



# MP3

## Il suono del 2000 è su Internet

Chi, come me, è appassionato di fantascienza, non può non conoscere un film che è diventato uno dei "cult" del genere. Sto parlando de "Il pianeta proibito" (Forbidden Planet - U.S.A. 1956). In quello storico film un'astronave terrestre, la United Planets Cruiser C-57-D, capitanata da un giovanissimo Leslie Nielsen (il comandante J.J. Adams), sbarca su Altair-4, un pianeta dove, anni prima, una precedente spedizione sull'astronave Bellephophon aveva fatto perdere le sue tracce. Su Altair-4 il comandante Adams incontra il Dottor Morbius (Walter Pidgeon) e la sua bella figlia Altaira (Anne Francis), oltre al simpati-

co robot Robby, il primo robot mai presente in un film di fantascienza, che segue le leggi della robotica di Asimov, tanto da autoinibirsi quando gli viene impartito l'ordine di sparare ad un uomo. L'enigmatico Dottor Morbius mostra al comandante Adams i resti di una avanzatissima civiltà, i Krell (nome ripreso poi da una famosissima casa costruttrice di elettroniche di altissima fedeltà). Tra le tante fantastiche ed incredibili invenzioni dei Krell che il Dottor Morbius mostra al comandante c'è un minuscolo riproduttore di musica: il Dott. Morbius prende un piccolo parallelepipedo nero, di due o tre centimetri di

lato, e lo inserisce in un altrettanto piccolo apparecchietto, sempre nero, posto sulla scrivania; subito l'ambiente dove si trovano i personaggi viene avvolto da una musica che definire strana è dir poco.

Come al solito gli autori di fantascienza hanno, anche in questo caso, previsto delle cose che poi si sono puntualmente avverate. La corsa alla miniaturizzazione ha toccato non solamente le apparecchiature elettroniche, ma anche i supporti sui quali registrare i dati e la nascita di Internet ha portato alla miniaturizzazione dei dati stessi! Così mentre solo una quindicina abbondante di anni fa per



La locandina de "Il pianeta proibito", con Walter Pidgeon, Anne Francis e Leslie Nielsen

ascoltare musica si doveva avere un giradischi ed un bel disco in vinile da 30 cm di diametro, il compact disc e, ormai anche le memorie flash RAM, unite alla compressione dei dati, hanno reso possibile il sogno di tanti anni: apparecchi musicali piccoli, portatili ed estremamente fedeli.

## L'audio per Internet

La nascita di Internet ha certamente favorito lo sviluppo della compressione dei dati. I programmi sempre più complessi, le applicazioni ed i file di dati sempre più grandi hanno reso necessario il ricorso a metodi di compressione sempre più sofisticati. E questo è ancora più valido quando si parla di file audio, dove l'ordine di grandezza è delle decine di megabyte. La sofisticazione raggiunta dagli algoritmi di compressione ha ormai raggiunto risultati ragguardevoli. I file MP3, che oggi chiunque può scarica-



Il Sito [www.RioPort.com](http://www.RioPort.com) è completamente dedicato alla musica in formato MP3 ed al lettore Rio. Da questo sito è possibile collegarsi direttamente ad altri siti che distribuiscono musica e software per lo standard MP3. RioPort, al quale collaborano molte aziende di spicco nel settore multimediale, è ancora in versione sperimentale. Supporter di questa iniziativa troviamo tra gli altri Acme Entertainment, Audible, Audio Explosion, AudioSoft, Liquid Audio, Mp3.com e Music Match.

re dalla Rete, sono in media lunghi un decimo della grandezza dei file originali, senza le perdite della qualità sonora a cui ci avevano abituato altri metodi di compressione. Ma cosa sono i file MP3? Non sono altro che file audio dove però le informazioni sono compresse. Il metodo di compressione adottato è l'MPEG Layer 3, dove MPEG sta per Moving Picture Experts Group, cioè un grup-

po di lavoro che si occupa di standardizzare la codifica e la compressione di filmati ed audio. Il decoder MP3 permette di ridurre di oltre il novanta per cento la lunghezza dei file audio utilizzando un metodo di codifica molto evoluto, come descritto più approfonditamente nel riquadro a parte. Vediamo di fare un po' di conti: senza la riduzione un secondo di dati audio con qualità CD stereo è costituito da ben 176.400 byte, con un bit-rate di circa 1,3 megabit al secondo, ed un file di 3 minuti occupa ben 31 megabyte. Utilizzando la codifica MP3 questi dati vengono ridotti di un fattore 12 con una perdita minima di qualità. O meglio, la perdita di qualità può essere scelta prima della codifica utilizzando bit-rate variabili tra 32 e 128 megabit/secondo. La tabella 1 mostra il rapporto tra bit-rate e prestazioni sonore, come si vede più è alto il bit-rate del segnale codificato e migliori sono le prestazioni audio, a discapito però di un minore rapporto di compressione. Alcuni programmi per la codifica MP3 (scaricabili da Internet, come vedrete più in là nel resto dello speciale) consentono inoltre di arrivare fino a 256 o 320 megabit/secondo, con conseguente minore compressione ma migliore qualità.

