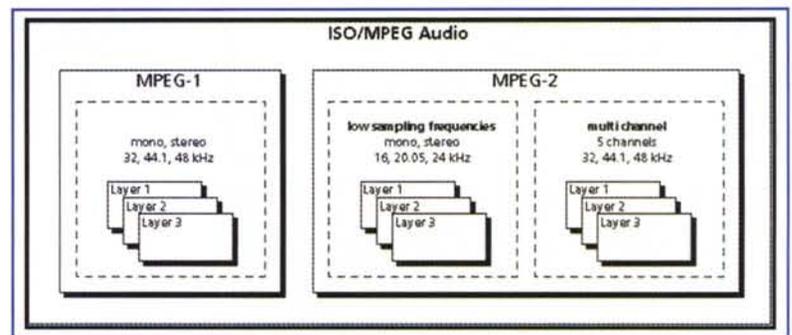


MPEG-1 Layer 3 alla scoperta dell'audio ad alta fedeltà

A volte ritornano. Dopo aver smesso di occuparci di MPEG, almeno per un certo periodo, ed aver sempre focalizzato il discorso sulla sua natura prettamente "video" e solo di riflesso audio, questa volta è d'obbligo concedere una rivincita proprio all'audio. L'MPEG-1 Layer 3 è (sembra) quanto di meglio si possa ottenere, sia in termini di flusso che di qualità intrinseca per la riproduzione sonora multimediale di elevata qualità. Provare per credere.

di Massimo Novelli



Quanti MPEG esistono?

La domanda non è casuale; infatti, allo stato delle cose, e nelle sue varianti, lo standard MPEG gode di una certa varietà di versioni, sia per ciò che riguarda la componente audio che quella video. Di conseguenza diverse sono le caratteristiche riguardanti le capacità di flusso ed i parametri-base, ma tutto sommato, le "feature" delle diverse "release" possono essere sintetizzate in una semplice elencazione.

MPEG-1 - "Coding of moving pictures and associated audio for digital storage at up to about 1.5 Mbit/sec".

MPEG-2 - "Generic coding of moving pictures and associated audio information".

MPEG-3 - In origine pianificato per ap-

plicazioni broadcast TV, ora fuso nelle specifiche MPEG-2.

MPEG-4 - "Coding of audio-visual object".

Come avete visto, abbiamo adottato le definizioni rigorosamente originali per consentire anche una più precisa ricerca di informazioni, magari via Internet. A seguire, non mancano una serie di sotto-insieme, sia nell'MPEG-1 che nell'MPEG-2, e che per dovere di cronaca vorremmo riportare sia per chiarire bene ciò di cui parliamo, sia per completezza, anche non solo didattica.

MPEG-1

ISO-11172-1 "System", descrive la sincronizzazione ed il multiplexing dei segnali audio e video.

ISO-11172-2 "Video", descrive la compressione del segnale video e la scansione dello stesso (principalmente riguardo

a produzioni Video-on-CD).

ISO 11172-3 "Audio", descrive una famiglia di codec audio generica, con tre membri gerarchicamente compatibili (chiamati Layer-1, Layer-2 e Layer-3).

ISO 11172-4 "Compliance Testing", descrive le procedure per determinare le caratteristiche dei flussi codificati, il processo di decodifica e parametri di test per la compatibilità con altri sistemi.

DTR 11172-5 "Software Simulation", rapporto tecnico sulla completa implementazione software delle prime tre parti dell'MPEG-1.

L'MPEG-2 consiste invece di nove parti, alcune già standard dal 1994, ed altre a differenti livelli di completamento.

MPEG-2

ISO-13818-1 "System", descrive la sincronizzazione ed il multiplexing dei segnali audio e video; è anche uno standard adottato dalla ITU-T come "H.222".

ISO-13818-2 "Video", descrive un generico tool set di codifica video, con supporto della scansione interlacciata; è anche uno standard adottato dalla ITU-T come "H.262".

