

XAMU – Xarxes Multimèdia - EPSEVG-UPC

Sessió Pràctica Unitat 4, Primavera 2024

Estudi Previ de Compressió de Vídeo

Ernest Anguera Aixalà
Nawal Bouallala Safyoun
Mariona Farré Tapias
Iván García Rodríguez

Índex

Índex.....	2
Exercici 1: Anàlisi d'un fitxer de vídeo.....	3
Apartat a.....	3
Apartat b.....	3
Exercici 2: Anàlisi d'arxius.....	5
Exercici 3: Estadístiques d'ExpertDVD.....	13
Apartat a.....	13
Apartat b.....	14
Apartat c.....	15
Apartat d.....	16
Exercici 4: FFmpeg.....	17

Exercici 1: Anàlisi d'un fitxer de vídeo

El fitxer aqua.avi conté un flux de vídeo i un altre d'àudio, sense comprimir.
L'utilitzarem durant aquesta sessió.

Apartat a

a) Visualitzeu-lo i comenteu les seves característiques respecte a freqüències visuals i moviments.

És un vídeo amb imatge i audio. La qualitat de les imatges no és molt bona, resalten els colors foscos i hi ha molts canvis d'escena. També cal resaltar que hi ha un efecte de difuminació especialment en objectes de fons, com ara l'aigua. En general en les escenes no hi ha alta freqüència, ja que hi ha canvis bruscos de colors, excepte en les escenes en la que apareix la noia, ja que hi ha un canvi brusc d'intensitat de llum, en la primera escena que apareix la noia podríem dir que hi ha un canvi en les zones verticals mentre que en la última escena hi ha més canvis en la zona horitzontal. Per tant en les escenes, en la que apareix la noia, hi ha alta freqüència, mentres que en les altres hi ha baixa freqüència.

La primera escena i els canvis d'escena son del tipus intra, mentres que les altres son del tipus inter.

Finalment cal resaltar que les imatges canvien molt ràpid.

Apartat b

b) Analitzeu el fitxer amb MediaInfo en mode text (mireu la Fig. 8 com exemple).

a. Sobre la pista de vídeo: codec, frame rate, resolució, profunditat de bit (bit depth) i bits/(pixel x frame). A quina qualitat típica es pot comparar?

Codec: 1, PCM

Frame Rate: 29.970 FPS (frame/s)

Resolució: 352 x 240 pixels

Profunditat de bit: 8 bits

Bits/pixel x frame: 24 bits

Amb els resultats obtinguts, podem comprar-ho amb un vídeo de VCD, ja que 29.970FPS es un frame rate estandar i 352*240 pixels es resolució baixa típica dels vídeo VCD

b. Justifiqueu numèricament el bitrate (bits/s) i mida (bytes) de la pista de vídeo, i compareu el vostre càlcul amb el que reporta MediaInfo.

Bitrate: $352 * 240 \text{ pixels} * 24 \text{ bits/pixel x frame} * 29.97 \text{ frame/s} = 60764774.4 \text{ bits/s} = 60.76 \text{ Mbits/s}$

Duració: 24,959s

Mida: $60764774.4 * 24.959 = 15166.280.042,496 \text{ bits} = 1895785005,312 \text{ bytes} = 180.796 \text{ MiB}$

MediaInfo:

Bitrate: 60.8 Mb/s

Mida: 181 MiB

Comentaris: Si arrondim tant el bitrate com la mida veiem que ens dona el mateix resultat

c. Sobre la pista d'àudio: codec, canals, freqüència de mostreig, bits/mostra A quina qualitat típica es pot comparar?

Codec:PCM

Canal:2

Freqüència de mostreig:44.1Khz

Bits/mostra:16 bits(8*2)

Amb aquests resultats podem dir que la qualitat típica es pot comparar amb un CD

d. Justifiqueu numèricament el bitrate (bits/s) i mida (bytes) de la pista d'àudio, i compareu el vostre càlcul amb el que reporta MediaInfo.

Bitrate: $44100 \text{ mostres/s} * 16 \text{ bits} * 2 = 1411200 \text{ bits/s} = 1411.2 \text{ Kb/s}$

Duració:24,959s

Mida: $1411200 * 24.959 = 35222140,8 \text{ bits} = 4402767,6 \text{ bytes} = 4.2 \text{ MiB}$

MediaInfo:

Bitrate: 1 411.2 kb/s

Mida:4.20 MiB

Comentaris:Com veiem ens dona el mateix resultat.

Exercici 2: Anàlisi d'arxius

Amb StreamEye Analyzer i MedialInfo, i per a cada un dels següents arxius:

- Max_Headroom.mpg
- tdt.m2t
- CableLabsHD.mpg
- Fixed2.mpg
- Trailer_Brave.mp4
- Trailer_Simpsons.mp4

Comenteu, com a mínim, els següents paràmetres (però discutiu també qualsevol altre paràmetre que considereu rellevant).

a) Còdec i contenidor.

b) Resolució, imatges/s, entrellaçat – i expresseu-ho en el format habitual (per exemple 576i25).

c) Paràmetres N i M, si són periòdics. Per què no ho són pel cas MPEG-4?

d) Perfil de taxa de bits (forma del gràfic bytes/imatge). Discutiu-ho i relacioneu-ho amb N i M (si és aplicable). Comenteu també si és CBR, VBR que és CBR en mitja, o VBR pur.

