

Tecnologie di riduzione «Data Rate»

di Massimo Novelli

Più le immagini ed i suoni sono compressi su un dato «media», più basso sta diventando il costo per immagazzinarli, ricercarli e trasportarli. Sfortunatamente, però, molte nuove tecnologie di compressione dati sono solo beneficio di fornitori di servizi, come compagnie telefoniche o banche dati, ma c'è un posto anche per il broadcast, sicuramente.

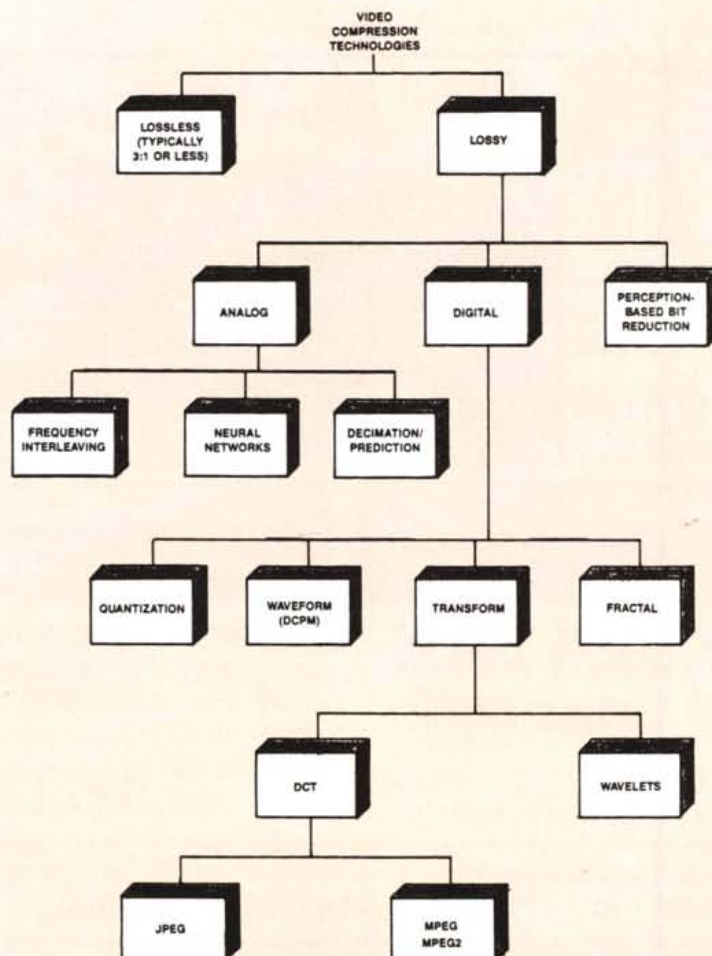
Se così queste tecnologie svilupperanno dei risultati come promettono, quantomeno l'exasperata frammentazione del mercato audio-video, soprattutto per quel che riguarda l'ambito registrazione, potrebbe essere solo l'inizio della rivincita dei broadcaster su tutti gli altri, aprendo le porte di un mercato potenziale che non si può, attualmente, nemmeno immaginare

Solo pochi anni fa, la parola compressione evocava, agli addetti ai lavori, memorie di notti insonni passate in compagnia di trasmettitori AM restii ad essere tarati come avrebbero voluto, ed in tempi più recenti attorno ai «colleghi» in FM, come il sottoscritto malvolentieri si adattava, per poter trasmettere «di più» costringendo la modulazione a rientrare nei canoni standard. È così sorta tutta una filosofia di comportamento o meglio di diverse filosofie su quali tagli/esaltazioni di frequenze o tassi di compressione fossero più adatti alla musica che si trasmetteva, fino alla compressione di tali scelte da dividere in «scuole», soprattutto quella europea che bada ad un più corretto rapporto frequenza/compressione o a quella americana che bada ad un rapporto «esplosivo» tra spettro e compressione. La compressione, quindi, nel broadcasting soffre della necessità di dover espandere in qualche modo il canale di trasmissione per poter far fluire più informazioni possibile.

