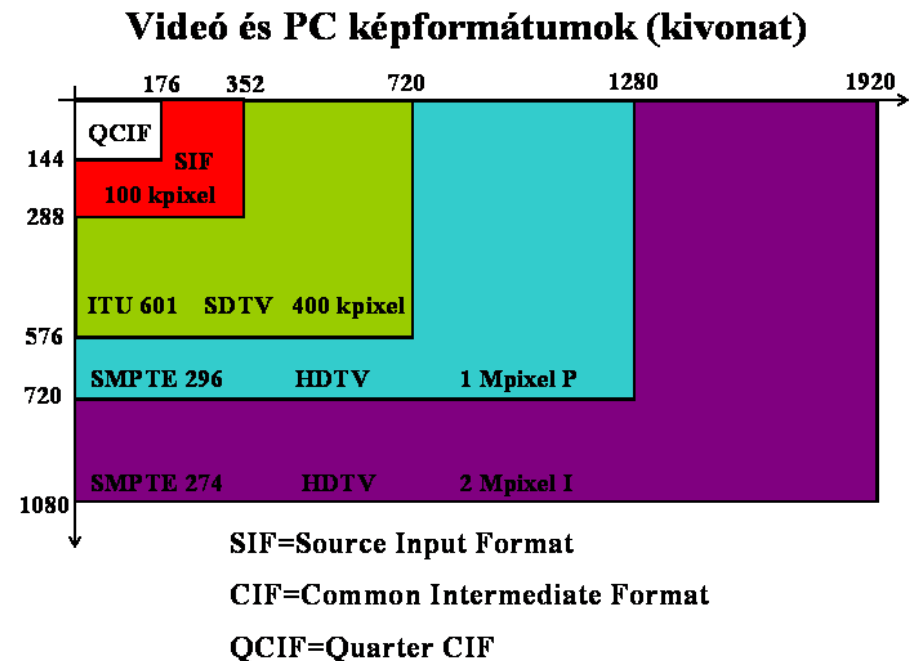


A videojel bitsebesség csökkentési eljárásai



Bitsebességek (európai), 8 bit / komponens:

- ITU-601/4:2:2: $(576 \times 720 + 2 \times 576 \times 360) \times 8 \times 25 = 166\text{Mbit/s}$
- ITU-601/4:2:0: $(576 \times 720 + 2 \times 288 \times 360) \times 8 \times 25 = 124\text{Mbit/s}$
- CIF: $(288 \times 360 + 2 \times 144 \times 180) \times 8 \times 30 = 37,3\text{Mbit/s}$
- QCIF: $(144 \times 180 + 2 \times 72 \times 90) \times 8 \times 30 = 9,3\text{Mbit/s}$

Szükséges bitsebességek és alkalmazások:

- ITU-601/4:2:2 : 25 - 50 Mbit/s, stúdiótechnika, MPEG-SP
- ITU-601/4:2:0 : 3 - 6 Mbit/s, műsorszórás, MPEG-2
- CIF : 64 - 1920 kbit/s, videokonferencia, H.261/3
- QCIF : 64 - 256 kbit/s, videotelefon, H.261/3
- SIF : 1,15 Mbit/s, videó CD, MPEG-1

Bitsebesség csökkentés kell, de milyen alapon

- A természetes mozgókép redundáns, statisztikai redundancia: pixelek adott környezetben belül és a képek adott időintervallumon belül hasonlóak
- Észlelési redundancia: a videojel a HVS számára lényegtelen részeket is tartalmaz, amely eltávolítható

A mozgóképben rejlő redundanciák

- Térbeli, ami lehet képen belüli (Intra-frame) és félképen belüli (Intra-field) hasonlóság
- Időbeli, képek közötti (Inter-frame) hasonlóság

Az emberi látás tulajdonságai (kompresszió)

- Világosságjelre látásunk felbontó képessége 3-5-ször nagyobb mint a színekre
- Álló és lassan változó képekre az emberi látás felbontás igénye lényegesen nagyobb, mint a gyorsan változó képszekvenciákra
- Hirtelen képváltás, vagy gyors mozgás esetén kevésbé zavaró a gyengébb képminőség
- Álló- és lassan változó képek esetén erősen érzékeny a képminőség változásra (térben és időben)
- Ellentétben a hanggal, a látásunk sokkal kisebb jel-zajviszony esetén is jónak tartja a képet (szubjektív/objektív képminőség)

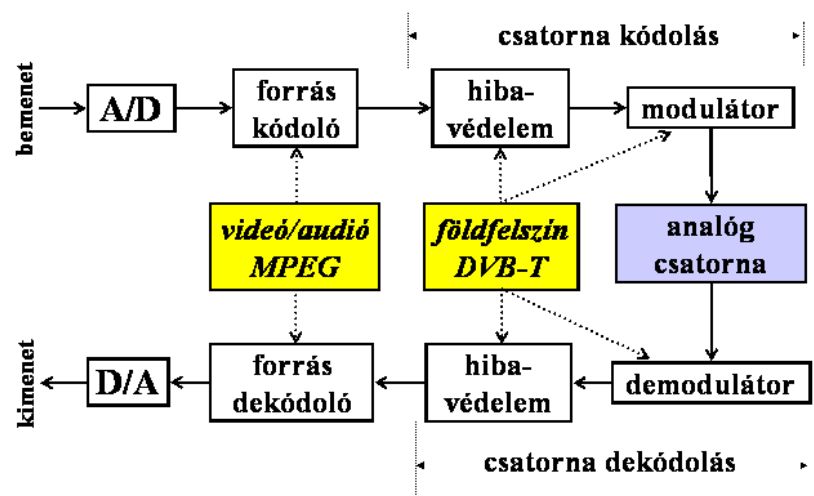
Bitsebesség (bitrate) definíciók:

1. **Bitsebesség** = a kódolt kép bitszáma osztva a képen belüli pixel számmal [bit/pixel, bpp], (döntően álló képre)
2. **Bitsebesség** = másodpercenként továbbított bitek száma [bit/sec], (mozgó képre)

Kódolási technikák:

- Veszteségmentes:
 - változó szóhosszúságú kódolás
 - futamhossz kódolás
- Veszteséges:
 - alul-mintavételezés
 - kvantálás, stb.

Digitális csatorna-modell

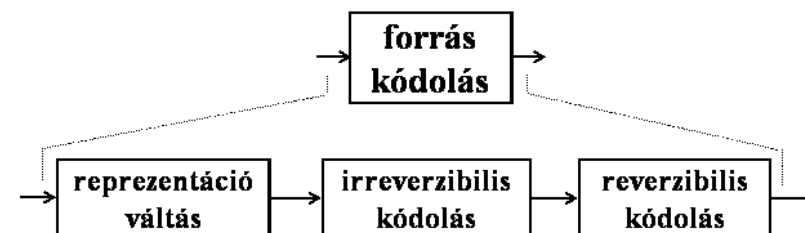


Forráskódolás

- Figyelembe vesszük a forrás és a nyelő tulajdonságait, eltüntetjük a forrás redundanciáját, a nyelő számára szükséges információt hagyjuk meg, cél: nyelő "ne vegyen észre semmit"

Csatornakódolás

- A csatorna átviteli tulajdonságait figyelembe véve választunk hibavédelmi algoritmust és modulációs eljárást



Reprezentáció váltás

- Új ábrázolási síkban kevesebb redundancia, formái: **predikció**, **transzformációs kódolás**, **mozgás kompenzáció**

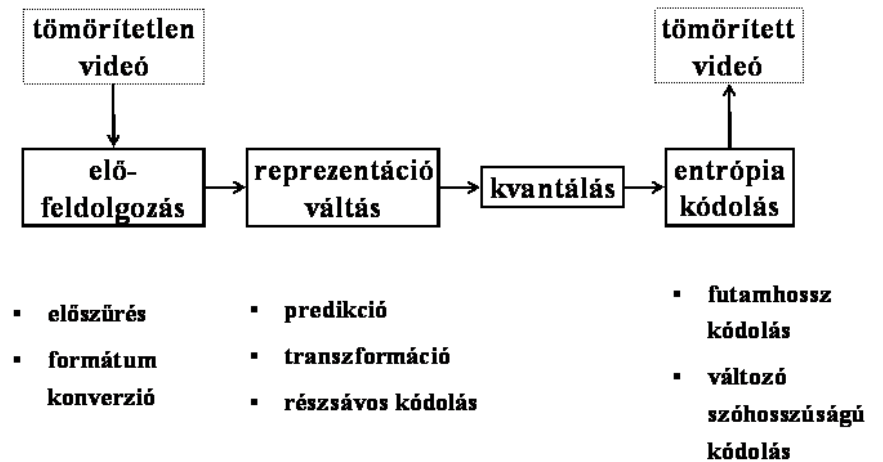
Irreverzibilis kódolás

- A pontosság csökkentése, a „lényegtelen” részek eltávolítása, megjelenik a kvantálási hiba, az elérhető kompressziós tényező nagy, formái: **kvantálás**, **térbeli**, **időbeli alul-mintavételezés**

Reverzibilis kódolás

- Hatékony kódszó kijelölés, limitált kompressziós tényező, a jel reprezentáció redundanciáját csökkenti, formái: **változó szóhosszúságú kódolás** (Variable Length Coding: VLC), **futamhossz kódolás** (Run-Length Coding: RLC)

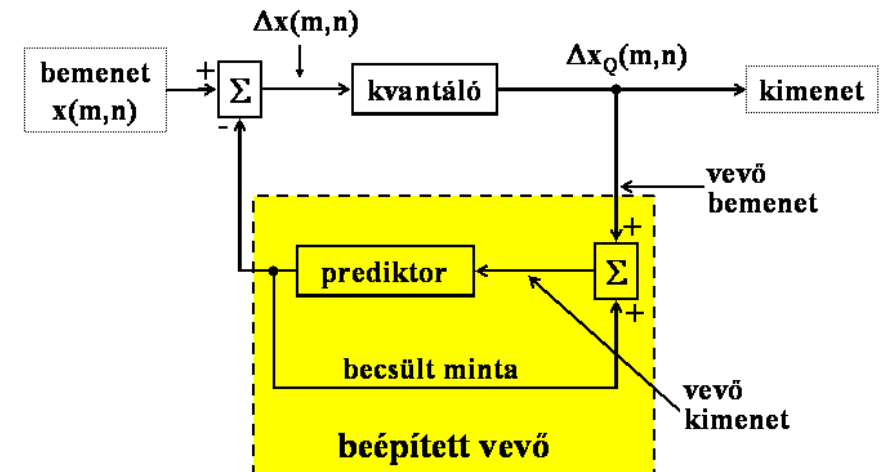
Általános forráskódoló videós megközelítése