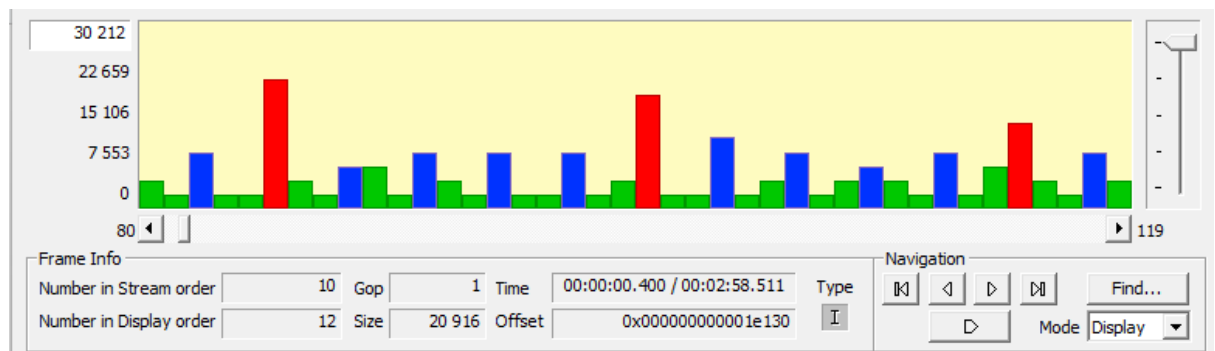
e) Perfil i nivell (recordeu les notes de classe i Wikipedia), i relacioneu-ho amb el que heu trobat a), b), c) i d), i el tipus de servei de vídeo. Per exemple: "És un 720p60 VBR MPEG-2, que sembla plausible que correspongui a un servei de televisió digital als Estats Units"

f) Només per als casos MPEG-1/2, a partir dels paràmetres N i M i les mides (en mitja) de cada tipus d'imatge (I, P, B) proporcionat per StreamEye, calculeu el bitrate mig, i compareu-lo amb el que reporten MedialInfo i StreamEye.

g) Pel cas CableLabsHD, busqueu un exemple de cada tipus de macrobloc, i discutir el valor dels seus paràmetres (reportat per StreamEye), depenent del tipus d'imatge i el contingut de la imatge.

Max_Headroom.mpg

- a) Contenedor → MPEG-PS
Còdec → MPEG-1 Video
- b) Resolució → 240p29,97
- c) M = 3, N = 15
- d)



Podem veure que cada 3 (M) macroblocs trobem un macrobloc forward i cada 15 macroblocs (N) trobem un macrobloc intra.

CBR a 1152 Kb/s

- e) És un 240p29,97 CBR MPEG-1, que sembla plausible que correspongui a arxius multimedia antics.
- f)

Info...	
File Picture Headers MB	
file type	MPEG2 (SS)
video stream type	MPEG1
resolution	352x240
profile:level	unspecified
aspect ratio	unspecified
interlaced	no
frames count	5 351
frame size max	30 212
avg	5 865
avg/max (I)	14 442 / 30 212
avg/max (P)	7 595 / 18 592
avg/max (B)	4 311 / 16 268
min	0
file size	31 457 280

framerate declared	29.97

real	29.97

bitrate declared	1 151 600

real max	1 541 896
real avg	1 409 309
real min	1 225 173

Sabem que per cada imatge I hi ha 4 imatges P i 10 imatges B. Per obtenir la mitja de bytes farem:

$14442 * 1/15 + 7595 * 4/15 + 4311 * 10/15 = 5862.133 \rightarrow 5862 \text{ bytes/frame aprox}$

$5862 * 8 = 46896 \text{ bits/frame}$

$46896 * 29,97 = 1405473,12 \text{ bits/s} \rightarrow 1405,47 \text{ Kb/s}$

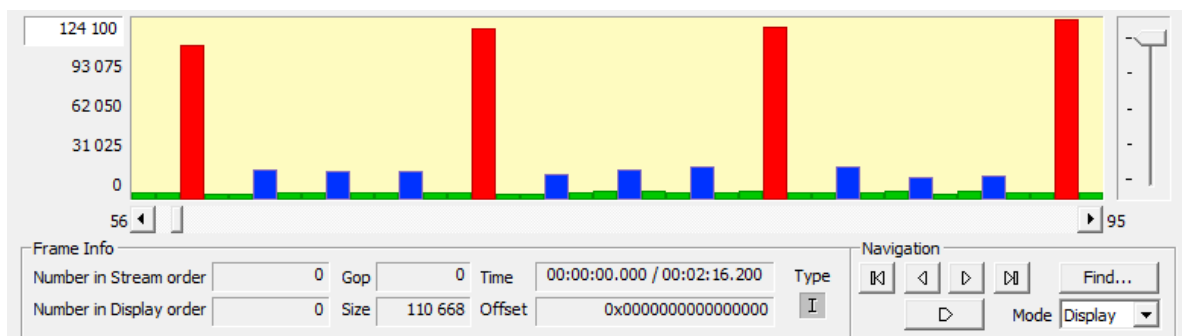
Mediainfo $\rightarrow 1409 \text{ Kb/s}$

StreamEye Analyzer $\rightarrow 1409 \text{ Kb/s}$

Com podem veure el resultat obtingut és molt similar al que obtenim mitjançant Mediainfo i StreamEyeAnalyzer.