### Bibliografia

ISO/IEC - "Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s" - Part 3: Audio - ISO/IEC 11172-3 International Standard, 1993.

E. Eberlein, H. Popp, B. Grill, J. Herre - "Layer III - A Flexible Coding Standard" - 94th Convention, Berlino 1993, Preprint 3493.

F. Baumgarte, C. Ferekidis, H. Fuchs - "A Nonlinear Psychoacoustic Model Applied to the ISO MPEG Layer 3 Coder" - 99th Convention, New York 1995, Preprint 4087.

K. Brandenburg, M. Bosi - "Overview of MPEG-Audio: Current and Future Standards for Low Bit-Rate Audio Coding" - 99th Convention, New York 1995, Preprint 4130.

U. Gbur, M. Werner, M. Dietz - "Realtime Implementation of an ISO/MPEG Layer 3 Encoder on Pentium PCs" - 101st Convention, Los Angeles 1996, Preprint 4386.

G. A. Soulodre, T. Grusec, M. Lavoie, L. Thibault - "Subjective Evaluation of State-of-the-Art 2-channel Audio Codecs" - 104th Convention, Amsterdam 1998, Preprint 4740.

## MP3 per tutti

Come c'era da aspettarsi la tecnologia applicata ai file MP3, nel momento stesso in cui è stata applicata ad Internet ha potuto godere della enorme diffusione che la Rete consente. Così sono ora disponibili numerosi siti dedicati all'MP3, dove si possono scaricare file musicali e programmi per crearsi da soli i propri MP3. Da poco sono poi stati messi in com-

### Prestazioni tipiche della codifica MPEG Layer 3

QUALITA' SONORA	LARGHEZZA DI BANDA	MODO	BIT-RATE	RIDUZIONE
migliore di una radio AM	7.5 kHz	mono	32 kbps	24:1
simile ad una radio FM	11 kHz	stereo	56...64 kbps	26...24:1
quasi CD	15 kHz	stereo	96 kbps	16:1
CD	>15 kHz	stereo	112...128kbps	14..12:1

mercio dei lettori portatili in prova proprio su questo numero. Con questi minuscoli apparecchi è possibile riprodurre i file scaricati dalla rete o,

sempre tramite computer, comprimere i file dei propri CD audio e andarsene a spasso sentendo la propria musica preferita.

## La codifica percettiva MPEG

un modello psicoacustico per "comprimere" la musica

La compressione dei dati permette, nel caso di dati "informatici" un notevole risparmio di sia spazio sugli hard disk che di tempo nei trasferimenti da Internet. Applicare dei metodi di compressione anche ai file audio ci permette allora di risparmiare spazio e tempo anche utilizzando questi file, di solito "enormi", se rapportati alla grandezza media di un file zippato. Basti pensare che la versione professional di Netscape, completa quindi di tutti i plugin, è un file EXE autoscompattante di poco più di 18 megabyte, mentre un brano musicale di appena quattro minuti e mezzo (una canzone di Sanremo, per intenderci) occupa sul nostro hard disk la bellezza di 47 mega e rotti.

Purtroppo nel caso dei file audio non è possibile applicare gli stessi algoritmi che vengono utilizzati per i file dati. Nel caso che abbiamo fatto prima, zippando il solito file da 47 mega otteniamo un altro file ZIP di quasi 46 megabyte. Questo avviene perché mentre i byte contenuti nei file dati sono "organizzati" (in una immagine possono esserci molti byte uguali, ad esempio quelli dello sfondo, mentre in un programma possono esserci molti byte nulli o sequenze ricorrenti) i dati contenuti in un file audio non lo sono, almeno non in una maniera utilizzabile per un algoritmo di compressione tradizionale. Facciamo un esempio: supponiamo di avere un segnale audio costituito da una sinusoide di frequenza 980 Hz, vi-