ISO-13818-3 "Audio", descrive un'estensione "backward" compatibile dell'MPEG-1 per codifiche audio multichannel (surround, multilingual) ed una estensione "no backward" compatibile a bassi "sample rate", adatta a supportare applicazioni sonore con limitate richieste di larghezza di banda.

ISO-13818-4 "Conformance Testing", descrive le procedure per determinare le caratteristiche dei flussi codificati, il processo di decodifica e parametri di test per la compatibilità con altri sistemi.

DTR-13818-5 "Software Simulation", rapporto tecnico sulla completa implementazione software delle prime tre parti dell'MPEG-2.

ISO-13818-6 "System Extension - Digital Storage Media Command and Control (DSM-CC)", descrive un set di protocolli per applicazioni client-server.

CD-13818-7 "Audio, No-Backward-compatible (NBC) coding", descrive più evoluti schemi di codifica audio per segnali mono, stereo e multichannel.

13818-8 "Video, extension to 10-bit input sample", è stato cancellato dal sotto-insieme, a causa dell'insufficiente interesse, dal gruppo di lavoro.

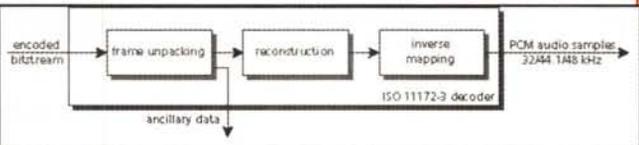
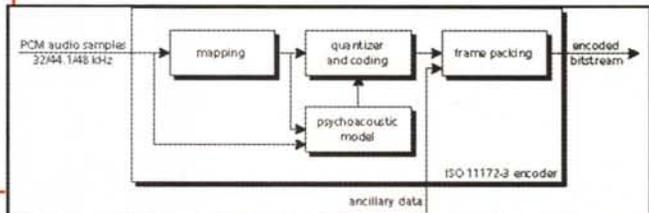
ISO-13818-9 "Real time interface specification for low-jitter applications", definisce le costanti di tempo per lo sviluppo di "trasporti" real time dei flussi MPEG-2.

WD-13818-10 "Conformance extension - DSM-CC", descrive aggiunte all'ISO-13818-4 per il DSM-CC.

Dopo questa sintetica descrizione delle specifiche MPEG, ciò di cui ci occuperemo riguarda la compressione audio MPEG-1, come componente base di un flusso di dati, e non a completamento di quella video.

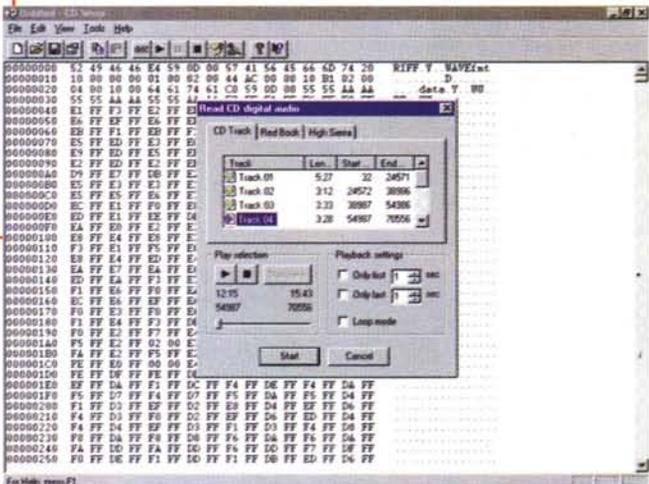
Senza una riduzione dei dati, come tutti sappiamo, i segnali audio digitali consistono tipicamente di campioni a 16 bit registrati a più del doppio della comune larghezza di banda audio (per esempio a 44,1 KHz per i compact disc), che, alla fine, hanno flussi di circa 1,4 MBit per ogni secondo di musica stereo in qualità CD. Una quantità difficile da trattare per tutta una serie di motivi che non torneremo a dire visto

Il sommario schema a blocchi di un encoder ISO-MPEG-1 Audio; mappatura dei dati, quantizzazione e codifica con aggiunta di un modello psicoacustico, impacchettamento dei dati in frame.