tdt.m2t

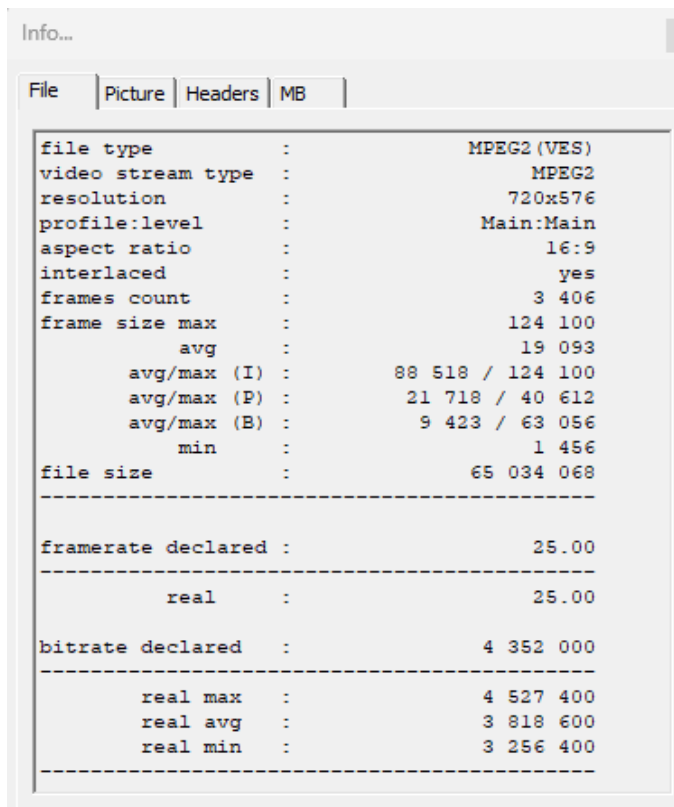
- a) Contenedor \rightarrow MPEG-2
Còdec \rightarrow MPEG-2 Video
- b) Resolució \rightarrow 576i25
- c) M = 3, N = 12
- d)



Podem veure que cada 3 (M) macroblocs trobem un macrobloc forward i cada 12 macroblocs (N) trobem un macrobloc intra.

VBR a 3819 Kb/s de mitja

- e) És un 576i25 VBR MPEG-2, que sembla plausible que correspongui a sistemes de televisió tradicionals.
- f)



Sabem que per cada imatge I hi ha 3 imatges P i 8 imatges B. Per obtenir la mitja de bytes farem:

$$88518 * 1/12 + 21718 * 3/12 + 9423 * 8/12 = 19088 \text{ bytes/frame}$$

$$19088 * 8 = 152704 \text{ bits/frame}$$

$$152704 * 25 = 3817600 \rightarrow 3817,6 \text{ Kb/s}$$

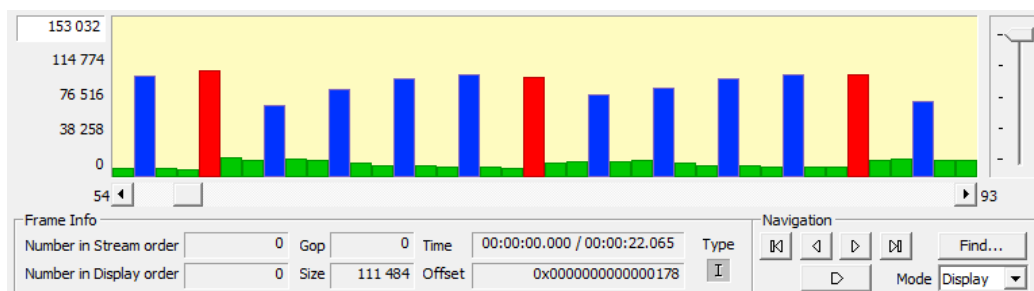
Mediainfo \rightarrow 3819 Kb/s

StreamEye Analyzer \rightarrow 3819 Kb/s

Com podem veure el resultat obtingut és molt similar al que obtenim mitjançant Mediainfo i StreamEyeAnalyzer.

CableLabsHD.mpg

- Contenidor \rightarrow MPEG-TS
Còdec \rightarrow MPEG-2 Video
- Resolució \rightarrow 720p59,94
- M = 3, N = 15
-

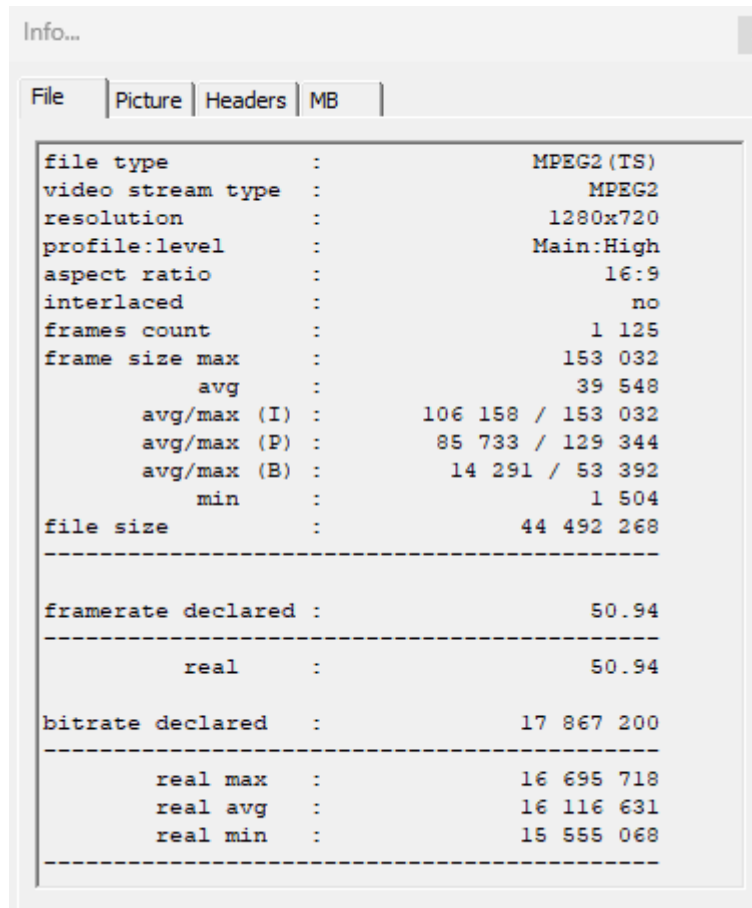


Podem veure que cada 3 (M) macroblocs trobem un macrobloc forward i cada 15 macroblocs (N) trobem un macrobloc intra.