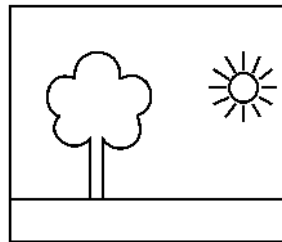
Prediktív kódolás

- Alkalmazható egy adott kép pixelein (térbeli predikció), illetve mozgóképek esetén az egyes képek között (időbeli predikció)
- Alapelv: az $x(n,m)$ minta közvetlen kódolása helyett egy becsült mintához képesti különbséget kódoljuk
- A becsült mintát az $x(n,m)$ minta környezetében lévő minták valamilyen súlyozott összegéből állítjuk elő
- Figyelembe kell venni, hogy a vevő oldalon csak a különbségi minták állnak a rendelkezésünkre
- Ezért a kódoló tartalmazza a feltételezett vevő oldali dekódert
- És a becslést is a vissza-dekódolt minták alapján végezzük

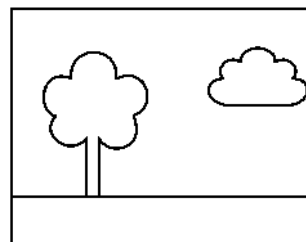
A predikciós kódoló tömbvázlata



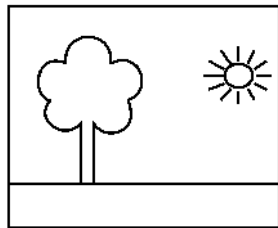
Képek közötti becslés (időbeli, inter)



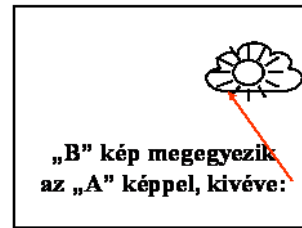
forrás „A” kép



forrás „B” kép



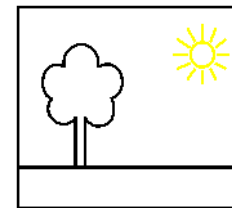
„A” kép önmagában kódolt



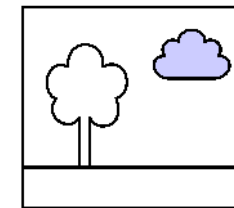
„B” képet becsljük az „A” képből

„B” kép megegyezik az „A” képpel, kivéve:

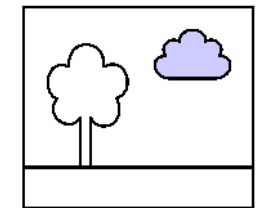
Kétirányú képek közötti becslés



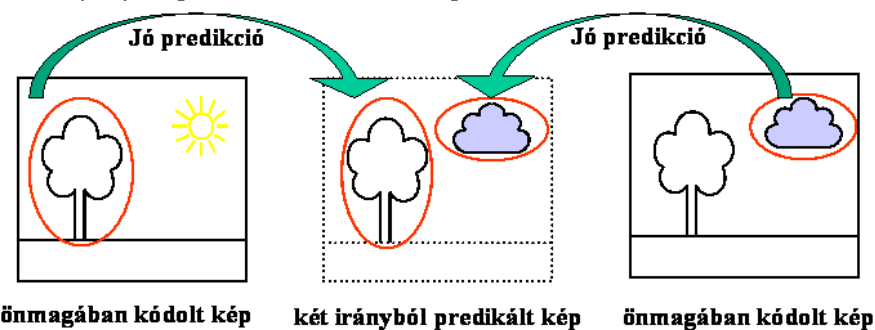
(N-1). kép



N. kép



(N+1). kép



Transzformációs kódolás

- A természetes képek pixelai közötti korreláció egy $N \times N$ -es pixel blokkot tekintve igen nagy
- A transzformációs kódolás ennek megfelelően blokk alapú
- Cél a jel felírása „alkalmasan” megválasztott koordináta rendszerben, melyben az hatékonyabban ábrázolható
- A transzformáció legtöbbször valamilyen mátrix művelet
- A mintákat nem átlapolódó szegmensekre osztjuk, majd a transzformáció mátrixával transzformáljuk
- A transzformáció után kapott együtthatók közötti korreláció általában kisebb, mint az eredeti minták közötti
- A kódolás további lépéseit e tartományban végezzük

Transzformáció típusok

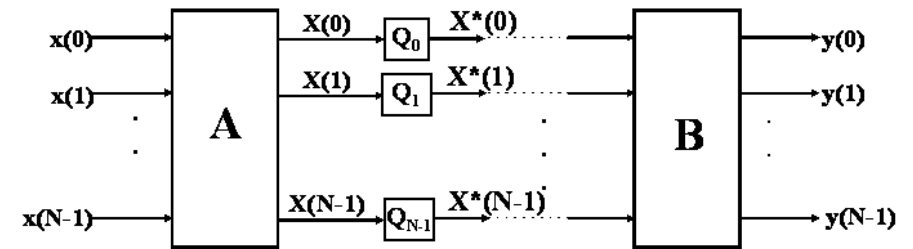
- Ha az együtthatók közötti korreláció a transzformált tartományban zérus, akkor a transzformáció optimális
- Gyakorlatban a fix bázisú transzformációkat alkalmazzuk
- Ezek csak szub-optimálisak, de sokkal kevesebb számítást igényelnek, ezért a valós idejű jelfeldolgozáshoz alkalmasabbak

Harmonikus bázisfüggvényű transzformációk

- DFT (Diszkrét Fourier Transzformáció)
- DCT (Diszkrét Koszinusz Transzformáció)
- DST (Diszkrét Szinus Transzformáció)

A tömörítés megvalósulása

- A transzformáció önmagában véve veszteségmentes
- Az oda- és vissza-transzformáció az eredeti jelet adja vissza
- Azonban míg a tértartományban általában minden elem kb. egyforma fontossággal járul hozzá a jelenergiához, addig a transzformált tartományban jóval kevesebb együttthatóba tömörül az energia nagy része
- A transzformáció "jóságát" éppen az dönti el, hogy a kapott együttthatók közül mennyi jelentős
- Általában a szegmens jó reprezentációját kapjuk, ha az inverz transzformációt csak a fontosabb együttthatókkal végezzük
- Ekkor az oda-vissza transzformáció után kapott kimeneti szegmens általában eltér a bemenetitől, az eltérés a torzítás



A képfeldolgozásban a DCT-t preferálják, mert:

- természetes képeket "majdnem" tökéletesen dekorrelálja
- közel optimális
- a jel energiáját a DC és néhány AC koeficiens hordozza
- a blokkon belüli korreláció jelentősen csökken
- számítástechnikailag attraktív, gyors algoritmusok

2-D DCT és inverz DCT (IDCT)

Kiszámítása:

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

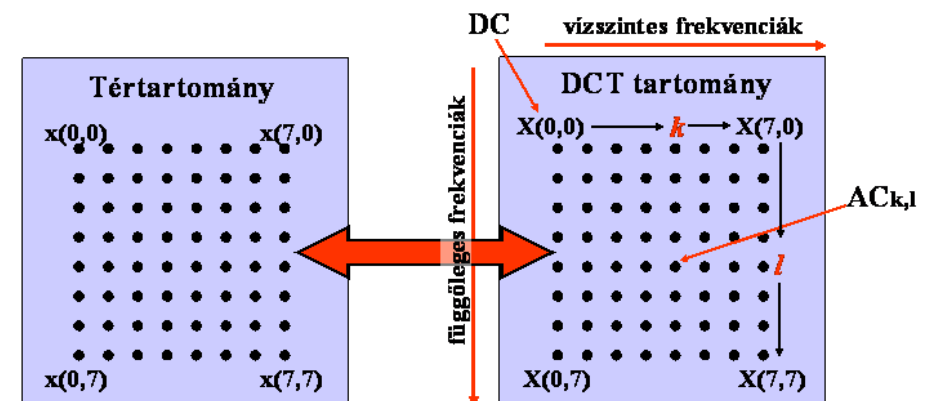
$$x(m, n) = \frac{2}{N} \sum_{l=0}^{N-1} \sum_{k=0}^{N-1} \alpha(k) \alpha(l) X(k, l) \cos\left(\frac{(2m+1)k\pi}{2N}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right)$$

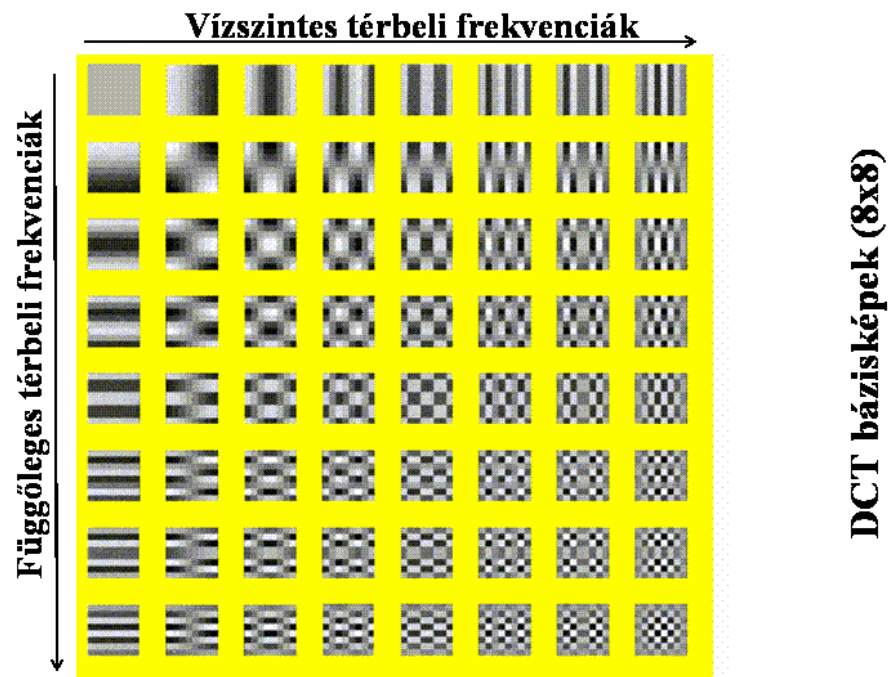
$$\text{És: } \alpha(k), \alpha(l) = \begin{cases} \frac{1}{\sqrt{2}} & \text{ha } k=0, l=0 \\ 1 & \text{ha } k \neq 0, l \neq 0 \end{cases} \quad k, l = 0, 1, \dots, N-1$$

Elnevezések: a DC koeficiens: $X(0,0)$

az AC együttthatók: $X(k,l)$ ha k és l nem nulla.