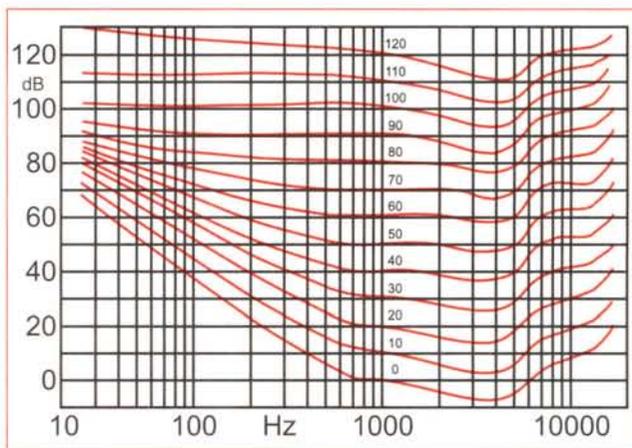


fig. 2

sibile in figura 1, e immaginiamo che questo segnale sinusoidale duri ad esempio 10 secondi. Se il segnale è campionato a 16 bit ad una frequenza di 44.100 Hz ogni secondo (immaginando un segnale stereofonico) è costituito da 44.100 x 4 byte (si moltiplica per 4 perché ci sono due byte per campione e due canali), quindi 176.400 byte. Il file sarebbe lungo quindi 1.764.000 byte, trascurando un eventuale header iniziale. Se consideriamo però un solo periodo della sinusoide in esso vi sono 44.100/980, e quindi 45 campioni. Dato che questo periodo poi sarebbe ripetuto per tutta la durata del brano, basterebbe registrare i campioni di un periodo e il numero di "repliche" del periodo. In pratica, in un secondo ci sono 980 periodi ed in dieci secondi 9.800 periodi. Quindi tale file potrebbe essere compresso in un altro file così costituito: 45 x 4 byte per periodo, più due byte per contenere il numero 9.800.

#### La soglia di udibilità

Il problema è che un file musicale è costituito da un numero infinito di sinusoidi, e per di più di durata diversa, e per ognuna di queste andrebbe applicato questo algoritmo, cosa chiaramente non attuabile. Nel caso dei file musicali è quindi necessario utilizzare una codifica di natura psicoacustica, che tiene conto cioè della maniera con cui l'orecchio percepisce i suoni. Le codifiche di questo tipo si basano sul fatto che, proprio per come funziona l'orecchio umano, alcune informazioni presenti nel segnale musicale non

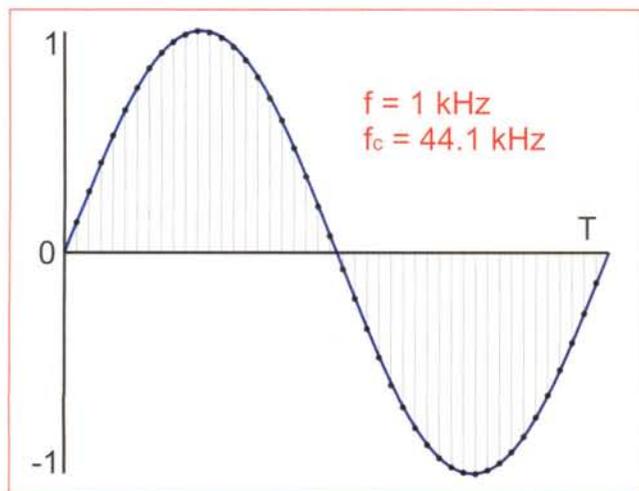
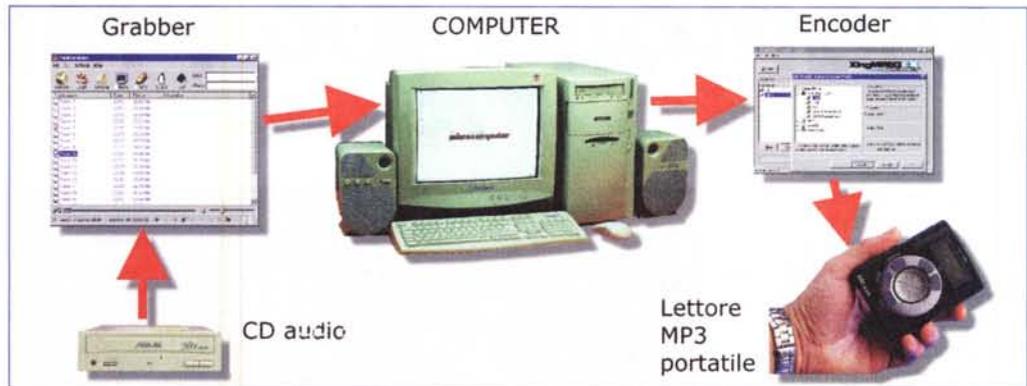


fig. 1

Schema di principio per la creazione di file MP3

Nelle pagine seguenti troverete la prova d'uso dei lettori MP3 portatili, il Rio della Diamond e l'MPMan della Saehan più una panoramica sui software e sui siti più interessanti dedicati all'argomento. Buona lettura!



vengono percepite, e quindi possono essere eliminate, risparmiando spazio prezioso. Facciamo un altro esempio: ammettiamo di avere una sinusoide a 15 Hz. Tale sinusoide non verrà percepita da nessuno, neanche dalle orecchie più allenate, perché è al di fuori dello spettro udibile. Chiaramente eliminando questa sinusoide, come tutti i segnali al di fuori della porzione di banda udibile, si risparmia un po' di spazio. Inoltre la soglia di udibilità dell'orecchio non è costante ma varia con il livello e con la frequenza del segnale, secondo delle curve, riportate in figura 2, dette di Fletcher e Munson, dal nome degli sperimentatori che per primi le rilevarono. Basta quindi eliminare tutti i suoni al di sotto della soglia di udibilità ad ogni frequenza per risparmiare ancora spazio.

