

VIDEOTECHNIKA

Előadásvázlat

Mócsai Tamás

BME Hálózati Rendszerek és Szolgáltatások
Tanszék

2015. szeptember

Előnyei

- A természetes képeket majdnem teljesen dekorrelálja
- Közel optimális
- A jelenergia nagy részét néhány kisfrekvenciás koefficiens hordozza
- Valós együtthatókat eredményez
- Számítástechnikailag hatékony algoritmusokkal számolható

1D DCT és IDCT

$$- X(k) = \sqrt{\frac{2}{N}} \alpha(k) \sum_{n=0}^{N-1} x(n) \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad k = 0, \dots, N-1$$

$$- x(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} \alpha(k) X(k) \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad n = 0, \dots, N-1$$

- ahol

$$\alpha(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{ha } k \equiv 0 \\ 1 & \text{ha } k \neq 0. \end{cases}$$

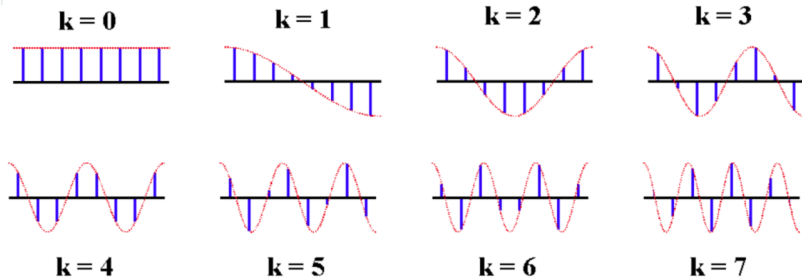


Elnevezések

- DC koefficiens: $X(0)$
- AC koefficiensek $X(1), \dots, X(N-1)$
- A k . bázisvektor: $\sqrt{\frac{2}{N}} \alpha(k) \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad n = 0, \dots, N-1$



$N = 8$ esetén a DCT bázisvektorok



- A bemenő jelvektort különböző harmonikus frekvenciájú elemi komponensek súlyozott összegeként (lin. komb.) állítjuk elő
- A súlytényezők a DCT együtthatók
- Ha a bemeneti vektoron belül a jel "lassan" változik, akkor a DC, illetve a kisfrekvenciás AC komponensek dominálnak

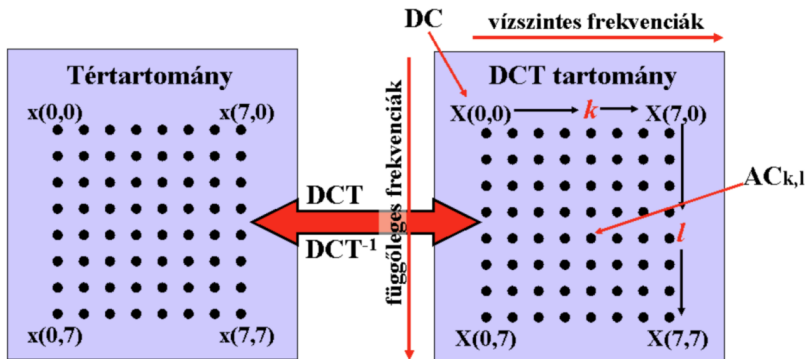


- $X(k, l) = \frac{2}{N} \alpha(k)\alpha(l) \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} x(n, m) \cos \left[\frac{\pi}{N} \left(m + \frac{1}{2} \right) k \right] \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) l \right] \quad k = 0, \dots, N-1$
- $x(m, n) = \frac{2}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \alpha(k)\alpha(l) X(k, l) \cos \left[\frac{\pi}{N} \left(m + \frac{1}{2} \right) k \right] \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) l \right] \quad k = 0, \dots, N-1$
- ahol

$$\alpha(k), \alpha(l) = \begin{cases} \frac{1}{\sqrt{2}} & \text{ha } k, l \equiv 0 \\ 1 & \text{ha } k, l \neq 0. \end{cases}$$

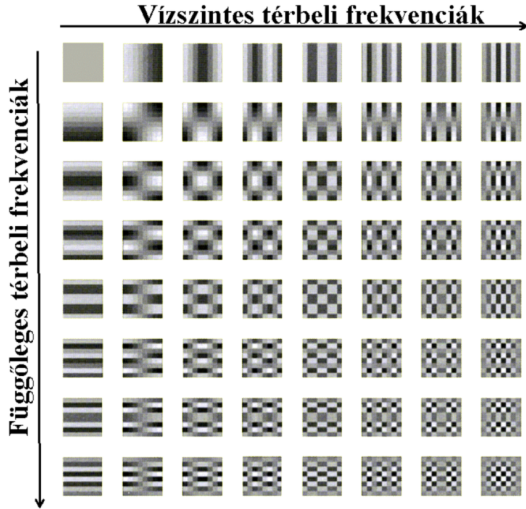


2D DCT illusztráció I.





2D DCT illusztráció II.



DCT bázisképek (8×8)



8 × 8-as pixel mátrix

83	87	92	90	89	91	47	95
98	81	77	96	71	44	58	49
27	43	65	40	64	99	61	55
83	21	45	51	59	80	48	63
87	94	56	62	41	74	75	57
98	82	68	79	54	46	52	60
72	42	70	84	69	50	97	67
76	53	85	88	73	66	78	86

DCT



8 × 8-as koefficiens mátrix

248	18	-12	2	3	-7	0	0
10	9	8	1	-4	1	0	0
-6	2	1	0	2	4	2	-1
2	1	-1	1	-3	0	2	0
0	0	0	1	2	0	0	-1
1	-2	1	0	1	-2	3	0
0	0	0	-1	1	2	0	0
0	0	0	0	0	0	0	0



1D DCT és DFT

- 1D DCT számítható az N hosszúságú $\mathbf{x} = x[n] \quad n = 0 \dots N - 1$ mintavektor szimmetrikus $2N$ hosszú kiterjesztésének DFT-jéből, ahol
- a szimmetrikus kiterjesztés $\bar{x}[p] = x[p]$, ha $p < N$, és $\bar{x}[q] = x[2N - 1 - q]$, ha $q > N$
- Ha $Y(k)$ jelöli az $\mathbf{x} = x[n]$ DCT együtthatóit,
- akkor $Y(k) = \Re \{ W F(k) \}$, ahol $F(k) \quad k = 0 \dots N - 1$ a $2N$ hosszúságú kiterjesztett mintavektor első N DFT együtthatója, és
- W egy jelfüggetlen skálatényező.



2D

- Ha $\mathbf{A} = A[k, m] = \sqrt{\frac{2}{N}} \alpha(k) \cos \left[\frac{\pi}{N} \left(m + \frac{1}{2} \right) k \right]$ az 1D transzformáció mátrixa, és
- $\mathbf{X} = x[m, n]$ a 2D kép mátrix, akkor $\mathbf{Y}_c = \mathbf{AX}$, majd $\mathbf{Y}_{c,r} = \mathbf{AY}_c^T$, akkor
- $\mathbf{Y}_{c,r} = \mathbf{A}(\mathbf{XA})^T = \mathbf{AX}^T \mathbf{A}^T$, amiből az eredeti kép DCT-je:
- $\mathbf{Y}_{k,l} = \mathbf{Y}_{c,r}^T$

Kvantáló mátrix

- A kvantálás blokkonként ún. kvantáló mátrixszal történik
- $X^*(k, l) = \frac{X(k, l)}{W(k, l)}$, ahol $W(k, l)$ jelenti a (k, l) együtthatóhoz tartozó kvantálási lépcsőt
- A kvantáló mátrix elemeinek megválasztása alapvetően a HVS tulajdonságainak megfelelően történik
- Maradék elhagyása

Kvantálás a HVS tulajdonságai alapján

- A HVS az apró részletekre (nagyfrekvenciás komponensek) kevésbé érzékeny
- A HVS szempontjából fontos együtthatókat (DC, néhány kisfrekvenciás AC) finomabban, a kevésbé fontos AC együtthatókat durvábban kvantáljuk
- Egy tipikus kvantáló mátrix pl. a következő

$$W(k,l) = \begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix}$$

Kvantálás a HVS tulajdonságai alapján

- A bitsebesség pl. a $W(k, l)$ konstanssal való szorzásával állítható
- A kvantáláson túlmenően a DC komponensek a szomszédos blokkhoz képest differenciálisan kódoltak (JPEG és MPEG)



DCT együtthatók kvantálása illusztráció I.

