**Le compressioni simmetriche e asimmetriche**

**Quesito 1**: Descrivere brevemente le differenze tra codifiche simmetriche e asimmetriche per la compressione video. Perchè la compressione H.263 è una compressione simmetrica, mentre la compressione MPEG è asimmetrica? Quale differenza significativa le fa appartenere a due categorie distinte di compressione?

**Quesito 2**: La compressione audio MP3 (MPEG layer 3) è simmetrica? Perchè? (10/15 fps per videoconf)

**Soluzione 1**: Per la compressione dei video si utilizzano solitamente tecniche di compressione con perdita, questo perché trasmettere un video nel suo formato originale può richiedere molto tempo. Infatti, per esempio, un video in HDTV richiede un bit-rate che può superare un Gbps, e se pensiamo ad un video di un’ora, oltre ai problemi di rete, vi sono sicuramente anche problemi di throughput sul disco, per cui diventa necessario comprimere. Gli algoritmi di compressione cercano di eliminare la ridondanza spaziale, come già visto per le immagini, ma anche quella temporale. Infatti, nella codifica dei fotogrammi di un video è possibile omettere molte informazioni, questo perché vi sono poche differenze tra due frame che distano di un breve lasso di tempo. Le differenze tra un fotogramma ed il successivo sono per lo più dovute a transazioni di parti di esso, quindi è possibile memorizzare solamente gli spostamenti di queste poche parti. Queste affermazioni valgono per due fotogrammi tra cui non è presente un cambio di scena, infatti se ci fosse un cambio di scena, i fotogrammi sarebbero completamente differenti tra loro.

Per eliminare la ridondanza temporale, gli algoritmi di compressione codificano interamente solo alcuni frame, mentre i rimanenti vengono codificati come differenze con il frame precedente o successivo. In questo modo, in fase di decodifica è possibile ricostruire l’intero video, nonostante siano state scartate molte informazioni del video originale. Per il calcolo delle differenze tra i vari frame, gli algoritmi di compressione eseguono due task:

1. Stima del movimento: consiste nella ricerca del blocco in cui è contenuto l’oggetto che ha subito uno spostamento rispetto al frame precedente. Trovato il blocco, ne viene calcolato il motion vector, ossia un vettore a due componenti che indicano di quanto il blocco si è spostato orizzontalmente e verticalmente rispetto al frame precedente;
2. Compensazione di moto: consiste nel ricostruire l’immagine applicando al blocco sul frame originale il motion vector precedentemente calcolato e nel ricavare e memorizzare le differenze tra l’immagine ricostruita e quella originale.

Grazie a questi due processi è possibile scartare molte informazioni del video originale e memorizzare solamente dei frame di riferimento, i motion vector e le differenze degli altri frame rispetto a quelli di riferimento.

La stima del movimento, ed in particolare la ricerca del blocco, è il processo più complesso della catena, ed è proprio nella ricerca del blocco che i vari standard si differenziano. Gli standard della famiglia H eseguono questo processo in maniera più semplice e producendo risultati qualitativamente peggiori rispetto agli standard della famiglia MPEG. In particolare, questo processo avviene solamente in fase di codifica poiché in fase di decodifica, per la ricostruzione dei frame, è sufficiente applicare ai frame di riferimento i motion vector calcolati durante la codifica e sommarci poi le differenze memorizzate. A seconda del tempo che il processo di ricerca del blocco richiede, è possibile ottenere una codifica asimmetrica oppure simmetrica:

1. Asimmetrica: il tempo di codifica è più lungo rispetto al tempo di decodifica (solitamente molto più lungo), questo perché la ricerca dei motion vector avviene in maniera esaustiva. In questo modo le differenze da memorizzare sono poche ed è meno probabile che vengano tagliate dalla compressione, per cui alla fine si ottiene un video di qualità migliore. La codifica asimmetrica viene utilizzata quando si ricerca la qualità video e non è interessante avere il video pronto in tempi brevi.
2. Simmetrica: il tempo di codifica è simile al tempo di decodifica, questo perché, a causa dello ristretto tempo a disposizione per la codifica, i motion vector vengono calcolati velocemente e in maniera approssimata. Questo causa però la memorizzazione di molte differenze tra i vari frame, per cui diventa probabile che esse vengano tagliate dalla compressione, comportando un risultato non ottimale in termini di qualità. La codifica simmetrica viene utilizzata quando il tempo di codifica deve essere breve e la qualità video non è importante, ad esempio nelle videoconferenze, dove non è ammissibile introdurre dei ritardi.

La compressione H.263 è simmetrica perché essendo progettata per applicazioni di videoconferenza e videotelefonia su linee telefoniche ISDN, la codifica e la decodifica avvengono real-time con un ritardo massimo di 150 ms. Questo perché nella videoconferenza non è possibile concentrarsi sulla qualità ma è necessario concentrarsi sulla riduzione del tempo di trasferimento.

