

Conceptes bàsics de so digital

Objectius

- Repassar els conceptes bàsics de la digitalització dels senyals analògics.
- Recordar algunes de les seves característiques més importants.
- Entendre com es converteix un senyal digital en analògic.

Digitalització del so

L'àudio digital és més immune a les degradacions –distorsions– i als sorolls que l'analògic perquè, encara que els polsos electrònics es deformin, es poden reconstruir i regenerar en cada etapa del procés. Amb el pas del temps, el so digital sempre conserva les característiques originals. No perd brillantor ni volum.

La qualitat de l'àudio digital només depèn de la freqüència de mostreig i del nombre de bits emprats. La qualitat només depèn de la codificació i la descodificació del senyal.

Els aparells digitals tenen una dinàmica més bona que els analògics perquè els suports no tenen soroll de fons. D'altra banda, la separació entre canals –diafonia– és molt elevada.

Gràcies als sistemes de compressió, el so digital de bona qualitat ocupa cada vegada menys espai. A més, es poden fer còpies d'un material original sense perdre generacions. En el fons no es fan còpies, sinó originals nous.

L'accés a la informació és ràpid i precís. Un únic aparell pot integrar la resta i fer-ho tot. És el que es coneix amb el nom d'**estacions de treball d'àudio digital (DAW)**. Sistemes com el Pro Tools HD o el Cool Edit Pro 2 ho són.

La informació digitalitzada, guardada a la memòria d'un ordinador, permet de treballar en xarxa. Es poden enviar fitxers a un ordinador que estigui al costat mateix o als antípodes. En el sistema Pro Tools, la transmissió de fitxers de so és la funció Digi Delivery.

El treball en xarxa també possibilita treballar simultàniament des de ciutats diferents en una mateixa sessió d'edició. L'aplicació DigiStudio de Pro Tools permet d'interconnectar dos editors que es trobin en estudis diferents per Internet i treballar en un entorn virtual comú.

Amb les xarxes informàtiques també es pot accedir a les bases de sons digitals de música o efectes que permeten als usuaris de comprar efectes de so o músiques de llibreria i carregar-los directament a l'ordinador –per exemple, la base de dades Digi Pro Net, de Pro Tools.

L'edició de so digital no és destructiva. Els talls es fan sobre còpies de l'original, de manera que, si hi ha algun error, sempre es pot recuperar l'original.

La gran capacitat de manipulació del so facilita la creativitat sonora. Amb el **processament digital del senyal (DSP)** es poden fer moltes transformacions d'un so real, i també es poden crear sons nous per síntesi electrònica.

Hi ha sistemes que poden analitzar el soroll de fons d'un enregistrament digitalitzat i crear-ne un d'igual però en contrafase, per a eliminar-lo. Amb els processadors psicoacústics d'harmònics es pot retornar la brillantor a un so apagat. Aquestes dues funcions permeten la reconstrucció sonora de materials antics.

Els ordinadors aplicats a la radiodifusió permeten d'automatitzar l'emissió totalment o parcialment.



Automatització total o parcial de l'emissió

En els ordinadors emprats a les radiofórmules, la combinatòria del sistema té en compte que un disc no soni sempre a la mateixa hora i minut, o seguit sempre d'un mateix tema.

Alguns dels programes dissenyats per a la ràdio informen dels anuncis emesos i en gestionen la facturació automàticament.



La digitalització del so ha portat la creativitat sonora a les llars. Qualsevol ordinador multimèdia actual permet a qualsevol persona d'experimentar l'enregistrament, l'edició i el processament del so.

Convertidor analògic/digital

Repassem breument les fases del procés de digitalització d'un senyal analògic: filtratge del senyal analògic, mostratge, quantificació i codificació, multiplexió dels canals, inserció de codis de detecció i correcció d'errors, i pas dels valors binaris a senyal electrònic modulad.



Lectura recomanada

Trobareu moltes dades tècniques sobre la digitalització dels senyals analògics i els seus paràmetres en el temari de l'assignatura de la UOC *Fonaments de tecnologia audiovisual*.

■ Filtratge

Abans de la digitalització del so, es filtra el senyal analògic per dues raons:

1. Per a processar només les freqüències que podem escoltar, és a dir, per sota dels 20 kHz, de manera que la resta no entrin a la cadena de la digitalització i no ocupin memòria o amplada de banda innecessària.
2. I perquè cal filtrar totes les freqüències per sobre de la meitat de la de mostratge per a evitar la generació de freqüències indesitjables en el marge audible –el que en anglès es coneix com a *filtre antialiasing*.