Il sommario schema a blocchi di un decoder ISO-MPEG-1 Audio; lo "spacchettamento" dei dati, la loro ricostruzione spettrale, la mappatura inversa.

Per produrre file .mp3 da sorgenti .wav non c'è niente di meglio che fornirci della massima qualità originale; un ambiente di cattura dati da CD audio è quello che occorre. In questo caso stiamo usando uno tra i migliori, il CD Worx.



che ne abbiamo già ampiamente discusso su queste pagine.

Il comitato dedito all'MPEG-1, quindi, ha scelto di raccomandare tre metodi di compressione di una sorgente sonora, nominandoli Audio Layer 1, Layer 2 e Layer 3. Il primo è il più semplice: è un codec sub-band aggiunto ad un modello psicoacustico, il secondo consta di

più sofisticate tecniche di "bit allocation" ed una più grande accuratezza, il terzo si delinea aggiungendo filtri di tipo ibrido e quantizzazioni non-uniformi, codifiche Huffman, una più alta risoluzione in frequenza (almeno 15-20 volte) e tecniche di "bit reservoir". I tre layer sono quindi, dal primo all'ultimo, in ordine di qualità, livello di compressione e in-

Qualità	banda	modo	bitrate	riduzione
telefonica	2,5 KHz	mono	8 Kbps	96:1
onde corte	4,5 KHz	mono	16 Kbps	48:1
radio AM	7,5 KHz	mono	32 Kbps	24:1
radio FM	11 KHz	stereo	56..64 Kbps	26..24:1
simile CD	15 KHz	stereo	96 Kbps	16:1
CD	>15 KHz	stereo	112..128 Kbps	14..12:1

Nota:
 128 Kbps = 15,6 KByte/sec; 96 Kbps = 11,7 KByte/sec;
 64 Kbps = 7,8 KByte/sec; 32 Kbps = 3,9 KByte/sec;
 16 Kbps = 2,0 KByte/sec; 8 Kbps = 1 KByte/sec.

cremento della complessità e delle risorse necessarie al trattamento dei segnali. Per quanto riguarda i decoder, le specifiche dettano una necessaria ed ovvia compatibilità all'indietro, dall'uno all'altro.

Usando codifiche MPEG, potremo "decimare" opportunamente, con tutta una serie di validi criteri, i flussi lineari generati dalle sorgenti digitali. Un esempio fra tutti proprio il flusso specifico dei compact disc, con il quale si può giungere ad ottenere fattori di riduzione da 1:10 fino ad 1:24 e mantenere una qualità sonora che è ancora significativamente migliore di quella che si otterrebbe solo riducendo il sampling rate e la risoluzione dei campioni in esame.

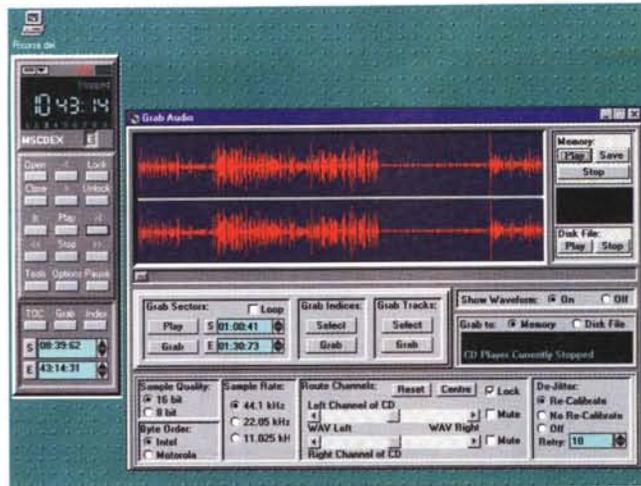
Tutto ciò è realizzato mediante tecniche di "codifiche percettuali", che adeguano la percezione delle onde sonore alle possibilità dell'orecchio umano.