Essa poi, in tempi più recenti, sta diventando il cuneo d'attacco di una strategia vincente con il suo uso, ancora di per sé abbastanza parsimonioso, nell'audio e nel video digitale. Solo per fare un esempio, il sistema ATV (Advanced TV), in particolare uno studio di TV di altissima qualità recentemente sviluppato in America, dipende fortemente da questa tecnologia. Infatti, tale sistema deve ridurre un segnale avente approssimativamente una larghezza di banda di luminanza di 20 MHz ed una di crominanza di 7 MHz nel richiesto canale di trasmissione di appena 6 MHz. Un problema di certo non facile da risolvere.

Compressione vs riduzione «bit rate»

La parola compressione, ed il suo parente più prossimo, riduzione «bit rate» sono spesso usate intercambiabilmente, la qual cosa non è corretta. Le due tecniche sono differenti. Nel broadcasting, comunque, esse servono entrambe allo stesso scopo — ridurre grandi quantità di dati audio-video a più piccoli e maneggiabili livelli —. La compressione riduce il rapporto dei dati, tra originale e «copia conforme», impiegando



L'albero genealogico delle tecnologie di compressione dati mostra chiaramente molte ramificazioni. Poche, comunque, sono state esaustivamente esplorate.

significati statistici matematici e di alto ordine per rimuovere informazioni ridondanti. Che cosa sarà mantenuto e cosa sarà gettato via è basato, principalmente quindi, sulla ridondanza, sulla dipendenza relazionale tra i dati, sulla predizione nel movimento degli stessi e sui livelli di entropia relativa. La riduzione «bit-rate», d'altro canto, riduce il data rate scaricando informazioni che sono senz'altro superflue o impercettibili, nella loro mancanza, sotto stabili condizioni di ascolto o di visuale. Anche se la compressione può essere usata per ogni tipo di flusso di dati, la riduzione «bit-rate» è di gran lunga più efficiente quando usata con sorgenti di informazione che sono a base percettiva come l'audio, il video o informazioni legate a rappresentazione d'immagine.

Metodo «lossy» oppure «loseless»?

Come abbiamo appena detto, la compressione o schemi di riduzione di dati sono richieste ogni qualvolta le informazioni desiderate sono in quantità maggiore della capacità del canale di trasmissione, qualunque esso sia. D'altro canto tali sistemi non sono molto consigliati se troppe informazioni sono distrutte dal processo di compressione/decompressione. In ogni caso, esisterà un ben circoscritto limite tra il grado di compressione, o tasso, e la qualità, cioè la simile fruizione dall'originale, dell'immagine o del suono ricostruito. La riduzione delle informazioni per processione sarà quindi di due tipi, principalmente: di tipo «lossy» (a perdere) e di tipo «loseless» (senza perdita). La compressione «loseless» implicherà che, quando avvenuto il processo di ristorazione dei dati, dette informazioni sono molto vicine ad una perfetta riproduzione dell'originale. Sistemi del genere usano tipicamente bassi rapporti di compressione, di solito entro 3:1 o meno e questo produrrà pochissimi percettibili scostamenti o errori nella ricostruzione. Tutto ciò quindi è altamente desiderabile per la preservazione di informazioni in un certo senso critiche, come il dettaglio nei particolari in una immagine oppure armoniche di ordine superiore nell'au-

dio, ma non abbastanza se una grossa quantità di dati deve essere inviata in un piccolo intervallo di tempo oppure attraverso un canale di trasmissione a lar-

ghezza di banda limitata. Sistemi a compressione «lossy» sono invece molto popolari, attualmente. Comunque, essi hanno vari e precisi limiti poiché per loro natura, possono produrre percettibili «artifacts» sotto processione. La reale sfida, di tali sistemi, è quella di rendere simili discrepanze il più possibile ridotte, camuffate o rese impercettibili. I sistemi di riduzione «bit-rate» potranno essere quindi di grande aiuto ai broadcaster, offrendo loro un enorme potenziale, e ciò perché la flessibilità estrema delle tecniche impiegate nel rimuovere componenti non necessarie, nell'audio o nel video, faranno sì che le informazioni trattate siano conformi il più possibile all'originale, tramite un processo che viene identificato come «conservazione di informazioni». Il significato poi dell'implementare una particolare tecnica di compressione è chiamato algoritmo. Un algoritmo, infatti, è un set di istruzioni che definirà come le informazioni d'ingresso possano essere disassemblate, compresse o troncate e codificate.