DCT koefficiensek

$$\begin{bmatrix} 284 & 18 & -12 & 2 & 3 & -7 & 0 & 0 \\ 10 & 9 & 8 & 1 & -4 & 1 & 0 & 0 \\ -6 & 2 & 1 & 0 & 2 & 4 & 2 & -1 \\ 2 & 1 & -1 & 1 & -3 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & -1 \\ 1 & -2 & 1 & 0 & 1 & -2 & 3 & 0 \\ 0 & 0 & 0 & -1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Kvantáló mátrix (W(k,l))

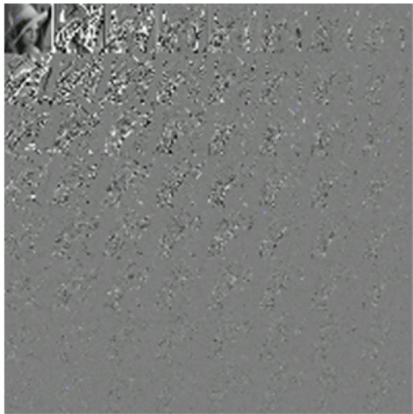
$$\begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix}$$

/

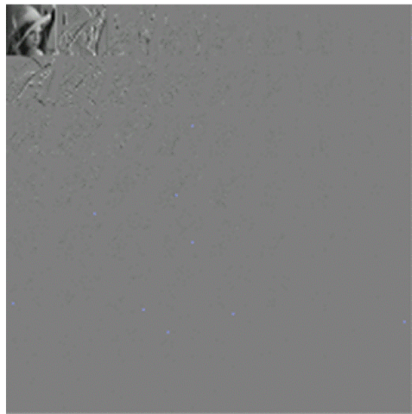
Kvantálás után

=

$$\begin{bmatrix} 35 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



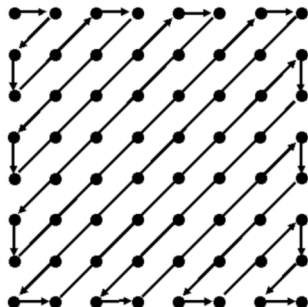
**DCT felbontás
eredeti DCT együtthatók**



**DCT felbontás
újrakvantált DCT együtthatók**

Cikk-cakk kódolás

- Kvantálás után az együtthatók nagy része zérussá válik
- Csak a nullától különböző együtthatókat kell tárolni/továbbítani, viszont ezek pozícióját ismerni kell
- Cikk-cakk kódolás: növekvő frekvenciák szerinti rendezés
- A DC együtthatót külön, általában veszteségmentesen, differenciálisan kódoljuk az előző blokk DC együtthatójával, mivel a szomszédos blokkok DC együtthatója hasonló
- Az AC együtthatók cikk-cakk rendezése miatt több nulla kerül egymás után
- A nullákra futamhossz kódot alkalmazunk
- (Zérusok száma, értékes amplitúdó) párokat (run,level) képzünk a cikk-cakk sorrend alapján
- A blokk végét EOB szimbólum zárja


 $X =$

$$\begin{bmatrix} 110 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Run-level párok

(0,-2)

(0,1)

(1,-3)

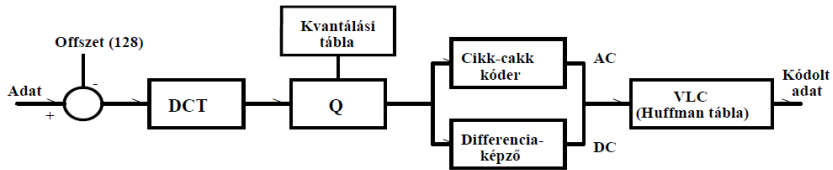
(2,1)

(0,-1)

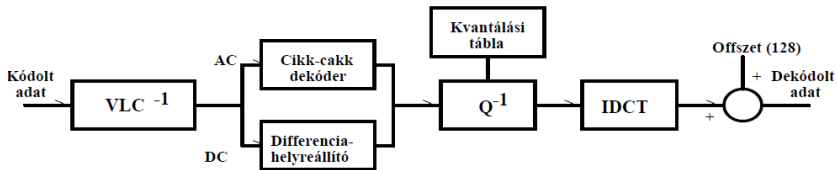
EOB



JPEG kódoló/dekódoló vázlata



DCT alapú szekvenciális JPEG kódoló



DCT alapú szekvenciális JPEG dekódoló



- képméret maximum 65535×65535 ,
- komponensek száma maximum 255,
- a színmérő-rendszer nem specifikált.
- a kóder és dekóder szimmetrikus felépítése,
- képminőség - bitsebesség (bpp) kompromisszum
- 4-féle üzemmód:
 - **DCT alapú szekvenciális kódolás,**
 - **DCT alapú progresszív kódolás,**
 - veszteségmentes DPCM alapú kódolás,
 - hierarchikus kódolás (DCT, DPCM).



DCT alapú JPEG kódolás lépései

- A kép 8x8-as, egymással át nem lapolódó blokkjainak transzformációja DCT-vel.
- Az együtthatók kvantálása blokkonként egy felhasználó által definiált súlyozó mátrix és egy kvantálási tényező segítségével.
- Maximum 4 súlyozó mátrix használható egyszerre (pl. komponensenként más és más).
- A DC együttható: az előző blokkhoz képesti különbség kódolása. (1-D veszteségmentes DPCM hurokkal).



DCT alapú JPEG kódolás lépései

- Az AC együtthatók cikk-cakk-ba rendezése, futamhossz kódolt párok 0-ák futási hossza, nem zérus amplitúdó képzése.
- A differenciális DC együtthatók és a futamhossz kódolt párok VLC kódolása.
- A blokk végét EOB jelzi.
- Maximum 2 DC és 2 AC Huffman tábla használható.
- A képkomponensek függetlenül kódoltak.
- A dekóder a bitfolyam fejlécéből nyeri ki azokat az információkat, amelyek a dekódolás vezérléséhez kellenek.
- A dekódolás számítás igénye kicsit kisebb mint a kódolásé.
- A rekonstruált kép minőségét a súlyozó mátrix, a kvantálási tényező és a DCT és IDCT pontossága határozza meg,
- A default-tól eltérő kvantálási és Huffman táblákat (8 bites 8x8 db. koefficiens) a fejlécben kell továbbítani.
- Nincs tényleges bitsebesség vezérlés.



DCT alapú JPEG kódolás lépései

- Az alapértelmezett (fix) kvantálási és Huffman táblákat használva a kóder és dekóder szimmetrikus a műveletigényt tekintve.
- A DC differenciák osztályozása amplitúdó szerint, osztály, amplitúdó párok képzése.
- Az osztály Huffman kódolása + az amplitúdó változó hosszúságú egészként való kódolása (VLI).
- Az osztály megadja a VLI bitszámát.
- A koefficiens osztályba sorolása és azon belüli érték meghatározása.
- VLI képzés: pozitív számoknál a szükséges bitszámon előjel nélküli bináris ábrázolása, negatív számoknál ennek az 1-es komplemente.



Osztály	Amplitúdó	VLI
1	-1, 1	0, 1
2	-3, -2, 2, 3	00, 01, 10, 11
3	-7...-4, 4...7	000...011, 100...111
4	-15...-8, 8...15	0000..0111, 1000..1111
5	-31...-16, 16...31
6	-63...-32, 32...63
7	-127...-64, 64...127
8	-255...-128, 128...255
9	-511...-256, 256...511
10	-1023..511, 511..1023
11	-2047...-1024, 1024...2047	00000000000...01111111111 10000000000...11111111111



JPEG hibák - ringing/blokkosodás





- A szabványok nem specifikálják a kódert (kódolási algoritmusokat) sem a videó, sem az audio részben! Csak a kódolás eszközeinek tárháza, és az azok által létrehozott kimeneti adatfolyam bitszintaxisa specifikált!
- MPEG-1 (ISO 11172 /1993/): Alacsony bitsebességű multimédiás alkalmazások (Video-CD, CD-I), kb. 1,5 Mb/s VHS képminőség, SIF felbontású kódolt videó.
- MPEG-2 (ISO 13818 /1994-95/): Műsorszórás (Digital Video Broadcasting, DVB), 2-8 Mb/s terjesztési minőség (distribution quality); Stúdiótechnika, 18-50 Mb/s újrafeldolgozási minőség (contribution quality); Digital Versatile Disc (DVD), 3-7 Mb/s (VBR!) jobb mint a PAL minőség; Az MPEG-2 kibővített MPEG-1, a kódolási elv mindkét eljárásban azonos.



- MPEG-4 (ISO/IEC 14496): Alacsony bitsebességű kódolások szabványának indult (1994), de jelenleg az interaktív multimédia szolgáltatások objektum orientált szabványává vált. *Ezen belül a két legfontosabb rész:*
 - MPEG-4 Part 2: Egyik szabványos változat a H.263, illetve implementációi a DivX, és XVID
 - MPEG-4 Part 10: H.264/AVC



- DCT alapú mozgáskompenzációt használó hibrid kódolás.
- A kódolás és dekódolás számításigénye különböző, a rendszer tehát aszimmetrikus.
- MPEG videókódolás egyik alapjellemezője a réteges szerkezete.
- A rétegszerkezet 6 egymásba ágyazott egységet tartalmaz, melyekben az alsóbb rétegek általában nem dekódolhatók a felsőbb szintek nélkül.

- A mozgókép soron egymás utáni képkockái erősen korreláltak
- A képváltásokat és a gyors mozgásokat (pl. sport) leszámítva a képtartalom képről-képre viszonylag kevésbé változik
- Mozgásbecslés alapú predikcióval a predikció hatékonysága javítható
- A redundancia csökkentés hatékonysága a mozgás meghatározás pontosságától függ
- A mozgásvektorok pontos meghatározása nehéz és időigényes
- Megfelelően definiált költségfüggvényekkel, kompromisszumokkal keressük a mozgásvektorokat (pl. átlagos abszolút hiba, vagy átlagos négyzetes hiba)
- Offline tömörítés esetén pontosabb mozgásbecslő algoritmusok használhatók



Képek közötti differencia (lassú mozgás)



N. kép



(N+1). kép



N. kép - (N+1). kép



Képek közötti differencia (képváltás)



N. kép



(N+1). kép

N. kép - (N+1). kép





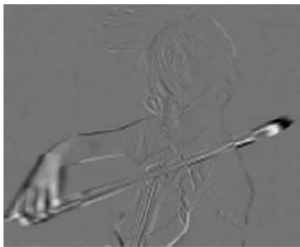
Képi pixelek "mozgása" - optical flow



Frame 1



Frame 2



Difference



Optical flow



Blokk alapú mozgásbecslés I.

- Az $N \times N$ -es blokk alapú transzformációs kódoláshoz illeszkedően blokk alapú mozgásbecslést és predikciót célszerű megvalósítani
- A cél az aktuális kép minden $N \times N$ -es blokkjához egy referenciakép egy-egy $N \times N$ -es területét rendelni, amely a kódolandó blokk prediktora lesz
- A blokkok pixel-koordinátáinak különbsége az adott blokk mozgásvektora
- Az optimális prediktor-blokk kiválasztása szélsőérték keresési feladat, melyben a kódolandó blokk körül kijelölt ablakon belül a referencia kép és a kódolandó blokk illeszkedést maximalizáljuk egy költségfüggvény szerint



Blokk alapú mozgásbecslés II.