Per això, es fa passar el senyal analògic per un filtre passabaix amb freqüència de tall de 20 kHz. Immediatament després, el senyal filtrat entra al circuit de retenció i mesura (*sample and hold*) per a ser digitalitzat.

■ Mostratge



El mostratge consisteix a prendre una mostra del senyal analògic cada cert temps, segons el que marqui la freqüència de mostratge, per a fer una exploració d'un senyal continu. Després de l'exploració, s'obtenen una sèrie de valors instantanis discrets. La freqüència de mostratge ha de ser com a mínim el doble (2,2) de la freqüència més alta de l'àudio que es digitalitzi (criteri de Nyquist).



Si no es fa així, la mostra analitzada no es correspondria amb la forma de l'ona original i sorgirien unes freqüències alienes al marge auditiu que provocarien distorsió harmònica. L'efecte d'un mostratge insuficient es coneix en anglès com a *aliasing* (*alias frequencies*).

La freqüència més alta per a codificar en àudio és de 20 kHz; aleshores, la freqüència de mostratge mínima ha de ser 44 kHz. Per aquest motiu, els aparells digitals de primera generació feien servir les freqüències de 44,1 kHz o 48 kHz. Per a usos de radiodifusió es va escollir una freqüència de 32 kHz, atès que en la difusió en freqüència modulada només es transmet àudio de baixa freqüència fins a 15 kHz.



Les freqüències de mostratge més emprades

L'abril del 1978, l'AES (Audio Engineering Society) va adoptar l'ús dels 44,056 Hz (44,1 kHz) com a freqüència estàndard de mostratge. El mateix any, l'Associació d'Indústries Electròniques Japoneses (EIAJ) va establir el format de quantificació de 14 bits per a usos domèstics.

Resumim les freqüències de mostratge més emprades:

- 11 kHz: fitxers de so dels primers CD-ROM multimèdia.
 - 32 kHz: dona una qualitat de so similar a la de la ràdio en freqüència modulada.
 - 44,1 kHz: emprada en els discos compactes.
 - 48 kHz: és un valor que s'utilitza molt en aparells de so professional dels noranta.
 - 96 i 192 kHz: valors que fan servir els equips més moderns de gamma alta –Pro Tools HD, Cool Edit Pro 2, etc. L'any 1994, Pioneer Electronics va presentar el Wide Band DAT-Recorder, que ja treballava a 96 kHz i 32 bits.
-

Com més alta sigui la freqüència de mostratge emprada, més exactitud s'obindrà en la codificació del so, la resposta transitòria serà més bona i, lògicament, la fidelitat sonora també. Però les freqüències de mostratge massa altes (96 o 192 kHz) tenen un altre objectiu: permeten de manipular repetidament un mateix senyal amb diversos efectes i amb la màxima precisió. Com més mostres es tinguin del senyal original, de més valors disposarà el sistema informàtic per a fer els càlculs dels algorismes amb més exactitud. L'augment de la velocitat dels processadors, que tracten milions de dades per segon, ha permès d'incrementar les freqüències de mostratge.



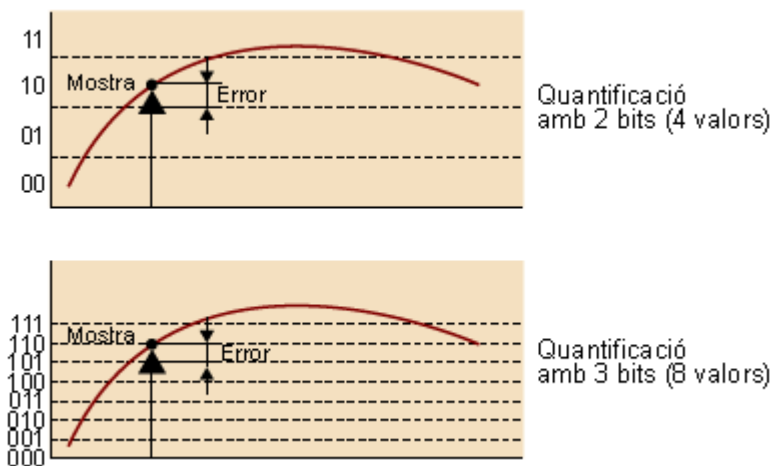
Escoltem el mateix so, sempre digitalitzat amb una quantificació de 16 bits per paraula, però amb tres freqüències de mostratge diferents:

- La primera: 44,1 kHz. El fitxer ocupa 126 kB.
- La segona: 8 kHz. El fitxer ocupa 23 kB.
- I la de pitjor qualitat: 2 kHz. El fitxer ocupa 5,86 kB.