Szemléltetés





Tipikus DCT transzformáció

83	87	92	90	89	91	47	95	DCT →	284	18	-12	2	3	-7	0	0
98	81	77	96	71	44	58	49		10	9	8	1	-4	1	0	0
27	43	65	40	64	99	61	55		-6	2	1	0	2	4	2	-1
83	21	45	51	59	80	48	63		2	1	-1	1	-3	0	2	0
87	94	56	62	41	74	75	57		0	0	0	1	2	0	0	-1
98	82	68	79	54	46	52	60		1	-2	1	0	1	-2	3	0
72	42	70	84	69	50	97	67		0	0	0	-1	1	2	0	0
76	53	85	88	73	66	78	86		0	0	0	0	0	0	0	0

Együtthatók kvantálása

- A kvantálás blokkonként ún. kvantálási mátrixszal történik
- Az alábbi képlet alapján:

$$X^*(k,l) = \frac{X(k,l)}{W(k,l)}$$

- ahol $W(k,l)$ jelenti a (k,l) pozícióbeli kvantálási lépcsőt
- A mátrix elemeit alapvetően kétféle módszerrel lehet megválasztani:
 - jel-zajviszonyra optimalizáltan
 - az emberi látórendszer (HVS: Human Visual System) tulajdonságainak megfelelően

Kvantálás a HVS-nek megfelelően

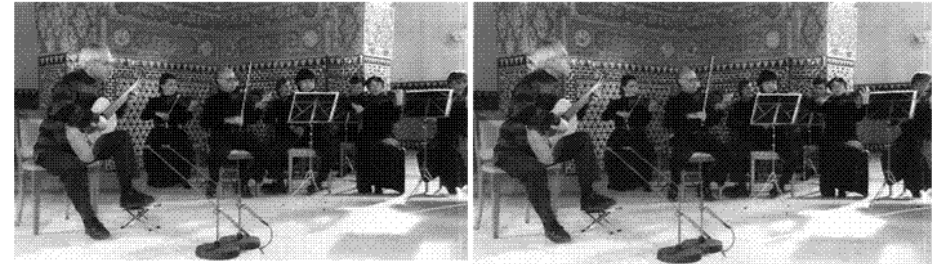
- A HVS nem képes észrevenni apró részleteket
- Ezért a fontos együtthatókat (DC, kisfrekvenciás) finomabb, a kevésbé fontosakat (nagyfrekvenciás) durvábban kvantáljuk
- Egy tipikus kvantálási mátrix pl. a következőképpen néz ki

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

- A bitsebesség $W(k,l)$ konstanssal való szorzásával állítható

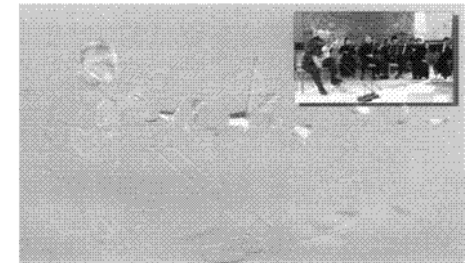
Mozgásbecslés (csak szemléltetés)

Képek közötti differencia



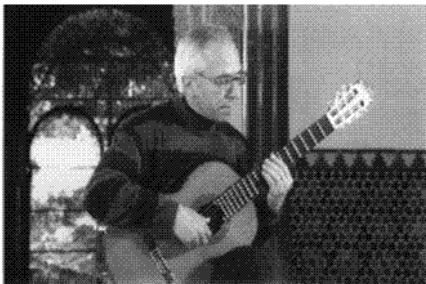
N. kép

(N+1). kép



N. kép - (N+1). kép

Képek közötti differencia



N. kép

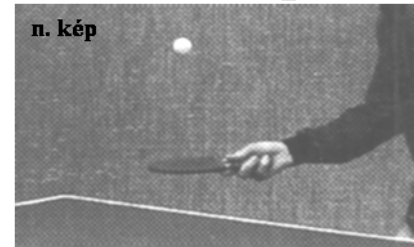


(N+1). kép

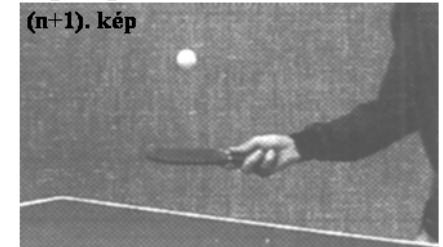
N. kép - (N+1). kép



Mozgásbecslés alapú predikció



n. kép



(n+1). kép



Képek közötti különbség
mozgás becslés nélkül



Képek közötti különbség
mozgás becsléssel

- Hatékonyság a mozgás meghatározás pontosságától függ
- A mozgás pontos meghatározása nehéz és idő igényes feladat

MPEG (videó) kódolások

MPEG: Motion Picture Experts Group

Az MPEG szabványok célja:
a videó és audio digitális, bitsebesség csökkentett
formában történő hatékony ábrázolása

A szabványok felépítése

- 1. rész a rendszer specifikáció: a kódolt videó, audio és adat összefűzésének szintaxisa a szinkronizált visszajátszáshoz
- 2. rész a videó specifikáció: a bitsebesség csökkentett videó bitfolyam szintaxisa és a videó modell dekóder specifikációja
- 3. rész az audio specifikáció: a bitsebesség csökkentett audio bitfolyam szintaxisa és az audio modell dekóder specifikációja
- 4. rész a Conformance test (tesztelés/megfelelőség)
- További részek egyéb kiegészítések (MPEG-2 esetében AVC)
- **A szabványok nem specifikálják a kódolókat!**
- **Csak a kimeneti adatfolyamok bitszintaxisa és a modell dekódoló specifikált**

MPEG fázisok és alkalmazások

- MPEG-1 (ISO 11172 /1993/): Alacsony bitsebesség, multimédia, kb. 1,5 Mbit/s, SIF képminőség, csak progresszív
- MPEG-2 (ISO 13818 /1994-95/):
 - Műsorszórás (DVB), 2-6 Mbit/s
 - Stúdiótechnika, 25-50 Mbit/s újrafeldolgozási minőség
 - Digital Versatile Disc (DVD), 3-7 Mbit/s PAL-nál jobb
 - Általában ITU-601 4:2:0, váltott-soros
- MPEG-4 (ISO/IEC 14496):
 - Alacsony bitsebességű kódolások (64 kbit/s) szabványának indult (1994), de az interaktív multimédia objektum orientált szabványává vált, később AVC

MPEG-1 szabvány felépítése

ISO/IEC 11172

- Information technology; Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s
- ISO/IEC 11172-1, 1993: Systems
- ISO/IEC 11172-2, 1993: Video
- ISO/IEC 11172-3, 1993: Audio
- ISO/IEC 11172-4, 1995: Conformance testing
- ISO/IEC 11172-5, 1995: Software simulation

MPEG-2 szabvány felépítése

ISO/IEC 13818

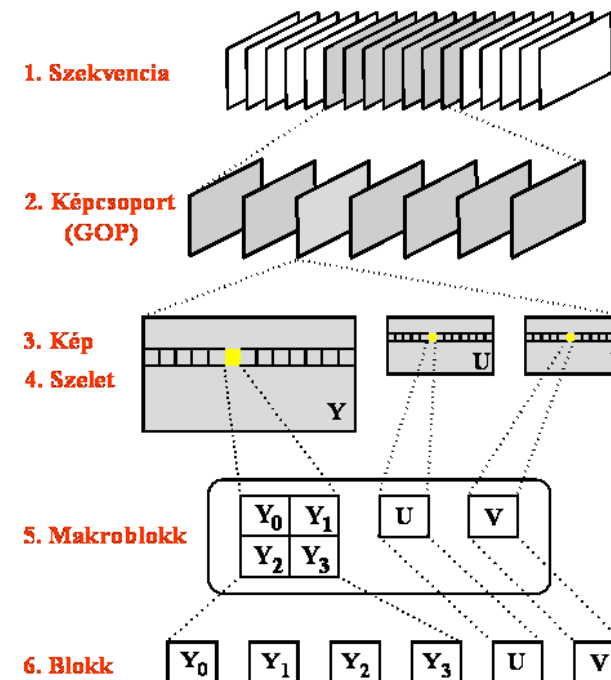
- Information Technology; Generic Coding of Moving Pictures and Associated Audio Information
- ISO/IEC 13818-1, Systems
- ISO/IEC 13818-2, Video
- ISO/IEC 13818-3, Systems
- ISO/IEC 13818-4, Compliance testing
- ISO/IEC 13818-5, Software simulation
- ISO/IEC 13818-6, DSM-CC
-
- ISO/IEC 13818-10, AVC

MPEG videókódolás

- Kifinomult, DCT alapú mozgáskompenzációt használó hibrid kódolás
- A kódolás és dekódolás számításigénye különböző, a rendszer tehát aszimmetrikus
- MPEG videókódolás egyik alapjellemzője a réteges szerkezete
- A rétegszerkezet 6 egymásba ágyazott egységet tartalmaz, melyekben az alsóbb rétegek általában nem dekódolhatók a felsőbb szintek nélkül
- A felső 4 egység egyedi startkóddal rendelkezik, így kezdetük dekódolás nélkül megtalálható