CBR → 17,9 Mb/s

e) És un 720p59,94 CBR MPEG-2, que sembla plausible que correspongui a televisió digital.

f)



Sabem que per cada imatge I hi ha 4 imatges P i 10 imatges B. Per obtenir la mitja de bytes farem:

$106158 * 1/15 + 85733 * 4/15 + 14291 * 10/15 = 39466.667 \rightarrow 39467$ bytes/frame

$39467 * 8 = 315736$ bits/frame

$315736 * 50,94 = 16083591,84 \rightarrow 16,08$ Mb/s

Mediainfo → 19 Mb/s

StreamEye Analyzer → 16,69 Mb/s

Com podem veure el resultat obtingut és molt similar al que obtenim mitjançant Mediainfo i StreamEyeAnalyzer.

g)

```

position      : 42x12 (672x192) [1002]
mb_type       : Intra
quantizer     : 6
size mb (in bits) : 229
  mode        : 2
  quant       : 0
  mv          : 0
  residual    : 227
cbp           : 1111 11 (63)
motion type   : MC_ZERO

```

```

position      : 42x9 (672x144) [762]
mb_type       : Inter Pattern Backward
quantizer     : 8
size mb (in bits) : 50
  mode        : 4
  quant       : 0
  mv          : 6
  residual    : 40
cbp           : 0011 00 (12)
motion type   : MC_FRAME
mv backward   : x:-001 y:0003

```

```

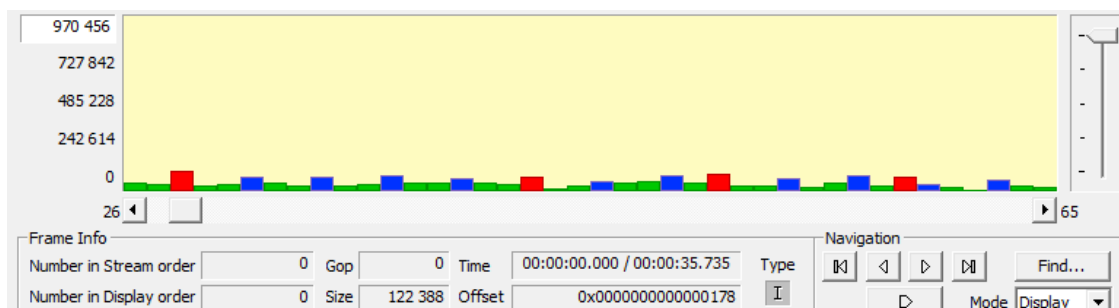
position      : 39x9 (624x144) [759]
mb_type       : Inter Pattern Forward
quantizer     : 6
size mb (in bits) : 76
  mode        : 2
  quant       : 0
  mv          : 2
  residual    : 72
cbp           : 1111 00 (60)
motion type   : MC_FRAME
mv forward    : x:0002 y:-001

```

Podem veure que la mida dels macroblocs intra és més gran que dels macroblocs forward i aquesta més gran que dels macroblocs backward. També podem veure que el macrobloc intra és MC_ZERO, en canvi els altres 2 són MC_FRAME

Fixed2.mpg

- Contenidor → MPEG-TS
Còdec → MPEG-2 Video
- Resolució → 1080i29,97
- M i N variables
- d)

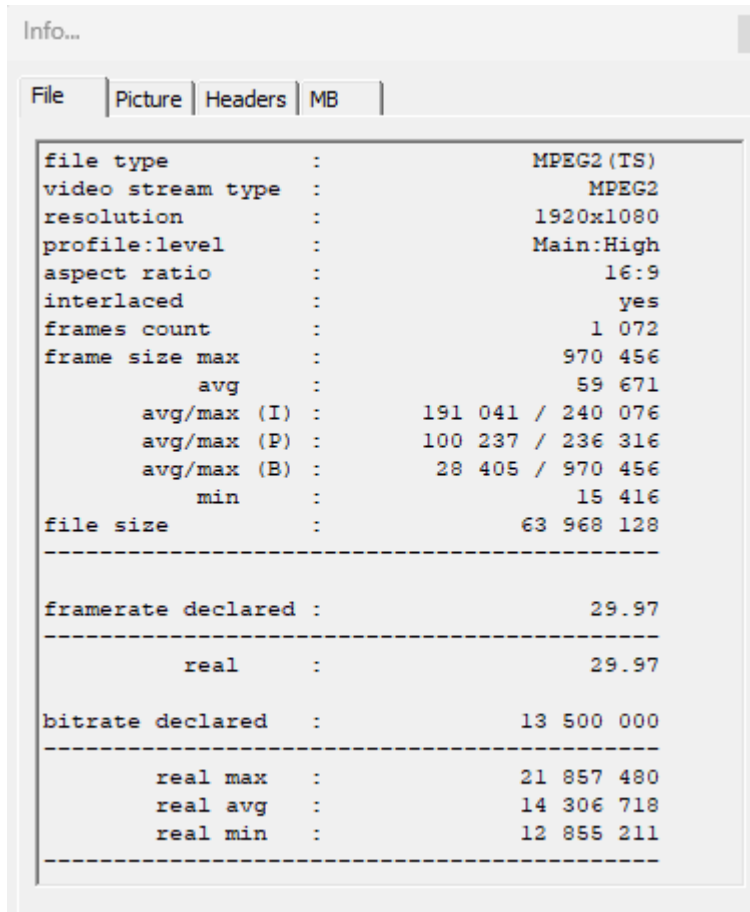


Podem comprovar que M i N són variables perquè al gràfic es veu que entre el 1r macrobloc intra i el segon hi ha 14 macrobloc i entre el 2n i el 3r i el 3r i 4t hi ha 7 macroblocs.

