

Codificació del senyal de televisió

Jorge Mata Díaz

PID_00196633



Els textos i imatges publicats en aquesta obra estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement-NoComercial-SenseObraDerivada (BY-NC-ND) v.3.0 Espanya de Creative Commons. Podeu copiar-los, distribuir-los i transmetre'ls públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>

Índex

Introducció	5
1. Estàndard de codificació DVB MPEG-2	9
2. Codificació de font	12
2.1. Entropia de la font	14
2.2. Codificació en seqüències de símbols	15
3. Codificació audiovisual MPEG	18
3.1. Compresió de vídeo MPEG	18
3.2. Estructuració de la seqüència de vídeo MPEG	20
3.3. Codificació <i>intra-frame</i> (imatges I)	21
3.3.1. Transformada cosinus (DCT)	22
3.3.2. Ordenació en zig-zag	23
3.3.3. Quantificació	24
3.3.4. Codificació a ràfegues RLC i codificació entròpica VLC	24
3.4. Codificació <i>inter-frame</i>	25
3.4.1. Codificació predictiva cap endavant (imatges P)	25
3.4.2. Codificació bidireccional-predictiva (imatges B)	27
3.5. Compresió d'àudio MPEG	28
4. Multiplexatge i <i>scrambling</i> en DVB	31
4.1. Empaquetament i multiplexatge MPEG	31
4.2. Informació específica d'un programa (PSI)	34
4.3. <i>Scrambling</i> i accés condicional	37
5. Codificació de canal	39
5.1. Aleatorització	40
5.2. Codificació de bloc Reed-Solomon	42
5.3. Entrellaçament de Fourny	43
5.4. Codificació convolucional o contínua i perforació	45

Introducció

En aquest mòdul s'expliquen els aspectes necessaris per a codificar la informació audiovisual de manera que es faciliti l'aprofitament dels mitjans de transmissió i es contrarestin les degradacions introduïdes per aquests mitjans.

En primer lloc, presentarem els conceptes fonamentals de la codificació de la font utilitzant tècniques de propòsit general i tècniques específiques per a la informació audiovisual. La teoria de la codificació de la font proporcionarà la base per a reduir el volum de dades que cal transmetre d'un contingut audiovisual.

En segon lloc, presentarem els mecanismes d'empaquetament de la informació i el seu multiplexatge per a un transport adequat de les dades en els mitjans de comunicació.

Finalment, introduïrem els conceptes bàsics de codificació de canal que protegiran les dades contra possibles distorsions durant la transmissió.

En la figura 1 s'indica el bloc de la cadena televisiva que es desenvolupa en aquest mòdul:

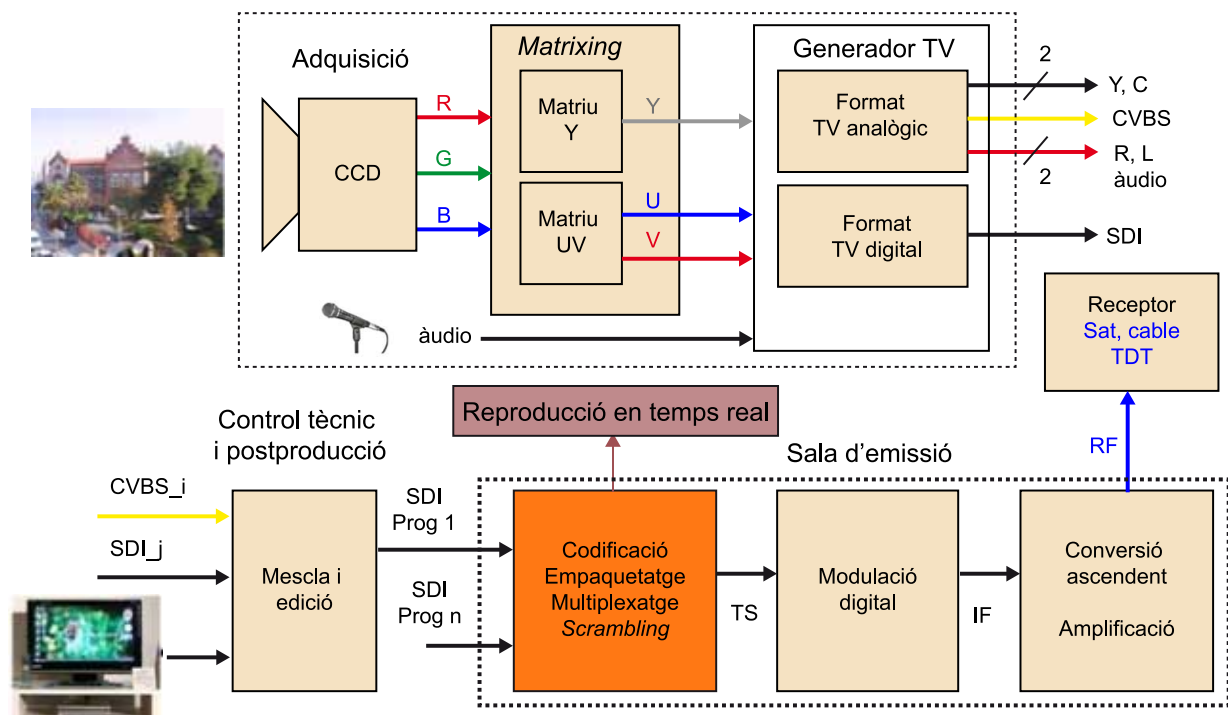


Figura 1. Cadena televisiva. En taronja s'indica el bloc que es desenvolupa en aquest mòdul.

El desenvolupament previ de la televisió analògica va predeterminar que els canals per a la transmissió de la televisió tinguessin una amplada de banda convencional de 6 MHz a 8 MHz en mitjans de cable o terrestre i de 27 MHz a 36 MHz per a satèl·lit. Amb aquests valors, és possible plantejar-se transmetre senyals digitals en aquests mateixos canals, les velocitats binàries dels quals siguin de prop de 30 Mbps a 40 Mbps, no obstant això, el senyal digital de vídeo SDI requereix una velocitat de 270 Mbps.

Per a poder convertir el senyal SDI (vegeu mòdul "Digitalització del senyal de televisió") en un senyal de televisió que pugui ser transmès pels canals de radiodifusió (*broadcasting*), s'han d'aplicar algorismes de compressió. Aquests es basen fonamentalment en tres aspectes:

- La **redundància espacial** (correlació entre píxels veïns) i les característiques de la percepció visual humana (baixa sensibilitat als detalls fins) d'una imatge fixa, que són la base dels mètodes JPEG.
- La **redundància temporal** entre imatges consecutives (base dels algorismes MPEG).
- Les **característiques de l'audició humana** que permet eliminar informació no audible (codis psicoacústics).

L'aplicació de tots aquests algorismes conforma el que es denomina **codificació de font** específica per al senyal audiovisual.

La transmissió de la televisió digital s'organitza en un multiplexatge de paquets que pertanyen a diversos programes de televisió, d'àudio o de dades que denominarem *flux de transport* (TS, *transport stream*). En aquests paquets s'inclou informació sobre els programes que el componen. La informació dels programes de pagament es troba xifrada mitjançant una tècnica de *scrambling* (o d'aleatorització impredecible) que impossibilita l'accés a aquests programes si no es disposa d'una clau per a desxifrar-los.

El flux de bits o *stream* que s'origina en la transmissió requereix d'algorismes eficients de detecció i correcció d'errors que configuren la denominada *codificació de canal*.

Tots aquests aspectes, que es poden resumir en la figura 2, s'explicaran en aquest mòdul.

Vegeu també

La conversió del senyal SDI a senyal de televisió s'explica en el mòdul "Digitalització del senyal de televisió"

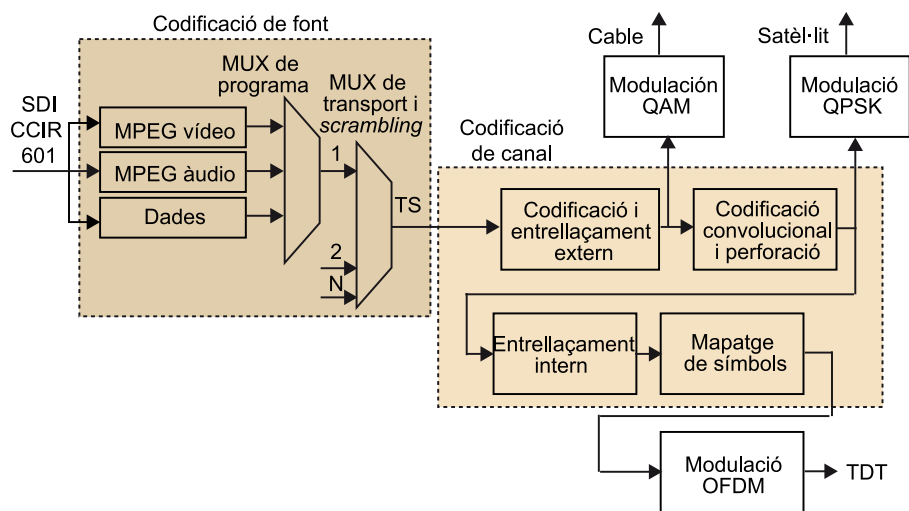


Figura 2. Diagrama de blocs de la codificació de la televisió digital

1. Estàndard de codificació DVB MPEG-2

Digital video broadcasting (DVB) és un sistema de transmissió basat en l'estàndard MPEG-2. La recomanació MPEG-2 especifica els mecanismes de compressió de vídeo i àudio, i també el seu multiplexatge per a ser emmagatzemats o transmesos. El DVB no és simplement un substitut de la difusió analògica de televisió, sinó que millora la qualitat audiovisual amb la possibilitat de veure imatges en format estàndard o alta resolució, juntament amb so mono, dual, estèreo o envoltant, i també permet una àmplia gamma de noves característiques i serveis que inclouen pistes de vídeo i àudio múltiples, subtitulació, continguts interactius, etc.

DVB és una iniciativa europea que ha estat acollida en tots els continents i que s'ha convertit en estàndard mundial per a la televisió digital. A més de DVB, als EUA s'utilitza un sistema derivat de l'*advanced television systems committee* (ATSC) que és una mica diferent del DVB. Entre altres coses, l'ATSC adopta un estàndard de codificació d'àudio diferent i banda lateral vestigial (*vestigial sideband modulation*, VSB) per a la modulació. No obstant això, les normes nord-americanes i europees tenen moltes coses en comú.

La clau del disseny del DVB és que no s'imposa cap restricció sobre la naturalesa del contingut enviat, solament se centra en el transport de la informació i, per tant, solament defineix la sintaxi del flux; és a dir, l'ordre i la diferenciació de les dades que s'envien.

Exemple

En el mateix flux de dades es pot transportar una gran varietat de serveis d'alta i baixa velocitat de dades: des de la difusió de continguts audiovisuals, amb les guies corresponents de programació (*electronic program guides*, EPG) i informació de teletext, fins al transport genèric de dades sobre el protocol d'Internet (IP).

El comitè d'estandardització Motion Pictures Experts Group (MPEG) va proposar inicialment MPEG-1 com a estàndard per a l'emmagatzematge de continguts audiovisuals en CD, que més tard es va normalitzar com a ISO/IEC 11172. El format triat per al codificador de vídeo seguia el format CIF amb submostreig 4:2:0. L'algorisme de compressió de so que es va especificar per a l'àudio partia d'una proposta denominada *MUSICAM*.

Posteriorment, el comitè MPEG va passar a especificar una nova norma per a la transmissió de continguts audiovisuals amb referència ISO/IEC 13818, que seria adoptada com a estàndard de codificació per DVB. Va treballar en la definició d'un estàndard flexible que preveia la possibilitat d'aconseguir resolu-

cions d'alta definició i imatge entrelaçada, i el resultat va ser l'MPEG-2, que va aparèixer el 1994. La normalització va donar lloc a tres especificacions diferenciades:

- MPEG-2 vídeo (ISO / IEC 13818-2). Especifica la sintaxi de la codificació d'un contingut elemental de vídeo i les operacions necessàries per a descodificar-lo.
- MPEG-2 d'àudio (ISO / IEC 13818-3). Especifica la sintaxi de la codificació d'un contingut elemental d'àudio i les operacions necessàries per a descodificar-lo.
- MPEG-2 del sistema (ISO / IEC 13818-1). Defineix la transmissió de les unitats elementals codificades i les relacions temporals que han de tenir per a ser desmultiplexades i descodificades.

L'estàndard MPEG-2 es va concebre per a abordar l'emmagatzematge i distribució de continguts audiovisuals per a un ampli ventall de terminals connectats a diferents xarxes. Tenint en compte les característiques dels terminals, es van especificar diferents perfils que es correspondrien amb les capacitats de procés i memòria del terminal al qual va destinada la codificació:

- **Perfil bàsic** (*simple*). Per a terminals amb poca capacitat de memòria i procés, per exemple un telèfon mòbil. El format de vídeo digital emprat per a la codificació segueix el patró de submostreig 4:2:0.
- **Perfil principal** (*main*). Per als terminals amb més prestacions, com televisors o ordinadors. També preveu com a format d'entrada de vídeo digital el 4:2:0.
- **Perfil alt** (*high*). Per a dispositius d'altres prestacions com els televisors intel·ligents o *set-top boxes* amb un maquinari dedicat a la descodificació. Permet l'ús dels formats 4:2:0 i 4:2:2 com a formats d'entrada de vídeo.

En la pràctica, els perfils que s'utilitzen són el perfil principal i el perfil alt.

Segons les xarxes de transport (satèl·lit, cable, Internet, etc.) i la resolució de pantalla establerta per als dispositius, es defineixen diferents nivells de codificació de vídeo. Quan l'amplada de banda no sigui una limitació, el factor determinant de la codificació a aplicar serà el perfil del dispositiu i la presentació de vídeo requerida. En la pràctica, moltes vegades serà l'amplada de banda del mitjà de transmissió el condicionant principal de la codificació. Depenent de l'amplada de banda disponible, es podrà aconseguir una taxa màxima de transmissió que limitarà la resolució de la imatge codificada a pesar que el receptor pugui arribar a presentar imatges de més qualitat.