Az MPEG videókódolás rétegszerkezete

- **1. Szekvencia réteg:** a kódolt szekvenciát azonosítja, a fejléc tartalmazza a rendszeradatokat (képméret, bitsebesség, stb.)
- **2. Képcsoport réteg (GOP=Group Of Pictures):** legalább egy önmagában kódolt (I) képet tartalmazó, bizonyos számú kép együttese, a véletlen hozzáférés egysége
- **3. Képréteg:** egy kép kódolt adatait tartalmazza
- **4. Szelet (slice) réteg:** MB-ok sorfolytonos csoportja, az újraszinkronizáció egység, a legalsó szint, amelyen a dekóder még képes feléledni bithiba esetén
- **5. Makroblokk réteg:** általában az Y 16x16-os, és a C_R, C_B 8x8 blokkjaiból áll, a mozgás-kompenzáció egysége
- **6. Blokk réteg:** a MB 8x8-as blokkjai, a DCT kódolás egysége

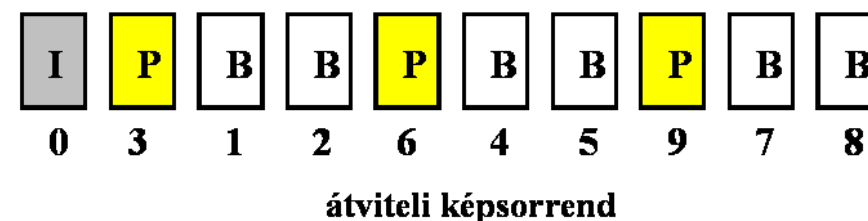
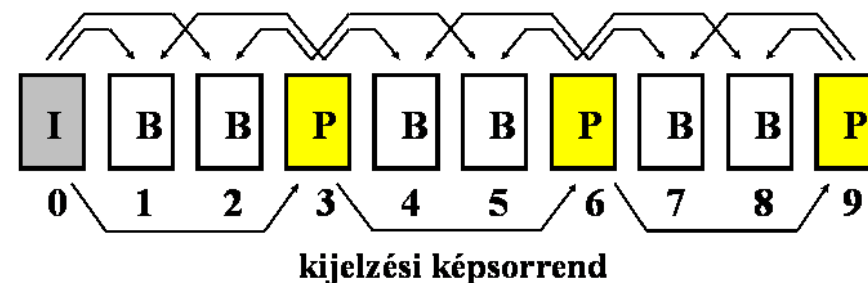


MPEG-videó réteg-szerkezete

MPEG képtípusok

- **I (Intra coded):** önmagában kódolt kép, a dekódoláshoz szükséges minden adatot tartalmaz
- **P (Predictive coded):** a múltból prediktíven kódolt kép, referenciája egy előző I vagy P kép, a dekódoláshoz a referencia szükséges
- **B (Bidirectionally coded):** két-irányból kódolt kép, referenciája az előző I/P (múltbeli referencia) és a következő I vagy P kép (jövőbeli referencia), a B kép nem lehet referencia
- Az I képek a legkevésbé, B képek a legjobban tömöríthetők

Átviteli és kijelzési képsorrend



MPEG képsorrend

- A műsorszórás gyakori kijelzési képsorrendjei:
 - Európa: IBBPBBPBBPBB (Long GOP)
 - Amerika: IBBPBBPBBPBBPBB (Long GOP)
- Stúdiótechnika kijelzési sorrendje: csak I képek (only I)

MB típusok és MB szintű predikció

- I képekben: minden MB önmagában kódolt, intra típusú
- P képekben: a MB lehet önmagában, vagy prediktíven kódolt
- B képekben: a MB lehet csak a múltból, csak a jövőből, a múltból és a jövőből jószolt, illetve önmagában kódolt

MPEG bitsebesség vezérlés

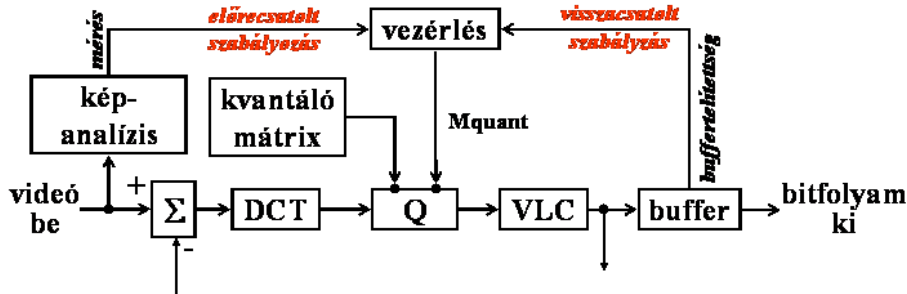
Konstans bitsebességű kódolás (CBR)

- A kódoló kimenetén egységnyi idő alatt keletkező bitek száma tág határok között változik, hiszen:
 - A különböző típusú képek (I, P, B) különböző bitszámmal kódolhatók még azonos minőség mellett is
 - Egy képen belül is változhat a képtartalom, a bonyolultabb struktúrájú területek leírásához több bitre van szükség
 - A változó szóhosszúságú kódolás is szerepet játszik a pillanatnyi bitsebesség ingadozásában
- A legtöbb átviteli csatorna **konstans bitsebességű** forrást igényel
- A probléma pl. a kódér kimenetére helyezett pufferrel oldható meg, melyből a csatorna fix órajellel olvassa a biteket

Változó bitsebességű kódolás (VBR: Variable Bit Rate)

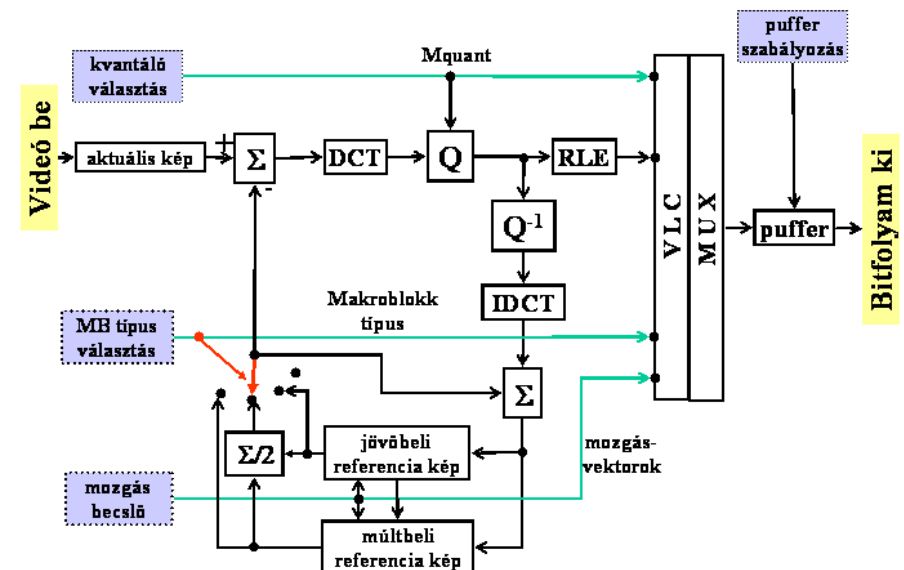
- Két alapvető típusa létezik:
 - **nyílthurkú** VBR: nincs puffer szabályzás
 - **visszacsatolt** VBR: van puffer szabályzás
- Nyílthurkú VBR, ha a bitfolyam olvasás vezérelhető eszközzel (pl. DVD) történik, ekkor puffer végtelen nagy
- Kétféle nyílthurkú VBR, attól függően, hogy a kvantálós skálafaktort, vagy a szubjektív képminőséget tartjuk állandón
 - **Konstans skálafaktorú kódolás** a kvantálási zajt egyenletesen teríti a képen (nem feltétlenül a legjobb képminőség)
 - **A képminőség konstanson tartásához** a skálafaktort a képtartalomhoz kell igazítani (DVD)

- A **visszacsatolt VBR** kódolás lényege, hogy a rendkívül változó pillanatnyi bitsebességet puffereléssel kisimítjuk mielőtt a bitfolyam a VBR csatornára kerül
- Így a sebesség a definiált csúcsebesség alatt marad
- A VBR kódolás kb. 20-30%-kal kisebb *átlagos* bitsebességet eredményez, mint az ugyanazon képminőségű CBR kódolás
- Kifinomult CBR kódolással azonban csökkenthető a hatékonyságbeli különbség:
 - Előreccatolt (feed-forward) bitsebesség szabályozás: globális (kép szintű) és lokális (MB szintű) szabályzás komplexitás analízissel.
 - Visszacsatolt (feed-backward) bitsebesség szabályozás: lokális szabályzás a buffer telítettség figyelésével.



- **Lépések:**
 - GOP-ra elosztható bitek száma adott, a cél bitszám meghatározása I, P, B képekre (globális képminőség)
 - A kvantálók tényezők meghatározása: MB komplexitástól és textúrától függő referencia kvantáló felvétele (lokális minőség)
 - A kvantáló módosítása a puffer telítettségnek megfelelően

MPEG videókódoló felépítése



MPEG-1 videókódolás alapparaméterei

- 4:2:0-ás mintastruktúra
- Képfrekvenciák: 24, 25, 29.97, 30, stb. kép/sec)
- Képméret maximálisan 4095x4095
- Letölthető kvantálási mátrixok
- I, P, B képek
- SIF (352x288x25, vagy 352x240x30) méret ajánlott
- Csak progresszív képek kódolása
- Elérhető tömörítés jó minőségű kódolás mellett: 25-30

MPEG-2 videókódolás

- Az MPEG-2 videó az MPEG-1-re épül, vele felülről kompatibilis
- Legfontosabb jellemzői:
 - Váltott-soros videó kezelés, kép/félkép alapú feldolgozás
 - Profile - level szerkezet
- Teljes képes (frame) kódoláskor két félképet együtt, míg félképes (field) kódoláskor a félképeket egymástól függetlenül dolgozza fel (DCT és mozgásbecslés is)
- Egy videó szekvencián belül a kép és félkép majdnem tetszőlegesen keverhető

Profile-level szerkezet

- Az MPEG-2 különböző osztályokat definiál, melyek meghatározzák az adott bitfolyamok dekódolásához szükséges dekóder képességet és kapacitást
- A **profile**-ok a mintavételi formátumokat, a képtípusokat, és az egyéb kódolási eszközöket definiálják és korlátozzák
- A **level**-ek a képméretet, a bitsebességet, a képfrekvenciát definiálják és korlátozzák
- A level-ek a következő alkalmazásokat célozzák meg:
 - Low : videokonferencia
 - Main : SDTV
 - 422P@ML : stúdiótechnika
 - High : HDTV