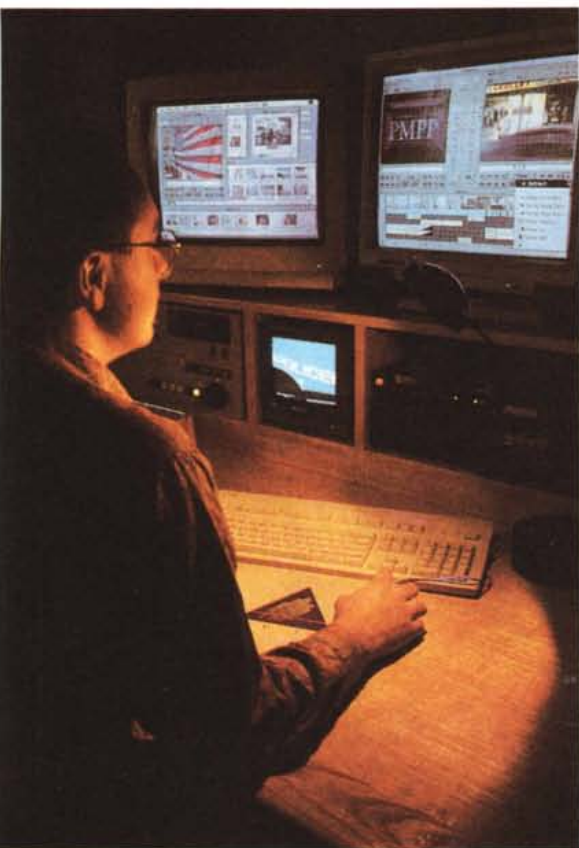
L'algoritmo in sé potrà anche comprendere un «error detection» oppure un sistema di correzione d'errore che impacchetterà i dati omogeneamente per la trasmissione o per la distribuzione. Il set di istruzioni complementari all'algoritmo sarà poi residente alla fine della catena, in pratica all'utilizzatore e sarà quindi usato per ricostruire le forme d'onda nel formato originale.

Tecniche di riduzione

La maggioranza degli schemi di riduzione dati usa una delle tre principali tecniche attuali: compressione digitale, compressione analogica e riduzione «bit-rate» su base percettiva. I sistemi di compressione digitale, per audio ad esempio, tipicamente usano processione di segnale con impiego di chip DSP e tecniche connesse, nella riduzione della ridondanza soprattutto. Tali sistemi sono attualmente ben sperimentati ed affidabili, consentendo una alta qualità del prodotto trattato. La riduzione dati per il video è, invece, significativamente più complessa. Esso, infatti, con-



Una delle ultime realizzazioni commerciali con inserite possibilità JPEG nel rappresentare funzioni di editing: la produzione Editing Machine Corporation, con i suoi modelli EMC 1 ed EMC 2 dà ampie possibilità in questo senso.



La più celebre delle implementazioni JPEG attualmente sul mercato: dalla Avid con la sua produzione Media Composer, nei modelli 2000 e 2001, ci si potrà attendere di fare editing non-lineare con una qualità, a detta della casa, paragonabile allo standard BetaCam SP. Semplicemente eccezionale.

sistendo in quadri individuali ma correlati tra loro potranno essere sì compresi, come pure la sua natura di essere una serie sequenziale di quadri, che altresì potrà essere ridotta in quantità. Ecco quindi che si dovrà aver bisogno di tecniche adatte al trattamento di un simile segnale, come per esempio l'«Intraframe coding» che opererà sulla ridondanza spaziale delle informazioni su un singolo quadro video oppure con l'«Interframe coding» che ridurrà la spaziale e temporale ridondanza di diversi quadri video alla volta, usando tecniche di compressione nello spazio e «motion prediction», ovvero interpretazioni temporali del movimento dei pixel nel tempo.