Usando l'MPEG Audio si possono avere, per esempio, riduzioni di data rate nei termini (a partire dal flusso lineare 16 bit, ossia a 1400 Kbps) illustrati di seguito:

1:4	Layer 1 (flussi a 384 Kbps stereo)
1:6/1:8	Layer 2 (flussi a 256..192 Kbps stereo)
1:10/1:12	Layer 3 (flussi a 128..112 Kbps stereo)

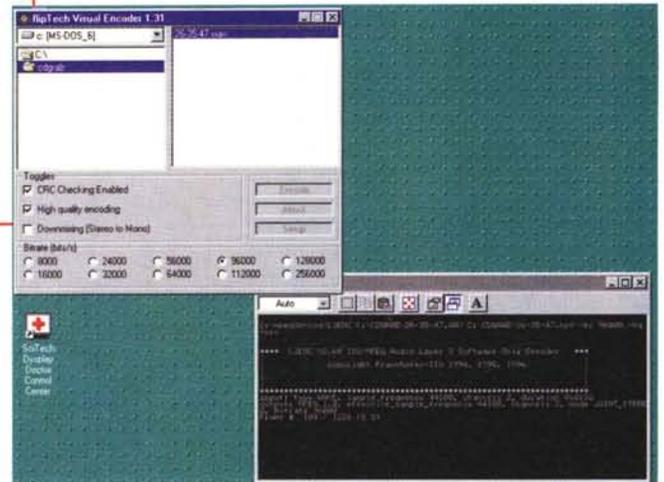
Con questi valori è ancora possibile mantenere una qualità paragonabile a quella dei CD audio.

Esasperando gli effetti stereo e limitando la larghezza di banda, gli schemi di codifica possono ottenere un'accettabile qualità audio anche a più bassi bit rate. Il Layer 3 è il più potente membro della famiglia, proprio dedito alla maggiore riduzione del bit rate con un dato



Un altro ambito di cattura dati CD audio è lo splendido CD Grab Pro for Windows che permette di catturare molto facilmente i dati audio, facilitati anche dalla splendida interfaccia grafica.

Il flipTech Visual Encoder è un ambiente che si lega in modo stretto al "L3ENC" del Fraunhofer Institut, governandone graficamente il suo funzionamento, con opzioni e parametri.



livello qualitativo in riproduzione, se vogliamo, stabilito un certo bit rate è in grado di fornire la più alta qualità rispetto ad altri codec audio.

Le performance tipiche di un codec Layer 3 possono essere riassunte, mettendo insieme le voci principali, così come potete vedere nella tabella pub-

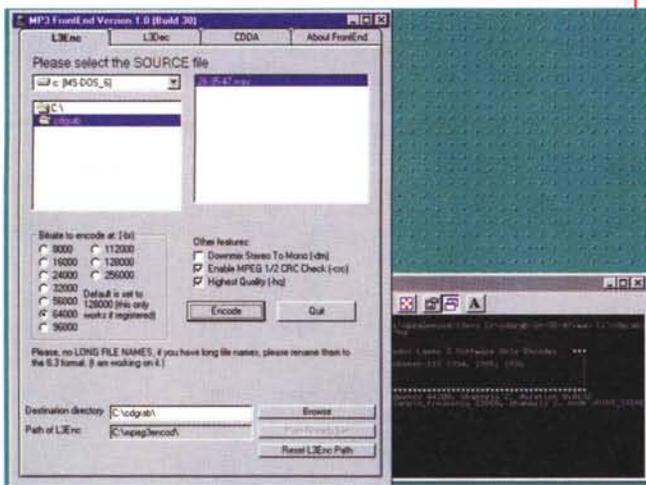
blicata in queste pagine.

Si può notare che anche a bit-rate di soli 50-60 Kbps avremo a disposizione una qualità audio più che dignitosa, senza apparenti pesanti "artifacts" e con una notevole riduzione di dimensioni dall'originale.

Ma quali sono le tecniche che permettono di ottenere simili risultati? Andiamo a vedere le varie fasi di codifica/decodifica.