Level-ek

	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

Profile-ok

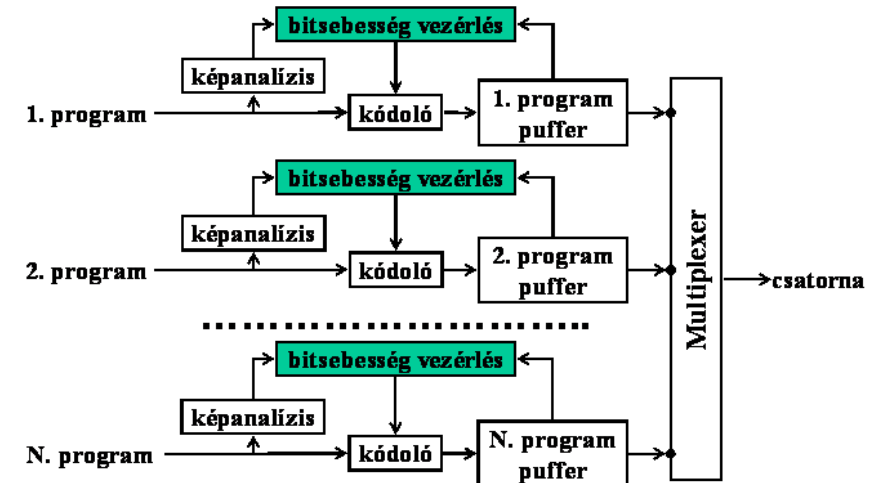
	Simple	Main	SNR	Spatial	High	422P
YUV	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0 4:2:2	4:2:0 4:2:2
Képtípus	I, P	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B
Skálázhatóság	nem	nem	igen	igen	igen	igen

MPEG-2 alkalmazhatósági felületei

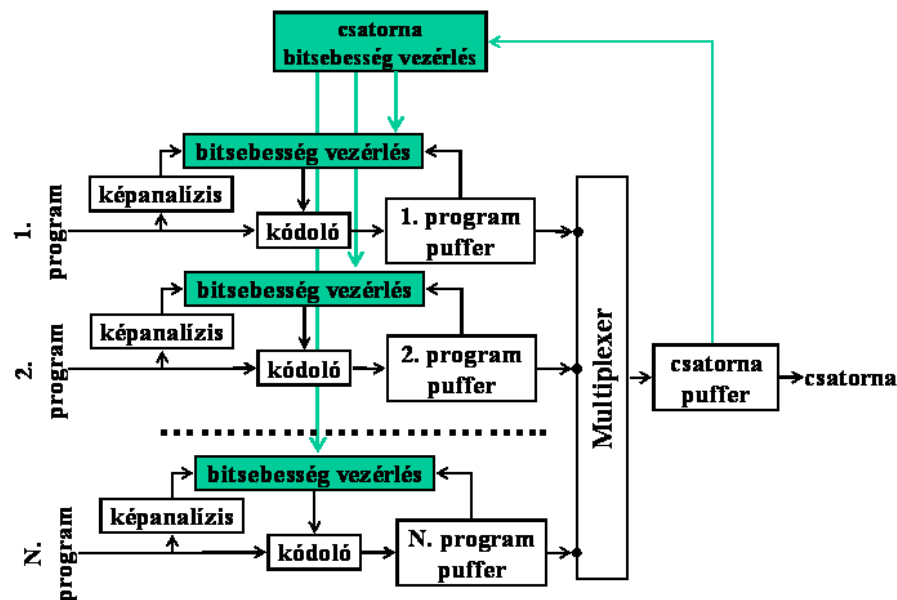
High-1920 level					AVC	
High-1440 level					VC-9	
Main level		720x576 15 Mbit/s				720x608 50 Mbit/s
Low level						
	Simple profile	Main profile	SNR scalable profile	Spatially scalable profile	High profile	422 profile

Programok független kódolása és multiplexelése

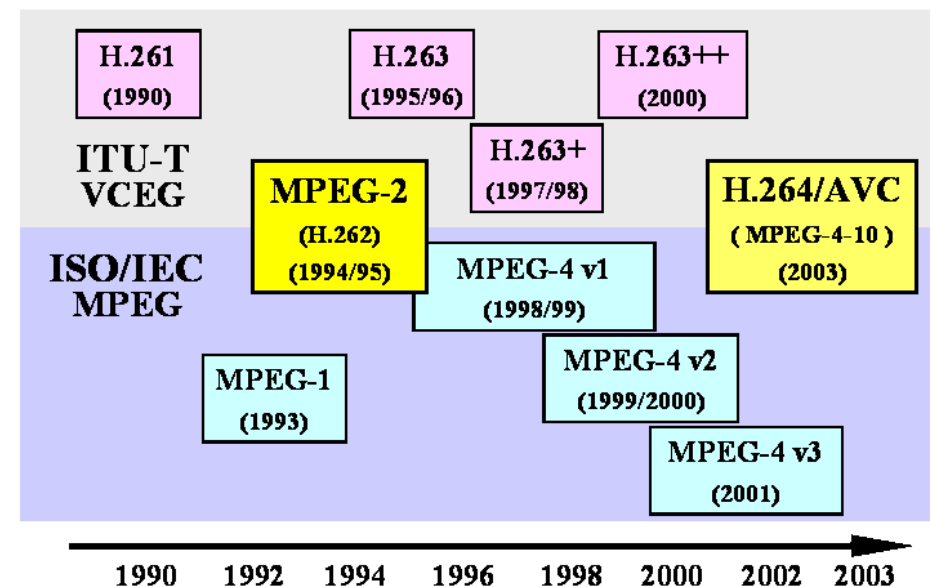
- Minden program előre definiált adatátviteli sebességet kap
- A bitsebesség vezérlés programonként külön-külön



A statisztikus kódolás és multiplexelés vázlata



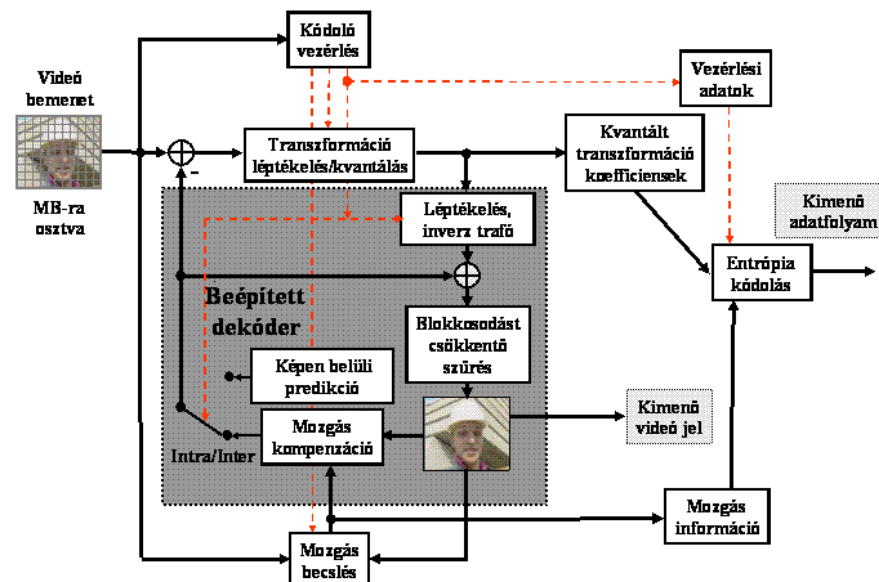
Videó kódolások időrendben



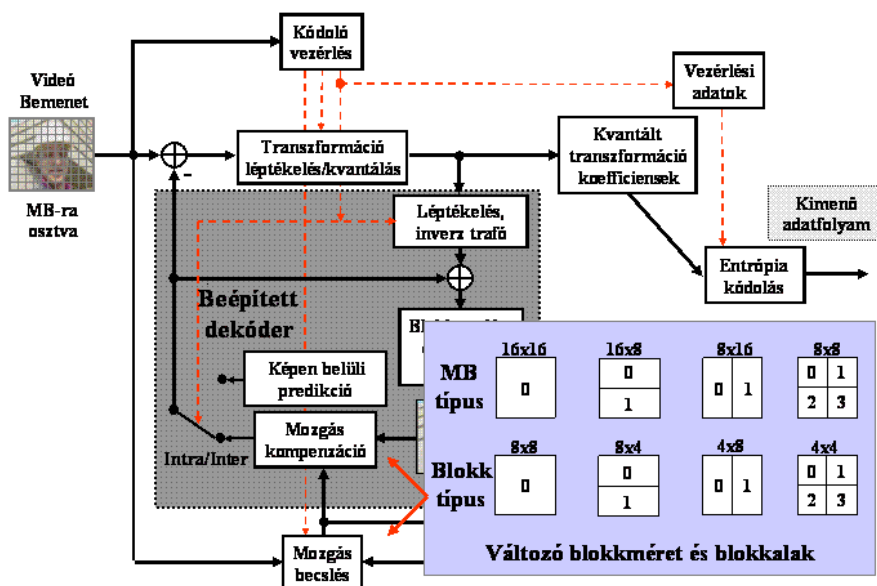
AVC (Advanced Video Coding)

- Lényegesen jobb videó minőség javított és bővített kódolási eszközkészlettel, ami jelentős adat kompresszió és kódolási hatékonyság növekedést eredményezett
 - Nem csak fix méretű mozgáskompensációs egység, hanem tartalom adaptív méret állítás (MPEG-2-ben nincs)
 - Képen belüli predikció (MPEG-2-ben nincs)
 - Nem csak két referencia képes időbeni predikció, hanem sok-képes predikció is
 - Blokkosodást gátló szűrés (MPEG-2-ben nincs)
 - Pontosabb mozgáskompensáció
 - Kétféle javított (15%) hatékonyságú entrópia kódolás

