e il valore 0 di ciascun segnale differenza di colore deve corrispondere al livello digitale 128.

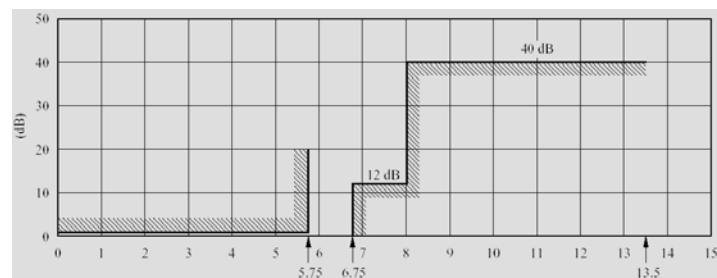
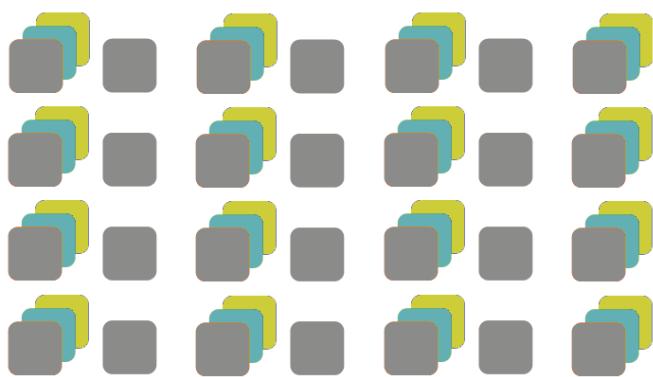


Fig. 2 - Specifiche per un filtro per i segnali di luminanza o RGB quando si utilizza la frequenza di campionamento 13,5 MHz: maschera per la caratteristica perdita d'inserzione in funzione della frequenza.

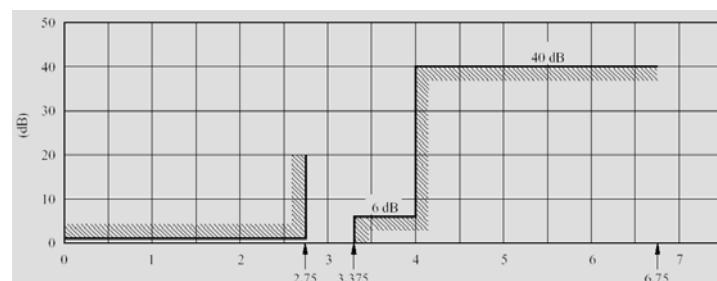


Fig. 3 - Specifiche per un filtro per i segnali di differenza-colore quando si utilizza la frequenza di campionamento 6,75 MHz: maschera per la caratteristica perdita d'inserzione in funzione della frequenza.

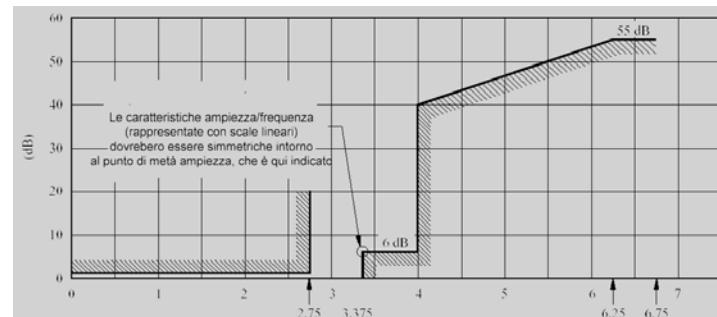


Fig. 4 - Specifiche per un filtro digitale per la conversione dalla frequenza di campionamento dei segnali differenza-colore per passare da 4:4:4 a 4:2:2: maschera per la caratteristica perdita d'inserzione in funzione della frequenza.

Fig. 5 - Nel caso di struttura di campionamento 4:2:2, il numero di campioni relativi alle componenti differenza-colore è metà rispetto ai campioni di luminanza. Entrambi i campioni differenza-colore sono co-posizionati con i campioni di ordine dispari relativi alla luminanza.

Le origini del video digitale

La raccomandazione ITU-R BT.601



Fig. 6 - Una ripresa può essere effettuata con formato d'immagine 16:9 (a) oppure 4:3 (c), in entrambi i casi può essere rappresentata da 720 campioni in orizzontale e 576 righe. Per visualizzare correttamente le immagini occorre avere display opportuni e adottare la corretta deflessione (nel caso di display analogico) o interpolazione (nel caso di display digitale). L'immagine 16:9 riprodotta ha ovviamente una definizione orizzontale inferiore. Le figure (b) e (d) rappresentano le immagini nel caso in cui si utilizzi un display a pixel quadrato, senza interpolazione, per visualizzare rispettivamente (a) e (c): vi è una deformazione spaziale dell'immagine particolarmente evidente nel caso del formato 16:9 (b), ma presente anche nell'altro caso (d) poiché 720 e 576 non sono in rapporto 4:3.