- A mozgásbecslést az Y komponensen végezzük, a kapott mozgásvektorokat a Cr,Cb blokkok mozgásvektorai is lesznek
- A mozgásbecslés/mozgáskompensáció alapegysége a **makroblokk** (MB)
- Minden nem intra blokkhoz tartozik minimum egy mozgásvektor, a blokkon belüli mozgás azonban nem detektálható (predikciós hibaként jelenik meg)
- A mozgásvektorok is differenciálisan kódoltak, a blokk-raszter sorrendjében a szomszédos MB mozgásvektorához képesti különbséget kódolják (hatékony, mert a szomszédos makroblokkok valószínűleg hasonlóan mozognak)
- A mozgásvektorok meghatározása 1, 1/4 ill. 1/8 pixel pontossággal történhet, a kódoló típusától függően.



Keresési eljárások

- Teljes/kimerítő keresés: pontos, de nagyon időigényes
- Pixel-rekurzív keresés: előzetes becslés pl. előző kép, szomszédos MB-k alapján, az előzetes becslés finomítása a mozgásvektor koordinátájának környezetében
- Logaritmikus keresés: Durvább, majd finomabb léptékű keresés a referencia illetve az adott kódolandó kép durvább majd finomabb felbontású változatain. Minden finomítás után pixel rekurzív keresés a korábbi durvább felbontású mozgásvektor eredmény környezetében



Blokk alapú mozgásbecslés, példa



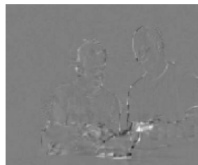
Frame 1



Residual : 16×16 block size



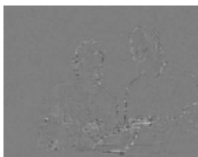
Frame 2



Residual : 8×8 block size

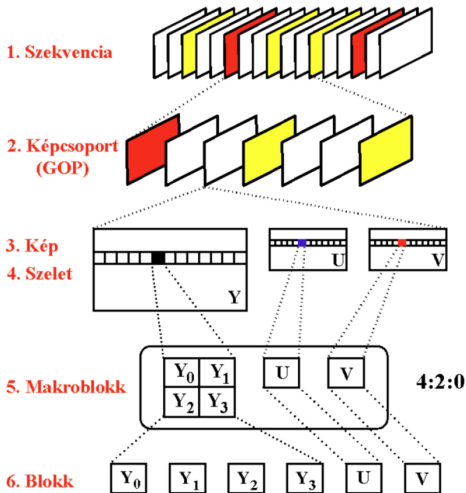


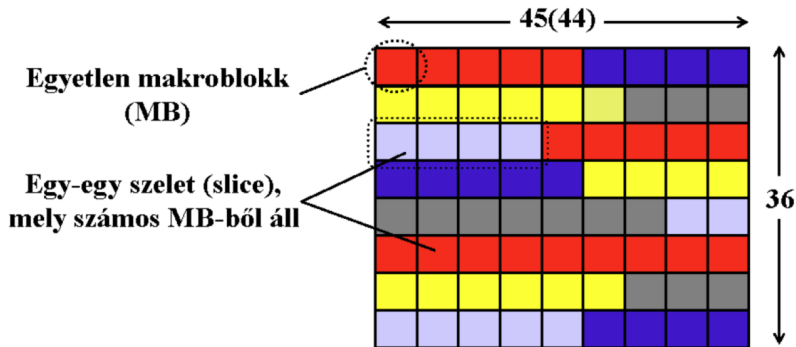
Residual : no motion compensation



Residual : 4×4 block size

- Szekvencia réteg** A kódolt szekvenciát azonosítja, a fejléc tartalmazza a rendszeradatokat (képméret, bitsebesség, stb.)
- Képcsoport réteg (GOP)** Legalább egy önmagában kódolt (I) képet tartalmazó, meghatározott számú kép együttese, a véletlen hozzáférés egysége
 - Képréteg** Egy kép kódolt adatait tartalmazza
- Szelet (slice) réteg** MB-k sor-folytonos csoportja, az újraszinkronizáció egysége (a legalsó szint, amelyen a dekóder még képes feléledni bithiba esetén)
- Makroblokk réteg** Az Y 16×16 -os (4 db 8×8 -as blokk), és a Cr, Cb 8×8 -as blokkjaiból áll (4:2:0 esetén 1 Cr 1 Cb, tehát 1 MB összesen 6 blokk), a mozgáskompenzáció egysége
- Blokk réteg** A MB 8×8 -as blokkjai, a DCT kódolás egysége





Egy MPEG-2 EU SD kép 45(44) × 36 MB tartalmaz

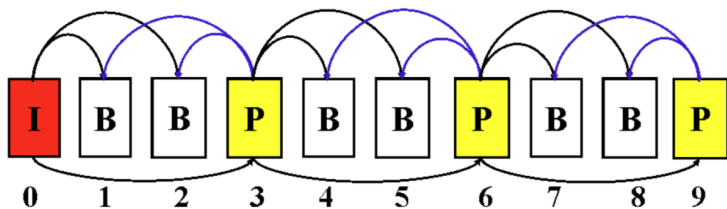


- I (intra) Önmagában kódolt kép, a dekódoláshoz szükséges minden adatot tartalmaz
- P (predictive) egy múltbéli referencia képhez képest prediktíven kódolt kép, referenciája egy előző I vagy P kép, dekódolásához a referencia kép szükséges
- B (bidirectional) két irányból kódolt kép, referenciája az előző I vagy P kép és a következő I vagy P kép. B kép nem lehet referencia

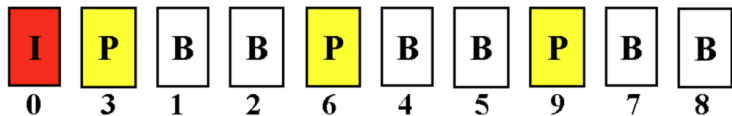


- A kijelzési és átviteli képsorrend általában különböző
- A műsorszórás leggyakoribb kijelzési sorrendje:
IBBPBBPBBPBB (Long GOP - EU)
- A stúdiótechnika kijelzési sorrendje általában: csak I képek
- MB típusok
 - I képekben: minden MB önmagában kódolt, intra típusú
 - P képekben: az MB lehet intra, vagy prediktíven kódolt
 - B képekben: az MB lehet intra, illetve a múltból, a jövőből, vagy múltból és jövőből prediktíven kódolt (múltból és jövőből történő predikció esetén a mozgás-kompenzált múltbéli illetve jövőbéli referenciaképek átlaga a predikció alapja)
- Frame és field alapú predikció (többfajta...)

Kijelzési képsorrend



Átviteli képsorrend



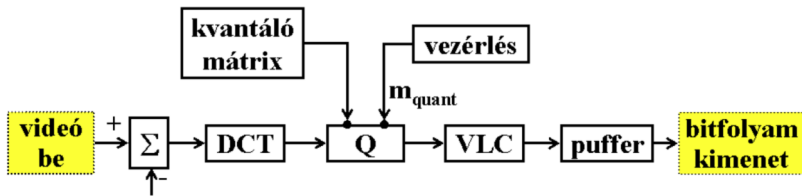


MPEG képtípus megfontolások

- A globálisan elérhető tömörítés 10:1 és 50:1 között van.
- B képek alkalmazásával növekszik a dekóder oldalon a számítás- és memóriaigény, és a kóder oldali a késleltetés.
- Az IPB struktúra nem specifikált, B képek használata nem kötelező: a szekvencia állhat csak I, vagy csak I és P képekből.
- Az I, P és B képek aránya kompromisszum kérdése:
- Gyors véletlen hozzáférés (szerkeszthetőség, dekódolás megkezdhetősége): sok I kép, kevés P és B kép.
- Jó tömöríthetőség: sok B, kevés I kép.
- Alacsony késleltetés, alacsony hardverköltés: nincsenek B képek.
- A B képek miatt kétféle képsorrend (kijelzési, átviteli) definiált, hiszen a jövőbeli referenciaképeket előbb kell a bitfolyamba tenni mint a belőlük jósolt B képeket.

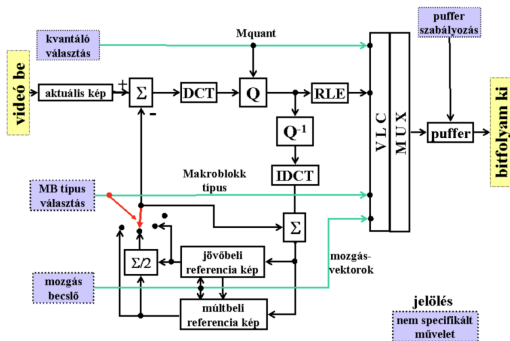
- Zárt képcsoport: a benne lévő első B képeknek nem referenciája az előző képcsoport utolsó I vagy P képe.
- Nyílt képcsoport: a benne lévő első B képeknek múltbeli referenciája az előző képcsoport utolsó I, vagy P képe.