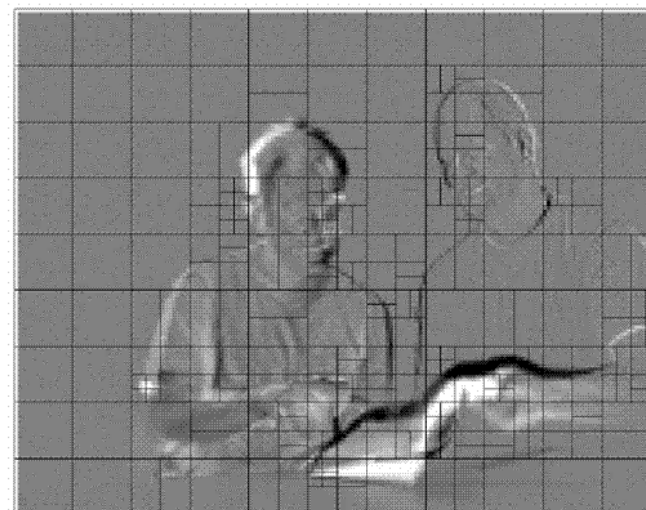
AVC kódolás alap struktúrája



Makroblokkra bontás

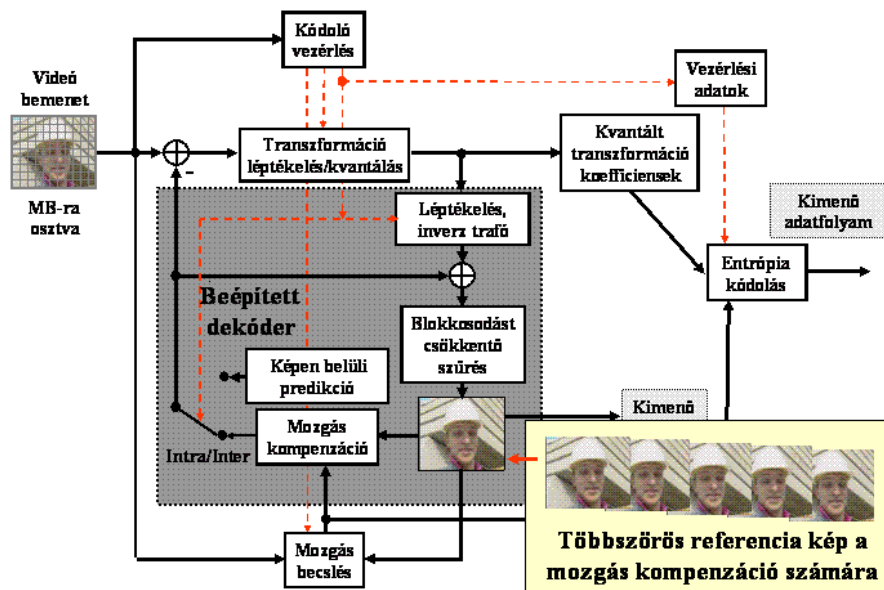


Mozgás kompenzáció blokk felbontása

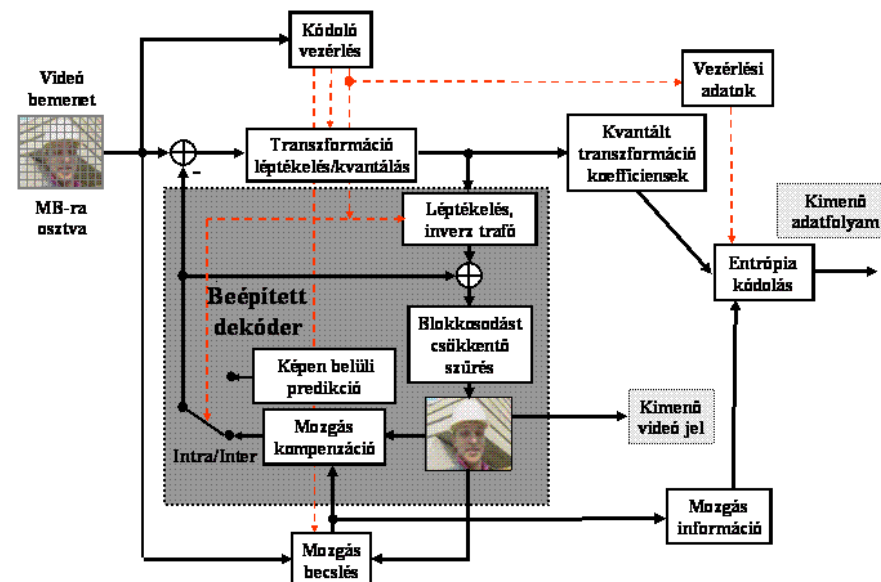


Plusz: pixel alatti (1/4 és 1/8) mozgásvektorok

Több referencia képes mozgáskompenzáció



Blokkosodást csökkentő szűrés



Blokkosodást megakadályozó szűrés

- Erősen kompresszált önmagában kódolt képekre
- Jelentősen csökkenti a predikciós hiba maradványok hatását



Szűrés nélkül

AVC szűréssel

Implementáció

- Mivel az AVC lényegesen nagyobb bonyolultságú, mint a megelőző kódolási módok, ezért nagyobb feldolgozási teljesítményt is igényel a kódolás, és a dekódolás is
- A következő táblázat tartalmaz egy összehasonlítást a megelőző és az AVC kódolási módok között a komplexitás és a hatékonyság tekintetében

AVC profilok	Megcélzott alkalmazás	Dekóder felépítés MPEG-2-höz képest	MPEG-2-höz képesti kódolás hatékonyság
Baseline	Kis késleltetés (videótelefon, mobil)	150%-kal komplexebb	50%-kal jobb
Extended	Mobil, streaming	250%-kal komplexebb	75%-kal jobb
Main	Váltott-soros videó, műsorszórás	300%-kal komplexebb	100%-kal jobb

AVC új profil igénye

- Az AVC 2003-ban elfogadott videó kódolási szabványa minden előzőnél nagyobb kódolási hatékonyságot biztosított
- A cél az SD és a kisebb felbontások támogatása
- Így 8 bit/mintára, és 4:2:0 mintastruktúrára optimalizálták
- Kezdetben 3 profil: Baseline, Main, és Extended
- Részben időhiány miatt a professzionálisnak számító képminőségeket nem különösebben jól kezelték
- A szubjektív képminőség vizsgálatok szerint az objektíven mért javulással nem egyenesen arányos a szubjektív képminőség javulás (különösen HD filmekre)
- Új megközelítés (pl. profil) kellett

AVC új profiljai (High profile: FReXt)

- **High Profile (HP):** 8 bites videó, 4:2:0 mintastruktúra, csúcs fogyasztói célokra, ha nem szükség a kiterjesztett szín mintastruktúra
- **High 10 Profile (Hi10P):** 4:2:0 mintastruktúra, 10 bit/minta
- **High 4:2:2 Profile (H422P):** 4:2:2 mintastruktúra, 10 bit/minta
- **High 4:4:4 Profile (H444P):** 4:4:4 mintastruktúra, 12bit/mintáig minden, járulékosan támogatható a hatékony veszteségmentes kódolás, és az integer transzformáció alapú színtér konverzió, mellyel a színtér transzformáció kerekítési hibája elkerülhető

A FReXt képminősége

- Inkább az érzékelt, mint az objektív képminőség javul
- A teszt szekvencia 24 képes HD (1920x1080) progresszív film
- A High Profile az MPEG-2-nél annak bitsebesség harmadánál már jobb képminőséget biztosított (8, 24 Mbit/s)
- A High Profile kb. 16 Mbit/s adatsebességnél nagyobb sebességek esetében transzparensnek volt minősíthető
- Az elfogadható minőséget jelenleg kb. 8 Mbit/s adatsebességnél biztosítja
- További optimalizálás után minden bizonnyal a 8 Mbit/s alatti sebességek is elegendők lesznek

Az AVC szubjektív tesztelése

- Mivel az AVC-t számos felbontás, félkép frekvencia és adatsebességre szánták, ezért azt legalább a három legfontosabb területen külön-külön kellett tesztelni:
 - Mobil és az Internet streaming
 - **SDTV**
 - **HDTV 720p, 1080i és 1080p**
- A szubjektív teszthez AVC bitfolyamot a Fraunhofer biztosította
- Sony adta a HD és a VideoTele az SD tartalmakat



„Husky” szekvencia nyitó képe



„Mobile” szekvencia nyitó képe



„Football” szekvencia nyitó képe



„Tempete” szekvencia nyitó képe



„Crew” szekvencia nyitó képe



„Harbour” szekvencia nyitó képe



„New mobile” szekvencia nyitó képe



„Stockholm” szekvencia nyitó képe



„Riverbed” szekvencia nyitó képe



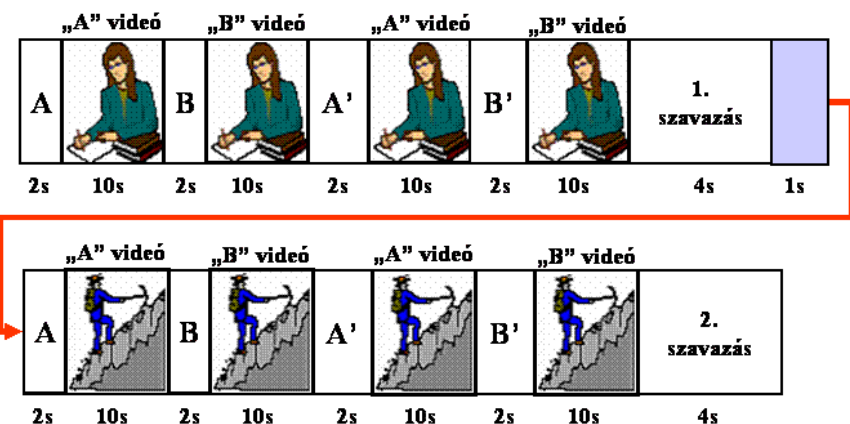
„Vintage car” szekvencia nyitó képe

A teszt menete

- A szubjektív teszt az ITU-R BT. 500 alapján zajlott
- A módszer a kettős gerjesztésű folyamatos minőségi skála szerinti (Double Stimulus Continuous Quality Scale: DSCQS)
- Egyetlen különbség a kijelző típusában van, ami a HD tartalmak esetében nem CRT, hanem digitális mikro-tükrös projektor (Digital Light Processing: DLP)
- A DSCQS teszt különlegessége a kódolt-dekódolt és az eredeti szekvencia kijelzési sorrendjében és a képminőség értékelési skálában van
- A DSCQS teszt minden egyes teszt cellája tartalmazza egyszer az eredeti és egyszer a kódolt-dekódolt szekvencia kijelzését