■ Quantificació i codificació en binari

Una vegada obtingudes les mostres d'un senyal, es quantifiquen. És a dir, es mesuren els volts de cada mostra i el valor del voltatge de cada mesura es codifica en forma de paraula binària. El bit de l'esquerra de cada paraula és el més significatiu (MSB) i el de més a la dreta, el que ho és menys (LSB).

Per a saber el nombre de nivells de quantificació que obtenim amb 8 o 10 bits, fem 2^8 (256 nivells) o 2^{10} (1.024 nivells). Com més nombre de bits per paraula, més exacta serà la mesura del valor de l'amplitud de la mostra.



La resolució també determina el valor de la relació senyal/soroll (S/N). Aproximadament, cada bit afegit fa pujar aquest paràmetre en uns 6 dB.



Nombre de nivells de quantificació	Nombre de bits	Relació senyal/soroll (dB) (aproximada)
64	6	37
256	8	49
1.024	10	61
65.536	16	97
1.048.576	20	121
16.777.216	24	145

Pel que fa al nombre de bits per mostra, les quantitats més habituals són:



Format de quantificació per a usos domèstics

L'any 1978, l'Associació d'Indústries Electròniques Japoneses (EIAJ) va establir el format de quantificació de 14 bits per a usos domèstics.

- 8 bits: so dels primers CD-ROM multimèdia.
- 16 bits: so professional dels noranta.
- 24 bits: so professional dels noranta de gamma alta.
- 32 bits: equips professionals més moderns.

Com més bits per paraula –o més freqüència de mostratge–, més espai de memòria ocupat, més gran ha de ser l'amplada de banda del canal de transmissió, o s'ha de poder treballar amb *bit rates* més elevades. Recordeu que vuit bits consecutius formen un byte (kB, quilobyte; MB, megabyte).

Si es treballa amb un nombre de bits baix, es poden produir indefinicions en els valors de les mesures de les parts més febles del senyal, que en el moment de la reproducció produïrien un efecte semblant a la distorsió harmònica. Per a evitar-ho, a l'entrada del convertidor d'analògic a digital, els senyals fluïxos se sobreposen a un soroll blanc aleatori que els fa de coixí. Aquest soroll aleatori, anomenat **dither**, millora la resolució del procés, atès que fa pujar el senyal analògic per sobre del nivell del bit menys significatiu i el manté sempre dins dels marges de quantificació. Això en redueix la distorsió harmònica i la d'intermodulació.



Escoltem el mateix so original sempre digitalitzat amb una freqüència de mostratge de 44,1 kHz, primer amb una quantificació de 16 bits per paraula. El fitxer ocupa 126 kB.

Ara el digitalitzem amb una quantificació de 8 bits per paraula. El fitxer ocupa, lògicament, la meitat, 63,3 kB. En aquest cas concret la qualitat de so no canvia gaire. Per tant, si l'haguéssim d'enviar, ens podríem estalviar temps de transmissió si quantifiquéssim la informació a 8 bits.

■ Multiplexió

Molts sistemes que treballen amb informació de dos canals, esquerre i dret, multiplexen les dades binàries. Això significa que cada bloc de mostres conté informació del canal dret i de l'esquerre, de manera que les dades dels dos canals podrien passar per una única línia.

Si multiplexéssim paraules binàries de quatre bits del canal dret "dddd" i de l'esquerre, "eeee", obtindríem aquestes noves ràfegues binàries: "ddee" i "ddee".

■ Detecció i correcció d'errors

En àudio digital hi ha dos tipus d'errors bàsics:

1. Els **errors de ràfega**, que provoquen la pèrdua de diverses mostres successives. A vegades els pot provocar l'absència temporal del senyal, un *drop-out* de les cintes, acumulació de pols a la superfície del CD, una interferència produïda per un pic de tensió, etc.
2. Els **errors aleatoris**, que provoquen la pèrdua de mostres aïllades. Acostumen a ser conseqüència d'un soroll o d'una qualitat de senyal insuficient.

