

Il formato AAC



Approfondimento del corso di Sistemi Multimediali

prof. Celentano

Outline

- Introduzione
- Formato AAC
- Cenni sulla codifica
- Diffusione
- AAC vs. MP3

Introduzione

AAC è l'acronimo per **A**dvanced **A**udio **C**oding ed è un formato di codifica per l'audio ad alta qualità.

È stato sviluppato dal gruppo MPEG (Moving Picture Experts Group) in collaborazione con aziende quali Apple, Fraunhofer, AT&T Bell Laboratories, Dolby, Sony Corporation e Nokia.

Dall'aprile del 1997 è divenuto uno standard la cui definizione si trova in ISO/IEC 13818 - 7.

AAC in origine veniva chiamato anche MPEG-2 NBC (Not Backward Compatible) in quanto non retro-compatibile con i formati già esistenti.

Caratteristiche principali

- Frequenze di campionamento da 8 a 96 kHz.
- Bitrate da 8 a 512 Kbit/sec.
- 48 canali full bandwidth
- 15 canali per la codifica Low Frequency Enhancement (i canali LFE forniscono informazioni solo per basse frequenze <120 Hz e non hanno quindi un effetto sulla direzione sonora percepita [4]). I canali LFE sono principalmente indicati per la codifica audio vocale “multi-lingua”.
- Codifica percettiva (utilizza il modello psicoacustico per la compressione del segnale audio).

Obiettivi

- Migliore compressione
- Migliore qualità audio anche a basso bit rate
- Flessibilità: la costruzione del codificatore/decodificatore è modulare.
- Profiling per adattare la codifica alla situazione sonora e alle risorse del sistema
- Pensato per l'utilizzo in streaming in tempo reale

Modello psicoacustico

La psicoacustica studia le relazioni tra i fenomeni acustici oggettivi e la percezione uditiva soggettiva

Tali studi sono estremamente importanti e coinvolgono tutte le codifiche audio “lossy” (con perdita di informazione) in quanto definiscono cosa è udibile/percepibile da cosa non lo è.



Sfruttando tali informazioni in fase di compressione audio possiamo tranquillamente deteriorare (anche pesantemente) le parti non udibili/percepibili senza perdere qualità nell'audio finale.

Studi dimostrano che l'orecchio umano riesce a percepire frequenze comprese tra i 20 Hz e i 20 Khz

Codifica AAC

La codifica AAC segue la struttura della codifica MP3.

Sono stati migliorati alcuni blocchi e aggiunte nuove funzionalità in modo da rendere qualitativa e performante la codifica anche a basso bit-rate [3].

AAC utilizza una codifica a tasso variabile (Variable Bit Rate, VBR), cioè una codifica che permette di adattare il numero dei bit usati per secondo di codifica del dato audio, in base alla complessità del flusso audio in un dato momento.

L'algoritmo usato è più performante rispetto a quello dell'MP3, il che permette di ottenere dei file più piccoli con una qualità migliore, avendo bisogno di meno risorse per la codifica o la decodifica.