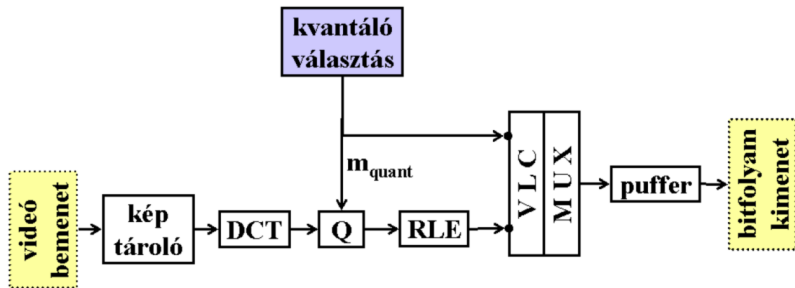
- A MB predikciós hibájának, vagy az intra MB kódolása DCT transzformációval (8 x 8 -as DCT)
- A DCT együtthatók újrakvantálása: a kvantálási lépcsőt a kvantálási mátrix megfelelő elemének és a kvantálási skálafaktornak a szorzata adja
- Kvantálási mátrix perceptuális alapú az I képekre, és általában konstans a P és B képekre
- A skálafaktor szeletenként, vagy MB-nként újradefiniálható (bitrate vezérlés)



MPEG-1 / MPEG-2 video kódoló

- 2 db képtároló: egy múltbeli referencia-, egy jövőbeli referencia-képtároló
- I képek esetén: megjelenítés, és átkerül a referenciatárolóba
- P képek esetén: inverz predikció, megjelenítés, referenciatároló
- B képek esetén: csak megjelenítés, utána eldobás







Konstans bitsebesség (CBR)

- A kódoló kimenetén egységnyi idő alatt keletkező bitek száma változó, mert
 - A különböző típusú képek (I,P,B) különböző bitszámmal kódolhatók azonos minőség mellett
 - Egy képen belül is változhat a képtartalom, az összetettebb blokkok leírásához több bitre van szükség
 - A VLC kódolás miatt is ingadozik a kimeneti bitmennyiség
- A legtöbb átviteli csatorna konstans bitsebességű forrást igényel
- A probléma a kóder kimenetére helyezett pufferrel oldható meg, melyből a csatorna fix ütemben olvassa a biteket



- GOP-ra elosztható bitek száma adott, a cél bitszám meghatározása I, P, B képekre (globális minőség biztosítása).
- A kvantálás meghatározása: MB komplexitástól és textúrától függő referencia kvantáló felvétele (lokális minőség biztosítása).
- és azok módosítása a puffer telítettségnek megfelelően.



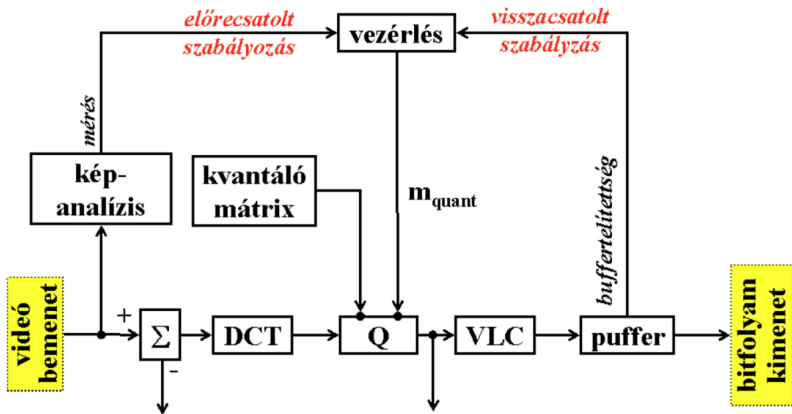
Változó (VBR)

- Két alapvető típusa:
 - Nyílt hurkú: nincs puffer szabályozás
 - Visszacsatolt: van puffer szabályozás
- Nyílthurkú VBR használható, ha a bitfolyam olvasás vezérelhető eszköztől történik (pl. DVD, memória, stb), ekkor a puffer gyakorlatilag végtelen nagy
- Nyílthurkú VBR-ek két alaptípusa:
 - Konstans skálafaktorú kódolás: a kvantálási zajt egyenletesen teríti a képen (nem feltétlenül a legjobb képminőség)
 - Konstans képminőségű kódolás: a skálafaktort a képminőséghez igazítja



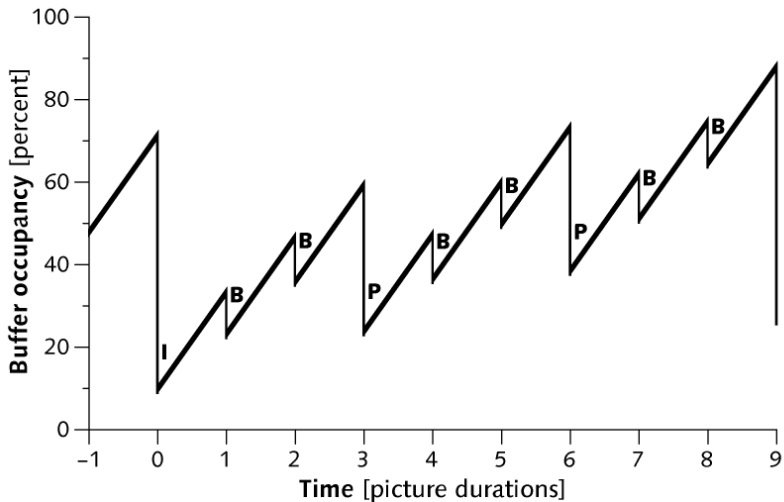
Változó (VBR)

- Visszacsatolt VBR: lényege, hogy a változó bitsebességet puffereléssel simítjuk mielőtt a bitfolyam a VBR csatornába kerül, így a sebesség egy előre definiált csúcssebesség (max. bitrate) alatt marad
- Előre és hátracsatolt vezérlés: Előrecsatolt bitsebesség szabályzás globális (kép) szintű, és lokális (MB) szintű komplexitás elemzéssel + visszacsatolt szabályzás a buffer telítettségének figyelésével (A VBR kódolás kb. 20-30 %-al kisebb átlagos bitsebességet eredményez, mint az azonos minőségű CBR kódolás)



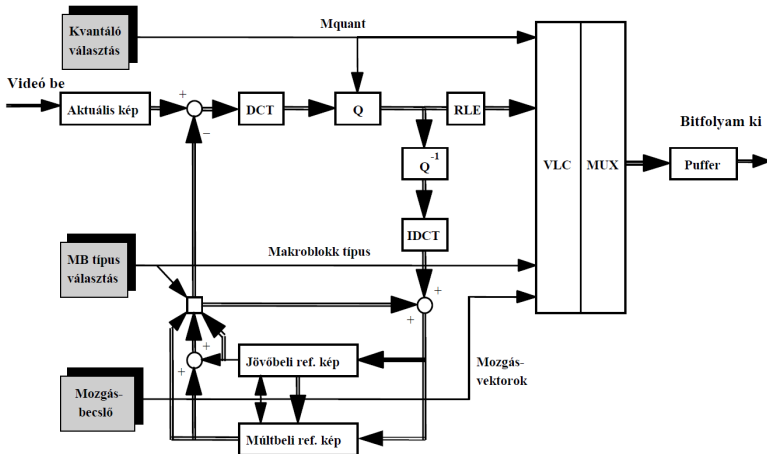


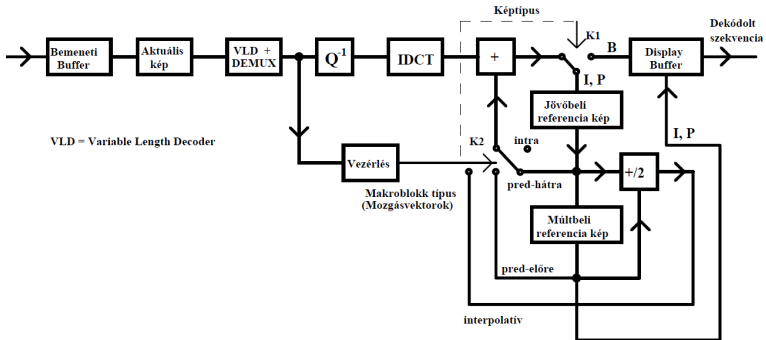
MPEG dekóder buffer telítettség (CBR)





MPEG-1/MPEG-2 kódoló







- 4:2:0-ás színkülönbségi jel mintastruktúra
- Képfrekvenciák: 24, 25, 29.97, 30 kép/sec
- Max. képméret: 4095x4095
- Csak progresszív képek kódolása
- Letölthető kvantáló mátrixok
- I,P,B képek
- Elérhető tömörtés: 1:25 - 1:30



- MPEG-1 -re épül, vele felülről kompatibilis
- Váltott-soros videó kezelés, félkép alapú feldolgozás (DCT és mozgásbecslés) is lehetséges
- Réteges kódolási módok (elavult)
- Profile-Level szerkezet
- Egy videón belül a kép és félkép alapú feldolgozás keverhető



MPEG-1 és MPEG-2 videó közötti különbségek

- Progresszív mellett váltottsoros képek kezelése is:
- kép-félkép alapú adaptív DCT
- félkép alapú mozgásbecslési algoritmusok:
 - félkép alapon félképre
 - félkép alapon képre
- Dual Prime predikció (félképre, ha nincs B kép)
- MPEG-1-nél magasabb bitsebességek, nagyobb képméret
- Csak félpixeles mozgásbecslés (egész pixeles már nincs), a cikk-cakk letapogatás helyett alternatív letapogatás is
- 4:2:2 és 4:4:4 mintavétel is (MPEG-1-ben csak 4:2:0)
- Profil és szint szerkezet
- Skálázható kódolás is: jel-zaj viszony és térbeli skálázás



- Különböző osztályokat definiál, melyek meghatározzák egy adott Profile/Level-nek megfelelő bitfolyam dekódolásához szükséges dekóder képességeket és kapacitást
- A **profile** a mintavételi formátumokat, a képtípusokat, és az egyéb kódolási eszközöket definiálja, illetve korlátozza
- A **level** a képméretet, a képfrekvenciát, és a bitsebességet definiálja, illetve korlátozza
- Célzott alkalmazási területek Level-enként:
 - Low: videokonferencia
 - Main: SDTV
 - 422P: stúdiótechnika
 - High: HDTV



MPEG-2 Profile-Level tulajdonságok

Minden, adott Profile/Level szintnek megfelelő dekóder a tőle balra álló Profile/Level szinteket is dekódolni tudja. Professzionális és konzumer elterjedés szempontjából a sárga területek jelentősek.