Per a resoldre o minimitzar l'acció negativa d'aquests errors, abans d'enregistrar o transmetre el senyal s'intercalen una sèrie de bits de control en les dades d'àudio original que alteren l'ordre de les mostres inicials. Un exemple de mètode de resolució d'errors aleatoris seria el codi Reed Solomon d'intercalació encreuada –*cross interleave Reed Solomon code* (CIRC)–, que és el que s'utilitza en els discos compactes.



El codi Reed Solomon d'intercalació encreuada (CIRC)

El CIRC fa sumes verticals, horitzontals i diagonals de files i columnes de les mostres quantificades, i pot corregir més d'una ràfega d'errors en una paraula.

A grans trets, aquest codi de correcció d'errors funciona així: el sistema agafa com a bloc inicial sis períodes de mostratge. Cada període de mostratge té 16 bits que corresponen al canal esquerre i 16 al dret. És a dir, cada bloc d'àudio original està format per dotze paraules de 16 bits i un total de 192 bits (32 bits \times 6 períodes).

Cada grup de 16 bits se separa en dues paraules de 8 bits. Per tant, els 32 bits de cada període de mostratge passen a estar expressats en quatre paraules de 8 bits. El bloc d'àudio original, els sis períodes, passen a ser vint-i-quatre paraules de 8 bits.

Aquestes vint-i-quatre paraules de 8 bits entren al codificador CIRC, que genera vuit paraules de paritat i una paraula de principi de bloc, totes de 8 bits. Surten, per tant, trenta-tres paraules de 8 bits (24 + 8 + 1).

A cada paraula de 8 bits s'hi afegeixen 6 bits d'intercalació per a evitar ràfegues seguides d'1 i 0. Això es coneix com a **modulació de bits 8 a 14** (*eight to fourteen modulation*, EFM). Ara tenim trenta-tres paraules de 14 bits.

El pas següent és intercalar un grup de 3 bits de separació (*merging*) per a unir cada paraula de 14 bits amb la següent. Les trenta-tres paraules tenen 17 bits cadascuna. Fan un total de 561 bits.

El processament acaba afegint 27 bits de sincronia al final del bloc. Per tant, surten 588 bits.

En resum, un bloc d'àudio original format per 192 bits acaba tenint 588 bits. O, dit d'una altra manera, 192 bits representen la codificació del so original i 396 bits s'encarreguen de la preservació i la seguretat.

Quan es perden mostres, aquestes dades redundants de seguretat permeten d'interpol·lar les dues mostres adjacents i obtenir uns valors aproximats que n'omplin el buit. Aquesta manera de dissimular els errors amb algorismes matemàtics es coneix com a *cancel·lació*. En un missatge musical, la interpolació momentània no provoca cap conseqüència perceptible. Altres sistemes més avançats fan una anàlisi en freqüència del so i quan un valor es perd, s'insereixen mostres que tinguin les mateixes característiques espectrals. Quan s'ha perdut molta informació i la cancel·lació no resol el problema, el sistema emmudeix.

■ El senyal digital electrònic

Després de multiplexar i aplicar els codis de seguretat, els 0 i 1 es transformen en polsos elèctrics que es poden emmagatzemar en un suport, enviar per una línia o transmetre radioelèctricament.

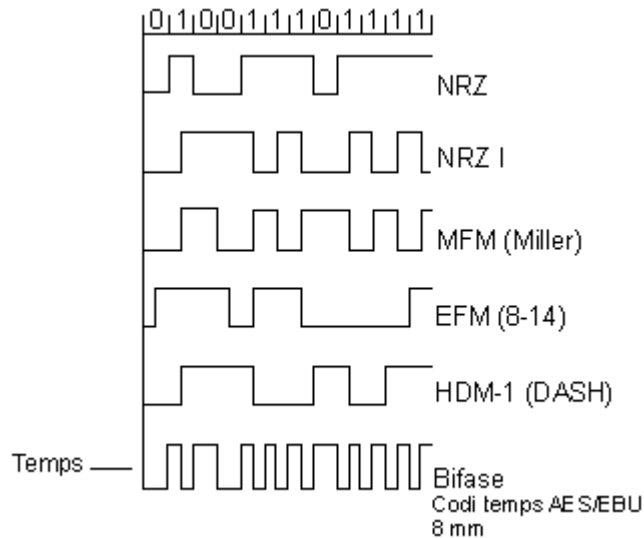


Alguns codis electrònics

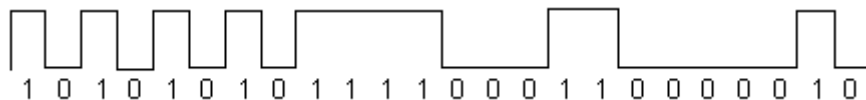
Es fan servir diversos codis electrònics per a representar els 0 i 1 del codi binari. N'esmentem tres com a exemple:

1. **Non return zero L:** dona un nivell de voltatge diferent en cada bit. 1, voltatge positiu. 0, absència de voltatge.
2. **Non return zero polar:** dona el mateix nivell de voltatge en cada bit, però de signe contrari. 1, voltatge positiu; 0, voltatge negatiu.
3. **Codi RZ binari:** els senyals de no-retorn a 0 (NRZ) tenen el problema que quan les sèries d'1 o 0 són molt llargues, no es produeixen transicions i això dificulta la recuperació del sincronisme. Per això, s'han pensat uns codis que facin tornar el senyal a 0 a mig interval. També n'hi ha dues variants, polar i unipolar.

Altres codis electrònics són EFM (*eight to fourteen modulation*), MFM (Miller) o HDM-1 (DASH).



El sistema electrònic digital té un llindar, de manera que si passa el senyal equival a 1 i, si no, a 0.



Finalment, s'apliquen unes informacions de sincronia i direcció als impulsos electrònics, equivalents a 1 i 0, que permeten de diferenciar el significat de les dades en cadenes de 0 i 1 consecutius. Aquest procés es coneix com a **modulació**.

Una de les més emprades és la **modulació per codi d'impulsos** o **PCM** –creada l'any 1967. Una altra modulació digital és la FSK o clau de desplaçament de freqüència.

Philips i Sony han desenvolupat per als sistemes més avançats, com el Super Audio CD, la tecnologia alternativa Direct Stream Digital, que enregistra els polsos del senyal digital bit a bit, aplicant una modulació delta-sigma i una freqüència de mostratge seixanta-quatre vegades superior als 44,1 kHz ($64 \times 44,100 \text{ Hz} = 2,822 \text{ MHz}$). Cada bit és una mostra.

Convertidor digital/analògic

Un convertidor digital/analògic restaura el senyal digitalitzat en el domini analògic. El convertidor digital/analògic dona un senyal continu a la sortida que canvia de tensió segons el valor de cada paraula binària.

Una vegada recuperats els impulsos electrònics, s'obté el 0 o 1 corresponent de cadascun. Per a recuperar el senyal, només cal detectar-ne el nivell; la forma d'ona no és rellevant. Així, encara que un impuls arribi deteriorat, si ultrapassa un cert nivell dona 1 i, si no, 0.

El que sí que pot afectar negativament el procés són els canvis en la durada dels impulsos, les variacions en l'amplada. Aquest error en la base de temps es coneix amb el nom anglès de **jitter**. Si aquesta fluctuació temporal és

elevada provoca clics o petites interrupcions en la reproducci3.

El problema es resol amb l'ajuda d'un circuit de sincronia. Les dades s'enregistren en una mem3ria RAM i despr3s es llegeixen a una cad3ncia fixa controlada per un rellotge de vidre de freq3ncia extremadament estable.

Quan ja s'han obtingut els bits, se'ls aplica el codi de correcci3 d'errors per a cancel·lar-ne les equivocacions detectades. Si el senyal estava multiplexat, se separa la informaci3 de cada canal. A continuaci3, cada paraula bin3ria entra en el convertidor digital/anal3gic i, una vegada descodificada, en surt el nivell del voltatge del senyal d'3udio original d'aquell instant.

El convertidor digital/anal3gic ha de ser el m3xim de prec3s possible. Per aix3, alguns sistemes fan servir convertidors de 20 o 24 bits, encara que les mostres processades siguin de 16 bits. Aix3 aconseguixen una resoluci3 m3s bona.

Tamb3 s'aplica un sobrestratge del senyal (*oversampling*) en el moment de la descodificaci3 per a tenir m3s precisi3 en la reproducci3 de les freq3ncies altes. Si no es fa el sobrestratge, cal que el senyal final passi per un filtre passabaix, molt prec3s i car –amb freq3ncia de tall de 20 kHz–, per a rebutjar-ne les freq3ncies d'imatge que apareixen entorn dels 20 kHz durant el proc3s.



El sobrestratge

El sobrestratge consisteix a descodificar les paraules bin3ries en r3fegues de bits m3s curtes, mitjançant l'increment de la freq3ncia d'exploraci3.