4.3 I membri della famiglia

La tabella 1 riporta i parametri per il membro 4:2:2 e frequenza di campionamento 13,5 MHz della famiglia di standard. Questi parametri si applicano sia nel caso in cui il formato d'immagine sia 4:3 sia nel caso in cui sia 16:9 (figura 6).

La scelta della stessa frequenza di campionamento per i sistemi a 625 righe e quelli a 525 righe facilita la realizzazione dei sistemi di instradamento e commutazione del segnale televisivo numerico negli studi e centri di produzione televisivi: il flusso binario complessivo è pari, in entrambi i casi, a 270 Mbps (l'interfaccia di interconnessione è specificata da una successiva raccomandazione [3] dove i campioni di luminanza e di differenza-colore sono multiplati e a ciascun campione corrispondono 10 bit).

Si osservi che il numero di campioni costituenti la riga attiva è uguale per i sistemi a 625 righe e quelli a 525 righe (720 campioni Y), malgrado

differiscano le durate totali delle righe (sono rispettivamente 864 e 858). Questa scelta ha consentito di facilitare la struttura dei sistemi in grado di elaborare o memorizzare il segnale video digitale. Non solo la quantità di dati per riga attiva è costante, ma anche la quantità di dati utili (parte attiva del segnale video) complessiva è costante: infatti, anche se questo parametro non è indicato nella [2], il numero di righe attive normalmente elaborato o memorizzato è rispettivamente 576 nel caso di sistemi a 25 quadri al secondo e 480 nel caso di quelli a 30 quadri al secondo e pertanto il numero totale di elementi d'immagine risulta essere in entrambi i casi 10 368 000 al secondo.

Nel caso del membro 4:4:4 (utilizzabile sia per R, G, B che per Y, C_R e C_B) valgono parametri del tutto simili a quelli indicati nella tabella precedentemente citata, tranne che per tutte le tre componenti valgono la stessa frequenza di campionamento (13,5 MHz) e lo stesso numero di campioni per riga totale e per riga attiva (720 per ciascuna componente).

Le origini del video digitale

La raccomandazione ITU-R BT.601

Parametri	Sistemi a 525 righe e 60 semiquadri	Sistemi a 625 righe e 50 semiquadri
Segnali codificati Y, C _R , C _B	Questi segnali sono ottenuti a partire dai segnali precorretti di gamma, vale a dire E' _Y , E' _R -E' _Y , E' _B -E' _Y	
Numero di campioni per l'intera riga - segnale di luminanza (Y) - ciascun segnale differenza-colore (C _R , C _B)	858 429	864 432
Struttura di campionamento	Ortogonale, ripetitiva a livello di riga, semiquadro e quadro. I campioni C _R , C _B sono coposizionati rispetto ai campioni dispari (1°, 3°, 5° ...) in ciascuna riga	
Frequenza di campionamento - segnale di luminanza - ciascun segnale differenza di colore		13,5 MHz 6,75 MHz
		La tolleranza per la frequenza di campionamento deve coincidere con la tolleranza per la frequenza di riga dello standard televisivo a colori corrispondente
Formato di codifica	PCM uniformemente quantizzato, 8 (opzionalmente 10) bit per campione, per il segnale di luminanza e per ciascuno dei segnali differenza di colore	
Numero di campioni per riga attiva digitale - segnale di luminanza - ciascun segnale differenza di colore		720 360
Relazione temporale orizzontale analogico-digitale - dalla fine della riga attiva digitale a O _H	16 periodi di clock per la luminanza	12 periodi di clock per la luminanza
Corrispondenza tra i livelli del segnale video e i livelli di quantizzazione - estensione della scala di quantizzazione - segnale di luminanza - ciascun segnale differenza di colore	Da 0 a 255 220 livelli di quantizzazione con il livello del nero corrispondente a 16 e il livello di picco del bianco corrispondente a 235. Il livello del segnale può occasionalmente superare 235 225 livelli di quantizzazione centrato rispetto alla scala di quantizzazione con il livello zero corrispondente a 128	
Utilizzo delle parole di codice	Le parole di codice corrispondenti ai valori numerici 0 e 255 sono usati esclusivamente per la sincronizzazione. I valori da 1 a 254 sono disponibili per il video	

Nel caso in cui il formato d'immagine sia 16:9 e si voglia conservare un'elevata risoluzione orizzontale, cioè non sia accettabile la perdita di risoluzione passando da 4:3 a 16:9 quando si utilizza uno dei due precedenti membri della

famiglia, si deve utilizzare uno dei due membri della famiglia caratterizzati dalla frequenza di campionamento pari a 18 MHz (in tabella 2 sono riportati i parametri relativi al numero di campioni per riga).