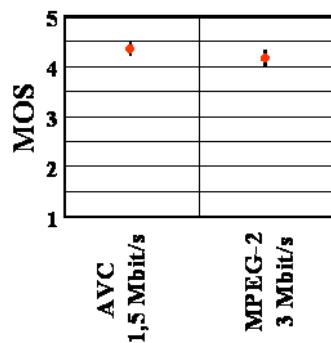
A teszt menete

- A kijelzési sorrend véletlen, az értékelő nem tudja melyik az eredeti és melyik a kódolt
- A kódolt-dekódolt és az eredeti szekvencia kijelzését megismétlik és csak ezután kérik a véleményt

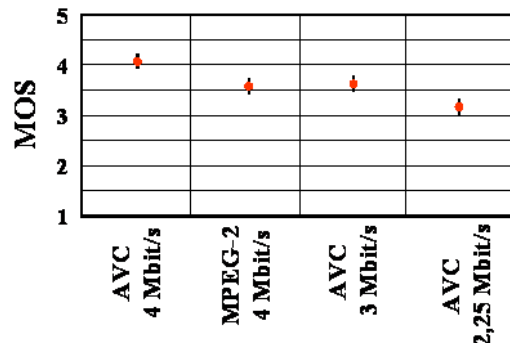


AVC teszt eredmények SD-re (példák)

720x486 60i „Tempete”

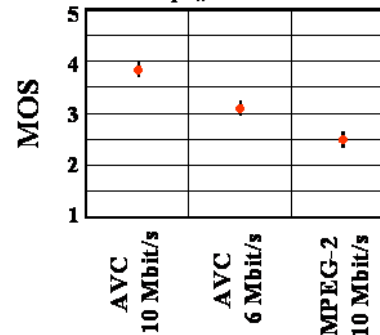


720x486 60i „Football”

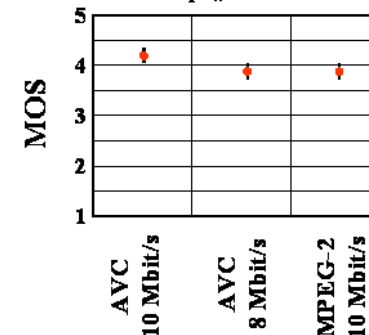


AVC teszt eredmények HD-re (példák)

1080p „Riverbed”



720p „Crew”



AVC kódolási hatékonysága

A vizsgálatok szerint az SD/HDTV alkalmazásokban az AVC 4/9 –2/5-ére csökkenti az MPEG-2-höz képesti adatsebességet, de a csökkenés képtartalom függő

Lehetőségek AVC kódolással

- Az AVC kb. 1,5-2 Mbit/s-val közel PAL minőség
- De mit jelent ez a küszöbön álló DVB-T rendszerrel
- Válasszuk adásparaméternek a 8k, 16-QAM, 2/3 kódarány, ¼ védelmi időt, a hasznos bitsebesség 13,27 Mbit/s
- A televíziós csatornák száma MPEG-2-vel maximum 4
- AVC-vel a csatornaszám multiplexenként 9-re is emelhető
- Az AVC-vel HDTV kb. 8 Mbit/s adatsebesség mellett
- DVB-H (Handheld), CIF méretű videó (200 kbit/s) mellett, egyetlen DVB-T csatornában $13/0,2 = 65$ program továbbítható

Audió bitsebesség csökkentés

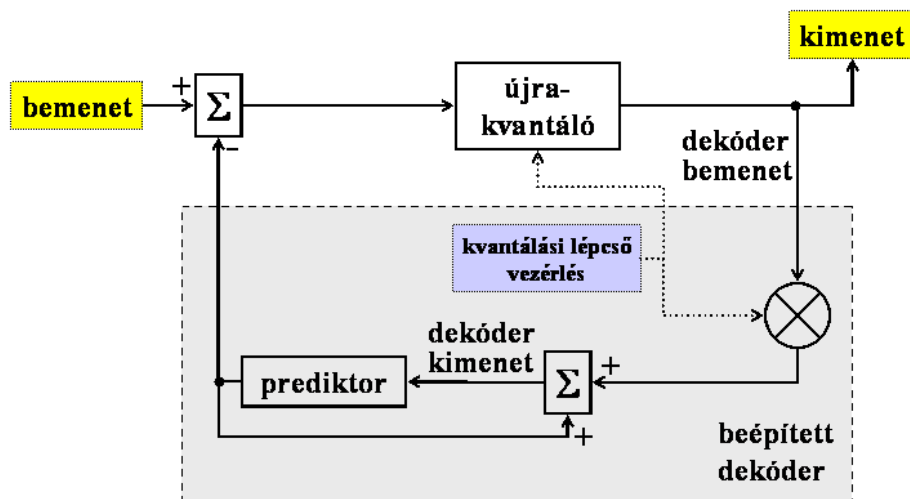
Legfontosabb kódolási elvek

- Prediktív kódolás, az időbeni redundanciát csökkenti
- Részszávos kódolás a hang spektrumának nem egyenletes voltát használja ki
- Transzformációs kódolás a transzformációs "síkon" megjelenő redundanciát csökkenti

Prediktív kódolás

- A becsült és az eredeti minta különbségének újrakvantálása
- A kódoló leutánozza a dekóderben zajló műveleteket
- A folyamat az időtartományban zajlik, tehát nagyon gyors
- Ezért kicsi a kódolási és dekódolási késleltetés

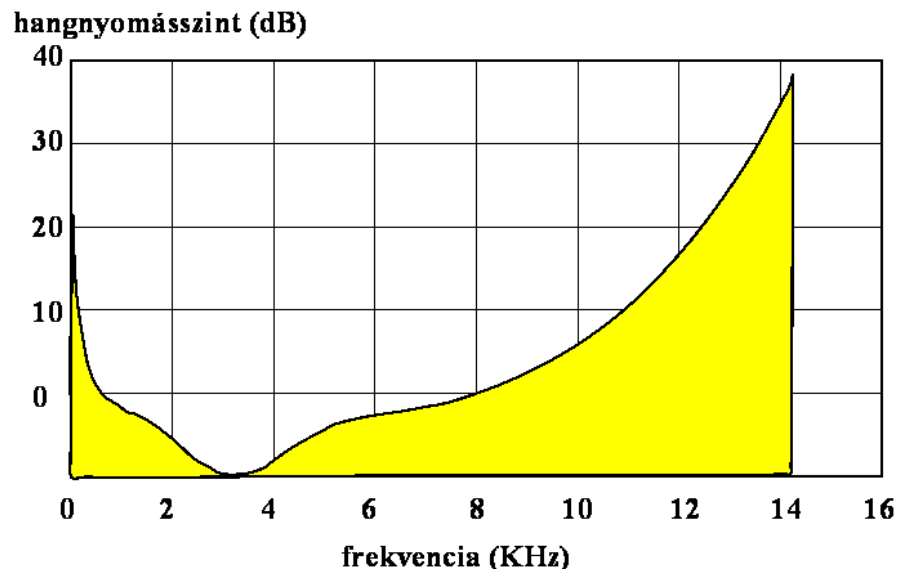
Prediktív kódoló vázlata



Pszicho-akusztikus kódolások

- Kihasználjuk az emberi hallás elfedési, maszkolási modelljét
- A hallás következő sajátosságait vesszük figyelembe:
 - Abszolút hallásküszöb
 - Kritikus sávok fogalma
 - Frekvencia tartománybeli elfedés
 - Dinamikus hallásküszöb
 - Idő tartománybeli elfedés
 - A hang tonalitás és nem tonális szerkezete
 - A frekvenciától függő hallás pontosság

Abszolút hallásküszőb



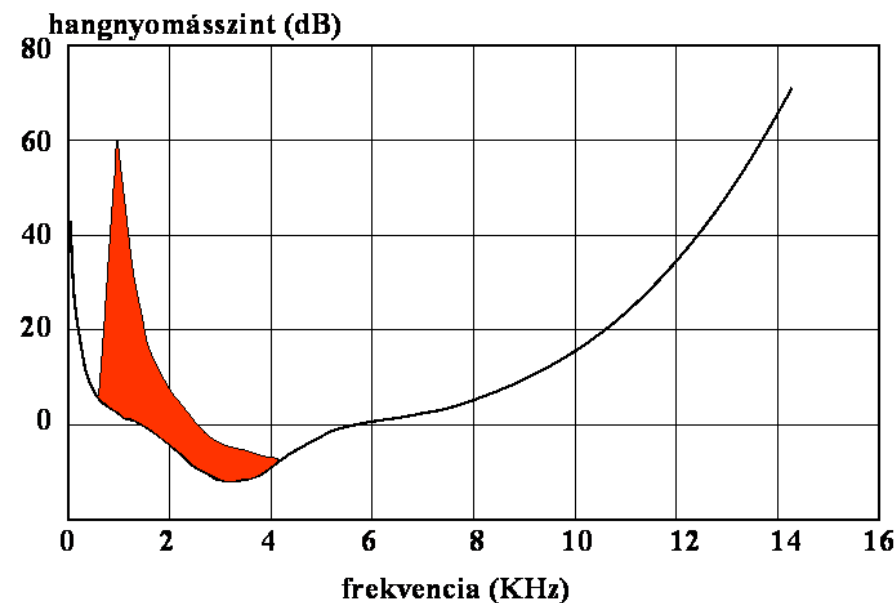
Kritikus sávok

- Hallásunk sajátosságai alapján a frekvenciatartományt fel lehet osztani az ún. kritikus sávokra (24 db. Bark)
- Ezek sáv szélessége a frekvencia növekedésével nő, (800 Hz alatt 100 Hz, 1 KHz-en 160 Hz, míg 10 KHz-en 2500 Hz)
- Több egyidejű komponens összegződése eltérő módon történik, ha azok egy kritikus sávon belül illetve kívül vannak
- Kritikus sávon belül a komponensek teljesítményben összegződnek
- Érzeti kódolóban elfedési jelenségeket vizsgálunk
- Célszerű a kritikus sávoknál keskenyebb sávokat alkalmazni, így az elfedési görbék egyszerűen, teljesítményben összegezhetők