Szintek (Levels)

	Low	Main	High-1440	High	422P
Horizontális méret	352	720	1440	1920	720
Vertikális méret	288	576	1152	1152	608
Képfrekvencia (Hz)	30	30	60	60	30
Bitsebesség (Mbit/s)	4	15	60	100	50

Profilok (Profiles)

	Simple	Main	SNR	Spatial	High	422P
YUV	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0 4:2:2	4:2:0 4:2:2
Képtípus	I, P	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B
Skálázhatóság	nem	nem	igen	igen	igen	igen



MPEG-2 Profile-Level tulajdonságok

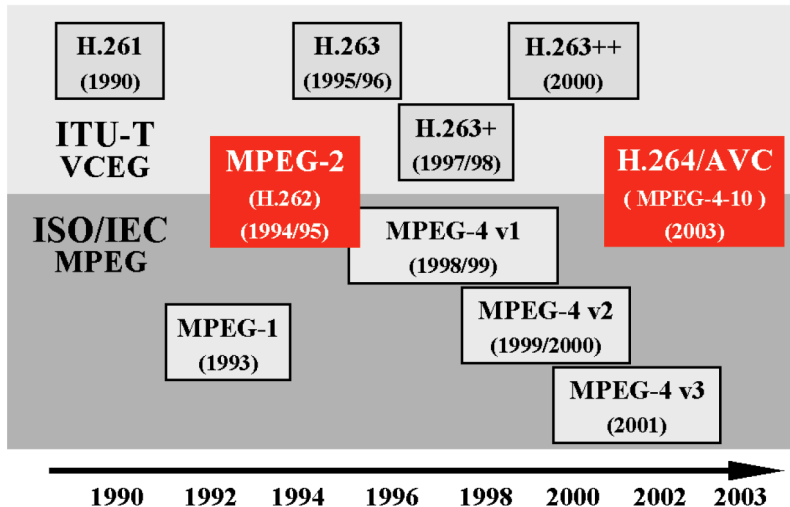
<i>Profile</i> @Level	MPEG-1 CPB	Simple (no B pictures)	Main (MP)	4:2:2 (422P)
High (HL)			1920×1152 60 Hz 80 Mb/s	1920×1088 60 Hz 300 Mb/s
High-1440 (H14)			1440×1152 60 Hz 47 Mb/s	
Main (ML)		720×576 30 Hz 15 Mb/s	720×576 30 Hz 15 Mb/s	720×608 30 Hz 50 Mb/s
Low (LL)			352×288 30 Hz 4 Mb/s	
MPEG-1 CPB † <i>max 99 Kpx</i>	768×576 [†] 30 Hz 1.856 Mb/s			



MPEG-2 Profile-Level tulajdonságok

<i>Profile@Level</i>	<i>Image columns (N_C)</i>	<i>Image rows (N_R)</i>	<i>Frame rate, Hz</i>	<i>Luma rate [samples/s]</i>	<i>Bit rate [Mb/s]</i>	<i>VBV size [KBytes]</i>
422P@HL	1920	1088	60	62,668,800	300	5,760
MP@HL	1920	1088	60	62,668,800	80	1,194
MP@H-14	1440	1088	60	47,001,600	60	896
422P@ML	720	608	60	11,059,200	50	1,152
MP@ML	720	576	30	10,368,000	15	224
MP@LL	352	288	30	3,041,280	4	58

Video tömörítés történelem 1990-2003 között





- Az MPEG-4 kiterjeszti az MPEG-2 korlátozott adat modelljét a multimédia alkalmazások számára is
- Az MPEG-4 olyan reprezentációs szabvány, mely lehetővé teszi az interaktivitást az audiovizuális objektumokkal, és
- közösen alkalmazható technikát biztosít számos konvergáló szolgáltatási formának
- A stúdiótechnika szempontjából kevésbé fontos



MPEG-4 (Part 2) és H.263 I.

- Számos profile és Level, melyek célul tűzték ki a teljes audiovizuális alkalmazási spektrum lefedését: alacsony bitsebességű alkalmazások (videokonferencia, zárt láncú biztonsági kamera hálózatok), közepes és magas bitsebességű alkalmazások (konzumer video, DVD, HDTV műsorszórás), és stúdió alkalmazások
- Az Advanced Simple Profile (ASP) implementált változatai a DivX (kivéve a Divx Plus HD, ami már H.264/AVC), és az XVID (open source)
- Az ITU által szabványosított kódolási változat a H.263, mely MPEG-4 video dekóder által dekódolható, és internetes videotelefon és internetes média alkalmazások (Flash video, stb.) legelterjedtebb kódolási módja volt a H.264 megjelenéséig
- Sokat vártak tőle, de csak részben váltotta be a hozzáfűzött reményeket



Kódolási elemek, eszközök

- Alapvető kódolás elemek, módszerek megegyeznek az MPEG-1/MPEG-2 - vel
- 1/4-pixel pontosságú mozgásbecslés
- Globális (frame) alapú mozgáskompenzáció (pan, rotate, zoom, és warp definiálása az egész képre), mely a teljes képtartalom mozgása esetén (pl. kameramozgás) esetén hatékonyabban leírja a mozgást, mint a blokk alapú mozgásbecslés - a gyakorlati implementációkban nem terjedt el, ezért a H.264-be már nem is került bele
- Átlapolt blokk alapú mozgáskompenzáció (blokkosodás csökkentése érdekében), az implementációja nem terjedt el
- Simple Studio Profile támogatja a 10/12 bit és 4:2:2, 4:4:4 mintaformátumokat (HDCAM SR elterjedt professzionális formátuma)



MPEG-4 (Part 10) H.264/AVC

- Az ITU-T Video Coding Experts Group és az ISO Joint Video Team, közös fejlesztése: ezért ITU-T H.264 szabvány, és ISO/IEC MPEG-4 AVC szabvány is
 - Jobb video minőség, mint az MPEG-2 -vel
 - Jelentős kompresszió növekedés
 - Javított predikció (intra- és inter-)
 - Javított mozgáskompensáció pontosság
 - Javított entrópia-kódolási hatékonyság
- Profile-Level szerkezet
- Széles alkalmazási spektrum (videotelefon, internet-média, digitális műsorszórás), kezdetben elsősorban "entertainment" (tehát SD és kisebb felbontású) alkalmazásokra
- A Fidelity Range Extensions (FREXT) kiegészítéssel professzionális (stúdió, digitális mozi, stb) alkalmazásokra is



MPEG-4 H.264/AVC Profilok

	<i>Profile</i>	<i>Baseline (BP)</i>	<i>Extended (XP)</i>	<i>Main (MP)</i>	<i>High (HiP)</i>
<i>Kódolási eszköz</i>					
Több referencia kép használata		•	•	•	•
Változó blokkméretű mozgásbecslés		•	•	•	•
I- és P-szeletek		•	•	•	•
1/4 pixel pontosságú mozgásbecslés		•	•	•	•
16 bites aritmetikájú integer transzformáció		•	•	•	•
UVLC		•	•	•	•
CAVLC		•	•	•	•
Deblocking szűrő		•	•	•	•
Flexibilis Makroblokk sorrend (FMO)		•	•		
Tetszőleges szelet sorrend (ASO)		•	•		
Redundáns szeletek (RS)		•	•		
Adat partícionálás			•		
SI és SP szeletek			•		
B-szeletek			•	•	•
Interlace kódolás (PICAFF, MBAFF)			•	•	•
Súlyozott (és ofszet) inter-predikció			•	•	•
CABAC entrópiakódolás				•	•
8x8 luma intra predikció					•
Nagyobb bitmélységek (8bitnél)					•
4:4:4 és 4:2:2 aluminávételezés					•
Veszteségmentes inter-predikció					•
Adaptív 8x8/4x4 transzformáció váltás					•
Kódoló által specifikálható kvantálási mátrixok					•
Cb és Cr független kvantálása (független Qstep)					•
Monkróm (alfa)csatorna kódolás					•



MPEG-4 H.264/AVC Szintek

<i>Level</i>	<i>Typ. image format</i>	<i>Typ. frame rate [Hz]</i>	<i>Max. bit rate [b/s]</i>
L1	176×144	15	64 k
L1b	176×144	15	128 k
L1.1	352×288 or 176×144	7.5 or 30	192 k
L1.2	352×288	15	384 k
L1.3	352×288	30	768 k
L2	352×288	30	2 M
L2.1	352×480 or 352×576	30 or 25	4 M
L2.2	SD	15	4 M
L3.0	SD	30 or 50	10 M
L3.1	1280×720	30	14 M
L3.2	1280×720	60	20 M
L4.0	1920×1080	30	20 M
L4.1	1920×1080	30	50 M
L4.2	1920×1080	60	50 M
L5	2048×1024	72 or 30	135 M
L5.1	4096×2048	30	240 M



FRExt - Fidelity Range Extensions

- Új profilkövetelmények:
 - 8 bit/mintánál nagyobb bitmélységek támogatása
 - 4:2:2 és 4:4:4 mintastrukturák támogatása
 - Video-kulcsolás, és transzparencia (alpha csatorna) támogatása a stúdióalkalmazások céljára
 - Nagyobb adatsebességek támogatása
 - Veszteségmentes tömörítés támogatása (lináris PCM, vagy transzformációs kódolás nélküli entrópiakódolás)
 - A színtér transzformációk numerikus hibáinak elkerülése (YCgCo színtér)
 - RGB (nem csak YCbCr) komponens reprezentáció támogatása
- 8×8 -as transzformáció
- Mára a FRExt High Profile-ja fontosabbá vált (mind konzumer, mind profi felhasználásra), mint a Main profile konzumer profilja, mert számottevően jobb kódolási hatékonyságot biztosít, miközben az implementálás komplexitása (pl. kódoló, dekódoló hardver) nem nőtt jelentősen