CBR → 13,2 Mb/s

e) És un 1080i29,97 CBR MPEG-2, que sembla plausible que correspongui a televisió digital d'alta definició.

f)



file type	:	MPEG2 (TS)
video stream type	:	MPEG2
resolution	:	1920x1080
profile:level	:	Main:High
aspect ratio	:	16:9
interlaced	:	yes
frames count	:	1 072
frame size max	:	970 456
avg	:	59 671
avg/max (I)	:	191 041 / 240 076
avg/max (P)	:	100 237 / 236 316
avg/max (B)	:	28 405 / 970 456
min	:	15 416
file size	:	63 968 128

framerate declared	:	29.97

real	:	29.97

bitrate declared	:	13 500 000

real max	:	21 857 480
real avg	:	14 306 718
real min	:	12 855 211

59671 bytes/frame

$59671 * 8 = 477368$ bits/frame

$477368 * 29,97 = 14306718,96 \rightarrow 14,3$ Mb/s

Mediainfo → 14,1 Mb/s

StreamEye Analyzer → 14,3 Mb/s

Com podem veure el resultat obtingut és molt similar al que obtenim mitjançant Mediainfo i StreamEyeAnalyzer.

Trailer Brave.mp4

a) Contenedor → MPEG-4

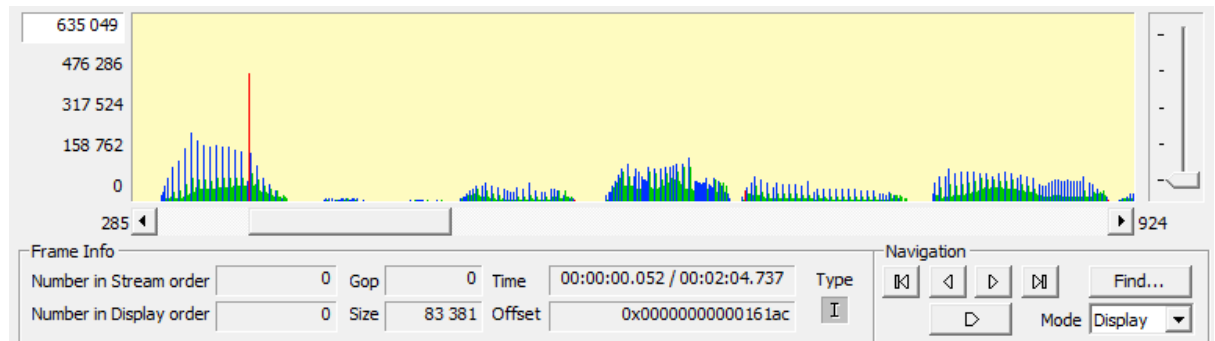
Còdec → AVC

b) Resolució → 800p23,952

c) En el cas de MPEG-4 M i N no són periòdics degut a que la codificació de MPEG-4 és més adaptativa i eficient que altres estàndards anteriors. Per maximitzar

l'eficiència de la codificació MPEG-4 pot canviar la freqüència dels quadres I i P per adaptar-se millor a les característiques del contingut del vídeo.

d)



Podem veure que M i N no són periòdics.

VBR → 5840 Kb/s

e) És un 800p23,952 VBR MPEG-4, que sembla plausible que correspongui a producció cinematogràfica.

Trailer_Simpsons.mp4

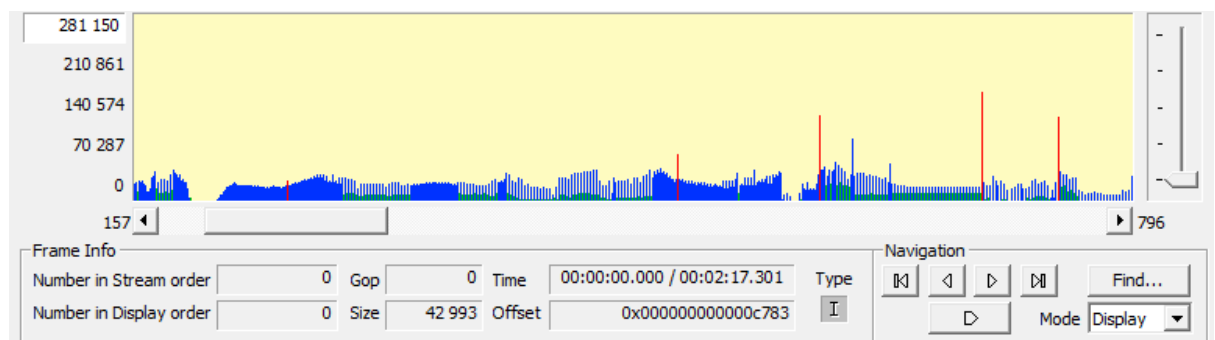
a) Contenedor → MPEG-4

Còdec → AVC

b) Resolució → 544p23,952

c) En el cas de MPEG-4 M i N no són periòdics degut a que la codificació de MPEG-4 és més adaptativa i eficient que altres estàndards anteriors. Per maximitzar l'eficiència de la codificació MPEG-4 pot canviar la freqüència dels quadres I i P per adaptar-se millor a les característiques del contingut del vídeo.

d)



Podem veure que M i N no són periòdics.

VBR → 4284 Kb/s

e) És un 544p23,952 VBR MPEG-4, que sembla plausible que correspongui a producció cinematogràfica.