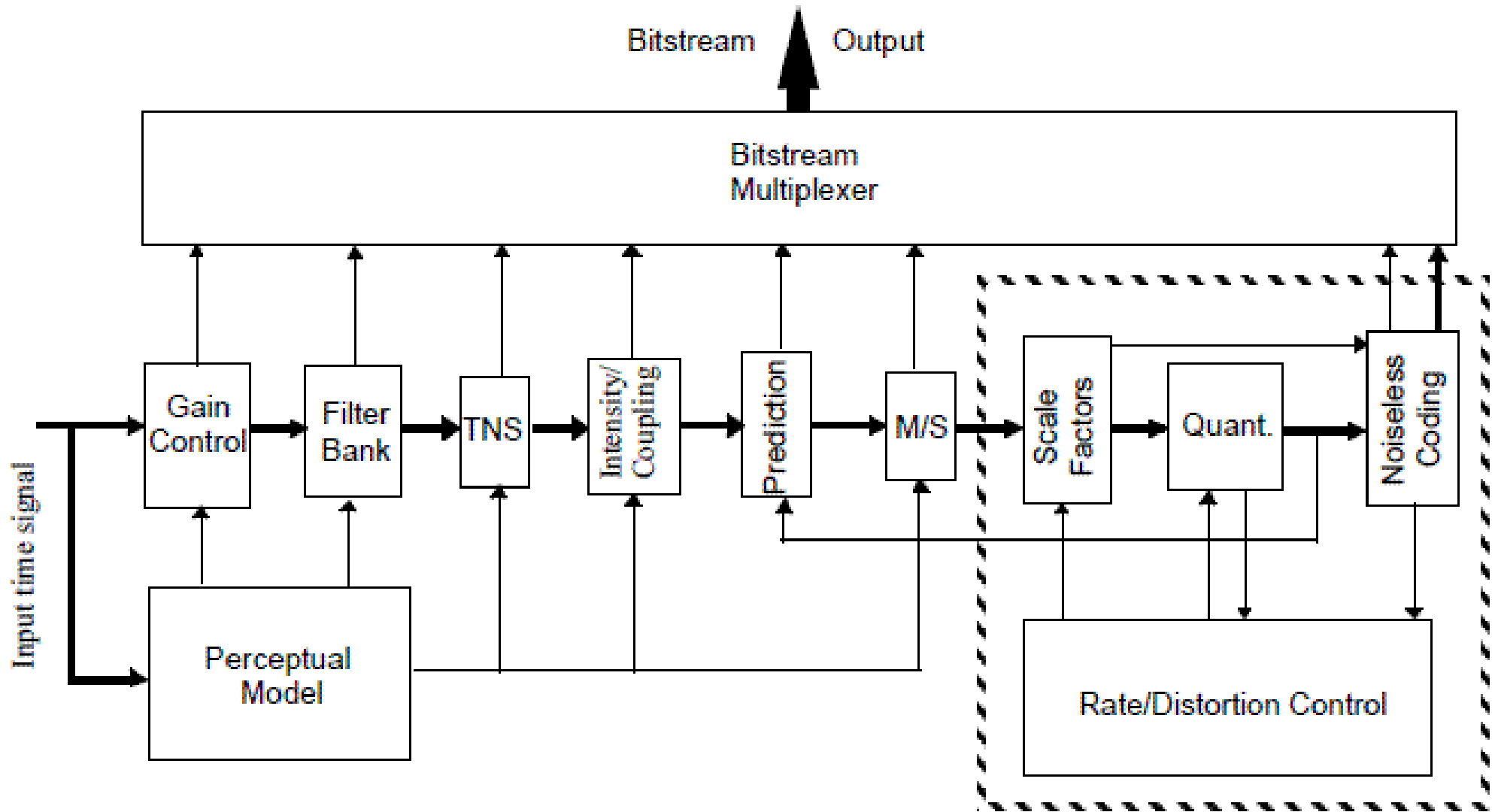
Codifica AAC

AAC, come i precedenti standard MPEG / Audio, segue uno schema di codifica percettivo eliminando tutte quelle informazioni ritenute non necessarie, in quanto non percepite dal nostro apparato uditivo.

Il segnale viene trasformato nel dominio delle frequenze mediante la trasformata di Fourier, suddividendolo per bande tramite il *filter bank*, e successivamente viene eseguita una quantizzazione non lineare guidata dal modello psicoacustico e da una serie di altri tool.

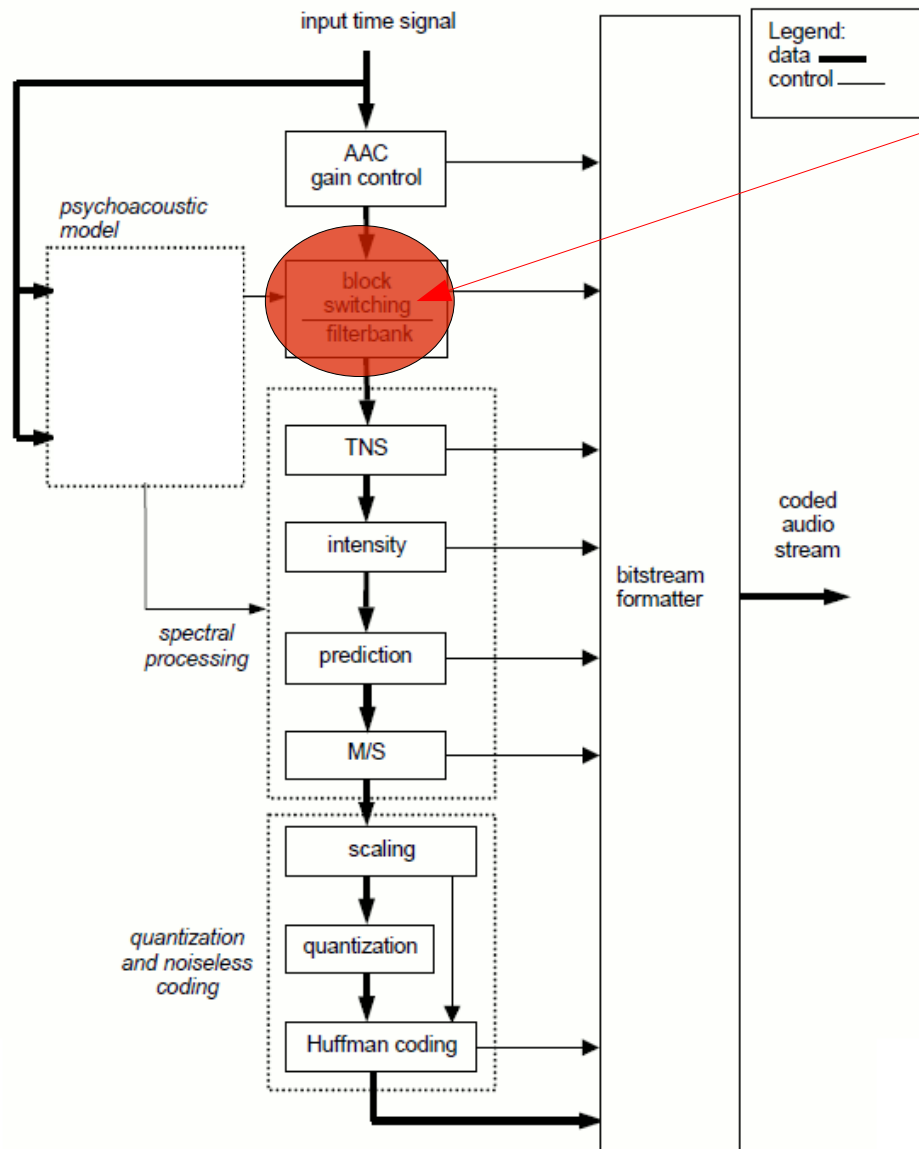
Vengono quindi associati pochi bit alle bande meno percepite ed un numero maggiore di bit a quelle maggiormente importanti.