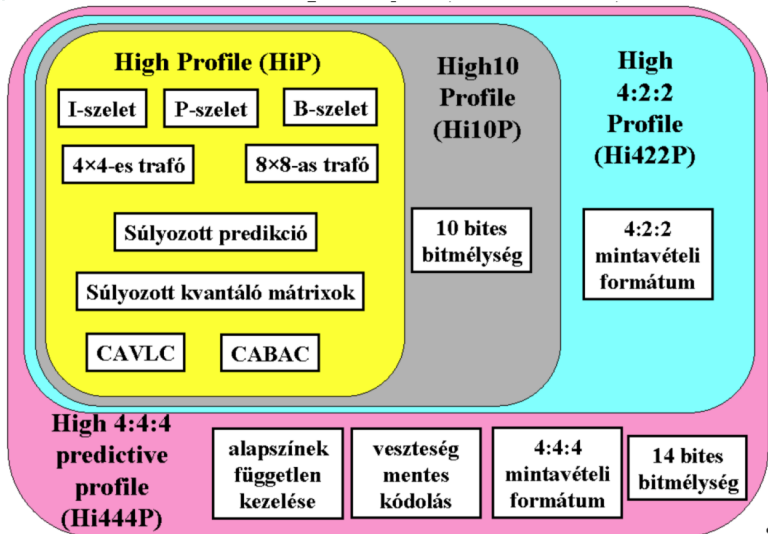


- High Profile (HP): 8 bit/minta, 4:2:0 mintastruktúra, csúcsminőségű fogyasztói célokra, ha nem szükséges a kiterjesztett szín mintastruktúra
- High 10 Profile (Hi10P): 10 bit/minta, 4:2:0 mintastruktúra
- High 4:2:2 Profile (H422P): 10 bit/minta, 4:2:2 mintastruktúra
- High 4:4:4 Profile (H444P): 12 bit/minta, 4:4:4 mintastruktúra, járulékosan támogatja a veszteségmentes kódolást

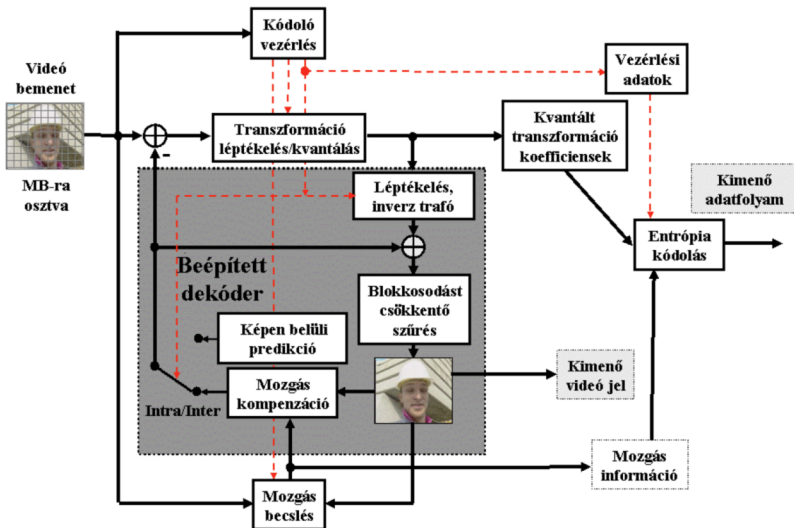
A H.264/AVC HD Intra kiterjesztésének profiljai (pl. High 10 Intra Profile) megegyezik a fenti profilokkal, de csak intra kódolási eszközt tartalmaz, professzionális kamera és editáló alkalmazások (stúdió) céljára



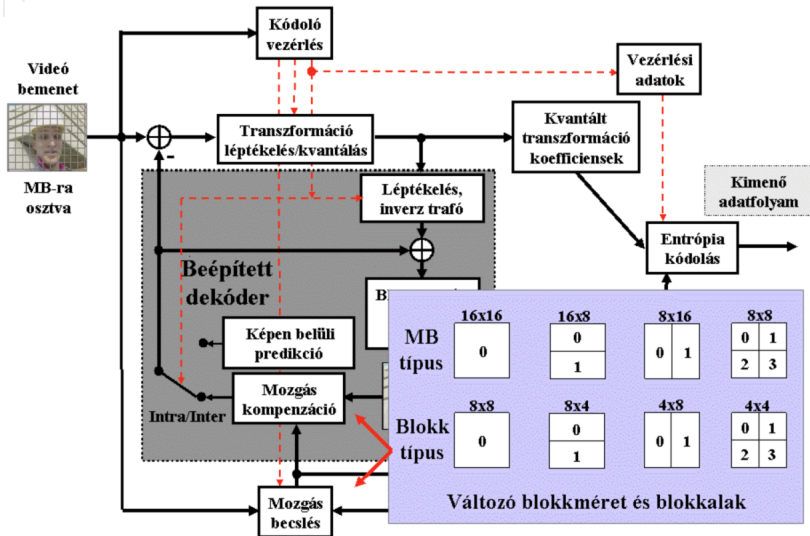
MPEG-4 H.264/AVC FRExt profiljai



MPEG-4 H.264/AVC kódoló vázlat



MPEG-4 H.264/AVC MB méretek





Blokk alapú mozgásbecslés (ismétlés)



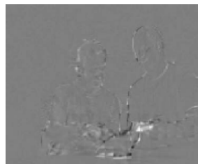
Frame 1



Residual : 16×16 block size



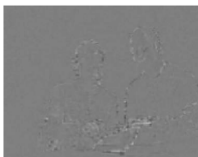
Frame 2



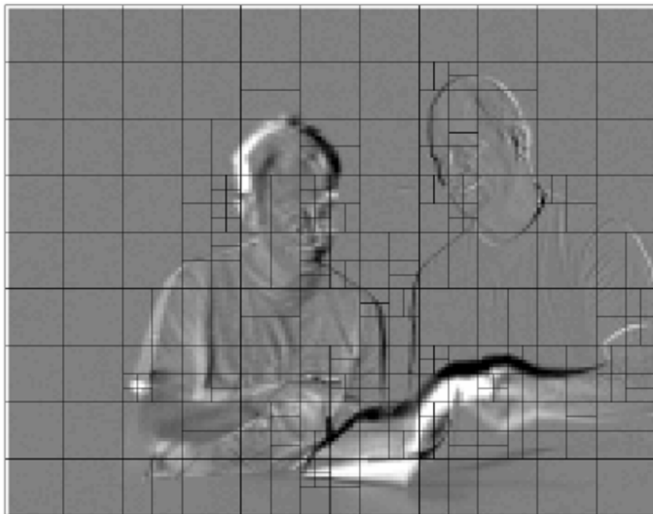
Residual : 8×8 block size



Residual : no motion compensation



Residual : 4×4 block size





- **Mozgáskompenzáció:** A MB-ok 16×16 -os méretűek, azon belül a mozgáskompenzáció szub-MB-on történik, melyek 16×16 , 16×8 , 8×16 , 8×8 , 4×8 , 4×8 és 4×4 méretűek lehetnek
- **Transzformációs blokkméret:** A predikciós hiba kódolása 4×4 -es vagy 8×8 -as (FRExt) blokkokon történik
- **Transzformáció:** Integer-DCT transzformáció a DCT helyett



- **Képfelépítés:** Szekvencia - Kép - Szelet - MB - szub-MB - blokk - minta
- **Blokkosodást csökkentő szűrés:** jelentősen csökkenti a predikciós hiba maradványait, a blokk határokon megjelenő élek simításával. Javul a dekódolt kép minősége, ezáltal az azonos minőséghez szükséges adatsebesség csökkenthető 5-10%-al
- **Kétfajta képtípus:** referencia és nem-referencia kép (predikció szempontjából)
- Képeken belül I,P,B típusú **képszeletek**
- I,P,B típusú MB-ok (pl. lehet intra MB egy P szeleten belül)



- Intra-frame predikció:
 - Intra predikció teljes MB-on, vagy a transzformációs blokkokon (4x4, vagy 8x8 (FRExt) blokkon)
- Inter-frame predikció:
 - Több referenciaképes mozgáskompenzáció
 - B képszeletek használata referenciaként is
 - Súlyozott predikció
 - A mozgásvektorok meghatározása 1/4 ill. 1/8 pixel pontossággal történhet.
 - Különböző méretű szub-MB alapú mozgáskompenzáció
 - Field/frame alapú predikció eldöntése képszinten (Picture-Adaptive Frame/Field coding), vagy MB szinten (Macroblok Adaptive Frame-Field Coding)



- Új entrópia kódolási módok (CABAC, CAVLC) - részletekkel nem foglalkozunk



Integer transzformáció

- Az MPEG-1/MPEG-2 DCT transzformációja idealizált trigonometrikus függvényeket használ.
- A dekódoló beépítése a kódoló hurokba megszünteti a DCT együtthatók újra-kvantálásából származó predikciós hiba-akkumulálódást, de nem garantálható, hogy a tényleges dekóder pontosan állítja elő a DCT bázisfüggvényeit (sin,cos).
- Az MPEG-1/MPEG-2 a DCT bázisfüggvények számítására vonatkozóan hibatoleranciákat definiál



Integer transzformáció

- A H.264/AVC 16 bites fixpontos aritmetikával implementálható transzformációja nem okoz eltérést a kódoló és dekódoló bázisfüggvényei között
- az integer transzformáció mátrixa a 4x4-es 1D DCT transzformációs mátrix megfelelően skálázott, legközelebbi egész értékekre kerekített, majd ortonormált elemeiből áll (közelítő DCT)
- Számítása emellett egyszerűbb, mint a DCT számítása (a 4x4-es transzformáció csak összeadást, kivonást, bit-shiftet igényel)
- A 4x4-es transzformáció kisebb bitsebesség esetén fellépő hibája (nem folytonos blokkhatárok: pl. blokkosodás és ringing) kevésbé látható (granuláris eloszlású), mint az MPEG 8x8-as blokkjai esetében