Encoder e decoder

Fermo restando, come dicevamo, che la potenza necessaria ad una codifica real time di flussi Layer 3 necessita di hardware molto veloce, dell'ordine dei 20 MIPS per canale; molte sono le case dedite allo sviluppo di hw specifico, come: Telos (Zephir) con tecnologia Audioactive, Dialog4, AVT, CCS. Abbiamo a disposizione, a titolo sperimentale (e non in real time), anche una serie di programmi (encoder più decoder) sviluppata dal Fraunhofer Institut Integrierte Schaltungen di Erlangen, in quel di Germania, serissima istituzione che

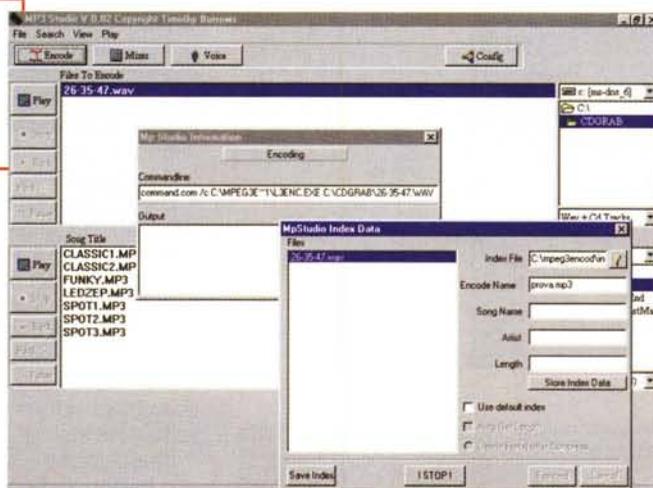


L'MP3 FrontEnd, offre una interfaccia grafica con molte più possibilità rispetto a "L3ENC". In un solo ambiente avviene la codifica, la decodifica e la cattura dati CD audio mediante interfaccia con CD-DA, semplice programma di cattura.



Continuando negli esempi, questo Ariès MP3 Manager offre anch'esso le solite cose, con in più possibilità di playback file .mp3 e midi.

L'MP3 Studio offre la possibilità di encoding di file, capacità di registrare la voce, per interventi immediati, e organizzare playlist del materiale.



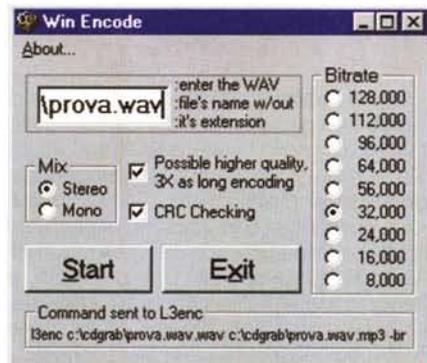
più di ogni altra ha contribuito allo sviluppo delle caratteristiche MPEG-1 Layer 3.

Tornando ad occuparci delle varie fasi di trattamento del segnale, nello schema a blocchi dell'encoder si può notare una prima "mappatura" dei dati, nel dominio della frequenza, ottenuta con banchi di filtri di tipo polifase ed integrata dalla quantizzazione e la codifica mediante modelli psicoacustici (per eliminare le consuete ampie ridondanze nel segnale). Il tutto è convertito in "frame" successivamente impacchettati in ordine (ferme restandone le differenze, il processo è simile al trattamento del video MPEG), ai quali si potranno poi aggiungere altri dati ausiliari. Infatti, il sistema di codifica divide lo spettro delle frequenze in 32 sub-bande ed applica filtri di tipo FFT in parallelo con il modello psicoacustico definito. L'encoder usa tali componenti spettrali come informazioni aggiunte alla codifica, al fine di ottenere una più alta risoluzione alle basse frequenze, dove l'orecchio umano è più sensibile.

Il decoder, invece, molto più semplice e di facile implementazione, esegue il

procedimento inverso, dallo "spacchettamento" (demultiplex) dei frame (con l'estrazione degli eventuali dati ausiliari), alla ricostruzione della trama sonora quantizzata, alla mappatura dei dati.

Come dicevamo, una codifica hardware real time ha ancora una diffusione ridotta ed un costo elevato che la rende adatta esclusivamente agli addetti ai lavori. Via Internet, e proprio al Fraunhofer Institut, è reperibile una versione software (peraltro per più piattaforme hardware) di encoder e de-



Ma a chi può interessare?