Codifica AAC



Schema del codificatore AAC

Codifica AAC

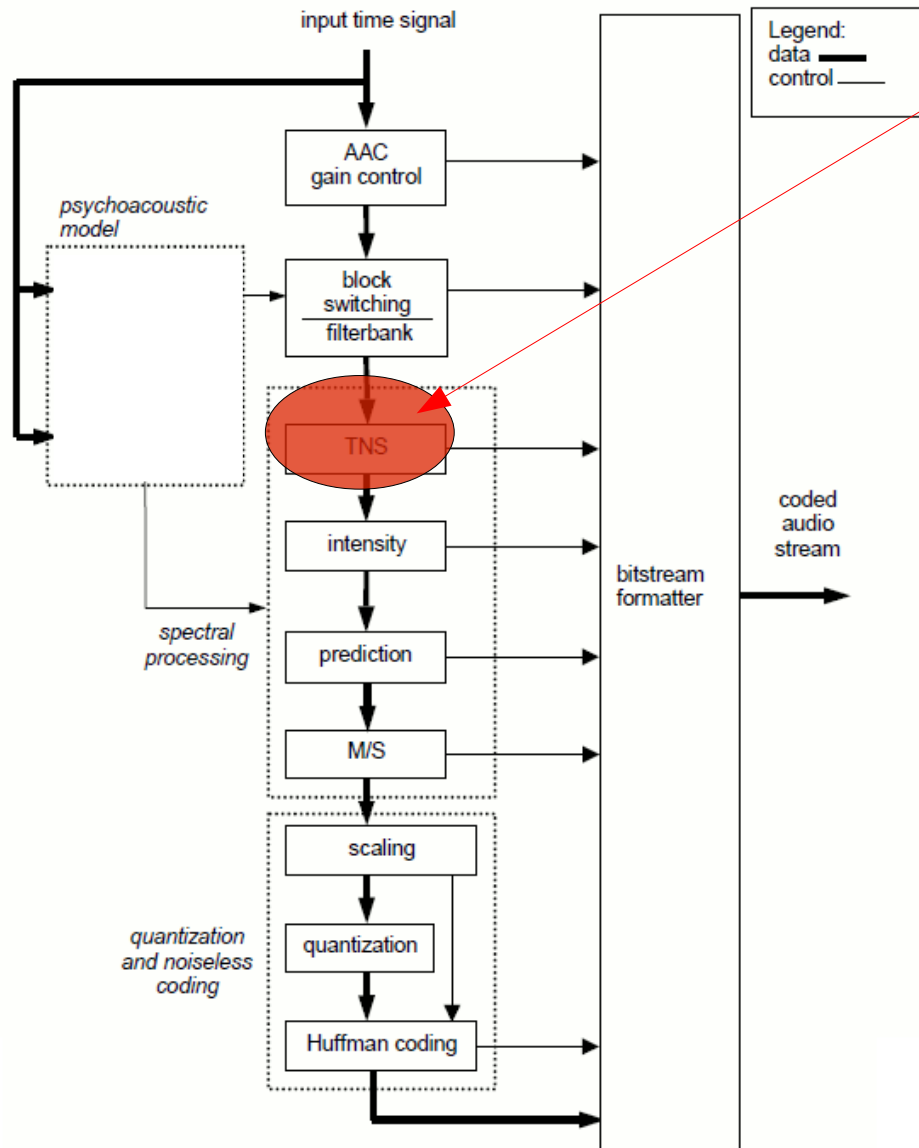


Banco di Filtri: Al posto del banco di filtri polifase presente in MP3, che introduceva una serie di imprecisioni, viene utilizzata direttamente la trasformata MDCT passando dalle 576 linee spettrali di MP3 alle 1024 di AAC.

Tale modifica rende definitivamente incompatibile il formato AAC con le precedenti versioni di MPEG.

Schema di un encoder AAC [1]