Exercici 3: Estadístiques d'ExpertDVD

Estadístiques d'ExpertDVD Estudieu les estadístiques de codificació que s'inclouen a l'annex, que ExpertDVD produeix després de la codificació d'un arxiu MPEG. Responen:

Apartat a

a) Analitzeu els paràmetres bàsics: velocitat de bits, resolució, velocitat de fotogrames, submostreig de crominància. ¿Es pot afirmar si és europeu?

Donades les estadístiques de l'ExpertDVD, podem observar els paràmetres bàsics utilitzats:

- **Velocitat de Bits(Bitrate):** La velocitat dels Bits estàndar en l'aplicació és de 6000000 bits per segon (o 6 Mbps), una taxa molt típica amb estàndards europeus i internacionals o estatunidencs.
FrameRateCode:4, BitRate:6000000.0
- **Resolució:** la resolució proporcionada és de 720x480, que és típica de l'estàndard de televisió utilitzada a Amèrica, entre altres regions. No és la resolució 720x576 de l'estàndard europeu.
Hor Size:720 Ver Size:480 AspectCode:2
- **Fotogrames:** El valor dels fotogrames, no està especificat en cap lloc, però es pot calcular a partir de la informació proporcionada de les estadístiques.
BitRate:6000000.0
Avg Bits per Frame: 200200 bits
 $6000000.0 / 200200 = 29.97$ fps FrameRate
- **Submostreig de crominància:** El submostreig no està específicament proporcionat en els paràmetres, però si tenim els valors de Y, U i V per cada frame, podem observar que la crominància per un frame és:
Y: variance= 3804, MSE= 140 (26.7 dB), SNR= 14.3 dB
v:3804.067792 e:139.902008
U: variance= 496.1, MSE= 42.6 (31.8 dB), SNR= 10.7 dB
V: variance= 152.2, MSE= 26.7 (33.9 dB), SNR= 7.55 dB
Si ens fixem amb els valors de SNR de tots els frames, podem observar que hi ha un ratio de 4:2:0, un estàndard internacional i s'utilitza independentment de la regió.

Saben aquestes característiques, podem basant-nos en la combinació de la resolució de 720x576 i la velocitat de fotogrames de 6 Mbps i el framerate de 29.97 fps, és probable que aquest contingut de vídeo estigui optimitzat per a l'ús fora d'Europa, concretament per a regions que usen l'estàndard Americà. No hi ha indicis clars que aquest sigui un format amb els estàndards de transmissió europea, ja que aquesta segueixen les característiques d'una transmissió amb una resolució de 720x576 píxels i una velocitat de fotogrames de 25 fps.

Apartat b

b) Pel que fa a les estadístiques de cada imatge

- Identifiqueu els frames, el seu número de seqüència, i ordre en què apareixen les imatges (i justifiqueu-lo!)

En les estadístiques donades surten els frames:

- Frame 0: És un frame I (Intra-coded). És el frame inicial de la seqüència, utilitzat com a referència per als següents frames predictius.
- Frame 1: És un frame P (Predictive-coded). Utilitza dades del frame I anterior per a la seva codificació.
- Frame 2: És un frame B (Bidirectionally-predictive-coded). Utilitza les dades dels frames I i P anteriors i/o posteriors per a la seva codificació.
- Frame 3: És un frame B (Bidirectionally-predictive-coded). Utilitza les dades dels frames I i P anteriors i/o posteriors per a la seva codificació.
- Frame 4: Un altre frame P, que segueix el mateix patró del frame 1, referenciant els frames I i B anteriors.

La seqüència dels frames segueix una ordenació lògica basada en la seva dependència en termes de la informació necessària per a la seva codificació.

- Determineu la seqüència IBP, i comenteu els valors de N i M.

La seqüència IBP (Intra-coded, Bidirectionally-predictive-coded, Predictive-coded) determina com els frames són organitzats per a la codificació eficient:

N = 15: Indica que cada 15 frames hi ha un frame I.

M = 3: Indica que els frames P s'inserten cada 3 frames.

Això significa que la seqüència seria d'estil I B B P B B P ... per a cada grup de 15 frames, amb un frame I iniciant cada seqüència nova.

- Calcular el retard de reordenació, tant al codificador i al descodificador, en unitats de nombre d'imatges i el temps. Podeu utilitzar aquesta configuració per a una videoconferència? I per TV? Justifiqueu-ho.

El retard de reordenació en el codificador es refereix al temps necessari per organitzar els frames B entre els frames I i P. A causa que els frames B són decodificats utilitzant un frame I o P com a referència però són presentats abans en la seqüència, es produeix un retard de reordenació de fins a 2 frames (donat que M=3). Això passa perquè els frames B han d'esperar que els frames P de referència estiguin disponibles.

A una velocitat de fotograma de 29.97 fps, la durada de cada frame és aproximadament $1/29.97=0.033$ segons.

En el descodificador, el retard és similar, ja que els frames B no poden ser presentats fins que els frames I i P de referència hagin estat processats. Seguint el patró de codificació IPBBP descrit, els frames B han de ser emmagatzemats en el descodificador fins que el frame P subsegüent sigui rebut i processat. Així, hi ha un retard de dos frames per al Frame 1, i un retard d'un frame per al Frame 2, esperant els frames P de referència per ser mostrats.

Pel Frame 1, el retard seria de dos frames: $2/29.97=0.067$ segons

Pel Frame 2, el retard seria d'un frame: $1/29.97=0.033$ segons

Per a una videoconferència, aquesta configuració no seria ideal a causa del retard de reordenació i a un alt bitrate de 6 Mbps, que podria ser problemàtic per a connexions amb amplada de banda limitada.