En MPEG-2 es defineixen tres nivells de codificació d'imatge associats a la resolució d'imatge i l'amplada de banda disponible que van des del format CIF fins a HD:

- **Nivell baix** (*low*). Es correspon amb la capacitat del descodificador MPEG-2 per a reproduir codificacions en resolució CIF, 352 mostres/línia × 288 línies/imatge.
- **Nivell principal** (*main*). Associat a la televisió estàndard, o qualitat de disc DVD, 720 mostres/línia × 576 línies/imatge.
- **Nivell alt** (*high*). Corresponent a la televisió d'alta definició. 1.920 mostres/línia × 1.152 línies/imatge.

Així doncs, un sistema combina un perfil amb un nivell i se simbolitza Perfil@nivell. En aquest sentit, la combinació més emprada ha estat el perfil principal amb el nivell principal, MP@ML, que correspon a la codificació de la televisió estàndard. La taxa binària que requereix la difusió d'un programa amb aquestes característiques és de fins a 6 Mbps.

S'ha de tenir en compte que l'amplada de banda ocupada amb la codificació MPEG-2 en un mitjà de transmissió és, per a la mateixa resolució d'imatge, la sisena part de la usada en transmissió analògica. En altres paraules, en l'amplada de banda on abans es transmetia un programa de televisió analògica, ara es transmeten sis programes, de qualitat equivalent, en MPEG-2.

2. Codificació de font

Per a reduir notablement la velocitat requerida per a transmetre un contingut audiovisual és necessari caracteritzar la font d'informació. La caracterització de la font es basa fonamentalment en la naturalesa de la informació que genera i en la seva estadística. Podrem aplicar tècniques de compressió específiques i generals sobre les dades generades per la font que ens facilitaran l'eliminació de gran part dels elements redundants, reduint notablement l'espai d'emmagatzematge o l'amplada de banda de transmissió.

En termes generals, per a qualsevol font d'informació es poden aplicar tècniques basades en l'entropia pròpia de la font. Per a això, inicialment, s'ha d'identificar el conjunt de símbols que pot emetre una font d'informació. Aquest conjunt de símbols definirà l'**alfabet de la font**.

Exemple

Per al cas audiovisual podríem considerar en vídeo els valors que podrien prendre els píxels d'una imatge, o en àudio, els valors de les mostres obtingudes per mitjà d'un micròfon. Aquests valors formarien, respectivament, l'alfabet dels símbols de la font de vídeo i l'alfabet dels símbols de la font d'àudio.

Una vegada determinat l'alfabet d'una font, es poden aplicar tècniques de codificació. La codificació és pròpiament l'assignació d'una paraula codi a un símbol de font o, de manera més general, a una successió de símbols de font. En la pràctica, les paraules codi sempre són successions de dígit binaris.

Exemple 1

Una font de vídeo en què els píxels prenen valors dins d'una escala de quatre grisos té per alfabet de font {0, 1, 2, 3}. Es podria codificar assignant respectivament les paraules codi de dos dígit binaris {00, 01, 10, 11}. Aquesta codificació seria una codificació de longitud constant, ja que tots els símbols de la font s'han codificat amb paraules codi de longitud 2.

Exemple de codificació de longitud constant per a un alfabet de quatre símbols

Símbol font	Codificació binària
0	00
1	01
2	10
3	11

Per a valorar si una codificació és millor o pitjor hauríem de tenir en compte quants dígit binari han calgut per a codificar un determinat missatge. En el nostre exemple, podríem considerar una imatge de 1.000 píxels com el missatge que s'ha de codificar. Aplicant a la imatge la codificació de longitud constant proposada, es necessitarien 2.000 dígit binari per a transmetre-la o emmagatzemar-la.

La codificació de font ens permetrà caracteritzar la font de manera que puguem trobar una codificació que redueixi les necessitats d'emmagatzematge o transmissió de la font. La reducció obtinguda es fonamenta a assignar paraules codi més curtes a símbols de font més probables i, viceversa, paraules codi més llargues a símbols de font menys probables. El resultat obtingut és, habitualment, una codificació de longitud variable.

La codificació de font tracta de determinar l'estadística subjacent en la generació dels símbols de l'alfabet de la font i, a partir d'aquest coneixement, aplica una codificació.

Exemple 2

Per a la font de l'exemple anterior, suposarem que les freqüències d'aparició dels símbols de l'alfabet de font {0, 1, 2, 3} són $1/4$, $1/8$, $1/8$ i $1/2$, respectivament. La codificació hauria de premiar amb menys dígit binari els símbols 0 i 3, mentre que hauria de penalitzar els símbols 1 i 2. Una possible codificació de longitud variable assignaria, per exemple, les paraules codi següents a cadascun dels símbols de font:

Exemple de codificació de longitud variable per a un alfabet de quatre símbols

Símbol font	Codificació binària
0	00
1	010
2	011
3	1

Com veiem, al símbol més probable s'hi ha assignat una paraula d'únicament 1 dígit binari i paraules codi de longitud 3 als menys probables. Per a aquest cas, la codificació binària d'una imatge de 1.000 píxels requerirà:

$$1.000 \left(\frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 + \frac{1}{2} \cdot 1 \right) = 1.000 \cdot 1.75 = 1.750 \text{ dígit binari} \quad (1)$$

Com veiem, s'han reduït les necessitats d'emmagatzematge o transmissió de 2.000 a 1.750 dígit binaris. És a dir, s'ha comprimit el 87,5%, per la qual cosa es requerirà un 12,5% menys d'espai d'emmagatzematge o amplada de banda per al mateix contingut.

Per a determinar la bondat d'una codificació, en comptes de fer els càlculs sobre un missatge compost per una quantitat de símbols de font, es determina el valor mitjà de la longitud de les paraules codi resultants. Així, per al primer cas, de longitud constant, totes les paraules codi són de 2 dígit. En el segon cas, el nombre mitjà de dígit binaris d'una paraula codi serà d'1,75 dígit. Òbviament, el segon codificador és més eficient perquè requereix menys dígit de mitjana per símbol de font.

El procés de codificació extreu les redundàncies de les dades fins a arribar a una representació de la informació que requereix un cert nombre de dígit binaris que s'aproxima, o fins i tot arriba, a un valor mínim. El valor mínim marca una cota en la qual es pot garantir que s'ha extret tota la redundància sense cap pèrdua d'informació. Aquest tipus de codificació rep el nom de *codificació sense pèrdues* i s'utilitza quan es vol que la codificació no introdueixi cap tipus de distorsió sobre les dades originals.

En cas que la codificació emprés menys dígit dels que imposa el valor mínim llavors, el procés de descodificació no podria garantir la recuperació de la informació original; és a dir, la informació descodificada podria ser diferent de la codificada. En aquest cas, la codificació admet pèrdues. En el camp audiovisual és habitual aplicar tècniques de codificació amb pèrdues que puguin introduir de manera ordenada un grau de distorsió entre la informació original i la descodificada, ja que la percepció humana admet certes degradacions durant la presentació de la informació.

2.1. Entropia de la font

La cota inferior que determina fins a quant es pot comprimir sense necessitat d'introduir pèrdues en els continguts està determinada per la mateixa font. L'estadística de la font defineix un grau o magnitud de desordre i variació dels valors dels símbols generats que es denominada *entropia*. La font amb un grau d'entropia més elevat és la que emet els seus símbols de manera equiprobable i incorrelada. L'emissió incorrelada fa que una font no tingui en compte les emissions anteriors (no té memòria) per a la generació següent. L'entropia d'una font (H) sense memòria es defineix per l'expressió següent:

$$H = \sum_{i=1}^K p_i \cdot \log_2(1/p_i) \quad (2)$$

Lectura de la fórmula

En què K és el nombre de símbols que pot generar la font, i p_i és la probabilitat d'emissió de cada un.

La unitat emprada per a quantificar l'entropia és el bit d'informació per símbol.

Si tots els símbols són equiprobables tindrem que $p_i = 1/K$, per la qual cosa veiem que el valor més alt possible d'entropia d'una font és $\log_2(K)$.

A priori, si es desconeix el comportament d'una font, la millor hipòtesi sobre la seva estadística és considerar-la equiprobable. Com a resultat, la codificació òptima assignarà paraules codi d'una longitud molt similar, o igual, a tots els símbols de font. Aquest és el procés que hem seguit per al primer exemple: teníem una font que podia emetre 4 símbols, per la qual cosa la màxima entropia que podria tenir aquesta font seria de $\log_2(4) = 2$ bits per símbol. No obstant això, l'estadística descrita per la font (segon exemple) no mostra un comportament equiprobable, per la qual cosa l'entropia hauria de ser menor. Així, en aquest cas l'entropia calculada és la següent:

$$H = \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8) + \frac{1}{2}\log_2(2) = 1.75 \text{ bit/símbol} \quad (3)$$

El procés de codificació sense pèrdues està limitat per l'entropia, de manera que la longitud mitjana de les paraules codi binàries mai no podrà ser inferior a l'entropia. En el nostre segon exemple, hem aconseguit una codificació sense pèrdues òptima, ja que la longitud mitjana obtinguda és d'1,75 dígits per símbol, que es correspon exactament amb l'entropia de la font. Per tant, en aquest cas, 1 dígit binari porta 1 bit d'informació.

L'algorisme de codificació ha d'intentar aconseguir apropar al màxim la longitud mitjana de la codificació binària a l'entropia de la font. En general, s'obté com a resultat una codificació entròpica de longitud variable denominada *variable length coding* (VLC).

2.2. Codificació en seqüències de símbols

A més d'estudiar la freqüència de generació dels símbols en una font, es poden tenir en compte les relacions en les seqüències de símbols consecutius. Moltes de les fonts audiovisuals presenten un grau alt de correlació en els símbols emesos, de manera que, després d'emetre un símbol, és molt probable que se n'emeti un altre d'igual o de valor molt semblant.

Quan en una font la probabilitat d'emetre un símbol depèn de les emissions anteriors diem que la font té **memòria**. Aquesta característica redueix l'entropia de la font, ja que l'emissió dels símbols està més ordenada que si la font tingués una generació incorrelada.

En fonts amb memòria sabem més sobre l'emissió, ja que la generació actual condicionarà les emissions futures. Aquest condicionament es pot aprofitar a l'hora de codificar. Les fonts amb memòria se solen codificar de manera que a una cadena de símbols de font consecutius s'hi correspon una única paraula codi.

Exemple 3

Per a la imatge amb 4 nivells de grisos estudiada anteriorment {0,1,2,3}, es podria fer una codificació de longitud variable per a cada parell de píxels. Així tindriem la codificació següent de parells de píxels:

Exemple de codificació de seqüències de 2 símbols

Cadena	Codificació	Cadena	Codificació
00	1011	20	11110
01	10100	21	111110
02	10101	22	111111
03	100	23	1110
10	11010	30	010
11	110110	31	0110
12	110111	32	0111
13	1100	33	00

En situacions de llargues generacions consecutives del mateix símbol s'utilitzen alternatives de codificació als símbols de font. És el cas de la codificació d'imatges per línies, en què els valors dels píxels en la mateixa línia d'imatge se solen repetir consecutivament un nombre elevat de vegades. Aquestes fonts molt correlades es caracteritzen per l'emissió de símbols iguals en forma de ràfegues. Una tècnica eficient de codificació es basa en la codificació de la longitud de la ràfega i el valor del símbol repetit.

Exemple

Si tinguéssim la cadena BBBBBAACCC es podria codificar com a 5B2A3C.

Per a augmentar encara més l'eficiència de la codificació de font, es pot caracteritzar la freqüència amb què apareixen les ràfegues d'una determinada longitud i la freqüència del valor que pren la ràfega. Una vegada establerta la codificació per a la longitud i la codificació per al valor del símbol, el codificador genera per cada ràfega un parell de paraules codi associades a la longitud i valor. Aquest mecanisme de codificació es denomina *run length coding* (RLC).

Com es pot observar en les tècniques anteriors de codificació, el coneixement de la naturalesa de la informació generada per una font facilita el desenvolupament de tècniques de codificació específiques per a una font. En molts casos, aquestes tècniques ens permeten aproximar-nos a les cotes marcades per l'entropia sense introduir pèrdues. En particular, per a la informació audiovisual es poden utilitzar tècniques molt específiques, que al costat de les de propòsit general, ens permetran reduir notablement els temps de transmissió de la informació. Atesa l'eficiència de les tècniques de codificació aplicades, considerarem que els dígit binaris resultants d'una codificació equivalen, pràcticament, a 1 bit d'informació. Per això, considerarem indistintament el bit i el dígit binari a partir d'aquest apartat.

3. Codificació audiovisual MPEG

La naturalesa de la informació audiovisual fa possible el desenvolupament de tècniques específiques per a la seva codificació. Pel que fa a la distribució de continguts audiovisuals, les tècniques emprades aprofiten la caracterització de la percepció humana per a crear estratègies que maximitzin la fidelitat dels senyals reproduïts en els aspectes en què l'ésser humà és més sensible.

Exemple

Si pensem en imatges, el sistema visual humà és més sensible a la luminància que a la crominància. Una cosa semblant passa en el cas de l'àudio, en què hi ha bandes freqüencials que estimulen més el sistema auditiu amb la mateixa potència que altres.

La compressió MPEG audiovisual preveu la possibilitat d'introduir pèrdues d'informació durant el procés de codificació que donaran lloc a una falta de fidelitat o discrepància entre el contingut original i el contingut descodificat. Per tant, en general, la compressió MPEG és una codificació amb pèrdues que permet dépassar el límit que marca l'entropia. No obstant això, la distorsió introduïda en augmentar la compressió pot tenir un impacte mínim en la percepció humana si els elements de la informació es codifiquen d'acord amb la sensibilitat de l'ésser humà.