Codifica AAC

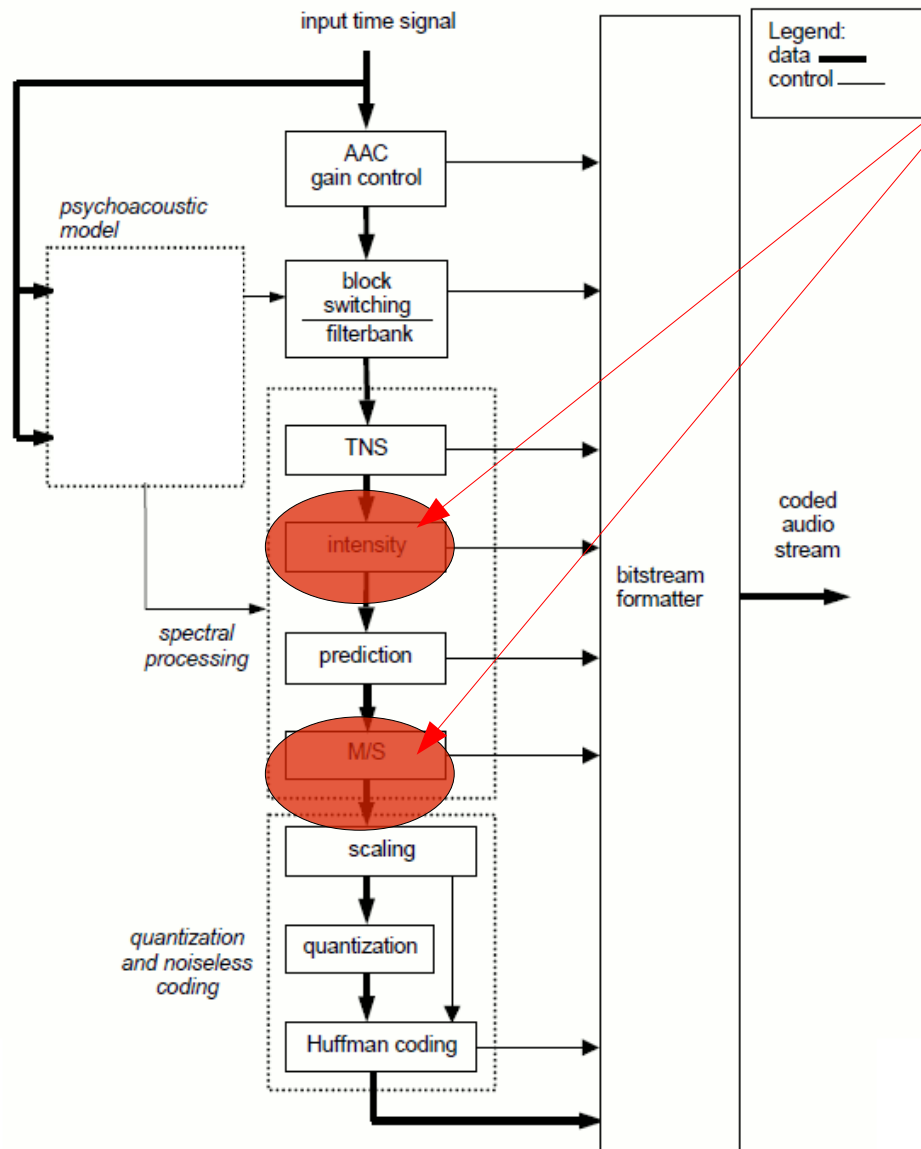


Schema di un encoder AAC [1]

Temporal Noise Shaping (TNS):
Esso predice la distribuzione del rumore di quantizzazione nel tempo lavorando nel dominio delle frequenze. Permette di ottenere evidenti miglioramenti nella codifica audio vocale [2].

Rumore di quantizzazione: nella quantizzazione ogni campione viene associato ad un valore che ne approssima il valore reale. È qui che l'operazione di conversione in digitale di un segnale analogico introduce una degradazione (rumore)

Codifica AAC

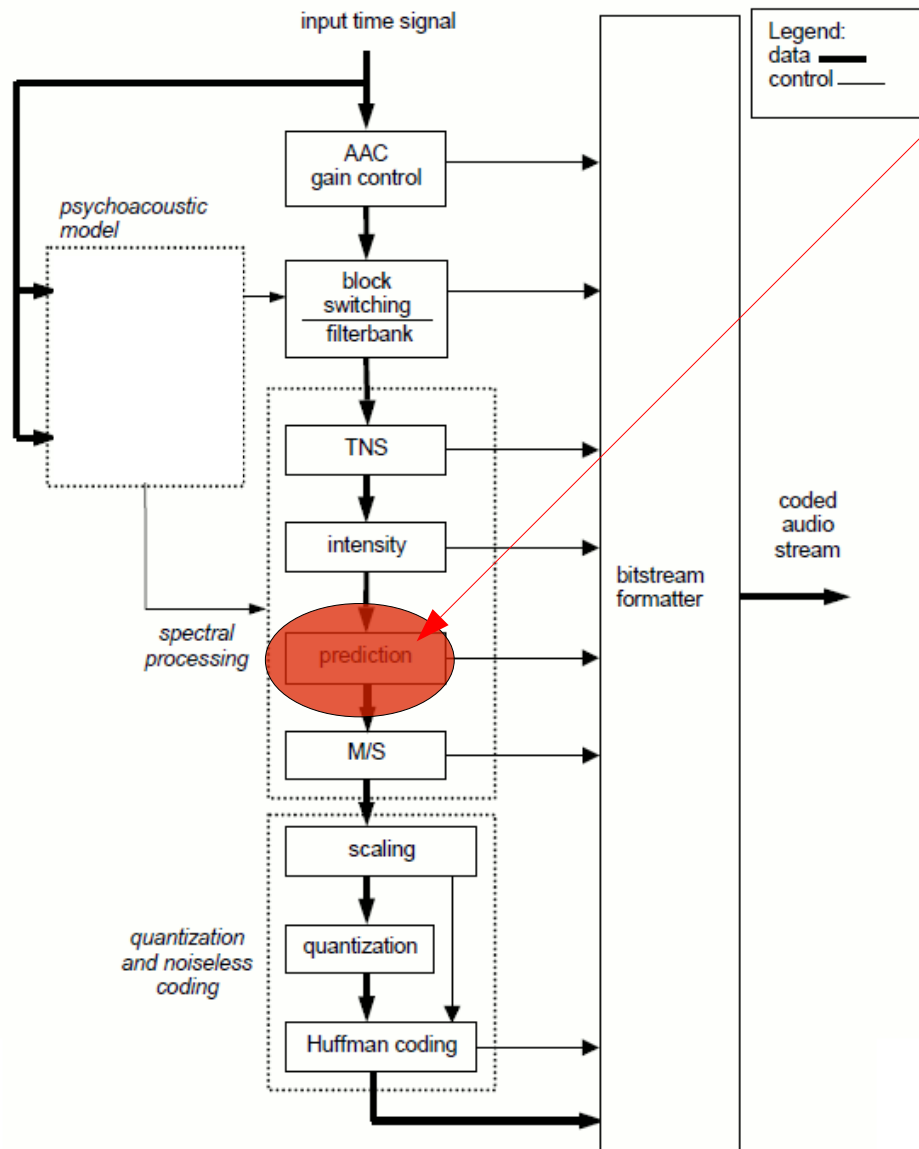


Schema di un encoder AAC [1]

M/S (Mid/Side) e Intensity:
Vengono utilizzati per la codifica stereo/multicanale in singola traccia. Anziché codificare i due canali separatamente si codifica la loro media e la loro differenza (left-right). In questo modo si risparmiano bit in fase di codifica.

Sono presenti diversi algoritmi di codifica stereo compressa e multicanale ulteriormente ottimizzati e migliorati rispetto a quelli presenti in MPEG Layer 3.