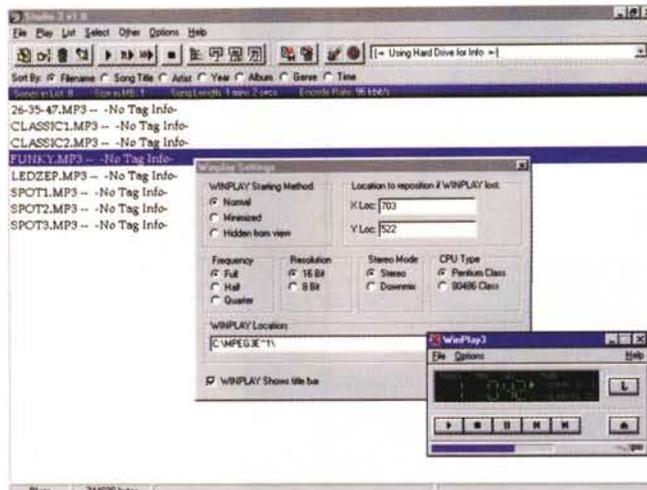
Se solo pensiamo alla ridotta larghezza di banda della quale necessita, pur rimanendo di ottima qualità sonora, potremmo dire senza smentite che l'MPEG Layer 3 potrebbe interessare qualsiasi applicazione legata all'obbligo di impiego di canali di trasporto a capacità limitata, vuoi in tempo reale che in differita, per ottimizzare archivi sonori o per fornire audio di ottima qualità in ambiti broadcast, LAN, ecc. Ed il discorso non poteva non cadere sul suo utilizzo all'interno di Internet, laddove già decine e decine di siti hanno in essere materiale Layer 3, sia per trasmissioni real time di materiale audio informativo, sia come archivio di produzioni commerciali (e non) in tale formato.

Anche qui, basta produrre una ricerca con chiavi come ".mp3", "Layer 3", "MPEG-1 Layer 3" e varianti, per ottene-

Il WinEncode, invece, offre lo stretto necessario, ma con una buona scelta di parametri.

re una miriade di siti, tra ortodossi e non (al limite dell'infrazione dei diritti di copyright), che vi potranno fornire tutte le informazioni del caso.

Ma se ambiti di tipo consumer ("la gente della rete") ne stanno proponendo a forza la sua adozione, in campo broadcast radio, MPEG Layer 3 è già una realtà consolidata; molti collegamenti ISDN tra enti radiotelevisivi avvengono in tale standard, altrettanto per i link audio tra sedi diverse, link tra postazioni mobili sul campo e gli studi



Lo Studio 3, sofisticato "playlist organizer", permette il playback di materiali .mp3 in modo evoluto e sofisticato, mediante tutta una serie di opzioni. In basso è presente il WinPlay 2.0, unico real time decoder MPEG-1 Layer 3.



Per il playback MPEG-1 Layer 3 esiste anche il Resonate, comodissimo player di materiale audio MPEG (per tutti e tre i layer), anche se la decodifica è "quasi" in real time.

Boss.net è uno dei più completi tra le decine e decine di siti Internet in cui trovare materiale MPEG-1 Layer 3, link verso altri siti, archivi di file in standard, utility per la loro gestione.



centrali. Tutta una serie di servizi di pubblica utilità stanno già usando le codifiche MPEG audio per i loro collegamenti.

Fra le decine di case produttrici coinvolte, l'artefice principale di questo proliferare di applicazioni è proprio la Telos, più nota tra gli addetti ai lavori per le proprie apparecchiature digitali broadcast in ambito audio-telefonico e per la sua tecnologia Audioactive, alla realizzazione della quale ha partecipato anche il Fraunhofer Institut.

Ma la concorrenza?