Exemple

El sistema visual humà presta molta menys atenció als contorns que a la part central de les imatges. Passa el mateix amb l'àudio, en què el sistema auditiu de l'ésser humà té molta menys sensibilitat en les freqüències més altes del so.

3.1. Compressió de vídeo MPEG

La compressió MPEG explota la naturalesa de la informació de vídeo per a definir una tècnica eficient per a codificar la majoria de les seqüències i es fonamenta en el fet que els píxels que estan propers espacialment en una mateixa imatge, o en imatges consecutives, en la majoria dels casos són de valors molt semblants.

Nota

Hem de tenir present que els objectes que apareixen en una imatge disposen d'una superfície que sol ocupar un nombre elevat de píxels contigus i que aquests objectes apareixen repetits en moltes imatges consecutives d'una mateixa escena. L'objectiu de la codificació serà extreure tota la redundància que hi ha dins d'una mateixa imatge i entre imatges properes temporalment.

Quan s'apliquen tècniques per a reduir la redundància espacial (semblança dins de la imatge) es du a terme una codificació que denominem *intraquadre* o *intra-frame* (entenen *frame* com a imatge). Quan la redundància extreta és temporal (semblança entre imatges) es du a terme una codificació que denominem *interquadre* o *inter-frame* o també *predictiva*:

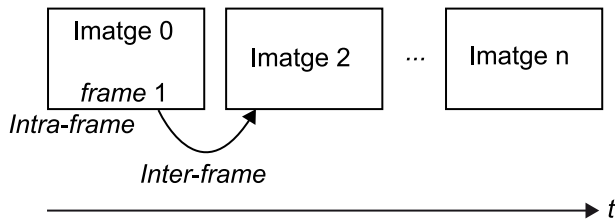


Figura 3. Tipus d'extracció de la redundància en els píxels de la imatge

En la codificació *inter-frame* es du a terme una codificació predictiva endavant (*P-frame*) que explota la semblança dels píxels de la imatge a codificar amb els de una imatge codificada prèviament. Es busca la millor predicció des de la imatge prèvia per a reconstruir l'actual, per a això el que es codifica és la diferència entre una zona de la imatge prèvia i una zona de la imatge actual. Les diferències codificades són un conjunt de residus o errors de predicció que, en principi, haurien de prendre valors petits i molt incorrelats quan la semblança entre les imatges sigui molt alta.

Per a millorar la codificació de la redundància temporal, augmentant la possibilitat de trobar zones d'imatge semblants entre imatges, es va incorporar una segona forma de codificació *inter-frame* denominada codificació **bidireccional-predictiva** (*B-frame*), que permet que la codificació d'un *frame* utilitzi com a referències porcions d'imatge d'un *frame* temporalment anterior i un altre de temporalment posterior al que es vol codificar. Aquesta opció de codificació comporta codificar *frames* en un ordre diferent de l'ordre temporal de presentació de les imatges d'una seqüència.

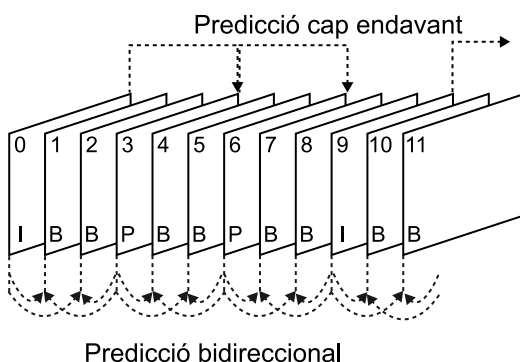


Figura 4. Exemple de codificació predictiva endavant i bidireccional

3.2. Estructuració de la seqüència de vídeo MPEG

Per a codificar la informació d'una seqüència de vídeo és necessari estructurar la informació que es codificarà. Per a això s'ha definit un conjunt d'unitats d'informació de vídeo que estructurin els elements a codificar:

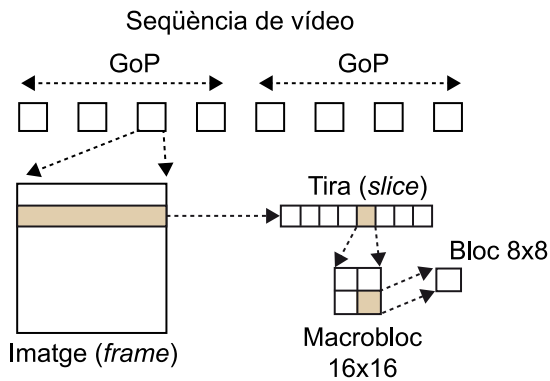


Figura 5. Estructuració de la seqüència de vídeo

En primer lloc, es defineix pròpiament la seqüència que estarà composta per totes les imatges del contingut audiovisual. La seqüència s'organitza en grups d'imatges anomenats GoP (*group of pictures*), que contindran un conjunt de *frames* que podran estar relacionats en la seva codificació.

Els *frames* s'estructuren en blocs de 8×8 píxels, que s'agrupen formant macroblocs. Els macroblocs estan definits per 4 blocs de 8×8 píxels de luminància. Així doncs, el macrobloc constitueix una zona quadrada de la imatge de 16×16 píxels de luminància i els valors associats dels components de cromina. De manera que si la imatge està submostrejada amb format 4:2:0, el macrobloc estarà compost per 4 blocs de luminància i 1 bloc de cromina per cada component de color.

Els macroblocs en posicions consecutives, d'esquerra a dreta i de dalt a baix dins d'una imatge, se solen agrupar lliurement per a determinar una tira d'imatges o *slice*. En la pràctica, la majoria dels codificadors utilitzen com a tira d'imatges tots els macroblocs que estan situats en la mateixa posició vertical de la imatge.

Per a codificar les imatges dins d'un GoP s'aplica un patró de codificació regularment, en el qual s'empren les tècniques *intra-frame* o *inter-frame* en les quals es basa l'estàndard MPEG-2. En un GoP podem trobar:

- **Imatges I o *intra-frames*.** No usen predicció, no depenen d'imatges anteriors. Serveixen de punts de resincronització en descodificació quan es produeixen pèrdues o errors en el canal de transmissió. Es poden emprar per a facilitar una visualització ràpida del contingut en la seqüència.

- **Imatges P o *predicted-frames*.** Imatges amb codificació predictiva cap a endavant. Usen informació d'una imatge I o P anterior.
- **Imatges B o *bidirectional-predicted-frames*.** Imatges que usen informació d'una o diverses imatges I o P tant anteriors com posteriors.

Com que el patró de codificació se sol mantenir constant durant tota la codificació, es defineixen dues magnituds per a determinar aquest patró. La distància entre imatges del tipus I, s'anomena N , mentre que a la distància entre imatges P o entre I-P s'anomena M . En la pràctica, un GoP s'inicia quan apareix una nova codificació d'una imatge I, per la qual cosa la grandària N també indica el nombre d'imatges que conté un GoP. El valor de N és un múltiple de M , d'aquesta manera es garanteix que el nombre d'imatges B consecutives és sempre igual i de valor $M-1$.

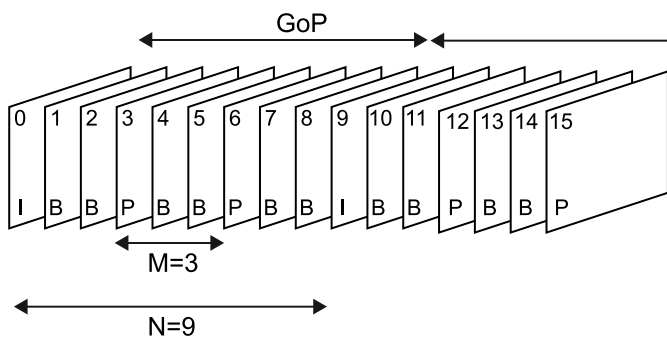


Figura 6. Exemple de codificació d'imatges en un GoP

3.3. Codificació *intra-frame* (imatges I)

La codificació *intra-frame* fa una codificació de tots els blocs de la imatge macrobloc a macrobloc. De manera que, per a una imatge en format 4:2:0, en cada macrobloc, primer es codifiquen els quatre blocs de luminància, després el bloc de cromina blava i després el de cromina vermella.

La figura següent esquematitza el procés seguit per a cada macrobloc de la imatge que es codificarà en mode *intra-frame*:

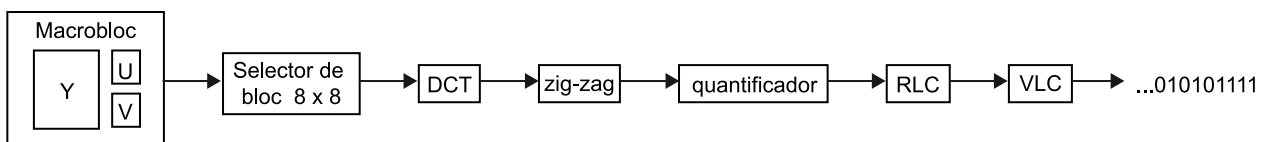


Figura 7. Esquema de codificació intra

Com es representa en la figura, els 6 blocs de cada macrobloc segueixen el mateix procés de codificació. El procés es pot descompondre en una sèrie d'operacions bàsiques que s'apliquen a cada bloc fins a codificar-lo. Aquestes operacions estan dissenyades amb els propòsits següents:

- Ordenar la informació del bloc (DCT i zig-zag) per a facilitar-ne la codificació.
- Dur a terme una codificació introduint pèrdues selectives (quantificador).
- Obtenir un grau elevat de compressió explotant una anàlisi prèvia del comportament estadístic de les fonts de vídeo (RLC i VLC).

3.3.1. Transformada cosinus (DCT)

Per a fer una codificació eficient d'una porció d'una imatge hem d'observar dues característiques fonamentals que apliquen en la majoria dels casos:

- Les imatges solen presentar una semblança alta dels píxels en una mateixa zona.
- La semblança dels píxels d'una zona d'imatge sol decréixer lentament.

Aquestes dues característiques dels píxels de la imatge es poden aprofitar per a desenvolupar una codificació òptima que intenti reduir al màxim el nombre de valors necessaris per a representar la informació continguda en la porció d'imatge. Per tant, el resultat és una representació equivalent de la informació denominada *transformada discreta cosinus* (DCT, *discrete cosine transform*). Aquesta tècnica s'ha utilitzat des del primer dels estàndards MPEG-1 i es continua aplicant en els actuals com l'MPEG-2 i l'MPEG-4.

La transformada DCT és un cas particular de la transformada discreta de Fourier, en què la informació es pot representar com una sèrie de cosinus en fase. Per al cas de les imatges serà una transformació bidimensional en què porcions d'imatge definides pel valor dels píxels seran representades pels seus coeficients transformats. La DCT és una funció reversible, per la qual cosa no introdueix cap distorsió. La DCT és simplement una manera equivalent de representar la mateixa informació i presenta com a avantatge principal respecte a la transformada de Fourier discreta que necessita la meitat de coeficients per a representar la mateixa informació.

Les porcions d'imatge que s'utilitzen per a l'aplicació de la transformada DCT són blocs rectangulars de píxels que en els estàndards de codificació MPEG solen ser de 8×8 píxels. En general, aquests blocs disposaran de valors de píxels més semblants com millor sigui la resolució de la imatge que s'ha de codificar. De manera equivalent, com més gran sigui el bloc emprat, més probabilitats hi haurà que els píxels prenguin valors més dispersos.

La transformació d'un bloc de 8×8 píxels dóna lloc a un conjunt de 8×8 coeficients. En la pràctica, els blocs i els coeficients s'organitzen matricialment. El primer element de la matriu s'indexa en la posició (0,0) i l'últim en la posició (7, 7).

En aplicar la DCT, cada coeficient indica la intensitat present d'una freqüència dins del bloc. El primer coeficient (0,0), o coeficient de contínua, representa el valor mitjà dels píxels del bloc. La resta de coeficients, o coeficients d'alterna, prendran un valor que dependrà de com siguin les variacions en el valor dels píxels de la mateixa posició horitzontal o vertical: si els valors de píxels consecutius en una direcció presenten variacions lentes, el resultat serà que els coeficients propers al de contínua, que denominarem *coeficients de baixa freqüència*, tindran un grau elevat d'intensitat; si les variacions dels píxels són ràpides o brusques, llavors seran els coeficients més allunyats del de contínua, o coeficients d'alta freqüència, els que presentaran valors més elevats.

En la majoria de casos, els blocs de 8×8 píxels d'una imatge tindran valors molt semblants i correlats, per la qual cosa els coeficients de contínua i de baixa freqüència seran els que presentaran una intensitat més alta. En els casos en què els píxels d'un bloc siguin molt dispers i incorrelats, aleshores tots els coeficients presentaran un grau d'intensitat semblant. Per tant, per al primer cas, l'energia del bloc està concentrada en les baixes freqüències, mentre que en el segon l'energia està distribuïda uniformement en totes les freqüències.

Diem que un bloc és més complex quan la seva energia està més dispersa entre les freqüències perquè els valors dels píxels estan més desordenats. En estar més desordenats presenten més entropia i, per tant, necessitarem més bits per a poder-los codificar.

3.3.2. Ordenació en zig-zag

En la majoria dels blocs, la DCT aconsegueix representar la seva informació amb un nombre reduït de coeficients de baixa freqüència, mantenint nuls o propers a zero els d'alta freqüència. En aquesta situació, els coeficients d'intensitat notable seran pocs i apareixeran ordenats de més a menys intensitat a mesura que ens allunyem del component de contínua.

A fi de facilitar el procés de codificació, s'ordenen les mostres seguint un recorregut en zig-zag que s'allunya gradualment del coeficient de contínua (0,0) fins al (7,7). Aquesta ordenació ens permet passar d'una representació bidimensional a una d'unidimensional de 64 elements. Una vegada ordenats els coeficients, i tenint en compte que la majoria estaran prop de 0, es podran aplicar fàcilment tècniques clàssiques RLC de codificació en ràfegues per a les seqüències de zeros consecutius.