Per a la televisió, aquesta configuració és acceptable, ja que la latència no és tan crítica com en les videoconferències i la qualitat de la imatge és prioritària, el alt bitrate és adequat per a la transmissió de contingut d'alta qualitat.

La justificació d'aquest retard és perquè en el codificador, els frames B han d'esperar que els frames P siguin codificats abans de poder ser processats ells mateixos, ja que depenen de les dades de referència d'aquests frames. Com en el descodificador, els frames B necessiten que els frames I i P estiguin disponibles per a la seva descodificació i presentació, provocant retards basats en el temps d'espera per aquests frames de referència.

En l'aplicació pràctica en una Videoconferència els retards i el alt bitrate podrien comprometre la viabilitat d'aquesta configuració per a videoconferències, especialment en connexions de baixa amplada de banda on la latència baixa és crucial. I en la televisió, la configuració és més adequada per a la televisió, on la latència és menys crítica i la prioritat es posa en la qualitat d'imatge i en mantenir un alt bitrate per assegurar una transmissió de qualitat.

Apartat c

c) Discutiu els percentatges de tipus de macrobloc utilitzats en cada imatge. Relacioneu-los amb el tipus i la posició de la imatge dintre de la seqüència.

Frame 0 (Tipus I)

of intra coded macroblocks: 1350 (100.0%)

Aquest frame és totalment codificat de manera intra, sense cap referència a altres frames. La codificació intra és la més bàsica i la que exigeix més dades perquè cada macrobloc s'ha de codificar de forma independent.

Frame 1 (Tipus P)

of intra coded macroblocks: 598 (44.3%)

of forw. pred. macroblocks: 565 (41.9%)

of backw. pred. macroblocks: 0 (0.0%)

Els frames P no utilitzen predicció enrere, combina macroblocs intra codificats amb macroblocs de predicció endavant, utilitzant dades del frame I anterior per minimitzar la quantitat de dades necessàries. Un alt percentatge de macroblocs intra podria indicar detalls nous en la imatge que no es poden predir des del frame anterior.

Frame 2 (Tipus B)

of intra coded macroblocks: 231 (17.1%)

of forw. pred. macroblocks: 405 (30.0%)

of backw. pred. macroblocks: 99 (7.3%)

Com a frame B, utilitza dades dels frames anteriors i posteriors, com serien els frames 0 i 1 per la seva predicció. Aquesta capacitat de utilitzar dades de múltiples fonts permet una major compressió i predicció, optimitzant la representació del moviment i els canvis a través dels dos frames de referència.

El frame B2 utilitza tant el frame I0 com el P3 com a referències per a la seva codificació. No obstant això, B2 depèn principalment de I0 perquè està temporalment més a prop i els canvis visuals des d'I0 a B2 són menys significatius comparats amb els de P3. Això significa que la informació necessària per reconstruir B2 és més similar a I0, fent que la predicció sigui més precisa i requereixi menys dades addicionals.

Frame 3 (Tipus B)

of intra coded macroblocks: 237 (17.6%)

of forw. pred. macroblocks: 165 (12.2%)

of backw. pred. macroblocks: 315 (23.3%)

Similar al Frame 2, el Frame 3 fa ús de dades de frames anteriors i posteriors per millorar la compressió. La combinació de diferents tipus de predicció permeten una reducció de la quantitat de dades necessàries, tenint una gestió eficient del moviment i altres canvis dinàmics que es poguessin produir en la imatge.

El frame B3, en contrast amb B2, depèn més de la referència de P3 que de I0. Això és degut a que P3 està temporalment més a prop de B3, i els canvis visuals entre P3 i B3 són menys intensos que els canvis entre I0 i B3. Aquesta proximitat temporal fa que la informació del frame P3 sigui més pertinent i eficaç per a la codificació de B3, permetent una millor predicció i una compressió més eficient en comparació amb la utilització de dades de I0.

Frame 4 (Tipus P)

of intra coded macroblocks: 593 (43.9%)

of forw. pred. macroblocks: 587 (43.5%)

of backw. pred. macroblocks: 0 (0.0%)

Aquest frame, semblant al Frame 1, presenta una alta proporció de macroblocs intra codificats juntament amb macroblocs de predicció endavant. Aquest patró suggereix que, igualment les referències al frame I anterior, hi ha una quantitat alta de contingut nou en la imatge que necessita una codificació intra per preservar la qualitat de la imatge.

Apartat d

d) Discutiu els valors de la "forward/backward search window" per a cada imatge, i compareu els valors de les dues imatges B. Què està passant?

Els valors de "forward/backward search window" per a cada imatge són els següents:

Frame 0 (Tipus I)

-

Com a frame I, no utilitza predicció basada en altres frames, així que no té finestres de cerca de predicció endavant o enrere. Els macroblocs són completament intra codificats, dependent exclusivament de la informació dins el mateix frame.

Frame 1 (Tipus P)

forward search window: -11...11 / -11...11

forward vector range: -16...15.5 / -16...15.5

Aquest frame utilitza una ampla finestra de cerca per buscar coincidències en el frame 0 de Tipus I, permetent vectors de moviment que poden ser bastant grans (fins a 16 píxels en qualsevol direcció). Sent un intent de capturar moviments amplis o ràpids dins del vídeo.

Frame 2 (Tipus B)

forward search window: -3...3 / -3...3

forward vector range: -8...7.5 / -8...7.5

backward search window: -7...7 / -7...7

backward vector range: -8...7.5 / -8...7.5

En aquest frame els valors són més restringits de la finestra de cerca endavant en comparació amb el Frame 1, però amb una finestra de cerca enrere més àmplia, indicant que el Frame 2 té la informació dels frames anteriors i posteriors. Els vectors de moviment són menys extrems que en el Frame 1, probablement perquè el moviment entre els frames I i P és menys pronunciat.