Quest'ultima potrà quindi essere in grado di dare risultati molto più accurati ed affidabili, e con più alti tassi di compressione, che non la precedente, anche se i costi per ottenerli sono molto più alti.

Sistemi di codifica digitale

Attualmente vi sono diverse variazioni delle codifiche digitali, la più semplice delle quali è la quantizzazione. Questi sistemi possono essere basati sia in grandezze scalari che vettoriali, e proprio quest'ultima, la «Vector Quantization» (VQ) sta diventando molto popolare nelle implementazioni. La VQ, infatti, agisce nel dividere ogni frame video in blocchi di dati, comparerà ognuno di essi ad una serie di blocchi sample in un «code book» dal quale otterrà gli indirizzi dei vettori immagine (Image Vector Addresses). Il sistema poi sarà in grado di trasmettere soltanto gli indirizzi. Il decoder VQ userà quindi gli indirizzi inviati per richiamare i blocchi da un identico code book, cioè istruzioni sulla codifica degli stessi. La fedeltà dell'immagine ricostruita dipenderà ovviamente da come gli indirizzi dei blocchi trattati incontrerà in similitudine gli stessi dell'immagine originale e questo, in ogni caso, è determinato da quanti blocchi di dati sono compresi nel code book iniziale. La «Waveform coding» meglio conosciuta come Differential Pulse Code Modulation (DPCM), è un'altra delle tecniche impiegate ed è abbastanza facile da implementare. Il sistema in questione lavora sull'assegnare un valore numerico a blocchi di pixel, basandosi sulla loro intensità. Essa poi comparerà il pixel corrente al precedente e ne trasmetterà solo la differenza. I sistemi DPCM sono spesso adattabili, nel senso che la maggioranza della loro potenza di processione potrà essere allocata su parti dell'immagine che ne ha più bisogno, offrendo poi una efficienza di data reduction comparabile alla codifica di trasformata (transform coding).

Transform Coding

Vi sono diversi tipi di codifica di trasformata, anche conosciuta come «function expansion». Essa infatti è basata su trasformate matematiche, come la «Discrete Fourier Transform» (DFT), la «Discrete Cosine Transform» (DCT) ed altre. Queste trasformate shiftano i pixel da un dominio, diciamo nella loro posizione spaziale, ad un altro, che può essere la distribuzione nel tempo, o in toto nel dominio del tempo. Tali riorganizzazioni di dati spesso fanno divenire la compressione molto più facile da implementare. Una analogia, per far capire la natura della cosa, potrebbe essere quella di considerare gli studenti in una classe ordinati affinché i più alti siedano in fondo, invece che magari in ordine alfabetico; questo per due ragioni all'apparenza futili ma indispensabili. La pri-

ma è che i più piccoli possano vedere comodamente e la seconda che sarà difficile passarsi foglietti o compiti in classe. Fine della digressione.

In modo simile, organizzare i pixel in una sequenza video sulla loro frequenza invece che sulla posizione nello schermo può renderli più facili da comprimere ma anche da confondere nel vederli. La tecnica DCT è particolarmente efficiente per le immagini e sta diventando largamente usata. L'esempio più celebre di codifica DCT è nelle mani dell'altrettanto celebre Joint Photographic Expert Group (JPEG) che con il loro algoritmo standard, giunto alla revisione 9.7, sono tra i più avanti nello sviluppo di tale codifica. Questo sistema, per chi ancora non lo sapesse, lavora sulla divisione dello schermo in blocchi di pixel e nel creare coefficienti che descrivono le relazioni di un pixel dall'altro. Occorrerà quindi una frazione del tempo nel trasmettere tali coefficienti, e non quindi i dati in sé, e tipicamente i tassi di data reduction saranno da 30:1 a 50:1 con punte di 70:1 in casi estremi.