Parametri	Sistemi a 525 righe e 60 semiquadri	Sistemi a 625 righe e 50 semiquadri
Numero di campioni per l'intera riga segnale di luminanza (Y) ciascun segnale differenza-colore (C _R , C _B)	1144 572	1152 576
Frequenza di campionamento segnale di luminanza ciascun segnale differenza-colore		18 MHz 9 MHz
Numero di campioni per riga attiva digitale segnale di luminanza ciascun segnale differenza di colore		960 480

Tab. 1 - Valori dei parametri di codifica per i membri 4:2:2, 13,5 MHz della famiglia. Sono utilizzati sia nel caso di televisione digitale con formato d'immagine 4:3, che nel caso di utilizzo con formato di visualizzazione 16:9.

Tab. 2 - In questa tabella sono riportati solo i valori dei parametri per i membri 4:2:2, 18 MHz della famiglia che differiscono da quelli in tabella 1.

5. I problemi di conversione

Inizialmente il video digitale veniva acquisito e utilizzato soprattutto nel formato 4:2:2, con formato di immagine 4:3. L'evoluzione delle tecnologie digitali in campo video è stata molto rapida e oggi è possibile produrre e post-produrre a costi limitati e con qualità elevata (in molti casi, sia per effetti speciali che per correzioni cromatiche è necessario operare in 4:4:4) ed è sempre più diffuso il formato d'immagine 16:9 (supportato dal DVB, dal DVD e fruibile mediante i nuovi televisori a grande schermo). Diventano sempre più frequenti le conversioni dei vari membri della famiglia, fra di loro e verso i formati di ripresa e di editing non lineare semiprofessionale e consumer.

Limitandosi ai soli ambiti di applicazione della Rac. ITU-R BT.601, si può comunque far cenno ai seguenti problemi che, solo se noti e tenuti in considerazione, consentono di sfruttare al meglio le caratteristiche di qualità dei segnali video digitali:

- La corrispondenza fra i livelli compresi fra nero e bianco è in ambito video, come si è visto precedentemente, è ristretta alle rappresentazioni numeriche fra 16 e 235, ma questa limitazione non è spesso presa in considerazione dai sistemi di grafica elettronica (in genere di derivazione informatica, per i quali al nero corrisponde 0 e al bianco corrisponde 255).
- L'insieme dei colori rappresentabili nella forma Y, C_R e C_B è più ampio di quello rappresentabile con R, G e B: ciò può avere delle implicazioni sulla generazione ed elaborazione di immagini in ambito digitale. E' opportuno limitare i segnali Y, C_R e C_B , sacrificando eventualmente la saturazione, prima che, convertiti in R, G e B, possano dare origine a degradamenti non desiderati e percettibili su luminanza e tonalità di colore.
- Nel caso di postproduzioni complesse le operazioni di interpolazione e decimazione necessarie per passare da 4:2:2 a 4:4:4 possono dare origine a degradamen-

ti percepibili. D'altro canto è possibile che si verifichino conversioni multiple di questo tipo, poiché i sistemi di videoregistrazione professionali sono in genere di tipo 4:2:2 e quelli di derivazione consumer riducono ulteriormente la banda dei segnali differenza-colore.

- Il rapporto fra il numero di campioni attivi per riga (720) e quello delle righe attive (576) non corrisponde a 4:3, ovvero 1,33, bensì a 1,25: si dice che il pixel non è "quadrato", a differenza di quanto avviene nel caso di sistemi fotografici o di grafica elettronica. Le conseguenze di questo fatto sono che un'immagine 4:3 video risulta lievemente distorta sullo schermo di un computer, come percepibile confrontando le figure 6c e 6d, oppure che immagini create in ambito informatico appaiono deformate sullo schermo televisivo. In particolare le forme geometriche (cerchi e quadrati) possono apparire visibilmente deformate (ellissi e rettangoli). Quando si produce grafica o titoli, per poi convertirli in quadri video, utilizzando sistemi e programmi grafici realizzati per applicazioni informatiche, è quindi buona norma realizzarli in formato 768 x 576 (rapporto 4:3) e trasformarli nel formato 720 x 576 al momento della loro conversione in immagini video digitale secondo la Rac. ITU-R BT.601.

Bibliografia

1. G.F. Barbieri: "Codifica Numerica del segnale video: standard per gli studi televisivi", Elettronica e Telecomunicazioni, n. 2, 1982, pp. 42-52
2. Recommendation ITU-R BT.601-5: "Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios" (1982-1986-1990-1992-1994-1995)
3. Recommendation ITU-R BT.656-4: "Interfaces for Digital Component Video Signals in 525-line and 625-line Television Systems operating at the 4:2:2 Level of Recommendation ITU-R BT.601-5 (Part A)", (1986-1992-1994-1995-1998)