Integer transzformáció

- A 16x16-os intra predikciós módban a transzformált blokkok DC együtthatóin egy további Hadamard transzformációt is végrehajtanak (ld. később)
- Szintén újdonság, hogy a kvantálási skálafaktor (bitrate vezérlés) nem lineárisan, hanem logaritmikusan növeli a kvantálási lépcsőket (8 bites esetben 52 kvantálási skálafaktor, ahol 6 érték növelés pontosan egy bittel növeli a kvantálási szóhosszt)
- A FRExt profil MB szinten adaptívan válthat a 4x4-es, illetve 8x8-as transzformáció között



4x4-es integer transzformáció mátrixa

A 4x4-es 1D DCT közelítő transzformációs mátrixa

$$\mathbf{T}_{4 \times 4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad \mathbf{D}_{st} = \begin{bmatrix} 1/2 & & & \\ & 1/\sqrt{10} & & \\ & & 1/2 & \\ & & & 1/\sqrt{10} \end{bmatrix}$$

ahol a diagonális \mathbf{D}_{st} mátrix-al való szorzás $\mathbf{T}_{4 \times 4}$ elemeit ortonormálja
 $(D_{st}[i, i] = \frac{1}{\sqrt{\sum_j (T[i, j])^2}})$

Ezek alapján a transzformáció felírható:

$$\mathbf{A}_1 = \mathbf{D}_{st} \mathbf{T}_{4 \times 4} \quad \mathbf{Y} = \mathbf{A}_1 \mathbf{X} \mathbf{A}_1^T$$



4x4-es integer transzformáció mátrixa (folytatás)

A gyakorlatban a \mathbf{D}_{st} diagonális mátrix-al való balról-jobbról szorzást a főátló elemeiből alkotott vektor diadikus szorzatából képzett mátrix-al való elemenkénti szorzással (\star) valósítják meg:

$$\mathbf{Y} = \mathbf{D}_{st} \mathbf{T}_{4 \times 4} \mathbf{X} (\mathbf{D}_{st} \mathbf{T}_{4 \times 4})^T = \mathbf{D}_{st} \mathbf{T}_{4 \times 4} \mathbf{X} \mathbf{T}_{4 \times 4}^T \mathbf{D}_{st}^T = (\mathbf{T}_{4 \times 4} \mathbf{X} \mathbf{T}_{4 \times 4}^T) \star (\mathbf{d}_{st} \mathbf{d}_{st}^T)$$

vagyis

$$\mathbf{Y} = (\mathbf{T}_{4 \times 4} \mathbf{X} \mathbf{T}_{4 \times 4}^T) \star \mathbf{S}_t$$

ahol

$$\mathbf{S}_t = \mathbf{d}_{st} \mathbf{d}_{st}^T$$



4x4-es integer transzformáció mátrixa (folytatás)

A transzformált együtthatók újrakvantálása a következőképpen valósul meg:

$$\mathbf{Y}_q = \frac{1}{2^{15}} \left[(\mathbf{T}_{4 \times 4} \mathbf{X} \mathbf{T}_{4 \times 4}^T) * \mathbf{W}_t \right]$$

$$\mathbf{W}_t = \frac{2^{15}}{Q_{step}} \mathbf{S}_t$$



4x4-es integer inverz transzformáció mátrixa

$$\mathbf{T}_{4 \times 4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1/2 & -1/2 & -1 \\ 1 & -1 & -1 & 1 \\ 1/2 & -1 & 1 & -1/2 \end{bmatrix}$$

$$\mathbf{D}_{si} = \begin{bmatrix} 1/2 & & & \\ & \sqrt{2/5} & & \\ & & 1/2 & \\ & & & \sqrt{2/5} \end{bmatrix}$$

$$\mathbf{A}_2 = \mathbf{D}_{si} \mathbf{T}_{4 \times 4} \quad \mathbf{X} = \mathbf{A}_2^T \mathbf{Y} \mathbf{A}_2$$



4x4-es integer inverz transzformáció mátrixa (folyt.)

Hasonlóan az előzőekhez:

$$\mathbf{A}_2 = \mathbf{D}_{si} \mathbf{T}_{4 \times 4} \quad \mathbf{X} = \mathbf{A}_2^T \mathbf{Y} \mathbf{A}_2$$

$$\mathbf{X} = (\mathbf{D}_{si} \mathbf{T}_{4 \times 4})^T \mathbf{Y} (\mathbf{D}_{si} \mathbf{T}_{4 \times 4}) = \mathbf{T}_{4 \times 4}^T \mathbf{D}_{si}^T \mathbf{Y} \mathbf{D}_{si} \mathbf{T}_{4 \times 4} = \mathbf{T}_{4 \times 4}^T (\mathbf{Y} \star \mathbf{S}_i) \mathbf{T}_{4 \times 4}$$

ahol

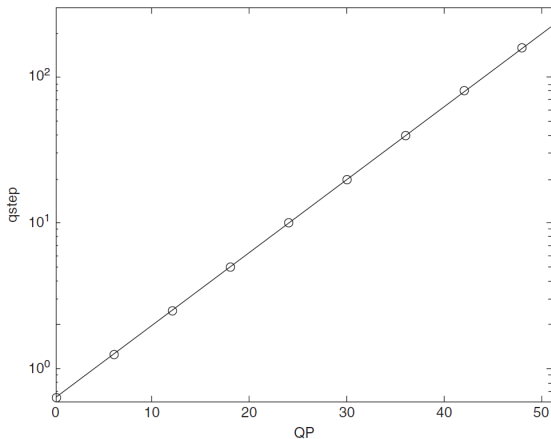
$$\mathbf{S}_i = \mathbf{d}_{si} \mathbf{d}_{si}^T$$

Az inverz kvantálás a következőképpen valósul meg:

$$\mathbf{X} = \frac{1}{2^6} [\mathbf{T}_{4 \times 4}^T (\mathbf{Y}_q \star \mathbf{W}_i) \mathbf{T}_{4 \times 4}]$$

ahol

$$\mathbf{W}_i = 2^6 Q_{step} \mathbf{S}_i$$



Q_{step} : effektív kvantálási tényező

QP : kvantálási paraméter (pl. a bitsebességvezérlés, vagy a felhasználó szabályozza)



Intra predikció

- I képszeleteken belül a teljes MB (16x16 pixel), vagy egy 4x4-es, vagy egy 8x8-as (FRExt) blokk prediktíven kódolt a szomszédos, már korábban kódolt/dekódolt blokkok pixeleiből. A predikció az adott blokk szomszédos pixeleiből történik többféle módon: a predikció iránya lehet vízszintes, függőleges, DC alapú (átlag) vagy síkbeli (lin. függvénykapcsolat)
- 4x4-es intra predikció esetén 9 predikciós irány közül választhat a kódoló, mely irányok a szomszédos pixelek különböző együtthatókkal képzett lineáris kombinációjaként állíthatók elő



Intra predikció

- A H.264/AVC intra predikciója jelentősen megnöveli a tömörítés hatékonyságát szemben az MPEG-1/2-vel, ahol csak a DC komponensek differenciális kódolását alkalmazzák az intra MB-k esetében
- 16x16-os intra predikció esetén a 4x4-es integer transzformációból eredő blokkok DC együtthatóit (16 DC eü.) egy 4x4 blokk alapú Hadamard transzformációval kódolják



Nem-harmonikus bázisfüggvényű

- Valós, ortogonális transzformáció
- 2×2 -es transzformációs mátrixa:

$$\mathbf{H}_{2 \times 2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

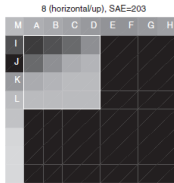
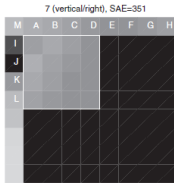
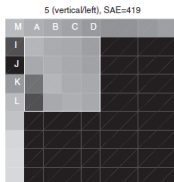
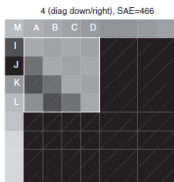
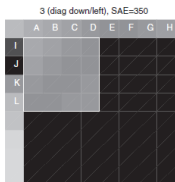
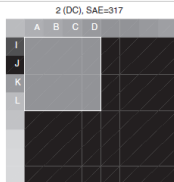
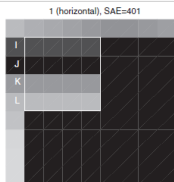
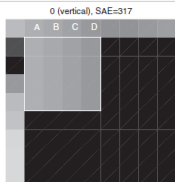
- Nagyobb blokkméretre \mathbf{H} rekurzívan számolható:

$$\mathbf{H}_{2N \times 2N} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{H}_{N \times N} & \mathbf{H}_{N \times N} \\ \mathbf{H}_{N \times N} & -\mathbf{H}_{N \times N} \end{bmatrix}$$

- Az 2D transzformáció kifejezése
 - $\mathbf{y} = \mathbf{H}\mathbf{x}\mathbf{H}^T$ és
 - $\mathbf{x} = \mathbf{H}^T\mathbf{y}\mathbf{H}$