3.3.3. Quantificació

Atès que l'MPEG considera la possibilitat d'introduir degradacions en la informació de vídeo, després de l'operació d'ordenament en zig-zag, s'aplica la quantificació dels coeficients de la DCT, que implica un arrodoniment dels coeficients a un valor més o menys proper. Si s'aplica un pas de quantificació petit, l'arrodoniment que pot arribar a experimentar un coeficient també serà petit i, per contra, si s'utilitza un pas de quantificació gran, l'error entre el valor original i l'arrodonit pot ser notable. Aquest error el denominarem *error de quantificació*.

El pas de quantificació, Q , pot prendre valors que solen anar d'1 a 31. Així, per a $Q = 1$ podem dir que els coeficients en general tenen un error de quantificació que afecta fonamentalment el bit de menys pes, mentre que per a $Q = 31$, l'error pot arribar a afectar els 5 bits de menys pes del coeficient. Per tant, quan s'empra un pas de quantificació relativament gran, la majoria dels coeficients de la DCT s'arrodoneixen al valor 0, la qual cosa ens porta a augmentar les seqüències de zeros seguits que hem de codificar després de la quantificació.

3.3.4. Codificació a ràfegues RLC i codificació entròpica VLC

L'ordenació freqüencial de menys a més freqüència feta amb el zig-zag i la quantificació posterior fan que la informació a codificar d'un bloc estigui representada per seqüències de 64 coeficients quantificats. Aquestes seqüències disposen de llargues successions de valors nuls que faciliten una codificació eficient. Per a aquest tipus de dades, el més apropiat és utilitzar la codificació per ràfegues RLC estudiada anteriorment. El codificador determina per a cada successió la longitud de les ràfegues de zeros i el valor que la segueix. Aquest parell d'elements (longitud, valor) disposen d'una codificació predeterminada que es troba en la taula de codis VLC (*variable length coding*).

Recorden

La taula de codis VLC s'haurà construït prèviament segons la probabilitat dels elements que s'han de codificar.

En aquest cas, els valors de codificació VLC es van obtenir analitzant l'estadística dels blocs transformats de moltes seqüències diferents. A partir d'aquest estudi es determinen les freqüències d'aparició de les ràfegues de zeros i els valors que les segueixen i defineixen la codificació dels elements (longitud, valor).

El coeficient de contínua es codifica d'una manera diferent, ja que depèn del valor mitjà dels píxels del bloc. La codificació aplicada és diferencial respecte al coeficient de contínua codificat per al bloc anterior. En aquest cas, es considera que blocs adjacents disposen de valors semblants en els píxels, per la qual cosa la diferència entre el component de contínua d'un bloc i d'un altre de proper ha de ser relativament petita. Per a iniciar la codificació del primer coeficient de contínua es parteix del valor central que es pugui codificar.

El valor de contínua se sol quantificar amb molt poc arrodoniment, de manera que no es comenten errors notables entre un bloc i qualsevol dels contigus. L'estadística dels valors diferencials quantificats determina les paraules generades en el procés de codificació VLC.

L'error introduït en les mostres de contínua provoca la percepció del denominat *efecte bloc* en les imatges presentades. Així, sobre la imatge es poden observar zones rectangulars amb diferents valors dels píxels. Els efectes de la quantificació en els coeficients d'alterna provoquen fonamentalment la difuminació dels contorns dels objectes dins de la imatge.

3.4. Codificació *inter-frame*

La codificació *inter-frame* intenta extreure de manera eficient la redundància temporal entre imatges consecutives de la seqüència. Per a extreure la redundància, primer es determina la imatge o imatges que s'empraran com a referència per a codificar la imatge actual.

3.4.1. Codificació predictiva cap endavant (imatges P)

Quan s'aplica una codificació predictiva, el *frame P* (*predicted*) a codificar utilitza el *frame I* o *P* immediatament anterior, que ja haurà estat codificat i emmagatzemat en memòria per a utilitzar-lo com a referència.

La porció d'imatge que s'empra per a fer la predicció és el macrobloc. La codificació predictiva d'una imatge es fa macrobloc a macrobloc en ordre seqüencial (d'esquerra a dreta i de dalt a baix). Per a cada macrobloc, es busca la porció d'imatge de la mateixa mida en la imatge de referència que tingui els píxels menys diferents respecte als píxels del macrobloc que s'ha de codificar. Aquesta cerca de la porció d'imatge més semblant en la imatge de referència no es fa sobre tota la superfície de la imatge. Es considera que la cerca s'ha de limitar a un determinat rang de desplaçament relatiu a la posició del macrobloc que s'ha de codificar que estarà determinat per la resolució de la imatge, de manera que per a imatges més grans s'hauria d'utilitzar un rang de cerca més gran. Per a fer la cerca s'apliquen algorismes experts que intenten fer el mínim nombre de càlculs. Una vegada seleccionada la porció de la imatge de referència, s'indica quin és mitjançant les seves coordenades, expressades amb el denominat *vector de compensació de moviment*.

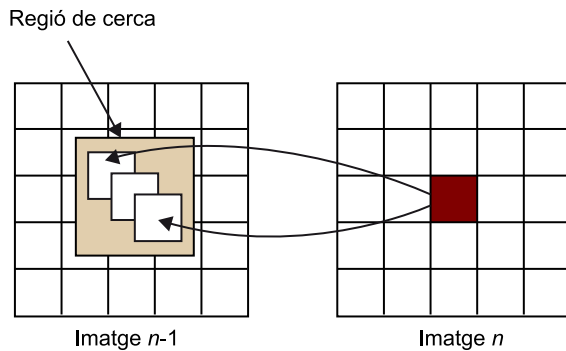


Figura 8. Rang de cerca per a la codificació predictiva

En la codificació d'un macrobloc predictiu, a més del vector de moviment, es pot codificar la diferència del macrobloc respecte a la seva referència. Per a això s'obté un macrobloc residual resultat de la diferència entre cada píxel del macrobloc que s'ha de codificar i de la porció d'imatge seleccionada com a referència. Si el grau de semblança entre tots dos conjunts de píxels és alt, el macrobloc diferencial disposarà d'un conjunt de valors amb un comportament molt aleatori i proper a 0. Si no és així, el codificador pot decidir en cada macrobloc de la imatge predictiva si renuncia a la codificació diferencial. En aquest cas, el codificador pot codificar el macrobloc de manera autònoma aplicant una codificació com es faria si fos un macrobloc d'una imatge en què s'apliqués la codificació *intra-frame* explicada anteriorment. Per això, per a cada macrobloc d'una imatge predictiva, el codificador senyalitzarà el tipus de codificació aplicat a cada un, tant si és predictiva com intra.

En els macroblocs diferencials s'aplica la mateixa metodologia que per a la codificació *intra-frame*: DCT, zig-zag, quantificació i RLC amb VLC per a cada bloc. La diferència principal entre tots dos casos és que es considera que l'energia de la informació diferencial està distribuïda uniformement entre totes les freqüències, la qual cosa porta a quantificar de la mateixa manera tots els coeficients freqüencials, fins i tot el coeficient de contínua. Per tant, la codificació de cada bloc després del zig-zag i la quantificació donarà lloc a ràfegues de zeros i valors amb una estadística que serà diferent de la dels *frames* I. Per això, les taules de codificació VLC seran diferents per a la codificació predictiva d'un macrobloc respecte a la codificació intra.

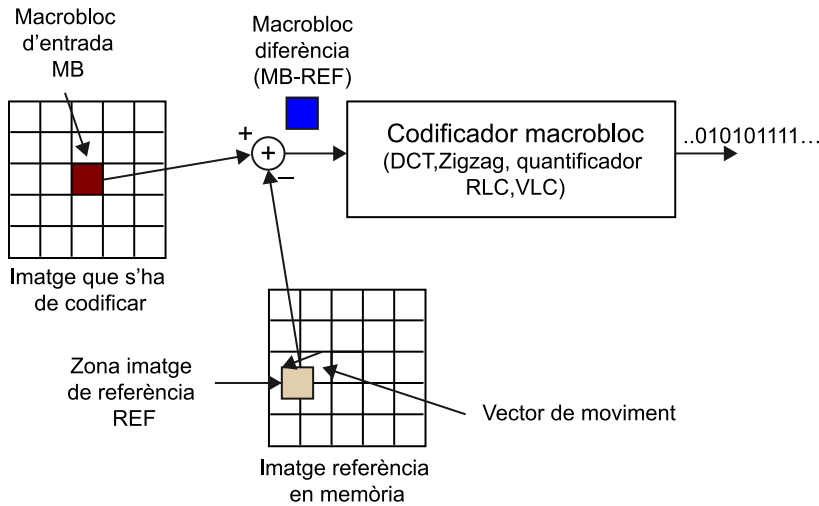


Figura 9. Esquema de codificació predictiva

3.4.2. Codificació bidireccional-predictiva (imatges B)

Per a facilitar l'extracció de la redundància temporal, els estàndards de codificació MPEG inclouen la possibilitat d'utilitzar dues imatges de referència en el moment de la codificació. Aquestes imatges de referència es corresponen una amb la imatge I o P anterior i l'altra amb la I o P posterior. Per a emprar-les en la codificació de la imatge actual, cal que totes dues hagin estat prèviament codificades, la qual cosa implica que l'ordre de codificació i de transmissió no és el mateix que l'ordre de captura i presentació.

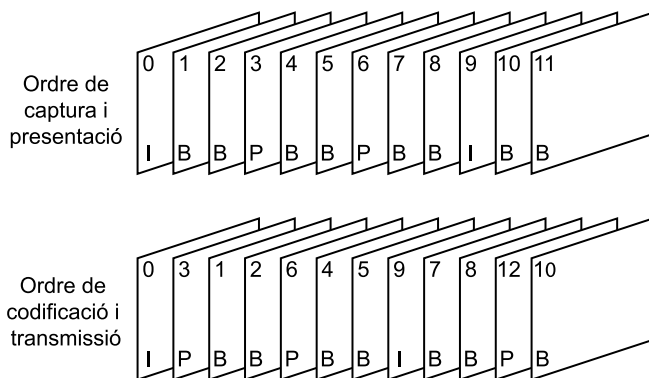


Figura 10. Modificació de l'ordre de transmissió-codificació respecte al de captura-presentació

Les imatges bidireccional-predictives, o imatges B, es codifiquen de manera molt semblant a la que s'ha explicat per a les predictives. En aquest cas, la generalització es basa en la selecció de la referència que emprarà un macrobloc; per a això, es fa una cerca d'una porció d'imatge semblant en la imatge de referència anterior, i també en la posterior. Cada porció quedarà definida per un vector de compensació de moviment. Una vegada trobades les porcions d'imatge, el codificador haurà de decidir si utilitzarà la referència cap enrere, cap endavant o un valor mitjà de totes dues. Per a això, compararà el valor dels píxels del macrobloc que s'ha de codificar amb els de la referència anterior, posterior o una semisuma de totes dues, i seleccionarà el que més s'hi assembla.

D'aquesta manera, la codificació del macrobloc podrà tenir un sol vector de compensació de moviment, per als dos primers casos, o els dos vectors si es decideix l'últim cas.

Una vegada fixada la millor referència, en última instància s'ha d'analitzar l'aleatorietat del macrobloc diferencial. Com en el cas predictiu, el codificador pot decidir utilitzar una codificació intra per a cada macrobloc si ho considera convenient. Si el grau d'aleatorietat és alt, es codificarà el macrobloc diferencial aplicant exactament el mateix procés que en el cas predictiu: DCT, zig-zag, quantificador i RLC amb VLC per bloc del macrobloc.

3.5. Compressió d'àudio MPEG

La compressió d'àudio que s'empra principalment per a la transmissió es deriva de MUSICAM (*masking pattern adapted universal subband integrated coding and multiplexing*), creat per a la transmissió de ràdio digital. L'estàndard MPEG-2 defineix a més altres tècniques de codificació que no s'utilitzen habitualment en la transmissió d'àudio en DVB. La recomanació que inicialment es va adoptar per a la televisió va ser la codificació d'àudio coneguda com a MPEG 2 àudio de capa 2.

MPEG va definir una codificació més simple que no s'ha emprat extensivament, l'MPEG-2 de capa 1 que du a terme les mateixes operacions que la versió de capa 2, però requereix menys memòria que la utilitzada en DVB. D'altra banda, MPEG també va definir la reeixida codificació MPEG-2 de capa 3, coneguda pels arxius MP3, que es fonamenta en la codificació de capa 2, però afegeix una transformació DCT addicional que millora la compressió. No obstant això, l'ús de la transformació freqüencial dificulta la interpolació de les mostres quan es produeixen errors en la transmissió, per la qual cosa es va optar per una decisió conservadora utilitzant l'MPEG-2 de capa 2.

Per a la codificació de l'MPEG-2 d'àudio de capa 2 s'especifiquen un conjunt de freqüències de mostreig (f_m) de valors: 16, 22,05, 24, 32, 44,1 i 48 kHz. Aquestes determinen l'amplada de banda màxima del senyal d'àudio que es reproduirà ($f_m/2$). La codificació d'àudio intentarà mantenir una velocitat constant que en l'estàndard s'especifica amb valors que van des de 32 kbps fins a 384 kbps.

Els codificadors MPEG-2 esmentats operen amb blocs de mostres obtinguts del filtratge i mostreig del senyal. Aquests blocs són de mida fixa i tenen un valor de 1.152 mostres per als codificadors de capa 2 i capa 3. Aquests blocs de mostres, denominats *trames*, constitueixen l'entitat bàsica dins de l'estructuració de la codificació d'àudio. Hem de tenir en compte que la durada temporal d'una trama d'àudio dependrà de la freqüència de mostreig.