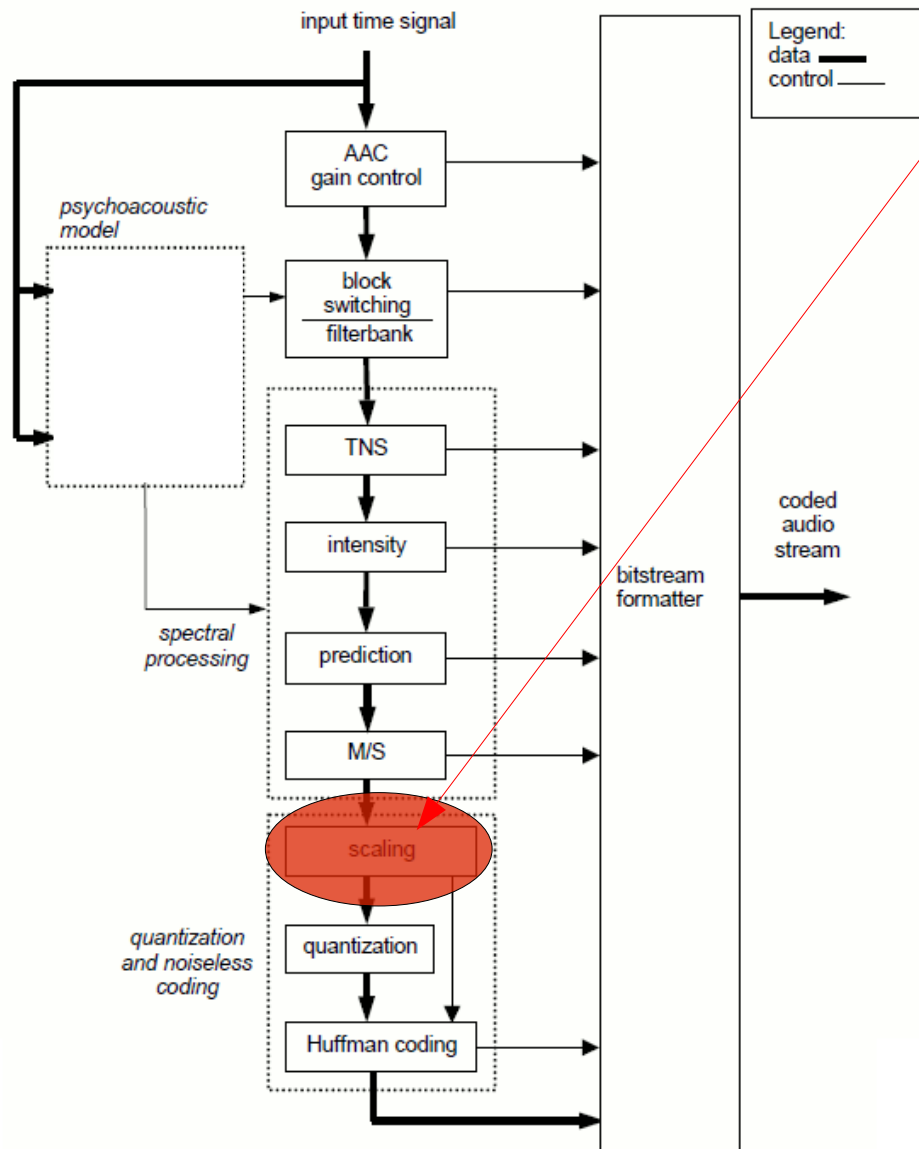
Codifica AAC



Prediction:
algoritmo di predizione, non presente in MPEG Layer 3, è principalmente indicato per la codifica di segnali vocali in quanto maggiormente stazionari nel tempo, e quindi, più facilmente predicibili.

Schema di un encoder AAC [1]