Il JPEG opererà però solo su quadri singoli; comunque, tale compressione potrà anche lavorare su video in movimento semplicemente nel sequenziare le immagini una dopo l'altra. Diverse case produttrici di apparecchi broadcast, soprattutto editor non-lineari, impiegano tali tecniche, e ne sono esempio produzioni JPEG su base Macintosh o IBM compatibili, adattati come VTR con tutte le funzioni standard di tali macchine e con una qualità adeguata a produzioni di buon livello, quindi ancora non eccezionalmente alta. Sta in ogni caso recuperando molto terreno, fino a giungere attualmente ad una bontà paragonabile al S-VHS.

I cugini del Motion Picture Expert Group (MPEG), altro strenuo gruppo di ricerca all'avanguardia, usano anch'essi diversi principi esposti nello standard JPEG, impiegando una combinazione di tecniche «motion-compensated» come predizione di movimento ed interpolazioni interframe, dividendo le immagini, come al solito, in blocchi o sub-band, per la codifica. Sono ancora alla ricerca di uno standard affidabile, anche se le ultime notizie ne danno il rilascio ufficiale tra breve.

Una nuova, promettente tecnica invece è conosciuta come «Wavelets» ed opera globalmente su diverse porzioni di immagine in modo simultaneo, non solo sui classici blocchi fissi campionati.

Frattali

Le trasformate frattali sono una nuova aggiunta nelle tecniche di com-

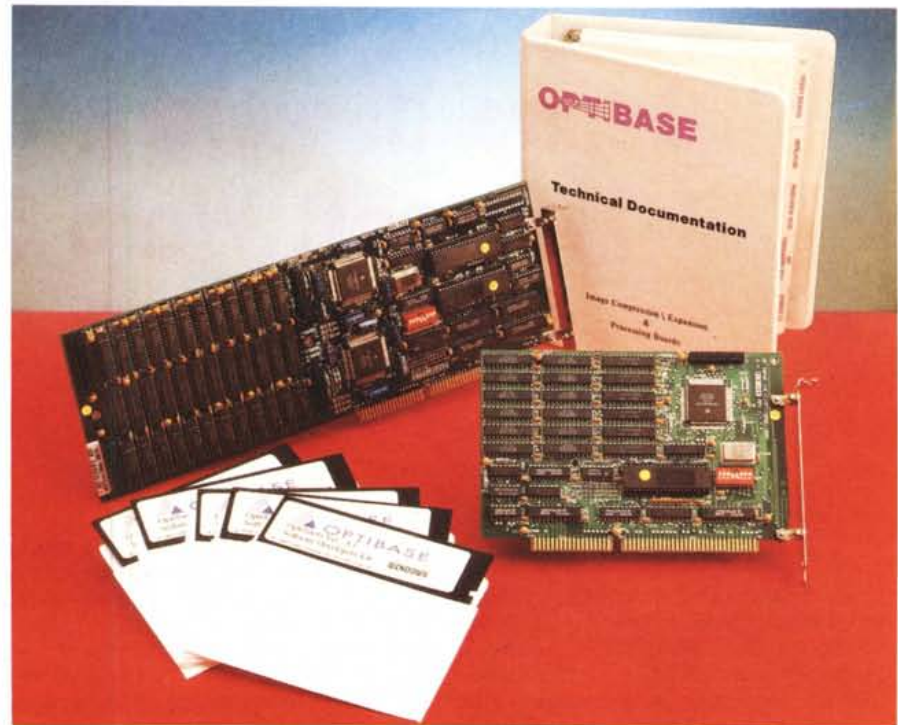
pressione in uso attualmente. Una efficiente codifica dipende fortemente dalla ridondanza nei dati dell'immagine che hanno una distinta e precisa relazione da ogni altro. Un frattale, come ben sapete, è una costruzione matematica che ha la curiosa abilità di ripetere se stessa e di crescere dimensionalmente. La sua codifica descrive i segmenti video in termini di «semi» matematici e di materiale residuo e sarà di tipo asimmetrico. Avrà bisogno di un mucchio di matematica per comprimere qualcosa ma sarà in grado, con poche operazioni invece, di riprodurre i dati trattati in modo egregio. Questo sicuramente farà della codifica frattalica un promettente strumento d'uso, soprattutto in sistemi di distribuzione video digitale poiché la grossa mole della tecnologia risiederà all'inizio della catena di trasmissione, quindi in modo distribuito ed univoco, e non alla fine. Non richiederà alcun code book nelle diverse fasi del processo, e sarà di natura scalabile, cioè adattabile a diversi formati ed indipendente dalla soluzione.