Exemple

Emprant la freqüència màxima de mostreig ($f_m = 48$ kHz) la durada de la trama serà de 24 ms, i utilitzant la freqüència de mostreig més lenta ($f_m = 16$ kHz) la durada de la trama serà de 72 ms.

La base fonamental de la codificació és explotar el coneixement de la percepció auditiva humana per a determinar quina informació és més rellevant i quina menys. Aquest coneixement determinarà quan s'ha de garantir un grau de fidelitat elevat entre la informació original i la descodificada o quan es poden introduir graus elevats de pèrdues en la informació sense que s'aprecii una distorsió important durant la reproducció del so.

Els tres fenòmens que es tenen en compte pel que fa a la percepció de l'àudio per a definir la codificació MPEG són:

- La **sensibilitat** depèn de la freqüència. De manera que en l'amplada de banda vocal de 300 Hz a 3.400 Hz l'oïda és més sensible. Aquesta característica s'associa a la generació de l'aparell fonador humà.
- L'emascament **freqüencial**. Un senyal de gran energia en una freqüència redueix la sensibilitat del sistema auditiu humà en les freqüències properes.
- L'emascament **temporal**. Una estimulació elevada redueix la sensibilitat del sistema auditiu durant un curt període de temps. En aquest cas, una gran energia durant un període de temps, a causa d'una gran amplitud del so, emmascara temporalment sons d'amplitud reduïda que la segueixen.

En principi, el procés de codificació consisteix a dividir la banda d'àudio en 32 subbandes de la mateixa mida mitjançant un banc de filtres. L'entrada al banc de filtres es realitza dividint les 1.152 mostres en grups de 32 mostres que es lliuren a tots els filtres simultàniament. Cada grup de 32 mostres genera una única sortida per cada filtre. D'aquesta manera s'obtenen, per a cada bloc d'entrada de 1.152 mostres, 36 mostres per a cada filtre que es quantificaran segons el model psicoacústic.

Com es mostra en la figura 11, a més de repartir la mostres de cada trama entre els 32 filtres, les trames s'empren per a explotar el model psicoacústic de percepció humana. En funció dels valors temporals que prenen les mostres, el conformador psicoacústic determina els llindars d'emascament temporal i defineix en quina escala de valors s'han de trobar les mostres per a ser percebudes. El conformador també determina els efectes de l'emascament freqüencial aplicant una transformada FFT de 1.024 punts sobre la trama d'àudio.

Com a resultat d'aquestes operacions s'escull el conjunt de quantificadors que s'ha d'aplicar al banc de filtres. L'elecció del quantificador dependrà de l'anàlisi psicoacústic i de la freqüència en què estigui centrat el filtre. D'aquesta manera,

les regions de freqüència en què l'oïda és més sensible es quantifiquen amb més precisió que les de menys sensibilitat auditiva, en altres paraules, el pas de quantificació és més petit per a les freqüències en què l'oïda és més sensible.

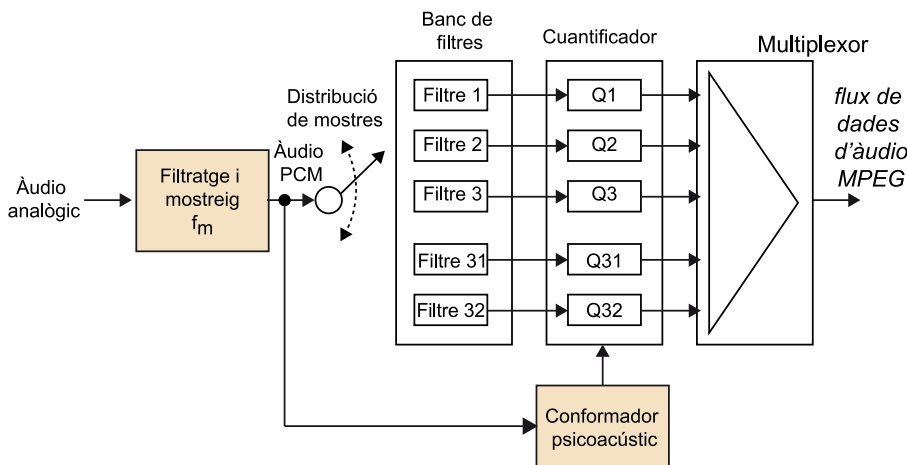


Figura 11. Esquema del codificador d'àudio MPEG 2

Les 36 mostres de sortida de cada quantificador són recollides per un multiplexor que les codifica per a formar un *stream* elemental d'àudio.

Quan es codifiquen diversos canals, com en el cas del so estèreo, es fa el mateix conjunt d'operacions per a cada canal. En alguns casos, s'explota la redundància de la informació entre canals i es desenvolupa una codificació psicoacústica conjunta.

Quan es codifiquen diversos canals, el multiplexor estructura la informació en un únic *stream* elemental en què les codificacions dels canals poden ser independents o no.

Per a DVB, inicialment s'han proposat quatre formats possibles de codificació MPEG-2:

- Format **mono**, en què es codifica un únic canal.
- Format **estèreo** amb canals codificats independentment que es descodifiquen i reproduïxen simultàniament.
- Format **estèreo de codificació conjunta**, en què s'explota la redundància entre els canals dret i esquerre. Aquesta codificació requereix una descodificació simultània per a la seva reproducció en paral·lel.
- Mode **dual**, en què hi ha dos canals que es codifiquen independentment, i solament se'n descodifica i reproduïxen uns quants segons la selecció de l'usuari.

4. Multiplexatge i scrambling en DVB

Les especificacions per al transport de la codificació MPEG per a àudio i vídeo donen lloc a la recomanació coneguda com a *capa de sistema*. En la capa de sistema s'indica com s'han d'agrupar els fluxos elementals (àudio o vídeo comprimits) i altres fluxos de dades de propòsit general per a formar el multiplex de transport MPEG. El multiplexatge dels fluxos audiovisuals ha de complir una sèrie de requeriments temporals per a garantir que el receptor disposarà de la informació necessària per a poder-la descodificar i presentar de manera síncrona.

4.1. Empaquetament i multiplexatge MPEG

Per fer el multiplexatge i la transmissió dels fluxos elementals (*elementary streams*, ES) el comitè MPEG va especificar el format de transport denominat *transport stream* (TS), que facilita a l'usuari conèixer els programes continguts en el flux d'informació i seleccionar-ne un. Per tant, en un TS es podran distribuir simultàniament la codificació audiovisual de diversos programes (per exemple, La 1, Telecinco, Antena 3, etc.). Cada programa de l'*stream* podrà disposar de diverses pistes o codificacions elementals d'àudio, corresponents, per exemple, a diferents llengües, i diverses pistes o codificacions de vídeo, corresponents, per exemple, a diferents angles de la mateixa presa.

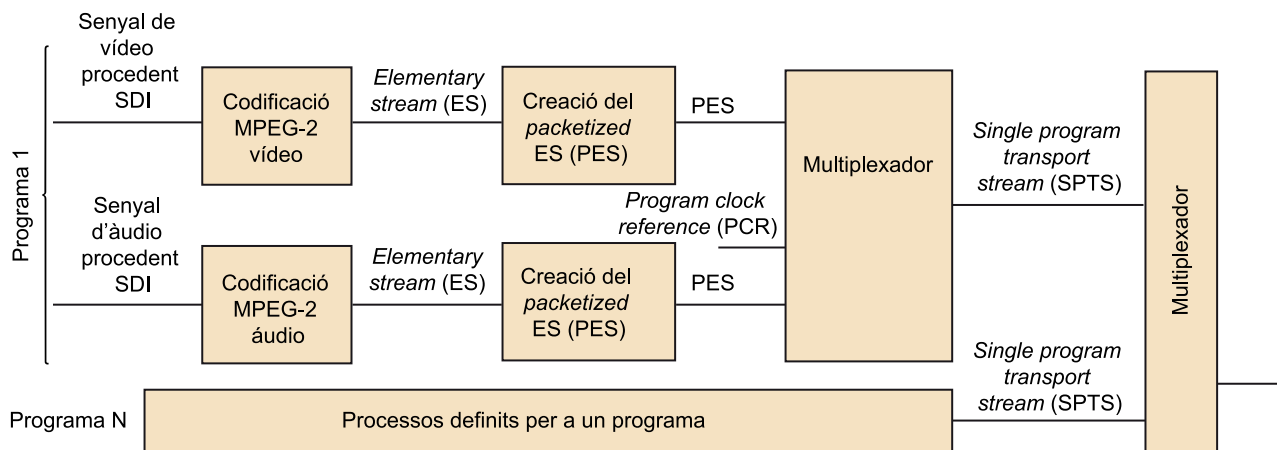


Figura 12. Esquema de la capa de sistema MPEG 2

En primer lloc, els fluxos binaris de vídeo i àudio (procedents del senyal SDI) es comprimeixen independentment seguint les codificacions de vídeo i àudio definides per l'MPEG-2 i cada un forma un *elementary stream* (ES), que també pot ser un senyal de dades (per exemple, per a aplicacions). Cada ES conté *access units* (AU), que solen ser un *frame* per a la imatge (de tipus I, P o B) i una trama de mostres codificades per al so.

A continuació, atès que un ES és un corrent de bits, cada ES del programa, és a dir, vídeo i àudio, s'estructura en paquets, i per a això, cada cert nombre de bits de l'ES (aquest nombre no està definit i depèn del dissenyador del sistema) s'afegeix una capçalera. Aquest nombre de bits sol estar determinat per la grandària de les AU. El resultat d'unir una capçalera a un conjunt de bits de l'ES es denomina *packetized elementary stream* (PES).

Atès que les AU són de longitud variable (no ocupa el mateix un *frame* I que un P o B), la capçalera incorpora un camp en què s'indica la mida del paquet PES. D'aquesta manera el receptor sap quan finalitza el PES i s'inicia la capçalera del paquet PES següent. Encara que la mida del PES pot ser la que el dissenyador del sistema vulgui, s'ha establert un màxim de 64 kB.

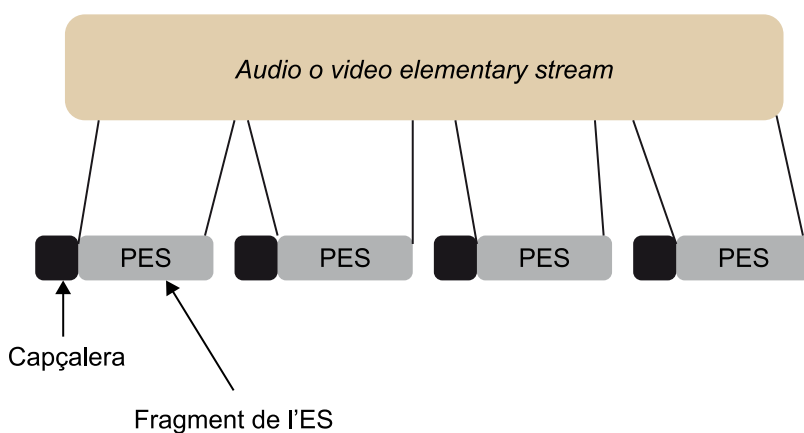


Figura 13. Empaquetatge d'ES en PES

En la capçalera del paquet PES s'han definit diferents camps per a cobrir les necessitats de la transmissió multiplexada:

- **Identificació.** Perquè el receptor seleccioni les AU d'un ES, s'ha proposat un camp d'identificació que té el mateix valor en tots els paquets PES del mateix ES.
- **Referències temporals.** Són necessàries per a garantir la descodificació i presentació sincronitzada dels ES d'un programa en el descodificador. Així, els fluxos audiovisuals d'un programa disposaran d'unes marques temporals que facilitaran el sincronisme durant la reproducció del programa. Aquesta marcació garanteix, per exemple, la sincronització de veu i imatge durant la reproducció del contingut.
- **Número de seqüència.** Facilita al descodificador la detecció de pèrdues d'informació que es puguin haver produït durant la transmissió.

Finalment, els paquets PES dels diferents ES d'un programa es fragmenten i originen paquets més petits de mida fixa, exactament de 188 bytes (4 de capçalera + 184 de dades), anomenats *transport stream packets* (TSP), que redueixen l'impacte dels errors de transmissió: si hi ha errors en la transmissió, el sistema haurà de protegir unitats de transmissió petites. La capçalera dels TSP

inclou un identificador PID (*transport packet identifier*) que indica de quin ES prové la informació. Un TSP solament pot portar parts d'un únic paquet PES i pot incorporar un farciment addicional (camp anomenat *adaptation field*), si és necessari, fins a completar els 184 bytes de dades que ha de contenir.

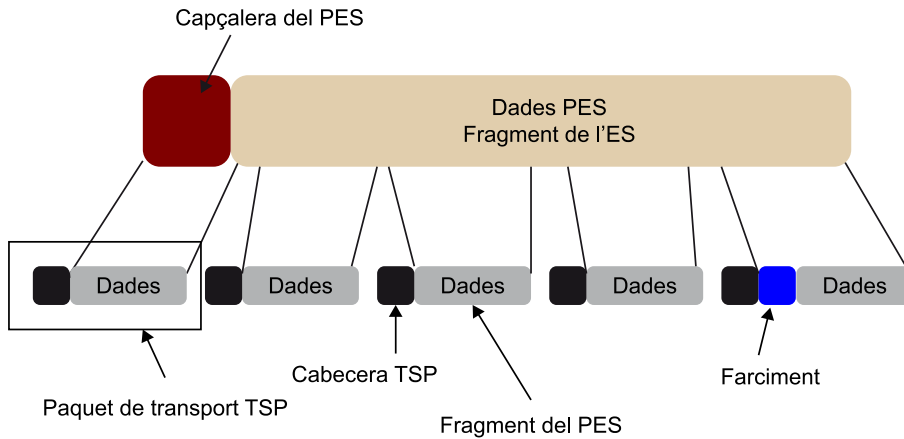


Figura 14. Fragmentació dels paquets PES en paquets de transport TSP

El conjunt de paquets TSP d'un programa es coneix com a SPTS (*single program transport stream*). Si per un mateix canal s'envien (com és habitual) diversos programes, és a dir, diversos SPTS, els diferents SPTS es multiplexen i originen un únic flux binari que es denomina MPTS (*multi program transport system*). En aquest cas cal tenir en compte que cada SPTS (programa) té la seva pròpia sincronització i, per a poder coordinar les temporitzacions dels ES d'un mateix programa, s'afegeix informació de sincronisme de manera independent per a cada programa multiplexat (cada SPTS). Aquesta informació temporal, denominada *program clock reference* (PCR), s'incorpora, quan apareix, en el camp opcional *adaptation field* dels paquets de transport (TSP).