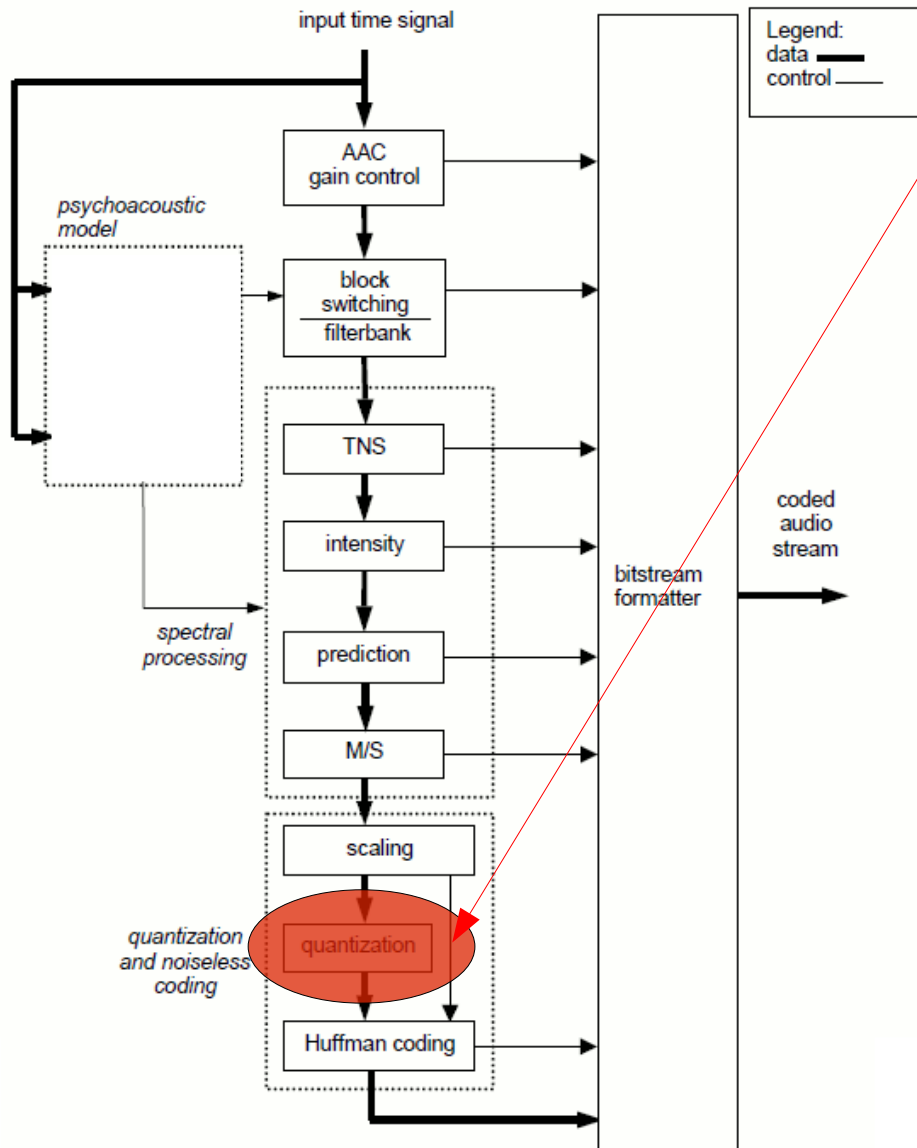
Codifica AAC



Scaling:
come in MPEG Layer 2 e 3 sono presenti dei fattori di scala in grado di pesare opportunamente le varie bande in funzione della loro importanza, raggruppandole per bande critiche (appartenenti allo spettro dell'udito).

Schema di un encoder AAC [1]

Codifica AAC



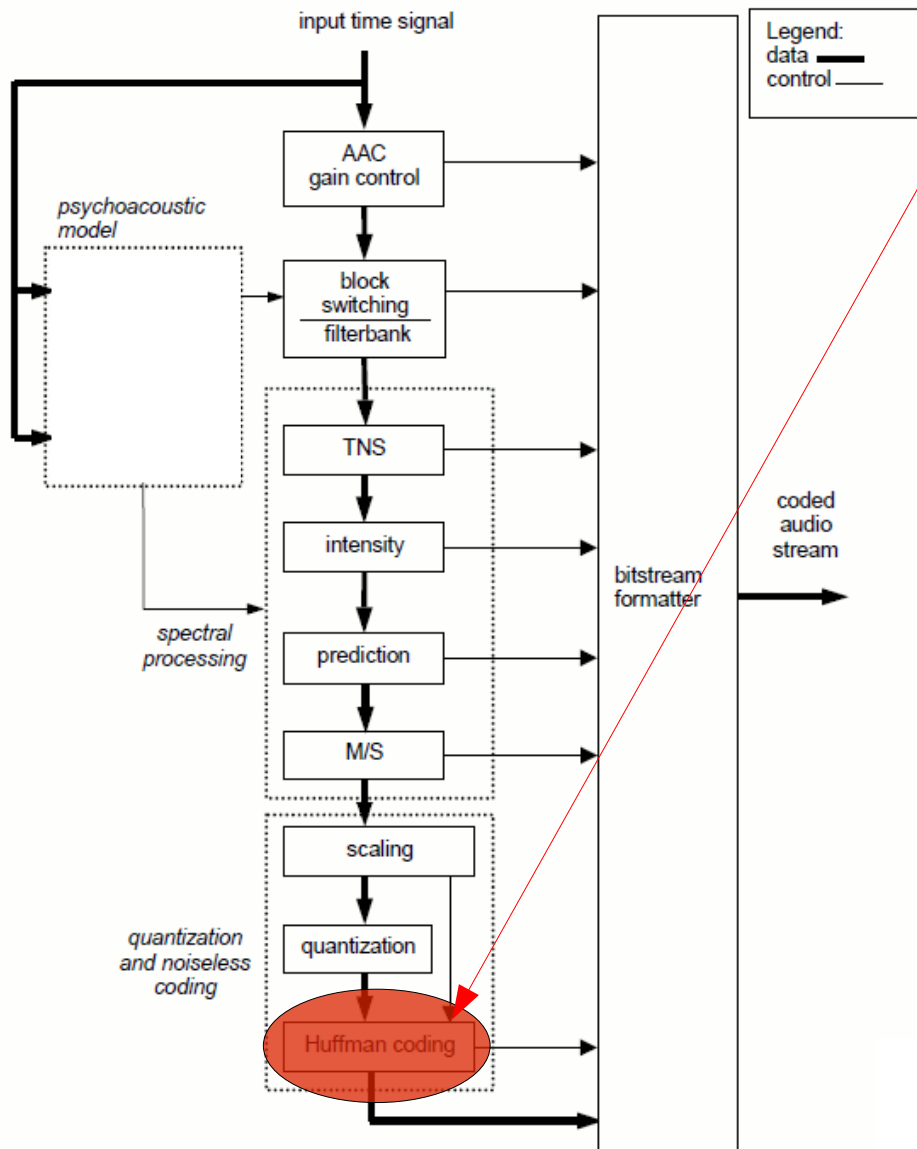
Schema di un encoder AAC [1]

Quantization:

Il sistema di controllo per l'allocazione dei bit è molto più fine in AAC e permette un utilizzo più efficiente e mirato del valore di bitrate.

L'obiettivo è quello di ottenere una quantizzazione dello spettro tale per cui il rumore di quantizzazione introdotto si trovi al di sotto della soglia di udibilità definita dal modello psicoacustico.