En un reproductor de disc compacte que fa un sobrestratge de vuit vegades, el convertidor de digital a anal3gic no processa 44.100 paraules de 16 bits per segon, sin3 352.800 r3fegues de 2 bits cada segon ($44,1 \text{ kHz} \times 8 = 352,8 \text{ kHz}$).

Es poden aplicar sobrestratges de 4, 8, 16, 32, 64, 128 o 256 vegades.

Hi ha dues tecnologies principals de convertidors digital/anal3gic: la multibit i la de nombre baix de bits (*low bit*).

Un **convertidor multibit** processa tots els 0 i 1 de cada paraula bin3ria en paral·lel. El circuit pot estar format per una xarxa de resist3ncies distribuïdes en branques paral·leles, tantes com bits tingui la paraula bin3ria. Cadascuna d3na una intensitat ponderada segons el pes del bit que li correspon. El valor del corrent de cada branca es duplica a mesura que el pes del bit augmenta. La suma de totes aquestes intensitats d3na el nivell del senyal anal3gic. Aquests convertidors tamb3 es denominen *R/2R* per la relaci3 que tenen els valors de les resist3ncies que es fan servir. Altres sistemes multibit fan servir circuits basats en condensadors o en amplificadors diferencials. A la sortida sempre s'obté una tensi3 proporcional als valors binaris de l'entrada.

Segui quina sigui la tecnologia, tots els valors dels elements d'aquests circuits han d'estar molt ben ajustats i els components han de tenir una toler3ncia el3ctrica molt petita per a garantir una distorsi3 harm3nica molt baixa i una resposta en freq3ncia lineal.

Quan els processadors inform3tics van començar a incrementar la velocitat de treball van sorgir els **convertidors de nombre baix de bits**. Aquests circuits no tenen un ajustament tan delicat com els anteriors.



El convertidor Philips Bitstream

El convertidor Philips Bitstream d'1 bit processa els 16 bits de cada mostra d'un CD un a un, en s3rie.

Els **convertidors de nombre baix de bits** aconseguixen resultats sonors m3s bons, sobretot quan el nivell del senyal 3s baix.



Configuracions t3piques i t3cniques principals dels convertidors de nombre baix de bits

Per a paraules de 16 bits, les configuracions t3piques que fan servir els convertidors de nombre baix de bits s3n: 1 bit amb una freq3ncia de treball d'11,3 MHz (sobrestratge de 256 vegades), i 4 bits a 2,8 MHz (sobrestratge de 64 vegades).

Dins dels convertidors de nombre baix de bits hi ha un parell de t3cniques principals: la Pulse Width Modulation, desenvolupada per Matsushita, i la Pulse Density Modulation, de Philips.

Aprofitant les elevades velocitats dels processadors informàtics des de mitjan anys noranta, la tecnologia dels convertidors de nombre baix de bits ha evolucionat i ha donat el sistema Direct Stream Digital (DSD). El DSD és la base del format Super Audio CD, de Sony i Philips, i processa el senyal digital amb una freqüència de conversió molt alta.

Per a evitar les discontinuïtats a la sortida del convertidor digital analògic, hi ha un circuit de retenció i mesura (*sample and hold*) que conserva el valor de l'última mostra descodificada fins que la paraula binària següent hagi entrat en el sistema.

Finalment, el senyal analògic passa per un filtre passabaix que en suavitza la forma d'ona.

Formats d'àudio

Hi ha molts formats d'àudio digital. Per a poder fer el màxim nombre de transferències o conversions, les estacions d'àudio digital tenen normalment una àrea anomenada *batch processing* o *batch converter*, que permet de convertir formats, freqüències de mostratge o nombre de bits de quantificació.

Els principals formats d'àudio són els que es presenten en la taula següent:

wav (wave)	Creat per Microsoft i IBM com una evolució del format RIFF 8 (<i>resource interchange file format</i>). Va ser el primer format d'àudio per a PC. El format wav/PCM (<i>pulse code modulator</i>) no està comprimit. Pot quantificar a 8 o 16 bits. La freqüència de mostratge pot arribar als 44,1 kHz. Per tant, té una qualitat igual a la d'un CD.
aiff (audio interchange file format)	Fitxer d'Apple per als seus ordinadors Mac. És una evolució dels fitxers IFF que feien servir els Amiga. El fitxer AIFF no comprimeix el so i treballa amb paraules de 8 bits. Té una qualitat similar a la d'un CD. També hi ha el format AIFF-C, que està comprimit.
au (audio)	Fitxer de Sun Microsystems que pot quantificar a 8 o 16 bits. És un dels formats més veterans dels fitxers de so per a Internet.
bwf (broadcast wave format)	La Unió Europea de Radiodifusió, l'any 1987, va definir el format <i>bwf</i> per a intercanvis de so entre ràdios (EBU document 3285). El <i>broadcast wave format</i> és un fitxer de so basat en el format <i>wave</i> que inclou una mínima informació sobre el contingut –metadata. La versió 1 del <i>bwf</i> destina 64 bytes dels 254 reservats a dades a la informació SMPTE UMID (<i>unique material identifier</i>).
D'altres	Altres extensions de fitxers de so que podem trobar són: voc (Soundblaster), snd (de Sun), sf (d'IRCAM), vox (dialògic), mat (<i>matlab</i>), SDD II (Sound Designer 2 de Digidesign) o raw (fitxer <i>audio frame</i> en brut sense cap format concret).

■ Formats de compressió d'àudio

Comprimir un so digital consisteix a reduir l'espai que ocupa per no haver de processar tants bits. Així s'aconsegueix més capacitat d'emmagatzematge i una reducció de l'amplada de banda de la transmissió d'àudio.

El procés de compressió consisteix a fer passar el senyal digitalitzat per un compressor de dades (*encoder*) abans de l'enregistrament o la transmissió i fer-lo passar per un descompressor (*decoder*) abans de reproduir-lo.

Hi ha dos **sistemes de compressió**:

1. **Sense pèrdues d'informació (lossy)**, que redueixen l'ocupació entre un 30% i un 50% amb mètodes estadístics
2. **Amb pèrdues d'informació (lossless)**, que apliquen criteris psicoacústics i aconseguen reduccions d'entre un 40% i un 80%. En aquest cas es tenen en compte les limitacions de l'oïda humana i s'elimina tota la informació que no és perceptible. El sistema ATRAC emprat per Sony en el Mini Disc o el sistema de reproducció multicanal SDDS en són exemples.



Tal com s'explica en el temari de l'assignatura de la UOC, *Fonaments de tecnologia audiovisual*:



"La compressió està basada en l'eliminació successiva d'elements redundants que apareixen en el senyal d'àudio i si s'eliminen en el codificador és possible reconstruir-los en el descodificador." "La redundància estadística consisteix a determinar quants valors de bits es repeteixen més en una seqüència de bits. La base de l'eliminació de la redundància estadística recau en el fet que hi ha paraules (agrupacions de bits) que es repeteixen constantment. Per tant, la idea és enviar aquestes paraules que es repeteixen amb menys bits."

Quan un senyal comprimit es recupera sense pèrdues, el senyal obtingut és idèntic a l'original. Per contra, si la compressió ha estat amb pèrdues, el fitxer descomprimit és, evidentment, diferent de l'original. La compressió sense pèrdues només és possible per a aplicacions que no siguin a temps real.

La compressió d'àudio digital comença l'any 1987 en els treballs de recerca del Fraunhofer Institut Integrierte Schaltungen d'Alemanya. Aquells algorismes de compressió i descompressió van ser acceptats finalment com a estàndard l'any 1992 per l'MPEG i l'ISO amb el número ISO/IEC 11172. Naixia la nissaga de sistemes MPEG.