Come abbiamo visto, ed altro ancora ci sarebbe da dire, le codifiche Layer 3 sono quanto di meglio ci si possa attendere da una compressione dati così stringente e "decimante", soprattutto in rete. Anche soluzioni di altro tipo possono aspirare ad avere un impiego comune, per materiali audio di ottima qualità. Una per tutte, e standard "de facto" anche se il livello qualitativo è nettamente diverso, la soluzione RealAudio, così come la quasi sconosciu-

ta Liquid Audio. Prove effettuate confrontando i diversi sistemi hanno evidenziato cose interessanti: con il bit rate di una comunicazione via modem 28,8 (16 Kbit/s), il materiale RealAudio 3.0 (codec, ricordiamo, basato su tecnologia Dolby-Net) suona in maniera molto simile all'MPEG-1 Layer 3 (pur se con qualche dropout di troppo). Tuttavia, quest'ultimo mostra una leggera preponderanza dei toni bassi, apparentemente meno presenti nel Layer 3, mentre altri test, basati su confronti tra i primi due ed il classico ADPCM Microsoft, hanno decretato, con inserita l'opzione "-hq" tipica del "L3ENC" Fraunhofer, una chiara vittoria di quest'ultimo, soprattutto nella stabilità spettrale del segnale.

La soluzione Liquid Audio, invece, codec basato su una variante del Dolby AC-3 con flussi di almeno 20 Kbit/s, è risultato il migliore fra tutti, soprattutto in test di musica leggera e pop, con performance anche migliori del Layer 3-hq a 24 Kbit/s.

Conclusioni

Il caro, vecchio, MPEG continua a stupirci quando ne vengono scoperte sempre nuove possibilità. Quella "only audio" ci sembra da tenere d'occhio attentamente, al di là delle classiche soluzioni video MPEG che abbiamo sempre apprezzato con le dovute riserve quand'era il caso.

La compressione dei dati, qualunque sia, è una soluzione per molti versi difficilmente evitabile, e necessariamente obbligata per altrettanti; se poi si riesce a realizzarla rispettando la massima qualità dell'originale, tanto meglio.

I concetti legati alle tematiche "lossless" sembra debbano ogni volta essere rimessi in discussione, specialmente quando accade che il segnale compresso sembra suonare, confrontato con l'originale, meglio di quest'ultimo o, quantomeno, inganna fortemente l'ascoltatore.

MS

I Professionisti del Controllo

Proxima

Lettori ad ampio spettro applicativo con carte di prossimità e per automazione degli autoparchi.



Sistema di Videotrasmissione HV

Minisistema personale di controllo di allarmi, immagini e suoni su linee commutate.



Small

Letture-Apriporta per scatole da incasso 503 a basso costo e con tutti i tipi di carte elettroniche.



Gate

Letture da PC formato 3,5" per il controllo degli accessi informatici con l'uso di carte elettroniche.



Galaxy

Lettori evoluti per il rilevamento delle presenze con tutti i tipi di carte elettroniche.



Micropos

Lettori-scrittori portatili di carte chip per applicazioni evolute (monetica, fidelity, ecc.)



Controllo ascensori CTA

Sistema di controllo degli ascensori, con dati per la manutenzione e con citofono in cabina, gestito a richiesta 24 ore su 24 dalla Europ Assistance.



Protel

Centrale multimediale evoluta per la gestione di allarmi, immagini e suoni, concentrabile su linee commutate.



SecurTech®
GRUPPO EMMEPI

THE FUTURE MAKER

Via Torino, 64 - 00040 Pavona (Roma) - Tel. 06/9311000 - Fax 06/9311033 (contatto Ing. A. Perla)

CONCESSIONARI - LOMBARDIA: TSI tel. 035/691892 - TRIVENETO: A&TEK tel. 049/8625885 - UMBRIA, VITERBO, RIETI: STEA tel. 0744/284700 - MARCHE, ABRUZZO, MOLISE: ASCANI tel. 0735/632145
CAMPANIA, LATINA, FROSINONE: VDB 081/7524455 - PUGLIA, BASILICATA: SDS tel. 080/5482886 - SARDEGNA: PORTA tel. 070/275777

CERCHIAMO AGENTI E CONCESSIONARI PER LE ZONE LIBERE