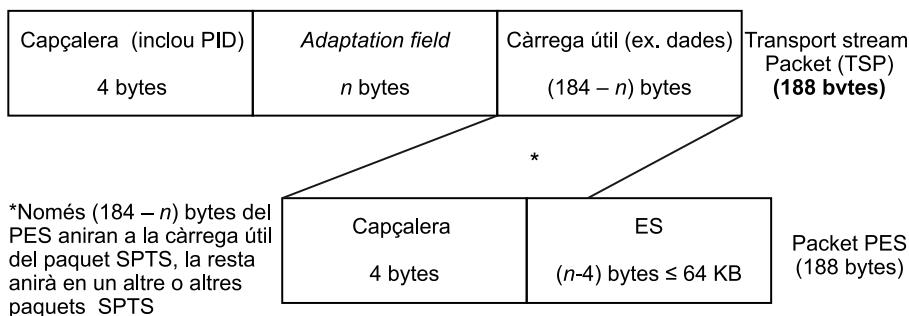
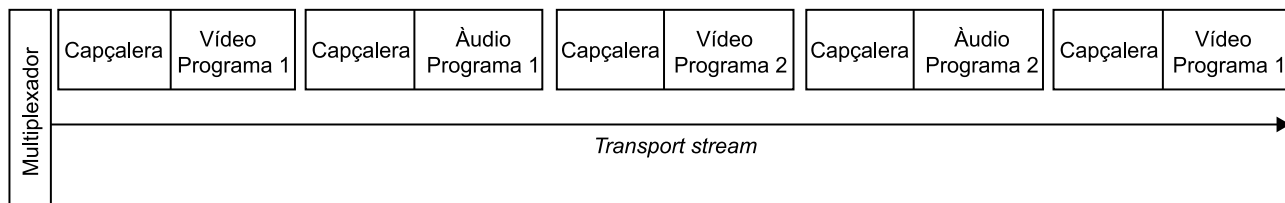


Figura 15. Format dels paquets de transport TSP

El resultat final de multiplexar diferents SPTS és que s'entremesclen els paquets dels diferents programes (SPTS). A continuació mostrarem un possible resultat d'un multiplexatge:

Figura 16. Exemple de *stream* de transport MPTS

A més dels *transport stream packets* (TSP) amb informació de vídeo/àudio/dades d'un programa, l'*stream* de transport (SPTS o MPTS) porta altres paquets de transport (TSP) que contenen metainformació de cada programa que la capa de sistema MPEG i l'estàndard DVB structuren en taules. La codificació d'aquestes taules s'envia mitjançant paquets de transport (TSP) els identificadors dels quals estan definits *a priori* o dins de les mateixes taules. A continuació veurem quin tipus de taules (o metainformacions) s'han definit. Aquestes taules es coneixen amb el nom de *program-specific information* (PSI).

4.2. Informació específica d'un programa (PSI)

El descodificador MPEG-2, a més de descodificar cadascun dels ES d'àudio i vídeo (i, de vegades, dades) que conformen un programa, ha de ser capaç de trobar-los dins d'un *stream* de transport de tipus MPTS. La informació específica d'un programa (*program specific information*, PSI) és el que permet que el descodificador faci aquesta tasca. Les taules principals en les quals s'organitza el transport d'informació en DVB són les següents:

- **Taula d'assignació de programes** (*program allocation table*, PAT). Aquesta taula és el primer element que necessita un descodificador per a poder determinar el conjunt de programes que s'estan transmetent en el múltiplex; és a dir, en el MPTS. Per tant, és una taula bàsica i obligatòria l'identificador associat de la qual en els paquets de transport és PID = 0. El seu propòsit és indicar per a cada programa transportat quin és l'identificador dels paquets de transport que contenen la taula PMT (vegeu l'element següent) que el descriu. Conceptualment, aquesta taula està composta per un conjunt de punters que indiquen on es pot trobar la descripció de cada programa transmès, com es mostra en la figura 17.
- **Taula del mapa de programa** (*program map table*, PMT). Aquesta taula detalla i descriu el conjunt d'ES que componen el programa. Es compon d'un conjunt d'identificadors de paquets de transmissió associats a cada ES i el tipus de flux audiovisual, mitjançant els quals, al costat de les descripcions de tipus, el descodificador podrà oferir a l'usuari diferents possibilitats de presentació del programa, per exemple, es pot seleccionar l'angle de la presa de vídeo i l'idioma de l'àudio. Una vegada feta la selecció, el receptor farà el desmultiplexatge dels paquets de transport i els muntarà en paquets PES abans de lliurar-los als descodificadors d'àudio i vídeo. Per a poder sincronitzar la presentació dels ES audiovisuals, en la PMT també

s'indica quin és el PID dels paquets de transport que portaran periòdicament les referències temporals del PCR.

- **Taula d'accés condicional** (*conditional access table*, CAT). Aquesta taula controla la possibilitat de descodificació d'alguns programes de l'*stream* de transport (TS). En altres paraules, permet el xifratge del senyal. Per tant, s'haurà de disposar d'una clau d'accés per a poder accedir als continguts dels programes referenciats en aquesta taula. Aquesta taula disposa d'un identificador per als seus paquets de transport de valor PID = 1.
- **Taula d'informació de la xarxa de comunicacions** (*network information table*, NIT). Aquesta taula transmet la informació relativa a la ubicació física a la xarxa de comunicacions del múltiplex de transport. Proporciona informació específica de la xarxa com freqüències i nombre de canals utilitzats. La informació s'envia en paquets de transport amb identificador PID = 16.
- **Taula de data i hora** (*time and date table*, TDT). Aquesta taula s'utilitza per a sincronitzar el rellotge del sistema del descodificador. Es transporta en paquets de transport amb identificador PID = 20.

Aquestes cinc taules s'envien en forma de paquets de transport (TSP) i es multiplexen amb els paquets de transport que porten la codificació d'àudio i vídeo dels programes. En la figura 17 es mostra un exemple de multiplexatge de diversos programes en un *stream* de transport. En aquest cas, la taula d'assignació de programes PAT especifica que hi ha més de 28 programes multiplexats simultàniament i per a cada un es defineix una referència o PID que indica on es detallaran els continguts (o *elementary streams*) que componen aquest programa. El descodificador utilitzarà aquesta referència per a accedir a aquesta informació analitzant els paquets de transport TSP l'identificador dels quals tingui per valor el PID associat al programa. Per exemple, en la figura es detalla que per al programa 1 s'utilitzaran paquets senyalitzats amb el PID 124, i per al 2 amb el PID 130.

Els TSP identificats amb el PID 124 i PID 130 contindran la informació detallada de quins continguts conformen el programa 1 i 2 respectivament. Com es mostra en la figura, aquesta informació està conformada segons una taula de mapatge de programes PMT que s'envia en un TSP per a cada un.

Els paquets de transport que contenen la informació d'una PMT també indiquen en quins paquets TSP s'inclouran les marques temporals del PCR que faciliten la descodificació i presentació sincronitzada dels components del programa. Com s'ha mostrat en la figura 15, els paquets TSP que porten les dades del rellotge PCR duen a terme una extensió de la capçalera amb un camp d'adaptació (*adaptation field*) que transporta aquesta senyalització. En la PMT s'indica el PID d'aquests paquets perquè el receptor els utilitzi com a font de

sincronisme per a la descodificació i presentació dels *streams* elementals (ES) del programa. Habitualment, el PID que transporta el PCR d'un programa coincideix amb l'*stream* de vídeo principal del programa.

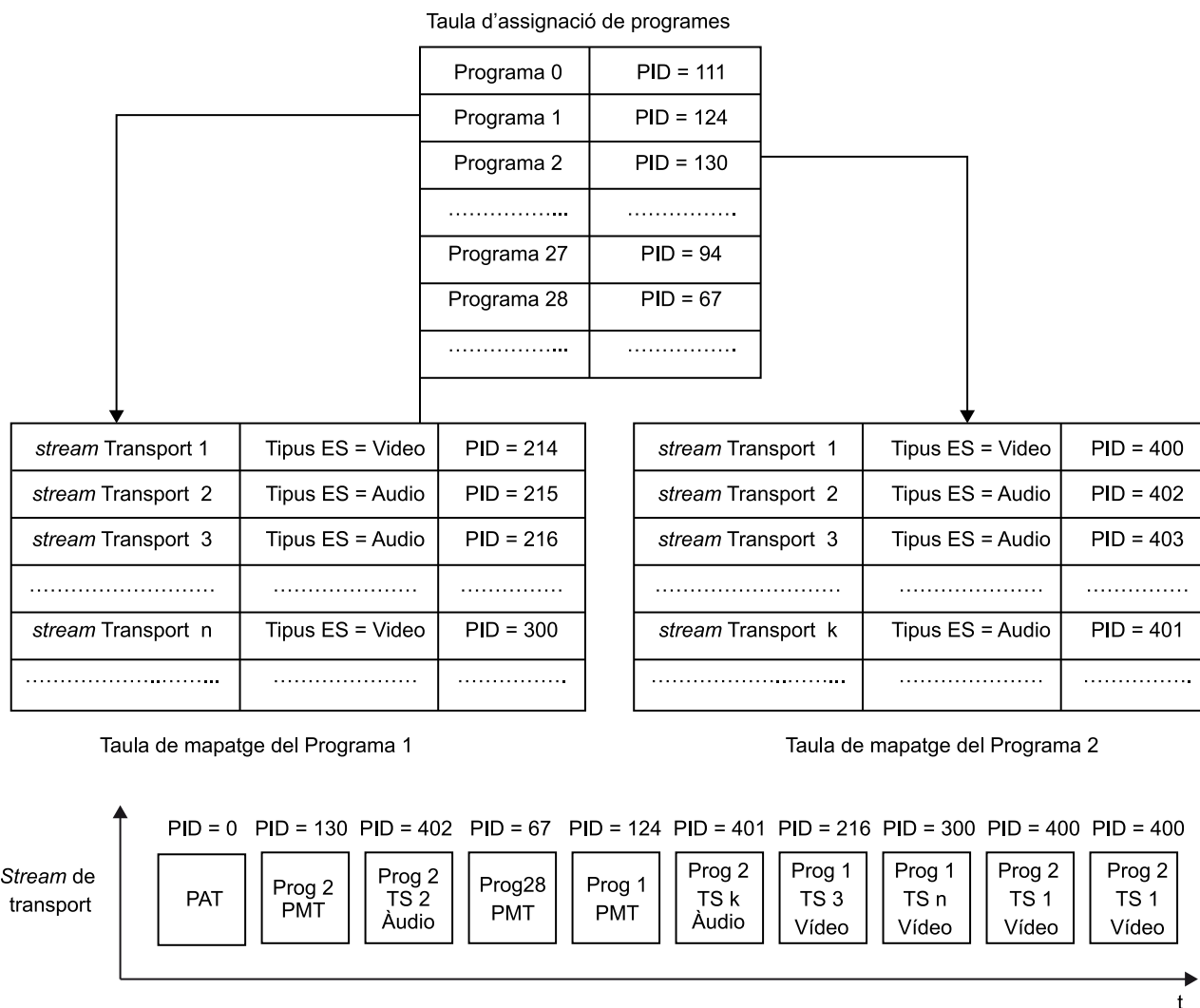


Figura 17. Exemple de taules i multiplexatge en un sistema amb múltiples programes

Per a accedir a un conjunt de *streams* elementals que formen part d'un mateix programa, el descodificador selecciona els paquets TSP de l'*stream* de transport que contenen els PID seleccionats d'un programa. La selecció dependrà de la configuració especificada per l'usuari.

Exemple

Un usuari escollirà del programa 1 un determinat angle de vista per a un pla en el vídeo (per exemple, PID = 300), i un idioma per a l'àudio (per exemple, el TSP amb PID = 215 podria ser àudio en espanyol i el TSP PID = 216 podria ser àudio en anglès).

Així doncs, a grans trets, els passos del descodificador a partir d'una selecció feta per l'usuari serien el següents:

- Buscar el PID = 0, que correspon a la taula PAT.
- Analitzar el PAT per a seleccionar el programa
- Buscar el PID de la PMT.

- Analitzar la PMT.
- Filtrar els paquets TSP del programa associats als *streams* elementals (ES) de vídeo i àudio seleccionats.
- Muntar els paquets de transport TSP en paquets PES per a lliurar-los al descodificador de vídeo o àudio.
- Descodificar cadascun dels fluxos elementals (ES).
- Reproduir sincronitzadament la informació audiovisual utilitzant el PCR.

A més de les taules de PSI, l'MPEG-2 defineix una estructura extensible que permet la transmissió de dades privades. Aquesta transmissió privada s'empra per a distribuir informació addicional per a la navegació sobre els programes, subtítols, dades addicionals per a l'accés condicional relacionat amb la distribució de claus de xifratge, actualitzacions de programari per als receptors, fluxos de dades encapsulades en IP, etc.

4.3. *Scrambling* i accés condicional

El DVB admet que qualsevol programa pugui disposar d'accés condicional, per exemple es requereix el pagament per visió (*pay per view*, PPV).

El conjunt de programes que es transmeten xifrats dins de l'*stream* de transport apareix en la taula d'accés condicional CAT del sistema d'informació. La taula CAT conté una llista d'identificadors PID iguals que els de la taula d'assignació PAT, que indiquen els programes la distribució dels quals està xifrada. Les tècniques específiques de xifratge i altres detalls es poden especificar en la CAT.