Riduzione «bit-rate» su base percettiva

Molto brevemente, tali tecniche lavorano sullo scaricare informazioni che non possono essere udite oppure viste. Il cervello umano, infatti, fa uso di tecniche di selezione delle informazioni molto modificabili ed adattabili alle circostanze, per discernere sempre cosa realmente abbisogna per la percezione di un fenomeno. Tali sistemi di riduzione infatti contengono accurati modelli di udibilità e di percezione visiva dando loro modo di eliminare informazioni che, in ogni caso, non possono essere riconosciute.

Nell'area della psicofisica acustica, diverse case produttrici stanno introducendo sul mercato schemi e tecniche di registrazioni digitali che operano sulla base del mascheramento acustico (aural-masking), fenomeno abbastanza noto da diverso tempo. Questi sistemi infatti hanno in sé un preciso modello della soglia di sensibilità dell'orecchio umano, basato sulle curve di Fletcher-Munson, come software principale da cui estrapolare i dati con i quali fare riduzioni.

Tali sistemi così divideranno l'intero spettro audio in sotto-bande, campioneranno simili componenti e con una successiva comparazione con le caratteristiche delle curve suddette, elimineranno tutte le informazioni al disotto della soglia di udibilità, ben precisa e sperimentata. In modo molto simile, nell'area della psicofisica visiva, ricerche con-



Una dignitosissima implementazione di standard JPEG alla portata di noi comuni mortali; è la produzione Optibase americana che offre un completo set di sviluppo, hardware e software a chiunque voglia cimentarsi nello sviluppare applicazioni in tale standard.

dotte in America ed in Europa hanno evidenziato risultati abbastanza confortanti; come per l'orecchio, anche l'occhio, o meglio quella parte di cervello che elabora i dati grezzi, ha una ben delimitata capacità di percepire fenomeni associati alla percezione visiva, soprattutto di immagini in movimento. Infatti si è determinato che è possibile ottenere una riduzione «bit-rate» video virtualmente senza perdita di informazioni con un rapporto di 50:1 e che si è anche stimato che essa può essere estesa a 150:1, approssimativamente, semplicemente nel rimuovere dati che l'occhio non potrà comunque percepire.

Il futuro della compressione

Applicare, per esempio, una riduzione «bit-rate» di compressione di 150:1 ad un segnale standard HDTV (alta definizione) con data rate di 1,2 Gbit/sec potrà quindi significare che tale segnale sarà in grado di fluire, su un dato canale, ad 8 Mbit/sec senza nessun difetto percettibile, pressappoco quello che è il normale canale televisivo attuale. Questo è certamente uno strumento più efficiente nel ridurre i dati che non indiscriminate o statistiche rimozioni di informazioni. Comunque, sarà senz'altro possibile che i futuri sistemi di trasmissi-

sione possano ottenere il meglio dei due mondi appena visti usando, per esempio, sistemi di compressione «lossless» all'inizio per poi ulteriormente condensare la riduzione di informazioni su base percettiva.

Conclusioni

Tutta la attuale tecnologia a disposizione è già in grado di assicurare un prodotto del genere ai massimi livelli. Simili implementazioni sono generalmente ben supportate sul mercato High-End broadcast ma la ritrosia, soprattutto delle industrie tradizionali nella produzione di apparecchiature broadcast, ancora non ne ha fatto un vero prodotto vincente.

Gli organismi preposti poi al dirimere gli standard, sottoposti dalle varie case produttrici, hanno anche dalla loro colpevoli ritardi nel decidere, mettendo in ansia un mercato già abbastanza difficile da organizzare. Se è vero che la qualità di un dato «media» di nuova concezione dovrà essere almeno all'altezza della qualità del miglior prodotto tradizionale in uso, è anche vero che una tecnologia non procede invano solo perché esiste. Ne sarà logico anche attendersi un loro completo sviluppo.

MC