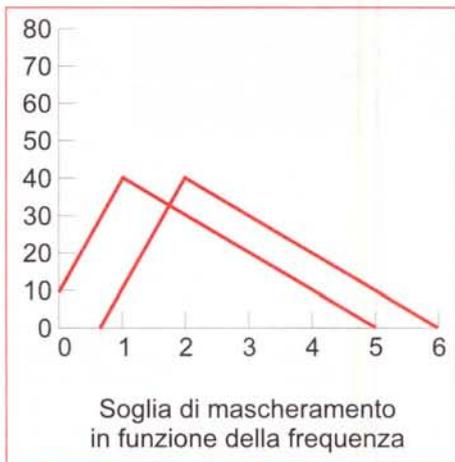


fig. 3

**L'effetto di mascheramento**

Ma la caratteristica psicoacustica dell'orecchio che se sfruttata

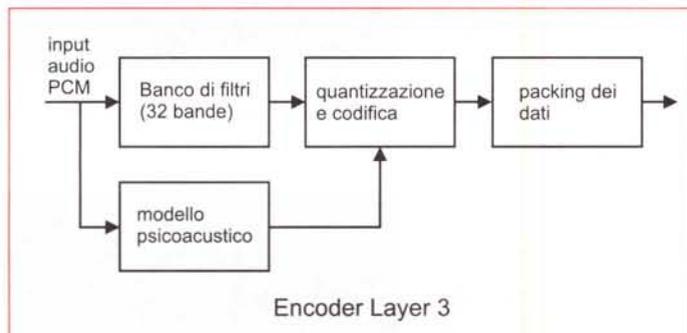


fig. 4

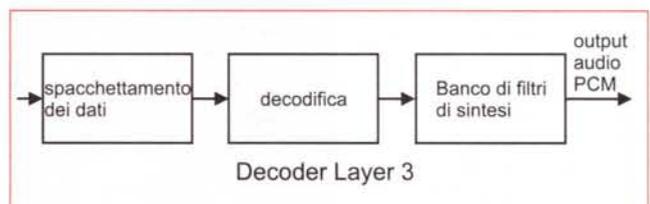


fig. 5

consente il maggiore risparmio di spazio è l'effetto di mascheramento. Questo effetto si manifesta quando nel segnale che viene inviato all'orecchio sono presenti più segnali: se abbiamo ad esempio un segnale sinusoidale ad una data frequenza ed un altro segnale, di livello più basso e ad una frequenza leggermente superiore, quest'ultimo viene "mascherato" dal primo e l'orecchio non lo percepisce. Per utilizzare positivamente questo effetto basterà "scovare" nel segnale musicale tutti i segnali di livello alto ed eliminare tutti i segnali immediatamente adiacenti e di livello più basso. In figura 3 vediamo il grafico di alcune soglie di mascheramento. Come si vede anche queste variano a seconda della frequenza.

**L'algoritmo di compressione**

In figura 4 vediamo lo schema di principio di un Encoder MPEG Layer 3. In pratica il segnale in ingresso viene inviato ad un banco di filtri che divide lo spettro in 32 bande equispaziate in frequenza. Contemporaneamente lo stesso viene portato nel dominio della frequenza tramite una FFT (Fast Fourier Transform) a 512 linee. Quest'ultimo segnale viene trattato dal blocco denominato "modello percettuale" che tiene conto di tutte le caratteristiche dell'orecchio prima descritte. Tale segnale viene poi utilizzato come maschera per la parte di codifica vera e propria. Questo blocco utilizza una quantizzazione non uniforme, codificando con un numero di bit differenti le varie porzioni più o meno importanti del segnale, ed in questo modo in pratica elimina la ridondanza dei dati. Infine i dati quantizzati, e le informazioni di canale vengono multiplexati in un flusso costante di dati. Lo schema del decoder è molto più semplice. Il primo blocco si occupa di estrarre i dati audio codificati dal flusso di dati impacchettati, il blocco centrale ricostruisce il segnale originale basandosi sulle informazioni fornite dal codificatore e impacchettate insieme ai dati. In pratica il decoder non fa altro che estrarre dal segnale codificato le varie informazioni sul numero di bit assegnati ad ogni porzione del segnale e a riassegnare ad ogni parte i canonici sedici bit. In questo modo la fase di decodifica è estremamente semplice e molto più veloce della codifica.

P.F.