L'accés condicional als programes i continguts genèrics de l'*stream* de transport es basa en l'aleatorització dels paquets d'informació mitjançant un xifratge en bloc. Aquesta funcionalitat s'especifica per a les aplicacions comunes de DVB de manera que es minimitzi la probabilitat de pirateria. Els detalls tècnics de l'algorisme de xifratge aplicat només són accessibles a desenvolupadors després de signar clàusules de confidencialitat.

L'algorisme de xifratge es basa en dos xifratges de bloc i utilitza una clau de 128 bits:

- *Advanced encryption standard* (AES128), de seguretat provada i d'ús extensiu.
- *Extended emulation resistant cipher* (XRC), desenvolupat explícitament per l'organització DVB.

El procés de xifratge es pot aplicar sobre els paquets de transport TSP o sobre els paquets PES, però no en tots dos nivells d'empaquetatge simultàniament. Per a això, les capçaleres dels paquets PES i dels paquets de transport TSP indiquen

si els paquets estan xifrats i, si ho estan, amb quina clau. En la pràctica es manegen dues claus simultàniament per a un programa amb la finalitat de dificultar-ne la pirateria.

L'accés condicional (CA) en si mateix no es defineix per la norma, ja que la majoria dels operadors han preferit utilitzar la seva pròpia tecnologia. L'estàndard DVB permet definir el control d'accés utilitzat en un programa mitjançant la taula d'accés (CAT). A fi d'evitar que l'abonat necessiti diversos descodificadors, es proporcionen targetes per a l'accés condicional que s'insereixen en les diferents ranures del descodificador. Els estàndards que segueixen les operadores en DVB són el simulcript i el multcript.

5. Codificació de canal

Una vegada ja es té el *transport stream* (en SPTS o MPTS), s'ha d'enviar per un canal per satèl·lit o terrestre o per cable. Per a això es defineixen unes característiques comunes per als tres canals, algunes són les següents:

- La codificació de vídeo és MPEG-2 (MP@ML).
- Multiplexatge de *streams*, és a dir, *transport stream*.
- Codis protectors d'errors basats en Reed-Solomon.
- Ampliació de les taules PSI. Això dóna lloc a DVB-SI (*service information*).

Per a cada canal variarà:

- La millor o pitjor protecció contra els errors.
- Modulació.

Ítem	Satèl·lit (DVB-S)	Cable (DVB-C)	Terrestre (DVB-T)
Amplada de banda	30 MHz	8 MHz	8 MHz
Potència/soroll	Poca potència, canal sorollós	Molta potència, canal poc sorollós	Canal sorollós amb trajectòria multicamí
Modulació	QPSK → molt robusta, poc eficient	QAM → poc robusta, molt eficient	COFDM → molt robusta, molt eficient

Comparativa entre mitjans de transmissió segons l'estàndard DVB

La codificació de canal aplicada sobre el *transport stream* (SPTS o MPTS), dependrà del mitjà de transmissió que s'utilitzi. La codificació de canal consisteix a afegir un conjunt de dígit binaris al *transport stream* que facilitin la detecció i correcció d'errors en el receptor. El conjunt de dígit binaris afegits constitueix un valor calculat redundat que consumeix capacitat del canal de comunicacions. Per això, la codificació de canal determina de manera eficient la quantitat de redundància que s'ha d'introduir en funció dels errors de transmissió que es generin en el canal. La codificació del canal també indica com s'ha d'introduir aquesta redundància per a aconseguir-ne la màxima capacitat correctora.

Hi ha tres blocs de correcció comuns en DVB per a la transmissió sobre els canals satèl·lit, terrestre i cable (aleatorització, Reed-Solomon i entrellaçament de Fourny). En l'estàndard DVB, a més, s'inclouen uns altres dos blocs específics (convolucional i perforat) per als canals amb més errors, que són el satèl·lit i el terrestre.

En la figura 18 es representa el conjunt de blocs que duen a terme les tasques de codificació de canal que es van enllaçant des de l'entrada del *transport stream* fins al lliurament del flux codificat a les etapes de modulació. Com podem veure, cada mitjà té els seus propis blocs de codificació i modulació.

Nota

En aquest mòdul expliquem els tres blocs comuns de codificació de canal i els blocs específics de codificació de canal compartits pels canals satèl·lit i terrestre.

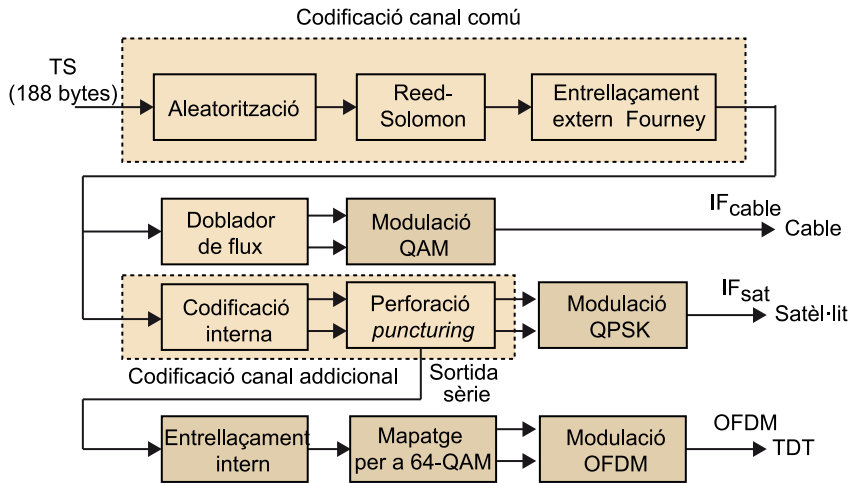


Figura 18. Diagrama de blocs de la codificació i formació dels paquets de transport en DVB-S, DVB-T i DVB-C

5.1. Aleatorització

El flux de bits generat per l'*stream* de transport (SPTS o MPTS) pot presentar un comportament poc aleatori, la qual cosa es tradueix en el fet que les probabilitats d'aparició de bits amb valor 0 i 1 poden ser significativament diferents o que seqüències de bits amb determinats valors de 0 i 1 no apareguin amb la freqüència esperada. Aquestes característiques d'aleatorietat són unes de les condicions de disseny especificades per a les etapes de modulació del nivell físic.

En el nivell físic es parteix de la hipòtesi que tots els punts d'una constel·lació es visitaran amb la mateixa probabilitat i de manera incorrelada, la qual cosa comporta que el senyal transmès no tindrà un component de contínua i ocuparà la màxima amplada de banda possible. D'aquesta manera, l'energia del senyal estarà repartida entre totes les freqüències menys la de contínua.

Per a garantir que aquesta hipòtesi de disseny sempre és certa, tots els sistemes de difusió DVB incorporen un aleatoritzador. Les condicions d'aleatorietat donaran com a resultat una distribució permanent de l'energia en tota la banda de freqüències ocupada per la transmissió del múltiplex de transport.

La implementació d'un aleatoritzador es basa en la generació d'una seqüència pseudoaleatòria mitjançant un registre de desplaçament de bits realimentat linealment, *linear feedback shift register* (LFSR).

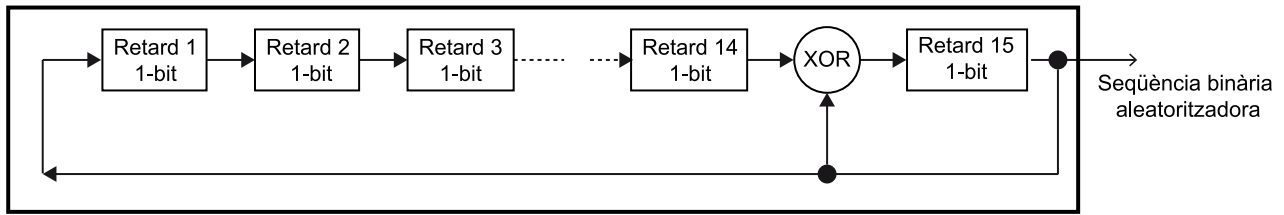
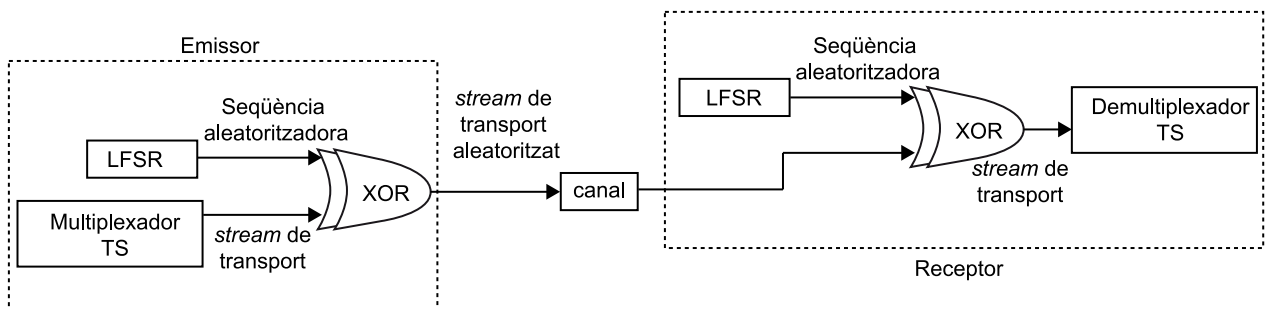


Figura 19. Registre LFSR de 15 elements

L'LFSR es dissenya perquè el període del senyal pseudoaleatori sigui màxim, triant una realimentació que ho garanteixi. El període màxim de la seqüència binària queda determinat per totes les combinacions diferents dels estats dels registres de desplaçament que no siguin nul·les. Així, un LFSR de 15 registres binaris amb la realimentació lineal adequada dóna lloc a una seqüència pseudoaleatòria amb període de 32.767 bits.

En l'emissor es genera una seqüència binària per mitjà d'un LFSR i es fa una combinació bit a bit amb l'*stream* de transport obtingut a la sortida del multiplexor. Aquesta combinació té lloc mitjançant una simple operació OR exclusiva com es mostra en la figura següent:

Figura 20. Esquema d'operació de l'aleatoritzador i desaleatoritzador per a la transmissió aleatoritzada de l'*stream* de transport

Per a procedir a la desaleatorització en el receptor es necessita un LFSR que apliqui sincronitzadament la mateixa seqüència aleatòria sobre el flux binari rebut. L'operació feta en el receptor torna a ser una OR exclusiva bit a bit entre el flux binari rebut i l'LFSR local sincronitzat. S'ha de tenir en compte que l'operació XOR és associativa i que en binari es verifica que:

$$(A \text{ XOR } B) \text{ XOR } B = A \quad (4)$$

El DVB especifica un aleatoritzador de 15 registres fàcil d'implementar i amb un cost computacional reduït en aritmètica sencera de 16 bits. Aquest dimensionament de l'LFSR permetria aleatoritzar completament fins a 21 paquets de transport de 188 octets. No obstant això, l'aleatorització només es fa en 8 paquets consecutius perquè es vol facilitar una resincronització ràpida de l'aleatoritzador davant esvaniments del senyal en el receptor. Per a això, la seqüència aleatoritzada incorpora punts de sincronisme en la capçalera dels paquets de l'*stream* de transport de manera regular.

5.2. Codificació de bloc Reed-Solomon

La codificació de font feta sobre els continguts audiovisuals extreu gran part de la redundància dels continguts i genera un flux binari amb una eficiència de codificació elevada. L'alt grau de compressió obtingut permet transferir els continguts audiovisuals amb una amplada de banda mínima. No obstant això, si els canals de transmissió introdueixen errors o pèrdues en el flux transmès, la descodificació dona lloc a notables distorsions en la reproducció dels continguts audiovisuals. En la pràctica, la codificació de font feta requereix que el receptor disposi d'un *stream* de transport lliure d'errors.

Per a poder corregir els errors i pèrdues que s'introdueixen durant la transmissió del flux binari s'ha d'incloure una redundància calculada mitjançant tècniques de codificació de canal. Aquesta redundància són dígit binaris addicionals que s'afegeixen a l'*stream* de transport aleatoritzat en blocs d'informació o intercalats amb els bits de dades. En el primer cas parlem de **codificació de canal en bloc** i, en el segon, de **codificació de canal convolucional o contínua**.

En la cadena de transmissió, primer s'aplica la codificació en bloc, ja que en el disseny de la codificació de canal es té en compte la pròpia estructuració de les dades generades. En aquest cas, l'estructura dels paquets de transport aleatoritzats definits en els apartats anteriors. Posteriorment veurem que en DVB de televisió terrestre i per satèl·lit també s'aplica la codificació de canal contínua.

La codificació de canal en bloc calcula un conjunt de dígit binaris, o redundància, que s'afegeix al final de cada paquet de transport. Aquesta redundància constitueix un *trailer* per al paquet. En DVB la redundància calculada per a cada paquet de transport és de 16 octets, per la qual cosa el paquet de transport amb la seva redundància assoleix una mida de 204 octets. La tècnica de codificació en bloc emprada en DVB es basa en la generació de codis de canal proposada per Reed i Solomon.

La codificació de Reed-Solomon opera amb agrupacions de M bits que defineixen un símbol de codificació. Per al cas particular d'una codificació binària M seria 1. En la codificació Reed-Solomon del DVB s'utilitza $M = 8$, per tant, el símbol emprat en DVB per a la codificació bloc és d'un octet, la qual cosa simplifica la implementació en aritmètica sencera de les operacions a fer. En general, la codificació Reed-Solomon ha de complir que la longitud N dels blocs de símbols codificats ha de seguir la relació $N < 2^M$, la qual cosa implica que els blocs de dades amb els quals s'ha de detectar i corregir errors en DVB seran com a màxim de 255 octets, incloent-hi la redundància.