Codifica AAC



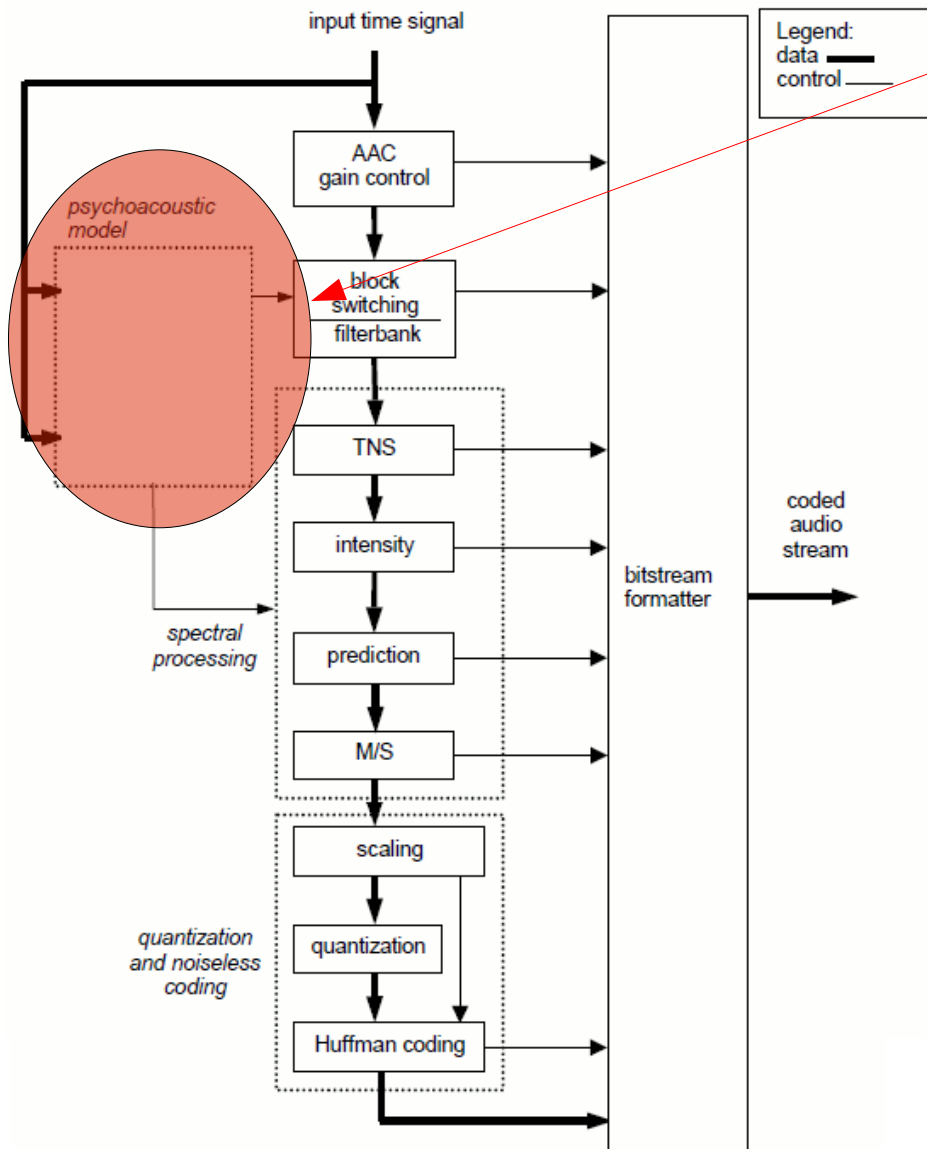
Schema di un encoder AAC [1]

Huffman coding:

Una volta ottenuta la quantizzazione numerica si procede alla codifica di Huffman in modo da minimizzare (e quindi comprimere) ulteriormente le ridondanze associando a elementi più frequenti una codifica più corta.

Per fare in modo che il segnale sia codificato più efficacemente si utilizzano diverse tabelle di codici di Huffman per le diverse parti dello spettro udibile [3].

Codifica AAC



Schema di un encoder AAC [1]

Modello psicoacustico:

L'output del modello psicoacustico è costituito dai parametri della soglia di mascheramento (Signal to Mask Ratio) per l'attuale campione in esame.

In MPEG Layer 3 le soglie di mascheramento sono simili alle bande critiche dell'orecchio umano mentre in AAC sono computate sui singoli campioni.

Se il rumore di quantizzazione viene mantenuto al di sotto della soglia di mascheramento, la compressione risultante sarà identica al segnale originale, in quanto si vanno a sopprimere elementi non udibili.

AAC Profiles

Il consorzio MPEG fornisce AAC di tre profili di default:

- **Main:**
Questo tipo di profilo viene utilizzato quando la quantità di memoria disponibile e le capacità di elaborazione sono elevate e fa uso di tutti i tool del codificatore.
- **Profilo LC (“Low Complexity”):**
Questo profilo viene utilizzato quando si hanno scarse risorse a disposizione
- **Profilo SSR (“Scalable Sample Rate”):**
Questo profilo vieta la codifica multicanale e limita l'uso del TNS.

Formato di trasporto di AAC

Lo standard definisce due formati per il trasporto dell'audio codificato con AAC:

- **ADIF** "Audio Data Interchange Format" tutte le informazioni che controllano il decodificatore (come la frequenza di campionamento), sono raccolte nell'header e subito dopo abbiamo lo stream di dati.
- **ADTS** "Audio Data Transport Stream" codifica pacchetti di dati AAC in frame con una intestazione simile a quella di MPEG – Layer 3. ADTS è ormai lo standard de facto per numerose applicazioni che usano AAC.

Metadati

AAC (come anche MP3) non prevede metadati [5] che ne descrivano testualmente il contenuto.

Viene dunque utilizzato lo standard ID3 che permette di “impacchettare” in un file unico informazioni quali:

Campo	Dimensione
Song Title	30 characters
Artist	30 characters
Album	30 characters
Year	4 characters
Comment	30 characters
Genre	1 byte

Dati raccolti in ID3 v1

Utilizzo in streaming

Come accennato tra gli obiettivi di AAC troviamo l'utilizzo in streaming. La rete è soggetta a perdita di pacchetti e trattando fonti audio una ritrasmissione (tipica del protocollo TCP) non risulta spesso conveniente. Pertanto sono stati implementati quattro metodi per il controllo degli errori (MP3 ne prevede solo uno, il CRC) [2]:

- *Error Concealment*: Attraverso il CRC vengono rilevati gli errori che poi vengono corretti attraverso la ricostruzione delle parti mancanti per approssimazione o semplicemente con l'aggiunta di silenzi.
- *Error Resilience*: questa tecnica ha l'obiettivo di rendere uno streaming AAC più robusto agli errori.