MPEG-1	MPEG-1 (<i>Moving Picture Experts Group</i>) apareix l'any 1991. Estàndard (ISO/IEC 11172) de compressió de vídeo i àudio dissenyat inicialment per a introduir vídeo i àudio en un CD-ROM.
MPEG-2	Establert l'any 1994. Estàndard (ISO/IEC 13818) de compressió de vídeo i àudio que dona més qualitat que el MPEG-1, però que necessita més amplada de banda, entre 3 i 10 Mbit per segon. Ofereix compatibilitat amb el MPEG-1. Permet un gran ventall de compressions, des de 20:1 fins a 80:1.
MP3	És el format de compressió més emprat a l'enregistrament d'àudio i la transmissió de fitxers de so per Internet. El MP3 o MPEG-1 Layer III és l'aplicació específica al so del sistema de compressió inicial dissenyat per a vídeo i àudio. La compressió és amb pèrdues, però amb relacions de 10:1 la qualitat del so final és prou bona, tot i que és inferior a la qualitat del CD Àudio. En un CD-ROM es poden tenir aproximadament unes cent cinquanta cançons en format MP3 en lloc de les deu o dotze en format WAV. Un fitxer WAV ocupa dotze vegades més que un de MP3. Per a fer les transferències de CD a fitxers MP3 hi ha diverses eines informàtiques. Els rippers són els programes que permeten salvar cançons de CD a l'ordinador en format de fitxers WAV i els encoders són els que comprimeixen els fitxers WAV a MP3. Però també hi ha els grabbers que simplifiquen el procés i transformen directament una cançó de CD en format MP3 sense haver de crear un fitxer WAV previ. Pel que fa al seu ús en Internet, el MP3 estava pensat inicialment per a distribuir so monofònic a 56 Kb/s. La millora del sistema va permetre transmetre so estèreo similar al de l'FM a partir d'uns 64 Kb/s i amb una qualitat propera al CD a més de 128 Kb/s.
MP3 Pro	Format dissenyat per Thomson Multimedia. És un nou algoritme que codifica els senyals d'àudio i es basa en un augment de la compressió superior a la del format MP3 tradicional. Aquest format és totalment compatible amb tots els aparells reproductors d'àudio MP3.
MPEG-4 o AAC	Establert l'any 1999. L'MPEG-4 o AAC (<i>Advanced Audio Code</i>) és un estàndard (ISO/IEC 14496) de compressió de vídeo i àudio dissenyat per a transmetre informació multimèdia interactiva –amb imatges, so i gràfics– per Internet. Té una gran flexibilitat pel que fa a la velocitat de transmissió de bits (<i>bit rate</i>) de treball, que va dels 10 Kbps als 10 Mbps. Incrementa la qualitat de l'MP3 i redueix encara més la mida dels arxius, en un 30% aproximadament. La qualitat de so és inferior, però, a la dels CD Àudio.

OGG Vorbis	Format de compressió d'àudio creat per la Xiph.org Foundation com un projecte de codi lliure, gratuït i sense patents. Vol ser una alternativa al MP3. Fins ara, els artistes que comercialitzen música en MP3 han de pagar un percentatge de les vendes al propietari de la patent. Els fitxers són més petits que els MP3. Les velocitats de transmissió de bits (<i>bit rates</i>) per canal poden anar des dels 16 als 128 Kbps. I admet fins a 256 canals d'àudio.
WMA	Format de compressió d'àudio de Microsoft. És l'evolució del format ASF i està pensat per al reproductor Windows Media Player per a escoltar música en línia a Internet mitjançant <i>streaming</i> .
RM-RA	Real Media (RM) i Real Audio (RA) són formats que fa servir el reproductor audiovisual Real Player (o Real One) per a llegir fitxers d'àudio en línia mitjançant <i>streaming</i> . Els fitxers RM o RA només es poden llegir, mentre que l'extensió RAM permet descarregar l'arxiu a l'ordinador.
AC-3	AC-3 ("Audio Compression 3") és un sistema de compressió desenvolupat per Dolby per al seu format de so envoltant (<i>surround</i>) digital de sis canals. Es fa servir en còpies de pel·lícules i DVD. Dóna una qualitat de so molt alta.

■ Formats de distribució musical

Hi ha diversos formats d'àudio dissenyats per a la distribució i venda de música per Internet de manera segura. Alguns han tingut una vida molt curta –com l'A2B d'AT&T– i d'altres encara es fan servir. Recollim alguns exemples d'un terreny en el qual treballen les discogràfiques per a trobar una nova manera de vendre els productes.

Liquid Audio (LQT)	Representa una alternativa al format MP3 en garantir més seguretat en la distribució d'àudio a les companyies discogràfiques. El reproductor de Liquid Audio es pot aconseguir de manera gratuïta en totes les pàgines web que utilitzen aquest sistema.
VQF	Desenvolupat per Yamaha. Els fitxers ocupen un 30% menys que els MP3. Es pot llegir amb el reproductor Winamp, afegint-hi un connector. Només permet d'accedir a les cançons si se'n coneix una clau.
SAF (Secure Audio Format)	L'estiu del 2000, BMG i Universal van crear amb InterTrust el sistema Digibox, que garantia el pagament i el respecte al <i>copyright</i> de les cançons o els discos sencers que el client volgués comprar. Finalment, el sistema es va dir SAF (<i>secure audio format</i>). Comprimeix l'àudio en format MPEG-4 o AAC (<i>advanced audio code</i>).
RMF (Rich Music Format)	Format de Beatnik de descàrrega d'àudio d'alta qualitat.