Frame 3 (Tipus B)

forward search window: -7...7 / -7...7

forward vector range: -8...7.5 / -8...7.5

backward search window: -3...3 / -3...3

backward vector range: -8...7.5 / -8...7.5

En aquest frame els valors s'inverteix respecte al Frame 2, la finestra de cerca endavant és més àmplia, mentre que la finestra de cerca enrere és més petita. Això pot indicar que el Frame 3 depèn més del frame anterior (El frame tipus B) que del posterior que és per la predicció de moviment (El Frame tipus P), mentre que el Frame 2 depèn més del frame posterior, que és quan té la finestra més àmplia.

Frame 4 (Tipus P)

forward search window: -11...11 / -11...11

forward vector range: -16...15.5 / -16...15.5

Aquest frame, com el Frame 1, utilitza una finestra de cerca endavant àmplia per maximitzar la capacitat de capturar moviments grans, utilitzant el frame I anterior com a referència principal.

Exercici 4: FFmpeg

1. `$ ffmpeg -i aqua.avi -an -vcodec mpeg2video -s 352x240 -r 30 -aspect 4:3 -deinterlace -b 6000000 m2s6Mb.mpg:`

Aquesta ordre pren aqua.avi com a fitxer d'entrada (-i aqua.avi), desactiva l'àudio (-an), estableix el còdec de vídeo a mpeg2video (-vcodec mpeg2video), canvia la mida del vídeo a 352x240 (-s 352x240), estableix la taxa de fotogrames a 30 fps (-r 30), estableix la relació d'aspecte a 4:3 (-aspect 4:3), aplica

desentrellat al vídeo (-deinterlace), i estableix la taxa de bits a 6000000 (-b 6000000). El fitxer de sortida és m2s6Mb.mpg.

2. `$ ffmpeg -s cif -r 30 -g 15 -pix_fmt yuv420p -i bus_cif.yuv -vcodec mpeg2video -f mpeg2video -b 1000k -bf 2 bus.m2v:`

Aquesta ordre estableix la mida del vídeo a cif (-s cif), estableix la taxa de fotogrames a 30 fps (-r 30), estableix l'interval de fotogrames clau a 15 (-g 15), estableix el format de píxels a yuv420p (-pix_fmt yuv420p), pren bus_cif.yuv com a fitxer d'entrada (-i bus_cif.yuv), estableix el còdec de vídeo a mpeg2video (-vcodec mpeg2video), estableix el format de sortida a mpeg2video (-f mpeg2video), estableix la taxa de bits a 1000k (-b 1000k), i estableix el nombre màxim de fotogrames B a 2 (-bf 2). El fitxer de sortida és bus.m2v.

3. `$ ffmpeg -i Wildlife.avi -an -vcodec mpeg2video -f mpeg2video -s 1280x720 -r 30 -aspect 16:9 -deinterlace -b 13500k -maxrate 13500k -minrate 13500k -bf 2 -bufsize 1835008 video.m2v:`

Aquesta ordre és similar a la primera, però amb diferents paràmetres. Pren Wildlife.avi com a fitxer d'entrada, desactiva l'àudio, estableix el còdec de vídeo a mpeg2video, canvia la mida del vídeo a 1280x720, estableix la taxa de fotogrames a 30 fps, estableix la relació d'aspecte a 16:9, aplica desentrellat al vídeo, estableix la taxa de bits a 13500k, estableix la taxa de bits màxima i mínima a 13500k, estableix el nombre màxim de fotogrames B a 2, i estableix la mida del búfer a 1835008. El fitxer de sortida és video.m2v.

Ara, anem a aprendre com pots usar FFmpeg per codificar fitxers de vídeo sense comprimir des de YUV o altres formats, i com codificar fitxers en CBR i VBR:

- **Codificar fitxers de vídeo sense comprimir des de YUV o altres formats:** Pots usar FFmpeg per codificar fitxers de vídeo sense comprimir desde YUV o altres formats. Aquí tens un exemple de com pots fer-ho:

Unset

```
ffmpeg -i input.yuv -c:v libx264 output.mp4
```

En aquest exemple, input.yuv és el teu fitxer d'entrada i output.mp4 és el teu fitxer de sortida. El còdec de vídeo s'estableix en libx264, que és un còdec popular per al format H.264.

- **Codificar fitxers en CBR (Constant Bit Rate):** FFmpeg permet la codificació en CBR. Aquí tens un exemple de com pots fer-ho:

Unset

```
ffmpeg -i input.mp4 -c:v libx264 -x264-params "nal-hrd=cbr"  
-b:v 1M -minrate 1M -maxrate 1M -bufsize 2M output.ts
```

En aquest exemple, `input.mp4` és el teu fitxer d'entrada i `output.ts` és el teu fitxer de sortida. El còdec de vídeo s'estableix en `libx264`, i la taxa de bits s'estableix en 1M. La taxa de bits mínima i màxima també s'estableixen en 1M, i la mida del búfer s'estableix en 2M.

- **Codificar fitxers en VBR (Variable Bit Rate):** FFmpeg també permet la codificació en VBR. Aquí tens un exemple de com pots fer-ho:

Unset

```
ffmpeg -i input.mp4 -c:v libx264 -crf 23 output.mp4
```

En aquest exemple, `input.mp4` és el teu fitxer d'entrada i `output.mp4` és el teu fitxer de sortida. El còdec de vídeo s'estableix en `libx264`, i el factor de taxa constant (CRF) s'estableix en 23. Un CRF més baix resulta en una major qualitat, però també en una mida de fitxer més gran.