Utilizzo in streaming

- *Error Detection*: questa tecnica permette di rilevare la presenza di errori nello streaming aggiungendo una serie di CRC associati a parti diverse del bitstream.
- *Error Protection*: Essa suddivide lo streaming MPEG-4 in parti cosiddette “sensibili” e ad ognuna di esse vengono associati dei bit di controllo che ne permettano il rilevamento di errori e l’eventuale correzione.

AAC vs. MP3

Numerosi test (test d'ascolto, metodi oggettivi di misurazione e tecniche di misura percettiva [3]) sono stati effettuati sullo standard AAC per capire realmente, se e di quanto migliorasse la qualità audio rispetto alle precedenti codifiche MPEG-1 ed MPEG-2.

A parità di condizioni (stesso valore di bitrate e frequenza di campionamento, medesimi brani audio, uguali caratteristiche dell'ambiente, stessi tester, ecc.), i risultati hanno dimostrato come le codifiche AAC mono, stereo e multicanale siano qualitativamente migliori rispetto alle stesse codifiche MPEG Layer 2 e Layer 3.

Più precisamente è stato dimostrato come una codifica audio AAC con bitrate a 96 Kbit/sec sia paragonabile, in termini di qualità, ad un MP3 a 128 Kbit/sec [3].

AAC vs. MP3

	AAC	MP3
Frequenza dei sample	8 – 96 kHz	16 – 48 kHz
Canali	1 - 48	2 (MPEG-1) – 5.1 (MPEG 2)
Banco dei filtri	MDCT puro	Ibrido per retrocompatibilità
Dimensione dei blocchi	1024	576
Moduli aggiuntivi	TNS Backwards prediction PNS	No
Profili	I moduli possono essere combinati per creare profili di codifica ad hoc	No

Diffusione

Grazie all'alta qualità audio prodotta e al basso bit-rate necessario viene designato come il successore dell'ormai noto MP3.

Ad oggi viene impiegato nelle piattaforme quali

- Apple iTunes, iPhone e iPod
- Nintendo DSi e Nintendo Wii
- Sony PlayStation 3
- Cellulari Nokia e Sony

Il formato comunque stenta a decollare, probabilmente a causa dell'enorme quantità di lettori MP3 già presenti sul mercato.

Diffusione

AAC viene anche utilizzato in una versione chiamata HE-AAC (High-Efficiency Advanced Audio Coding).

Tale versione è divenuta uno standard ISO/IEC 14496-3 e rappresenta un'estensione al profilo LC-AAC per l'utilizzo a bassi bit-rate.

HE-AAC si presta dunque particolarmente bene all'utilizzo in streaming e di fatto viene utilizzato in questo contesto.

Test scientifici condotti dall'European Broadcasting Union hanno indicato HE-AAC a 48 kbit/s ad un livello di qualità "Excellent" mentre il formato MP3 non raggiunge neanche la metà della scala di valori.

Programmi



Mac OS

- ▶ Apple iTunes
- ▶ VideoLAN VLC media player
- ▶ Roxio Toast 10 Titanium
- ▶ RealNetworks RealPlayer
- ▶ MPlayer



Windows

- ▶ Apple iTunes
- ▶ Media Player Classic
- ▶ KSP Sound Player
- ▶ VideoLAN VLC media player
- ▶ Roxio Creator 2010
- ▶ RealNetworks RealPlayer
- ▶ MPlayer



Linux

- ▶ VideoLAN VLC media player
- ▶ RealNetworks RealPlayer
- ▶ MPlayer
- ▶ XMMS with AAC plugin

Conclusioni

Utilizzando un encoder con buone performance entrambi i formati (MP3 ed AAC) possono comprimere segnali audio musicali rimanendo ad una qualità prossima al CD.

AAC è il successore di MP3 in quanto fornisce una qualità audio molto vicina al CD ma con fattori di compressione superiori. Questo permette di incrementare del 50% la vita delle memorie flash dei lettori musicali.

AAC permette una qualità superiore sconfinando nell'alta definizione grazie al sampling rate a 96 kHz [3].

Bibliografia

- [1] *Generic coding of moving pictures and associated audio information Part 7: Advanced Coding (AAC) - ISO/IEC 13818-7* - MPEG Group (2004) Audio
<http://jongyeob.com/moniwiki/pds/upload/13818-7.pdf>
- [2] *AAC - MP4 OVERVIEW*
Giancarlo Vercellesi – Laboratorio informatica Musicale -Dipartimento di Informatica e Comunicazione Università degli Studi di Milano (2003)
http://www.lim.dico.unimi.it/teaching/materiali/mp4-aac_infoMusicale.pdf
- [3] *MP3 AND AAC EXPLAINED*
Karlheinz Brandenburg - AES 17th International Conference on High Quality Audio Coding (2006)
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.26.5956&rep=rep1&type=pdf>
- [4] *WHAT IS LFE CHANNEL ?*
Dolby
http://www.dolby.com/uploadedFiles/zz-_Shared_Assets/English_PDFs/Professional/38_LFE.pdf
- [5] *Tecniche di codifica musicale*
F. Benedetto, G. Giunta - Elaborazione Numerica dei Segnali, Università di Roma Tre.
<http://www.comlab.uniroma3.it/ens/Nota%20su%20Codifica%20musicale.pdf>
- [6] *L'audio digitale*
A. Celentano – Corso di Sistemi Multimediali – Dipartimento di Informatica dell'Università Ca' Foscari (2009)
- [7] *Scenario e considerazioni per la transizione verso la radiofonia digitale*
AICT (2005)
<http://net203-143-025.mclink.it/files/000037/00003749.pdf>