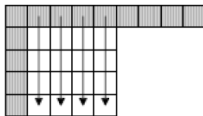


H.264/AVC intra-predikció példák

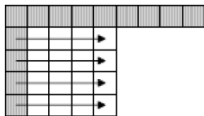


Intra predikció példák

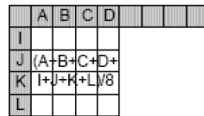
Mode 0: Vertical



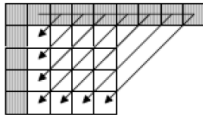
Mode 1: Horizontal



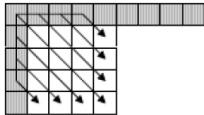
Mode 2: DC



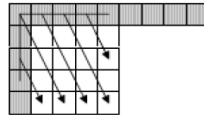
Mode 3: Diagonal-Down-Left



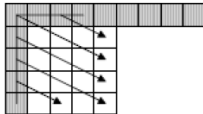
Mode 4: Diagonal-Down-Right



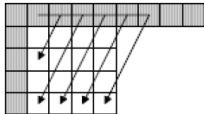
Mode 5: Vertical-Right



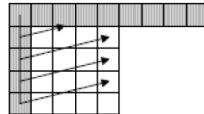
Mode 6: Horizontal-Down





Mode 7: Vertical-Left



Mode 8: Horizontal-Up



 Already coded and reconstructed pixels of neighboring blocks in the same frame

 Pixels of the current 4x4 block



4x4-es intra predikció példák

Intra_4x4_Vertical prediction mode

$\text{pred4x4}[x, y] = p[x, -1]$, with $x, y = 0..3$

Intra_4x4_Horizontal prediction mode

$\text{pred4x4}[x, y] = p[-1, y]$, with $x, y = 0..3$

Intra_4x4_DC prediction mode

$\text{pred4x4}[x, y] = (p[0, -1] + p[1, -1] + p[2, -1] + p[3, -1] + p[-1, 0] + p[-1, 1] + p[-1, 2] + p[-1, 3])$

Intra_4x4_Diagonal_Down_Left prediction mode

$\text{pred4x4}[x, y] = (p[x + y, -1] + 2 * p[x + y + 1, -1] + p[x + y + 2, -1])$

Intra_4x4_Diagonal_Down_Right prediction mode

If x is greater than y ,

$\text{pred4x4}[x, y] = (p[x - y - 2, -1] + 2 * p[x - y - 1, -1] + p[x - y, -1])$

Otherwise if x is less than y ,

$\text{pred4x4}[x, y] = (p[-1, y - x - 2] + 2 * p[-1, y - x - 1] + p[-1, y - x])$



Veszteségmentes MB mód

- Nagy bitsebességek mellett (kis kvantálási lépcsőknél) a transzformációs kódolás, predikció, mozgáskompenzáció adat overheadje miatt a kódoló tetszőleges MB-ra dönthet úgy, hatékonyabb a MB PCM kódolása (csak entrópiakódolás). A PCM makroblokk mód adatsebessége definiálja egyúttal a dekóder számára az adott képméret, képfrekvencia melletti legnagyobb kezelendő bitsebességet is.
- A FRExt definiál továbbá egy transzformációs kódolás nélküli prediktív kódolást+entrópiakódolást is.



Súlyozott predikció és mozgáskompenzáció I.

- Több fajta sub-MB méret alapú mozgásbecslés (minden szub-MB-hoz egyedi mozgásvektorok).
- A H.264/AVC lehetőséget biztosít több referenciakép használatára a P és B típusú predikciókhoz
- Ez azt jelenti, hogy míg az MPEG-2-nél csak az adott prediktívan kódolt képet közvetlenül megelőző I, vagy P kép (illetve kétirányú predikciónál az adott prediktívan kódolt képet közvetlenül megelőző, illetve közvetlenül követő I, illetve P képek) lehet a predikció alapja, a H.264 kódoló több, az adott képet megelőző, illetve követő kép közül választhat (és B kép is lehet referencia !), hogy melyiket használja referencia képként az adott kép mozgáskompenzált prediktív kódolásához.
- A felhasználható referenciaképek számát (a referenciaképtároló méretét) az adott Level szabályozza.
- A referenciaképek kiválasztása viszont MB alapú.



Súlyozott predikció és mozgáskompenzáció II.

- Jelentős újdonság, hogy a referencia kép(ek) tetszőleges skála és offset tényezőkkel súlyozhatók (explicit mód), majd ezen skálázott referenciaképeken hajtódik végre a mozgáskompenzált predikció.
- Implicit súlyozott predikció esetén a referencia képek egymáshoz képesti képtávolsága alapján kerülnek kiszámításra a súlyozótényezők .
- Egyirányú predikció (P képek) esetében ez hatékony kódolási lehetőséget kínál pl. a Fade-in, Fade-out típusú effektek kódolásához, míg kétirányú predikció esetén a crossfade jellegű effektusok hatékony kódolásához



Súlyozott predikció és mozgáskompenzáció III.

- A P és B típusú kép(szelet)ek esetén a mozgásvektorok is differenciálisan kódoltak
- A mozgásvektorok a 3 szomszédos MB mozgásvektorainak mediánjához képest differenciálisan kódoltak, vagy skálázott predikcióval kódoltak
- A kódoló el is hagyhatja a mozgásvektor átvitelét, és utasíthatja a dekódert, hogy az adott blokk környezetének mozgásvektorai alapján becsülje a hiányzó mozgásvektort



De-blocking filter

- A prediktív kódoló hurokba épített deblocking szűrő a blokkosodást csökkenti a szomszédos blokkok pixeleinek térbeli aluláteresztő szűrésével.
- A szűrőegytípusokat a szabvány specifikálja
- Természetesen a szűrt képek a további predikció alapja (predikciós hibaterjedés ellen)
- A szűrő erősségét a MB típusa (I,P,B), a kvantálási súlyozótényező, és egyéb tényezők szabályozzák
- Alapvetően minél nagyobb kvantálási lépcsővel történik a MB kódolása, annál erősebb a szűrő hatása



QP=32, no filter

QP=32, with filter



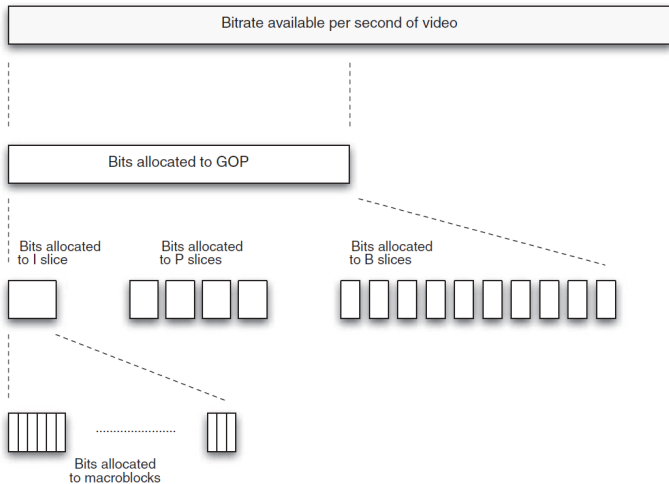


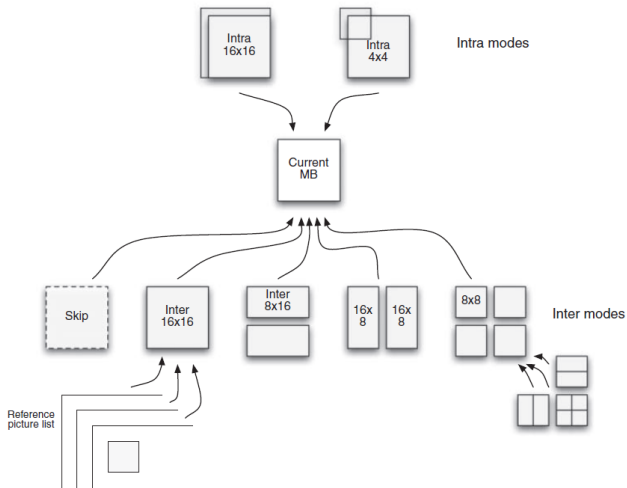
Új színtér: YCgCo (FRExt)

- Az YCbCr komponensek kiszámítása pl. az RGB komponensekből lineáris transzformációt igényel, melyben az együtthatókkal való szorzás lebegőpontos műveleteket igényel, ez nem mindig hajtható végre ideális pontossággal
- Az integer transzformáció alapú kódolás esetén ezen lebegőpontos YCbCr komponensek kódolása elkerülhetetlenül kerekítési hibák forrása
- Az új színtér egy új Y komponens mellett a Cg (green chroma) és Co (orange chrome) komponensekből áll
- Számításuk : $Y = \frac{1}{2} \left(G + \frac{(R+B)}{2} \right)$, $Cg = \frac{1}{2} \left(G - \frac{(R+B)}{2} \right)$,
 $Co = \frac{(R-B)}{2}$



Bitkiosztás, kódolási módok választása I.







- Aktuális kép interpolálása $1/4$ ($1/8$) pixel pontosságra
- Inter-predikciós mód választása az alábbiak közül:
 - Referencia képek kiválasztása
 - P vagy B típusú predikció választása (B esetén, predikciós irány megválasztása)
 - sub-MB méret megválasztása
 - Súlyozott predikció ? (kódoló által meghatározott, vagy képtávolság alapján ?)
- Mozgásvektorok meghatározása
- Mozgásvektorok predikciója (direkt mód ? mozgásvektor elhagyása)
- Anti-blocking filter alkalmazása



Rate-Distortion Optimization (RDO)

- $J = D + \lambda R$
- J : költség (cost)
- D : hiba/torzítás mérték (pl. MAD) - (distortion)
- R : bitmennyiség (rate)
- λ : Lagrange-multiplikátor
- PI: $\lambda = 0.852^{(QP-12)/3}$
- Az adott MB minden lehetséges kódolási módjára:
- Az MB kódolása az adott kódolási móddal, R (bitigény) kiszámítása
- MB dekódolása, az eredeti és a dekódolt MB közötti hiba (D) kiszámítása
- J meghatározása az adott kódolási módra
- Az összes lehetséges mód kiértékelése után a minimális költséget (J) eredményező mód választása a kimeneti bitfolyamban



Keresési opciók csökkentése:

- korábbi képek, illetve egyéb statisztikák felhasználása,
- képi textúra elemzés a predikciós blokkméret meghatározásához,
- DCT együtthatókból következtetés az optimális intra-predikciós irány választására,
- költség-monotonitási feltételezések felhasználása a blokkméret választásnál
- stb ...