La compressione MPEG è asimmetrica perché è stata progettata per la memorizzazione su disco di video di buona qualità. Per memorizzare su disco è tollerabile un tempo di compressione lento ma occorre un tempo di decompressione real-time.

**Differenze tra H.263 e MPEG**

H.261 prevede due tipologie di frame, gli intra-frame (I-frame) vengono codificati interamente, mentre gli inter-frame (P-frame) sono frame calcolati per differenza a partire dal frame precedente. Il calcolo dei motion vector avviene in uno spazio di ricerca molto piccolo (più o meno 15 px), questo per privilegiare il tempo di trasmissione rispetto alla qualità video. Nella codifica ci sarà quindi un I-frame e poi tanti P-frame, il primo codificato a partire dall’I-frame e i successivi a partire dal P-frame precedente. Ogni immagine è quindi codificata a partire dalla precedente, questo perché la distanza temporale breve mi assicura che le differenze memorizzate siano poche (movimenti piccoli). Il problema più grave accade quando si ha un errore di trasmissione sull’I-frame, il quale verrà propagato fino al prossimo I-frame. H.263 mantiene le stesse caratteristiche ma aggiunge:

* un’altra tipologia di frame (PB-frame), che serve ad aumentare il frame-rate tenendo basso il throughput finale. Il PB-frame prende le righe pari del P-frame precedente e quelle dispari del successivo;
* ricerca di un vettore di movimento senza restrizioni: permette al vettore di moto di riferirsi a pixel al di fuori dei confini dell’immagine.

MPEG utilizza la ridondanza spaziale tramite la codifica JPEG e la ridondanza temporale tramite codifica differenziata con tre tipi di frame diversi:

* Index-frame (I-frame): sono frame interamente codificati come se fossero delle JPEG, per cui richiedono maggiore spazio di memorizzazione rispetto agli altri. Essi bloccano la propagazione degli errori, perché se avvenisse un errore su un I-frame, questo si ripercuoterebbe fino al prossimo I-frame. Infine rendono possibile l’accesso casuale;
* Predictive-coded-frame (P-frame): sono frame codificati tramite differenze rispetto al I oppure P-frame precedente;
* Bidirectionally-predictive-coded-frame (B-frame): sono frame codificati a partire da due frame diversi, uno che lo precede e uno che viene dopo. Possono essere costruiti a partire da I e P-frame, ma non da altri B-frame.

La compensazione di moto nei B-frame avviene ricercando il motion vector sia nel frame precedente che in quello successivo, viene poi fatta la media dei due macro-blocchi identificati e alla fine vengono calcolate le differenze. Viene fatto questo perché è probabile trovare nel frame successivo le differenze che non ho trovato nel frame precedente, questo perché magari nel frame precedente erano coperte da un altro oggetto. Si fa notare come le differenze vengono calcolate in modo esaustivo, ed è questo che rendere l’intero processo di codifica asimmetrico.

Tutte queste osservazioni fatte sui due metodi permettono di capire come la dimensione dei macro-blocchi influenzi notevolmente il processo di codifica, perché più grandi sono e meno si può essere precisi, mentre più piccoli sono è più aumenta la complessità della codifica e quindi la sua asimmetria.

Per cui H.263 e MPEG appartengono a categorie di codifica diverse a causa della differente complessità nel processo di ricerca del blocco e nel processo di calcolo dei motion vector (processi descritti precedentemente).

**Soluzione 2**: MPEG Layer 3 presenta una codifica asimmetrica, nel senso che la complessità del codificatore è molto più grande rispetto alla complessità del decodificatore. MP3 ripartisce lo spettro di frequenze in sotto-bande di ampiezza non uniforme. In particolare, vengono utilizzate sempre 32 bande ma la loro ampiezza è simile alle bande critiche nelle frequenze basse. Il modello psico-acustico di MP3 comprende il mascheramento temporale, in maniera più complessa rispetto a MPEG Layer 2. Inoltre, esso considera la ridondanza stereo e implementa un bitrate variabile. In particolare, viene utilizzata una riserva di bit per inserire i bit avanzati durante la codifica di un frame, i quali verranno utilizzati per le codifiche successive (questo è il processo che introduce più complessità nel codificatore). Infine esso utilizza la quantizzazione non uniforme, questo perché l’orecchio umano è più sensibile ai cambiamenti di volume quando il volume è basso rispetto a quando il volume è alto. Grazie a questa tipologia di quantizzazione l’errore è variabile, nel senso che gli verrà data una penalità maggiore dove la nostra capacità di distinguere le differenze di ampiezza non è buona, mentre sarà più piccolo dove la nostra capacità di distinguere le differenze in termini di volumi è buona. Quindi, globalmente l’errore introdotto è lo stesso, ma è distribuito diversamente, in maniera da inserire più errore dove l’orecchio non è in grado di percepirlo (suoni forti) e meno errore dove invece l’orecchio è in grado di percepirlo (suoni deboli). Tutte queste novità e perfezionamenti rispetto agli standard precedenti fanno sì che la complessità del codificatore aumenti.