La longitud i el valor de la redundància depenen de la codificació Reed-Solomon disponible. Per a una mida de bloc de 255 octets la redundància necessària d'una codificació podria ser de 16 o 32 octets. La capacitat correctora

d'aquests codis sempre és la meitat de la redundància introduïda, de manera que depenent de la quantitat màxima d'errors que puguem esperar decidirem un cas o un altre. En DVB, com ja hem comentat, s'afegeix una redundància de 16 octets a cada paquet de transport. D'aquesta manera, es pot corregir qualsevol combinació de 8 o menys octets erronis que es produeixin durant la transmissió del paquet i el seu *trailer*.

16 octets sobre 204 és una sobrecàrrega de codificació molt reduïda per als resultats que es poden obtenir en la correcció d'errors de transmissió.

5.3. Entrellaçament de Fourny

L'entrellaçament és una tècnica que varia l'ordre d'una seqüència de bits. L'ús clàssic de l'entrellaçador és per a separar els errors introduïts durant la transmissió del senyal al canal de comunicacions. Això es deu al fet que la transmissió als canals de comunicacions sol introduir errors en ràfega, o consecutius, a causa dels esvaniments o interferències momentànies que pateix el senyal en propagar-se pel mitjà de transmissió.

Per a protegir-se de múltiples errors consecutius en els bits transmesos s'ha d'emprar una codificació de canal molt poc eficient, ja que requereix una gran redundància. Com a solució, la tècnica d'entrellaçament el que fa és desordenar els bits, de manera que, en tornar-se a ordenar en el receptor, els errors quedin prou repartits per a considerar-los independents entre ells.

L'entrellaçador desordena d'una manera predeterminada els símbols de canal, definits com a M bits consecutius en la codificació Reed-Solomon de *stream* de transport. El desentrellaçador du a terme l'operació contrària perquè els blocs arribin al descodificador de canal amb l'ordre original. Aquesta operació fa que la modificació del valor d'un conjunt consecutiu de bits transmès durant la transmissió no afecti un conjunt consecutiu de símbols de canal. En la pràctica, el que s'aconsegueix amb l'entrellaçament és distribuir els errors consecutius que podrien afectar un únic bloc entre dos blocs de codificació i, per tant, es redueix la quantitat d'errors que s'han de corregir en un únic bloc de codificació, la qual cosa permet introduir menys redundància en la codificació bloc perquè el nombre de símbols erronis que podem esperar en cada bloc és més petit.

Per a determinar fins a quant cal desordenar els símbols de codificació de M bits de les paraules codi abans de transmetre'ls, s'ha d'analitzar l'estadística de la longitud de les ràfegues d'errors. Tenint en compte la distribució de les longituds, determinarem quina és la profunditat de l'entrellaçament, magnitud que representarem amb D . En la pràctica, el valor de D serà superior o del mateix ordre de magnitud que tingui el valor mitjà de la longitud de la ràfega.

L'entrellaçament de Fourny es basa a anar distribuint els símbols de canal entre diferents registres de desplaçament. Cada registre de desplaçament emmagatzema símbols de canal, que en el cas de DVB són de 8 bits. El nombre de registres de desplaçament que s'utilitzen es correspon amb la profunditat de l'entrellaçament que s'emprarà. El desentrellaçament fa l'operació inversa, emprant també el mateix nombre de registres de desplaçament que es van alimentant amb els símbols de canal rebuts.

L'estat inicial dels registres de desplaçament és diferent per a cada registre emprat en l'entrellaçador i en el desentrellaçador. En l'entrellaçador, el primer registre de desplaçament està inicialment buit; el segon disposa d'un sol element d'inicialització; el tercer, de dos elements d'inicialització, i d'aquesta manera es continua fins al registre D , que disposarà de $D-1$ elements d'inicialització. La diferent precàrrega dels registres de desplaçament donarà lloc a la separació dels símbols que s'havien generat consecutivament a la sortida del codificador de canal.

El desentrellaçador disposa d'un estat inicial invers a l'entrellaçador. Inicialment, el primer registre disposa d'un conjunt de $D-1$ símbols; el segon, de $D-2$, i es continua reduint el nombre de símbols en una unitat fins a arribar al registre D , que no té cap element.

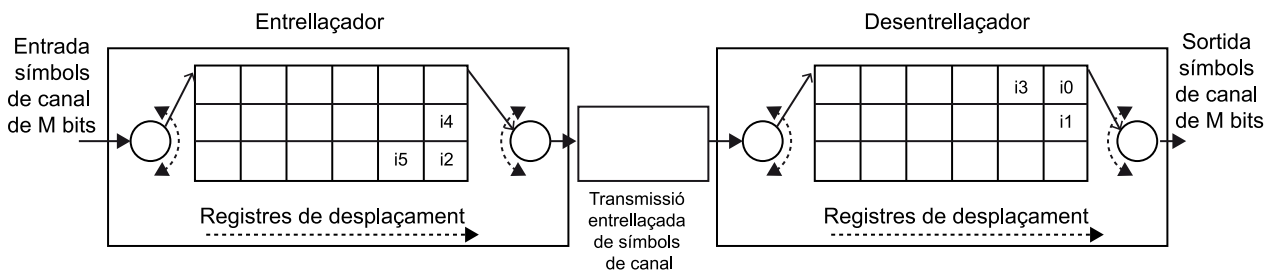


Figura 21. Exemple de l'estat inicial d'un entrellaçador i un desentrellaçador de Fourny per a $D = 3$

A partir de l'estat de precàrrega representat en la figura anterior, l'operació feta en cada grup de registres és escriure un símbol en un registre diferent cada vegada i, simultàniament, extreure un símbol en el mateix ordre d'escriptura. D'aquesta manera, el símbol que s'escriu en el primer registre de l'entrellaçador és immediatament retransmès cap al canal de comunicacions.

En el desentrellaçador es fa la mateixa operació. El primer símbol que arriba s'escriu en el primer registre, però la primera extracció en aquest cas és d'un valor d'inicialització. D'aquesta manera es compensen les esperes dels símbols en l'entrellaçador i el desentrellaçador, i s'aconsegueix la reordenació a la sortida del desentrellaçador.

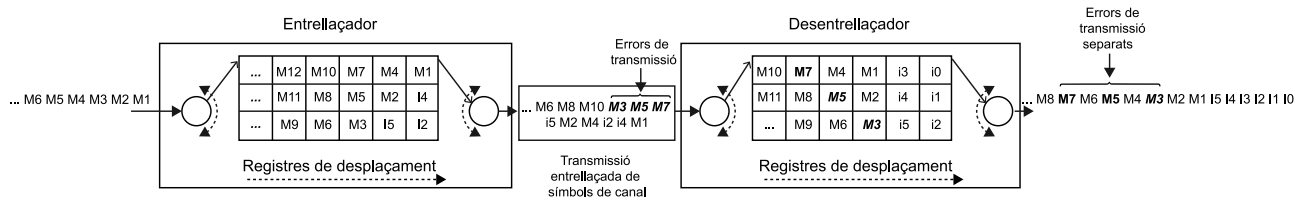


Figura 22. Exemple d'operació d'un entrellaçador i un desentrellaçador de Fourny per a $D = 3$

La figura anterior mostra el procés dut a terme durant la transmissió dels símbols M_i . Cada símbol s'escriu cada vegada en un registre diferent i , quan s'arriba a l'últim registre, es torna a alimentar el primer, per la qual cosa es fa una operació circular d'escriptura sobre els registres. A mesura que s'escriu, es llegeixen els primers elements que hi ha en els registres de desplaçament i s'envien al canal.

En el canal es poden produir errors en els bits que s'han transmès. En l'exemple de la figura considerem que es produeix una ràfega de 15 bits erronis que afecten als bits transmesos dels símbols de canal M_7 , M_5 i M_3 .

En desentrellaçador fa l'operació d'escriptura en els seus registres de la manera circular esmentada anteriorment. Així, els símbols erronis que arriben consecutivament al receptor se separen en registres diferents. La lectura del desentrellaçador retorna l'ordre inicial als símbols perquè el descodificador de canal els pugui corregir.

L'operació de l'entrellaçador i el desentrellaçador de Fourny dona lloc a un retard des que arriba un símbol M_i a l'entrellaçador fins que surt del desentrellaçador. El retard es correspon amb el nombre de símbols de precàrrega que tenen l'entrellaçador i el desentrellaçador, el valor dels quals és de $D \cdot (D-1)$ símbols. En l'exemple veiem que l'operació de l'entrellaçament té una sortida encapçalada pels sis símbols d'inicialització que determinen el retard introduït per aquesta operació i que són descartats pel receptor abans de procedir amb la descodificació de canal.

L'entrellaçament de Fourny especificat pel DVB disposa d'una profunditat d'entrellaçador $D = 12$, la qual cosa comporta que les ràfegues d'errors de bits previstes seran d'entorn de les desenes de bits.

5.4. Codificació convolucional o contínua i perforació

L'estàndard DVB afegeix més redundància específicament per als canals amb una taxa d'errors més elevada: per als mitjans de transmissió per satèl·lit i terrestres s'incorpora una codificació contínua amb un grau de protecció elevat que resulta molt eficaç per a fer front als errors aleatoris introduïts pel canal durant la transmissió de les dades. No obstant això, aquesta codificació no és gaire efectiva contra grans ràfegues consecutives d'errors degudes a interferències cocanal, interferències destructives multicamí o sorolls impulsius. Per

això, com ja hem comentat, prèviament la codificació de canal en DVB du a terme una codificació de bloc Reed-Solomon i un entrellaçament de Fourny que faciliten la correcció de les ràfegues d'errors.

A diferència dels codis de bloc, els codis continus operen directament amb el flux de bits que, en aquest cas, s'obté de la sortida de l'entrellaçador de Fourny. El codificador continu treballa en paral·lel amb un conjunt reduït de bits, fent operacions elementals basades en retards i OR exclusives. En el cas de DVB el codificador genera 2 bits de sortida per cada bit d'entrada.

La codificació contínua dissenyada depèn dels valors que hagin anat arribant al codificador. El codificador incorpora un conjunt de registres de desplaçament per anar emmagatzemant aquests valors entrants de manera que la combinació de l'entrada en un instant, i els seus valors en instants passats, determinarà els valors de la sortida en aquest instant. Les combinacions aplicades sobre els valors entrants es dissenyen heurísticament.

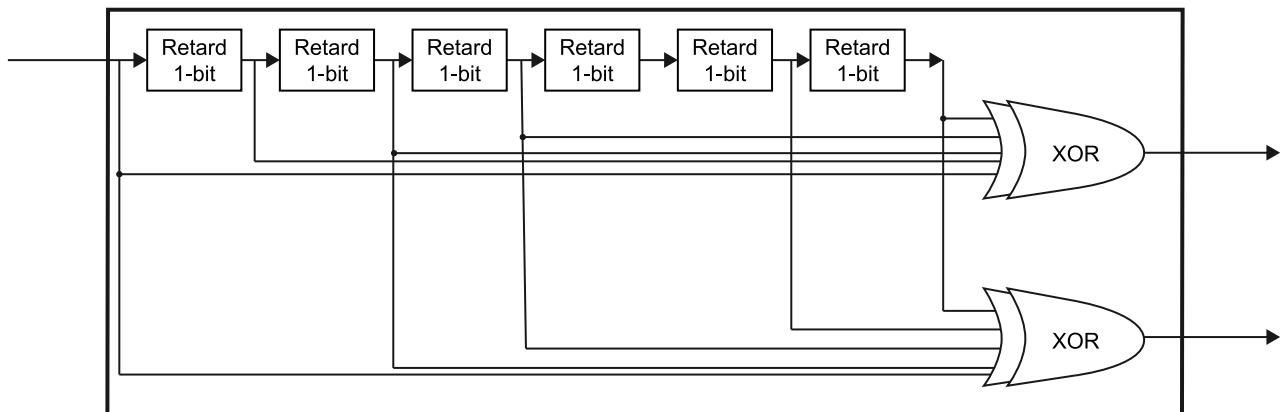


Figura 23. Codificador continu de DVB-T i DVB-S

La descodificació es du a terme utilitzant l'algorisme de Viterbi, el propòsit del qual és reduir el nombre d'operacions per a determinar quina és la seqüència més versemblant respecte a la rebuda en el receptor. Si es fes de manera exhaustiva, s'hauria de trobar la seqüència vàlida, o generable pel codificador, que discrepa en menys bits amb la rebuda. L'algorisme de Viterbi evita fer comparacions innecessàries accelerant el procés de decisió.

Atès que la redundància introduïda pel codificador és molt elevada, 2/1, 2 bits de sortida per cada 1 d'entrada, el DVB va proposar de manera opcional fer una perforació del codificador continu. La perforació és un procés de delmació controlada dels bits a la sortida del codificador continu. La perforació implica no transmetre tots els bits de la sortida del codificador continu.

Exemple

Per 4 bits d'entrada, el codificador continu genera 8 bits a la sortida. Si d'aquests 8 bits només se'n transmeten 6, el perforador haurà delmat en 2 bits la sortida del codificador continu.

En aquest exemple, la relació entre bits de sortida del perforador respecte als bits d'entrada en el codificador continu és de 6/4.

Les possibilitats de perforació ofertes a les operadores de DVB són 2/1 (no perforada), 6/4, 4/3, 6/5 i 8/7. Per a la codificació perforada triada haurem de considerar que la taxa de codificació resultant es derivarà inversament de la relació entre bits de sortida del perforador i bits d'entrada en el codificador continu.

Exemple

Per a una relació de bits de sortida de 6 per 4 d'entrada la taxa del codificador serà 2/3, la qual cosa implica 1 bit de redundància per cada 3 bits transmesos (33% de redundància).

L'elecció d'una taxa de codificació dependrà, en última instància, de la potència que rebí un receptor en la zona de cobertura: com menys potència més redundància.

La perforació millorarà l'aprofitament de l'amplada de banda disponible a costa de reduir la capacitat de correcció d'errors.