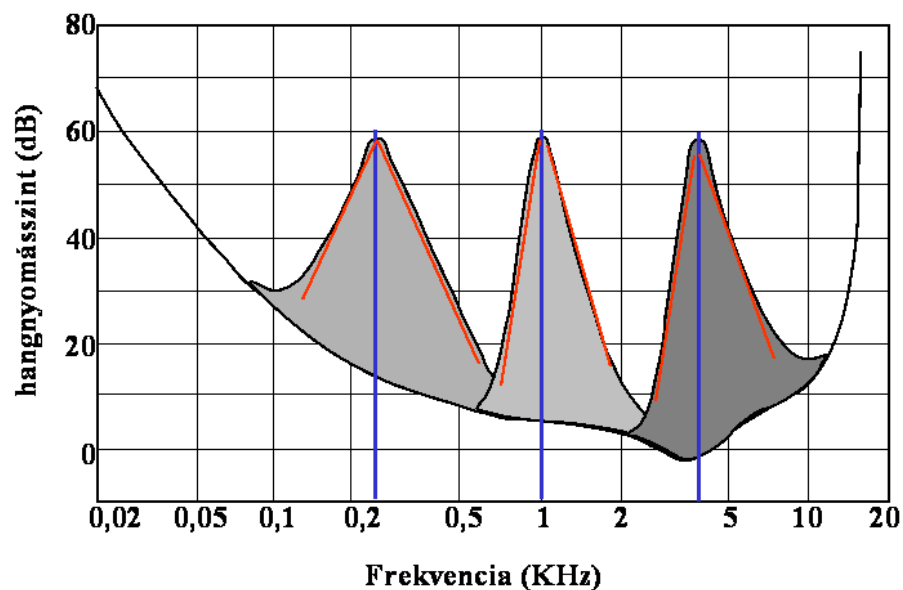
Elfedés a frekvencia tartományban

- A spektrumból kiemelkedő tonális, vagy keskenysávú zaj jellegű komponensek megemelik frekvenciatartománybeli környezetükben a hallásküszőböt
- Kialakul a dinamikus hallásküszőb, vagy maszk
- Ami a maszk alatt van az nem hallható
- Ez az elfedés frekvencia- és szintfüggő
- Az elfedési görbék frekvenciában aszimmetrikusak
- Az elfedés az elfedő jel szintjének növekedésével egyre szélesebb, de a görbék alakja és jellege nem változik

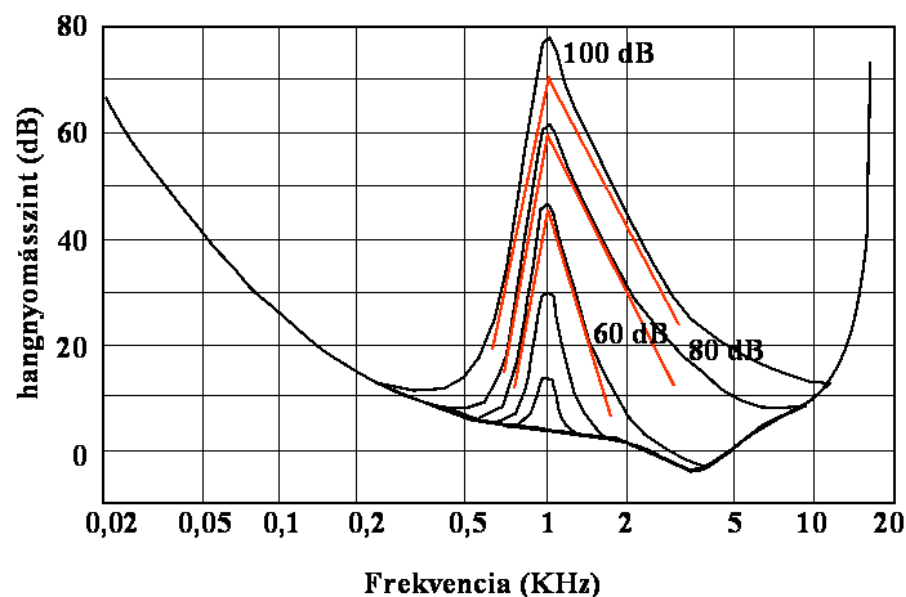
1 KHz-es tiszta szinusz elfedési görbéje



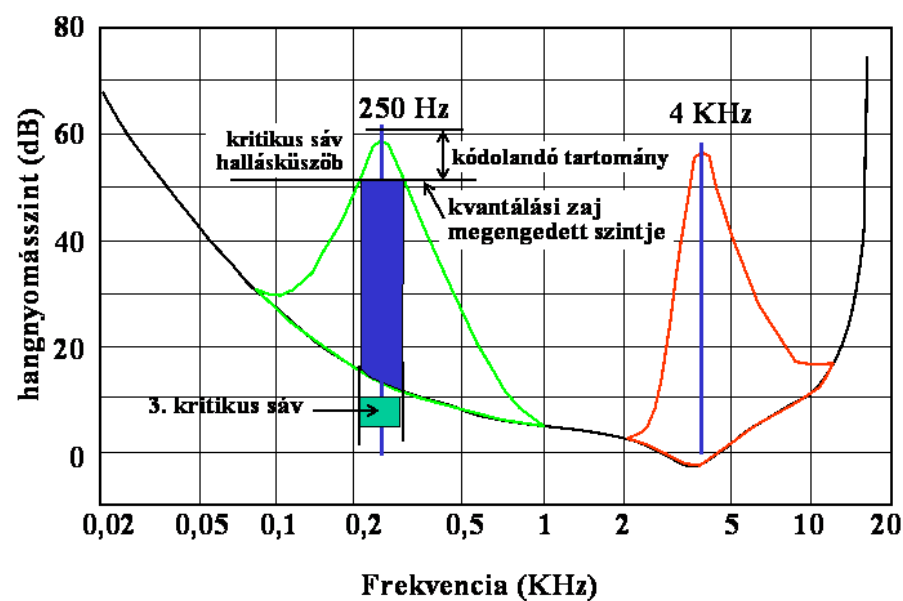
A 60 dB-es 250 Hz, 1 és 4 KHz elfedési görbéi



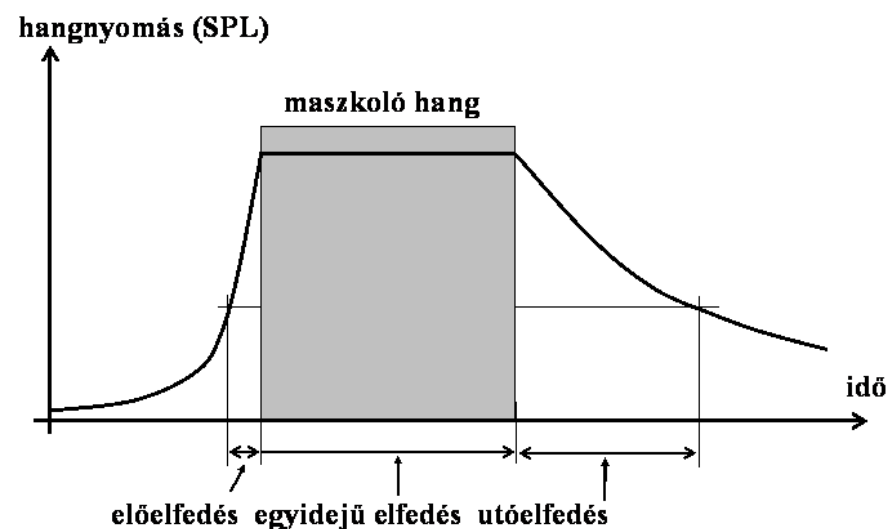
Különböző szintű 1 KHz-es szinuszos jelek elfedési görbéi



A kritikus sáv és az újrakvantálás kapcsolata



Időtartománybeli elfedés jellege



Tonalitás és a frekvencia pontosság

- Hallásunk a harmonikus hangok tisztaságára érzékenyebb, mint a zajjellegű hangokéra
- Ezért a tiszta (harmonikus) hangok spektrumvonalait pontosabban kell átvinni
- Ehhez azonban meg kell határozni, melyek a tonális illetve a zajjellegű komponensek
- Az emberi hallás a mélyhangok tekintetében nagyobb pontosságot igényel, mint a közepes és magas hangoknál
- Ezt a bitsebesség csökkentési eljárásokban figyelembe kell venni

MPEG kódolás

- MPEG-1: “ISO/IEC 11172 Information Technology Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mb/s”
- Hangkódolási eljárások: Audio csoport
- A következő fő szempontokat vették figyelembe:
 - kiindulási hangminőség a CD
 - különböző tömörítési fokok és hangminőség
 - az átviteli hibákkal szembeni védetség
 - kódolás/dekódolás bonyolultsága, késleltetési idő
 - kaszkádosisíthatóság

- Munkájuk eredménye az MPEG szabványok 3. része
- Először az MPEG-1-ben, majd továbbfejlesztett változatai az MPEG-2 és MPEG-4 Audióban
- Az MPEG-1-ben két fő vetélytárs:
 - MUSICAM (Masking pattern adapted Universal Sub-band Integrated Coding And Multiplexing), Eureka 147, Philips, CCETT, IRT, első sorban DAB célokra
 - ASPEC (Adaptive Spectral Perceptual Entropy Coding), AT&T Bell Lab, Thomson, Fraunhofer Society elsősorban ISDN célokra
- Összehasonlítás után a következtetés: célszerű bevezetni eltérő bonyolultságú és eltérő minőséget szolgáltató kódolási rétegeket

MPEG-1 Audió általában

- 48, 44,1, 32 KHz-es mintavételi frekvencia
- Kimeneti adatsebesség csatornánként: 32-384 Kbit/s között diszkrét lépésekben állítható
- Mono, sztereo, kettős, kapcsolt-sztereo
- Csak a kimeneti adatfolyam szabványos
- 3 réteg definiált: különböző alkalmazások és adatsebesség tartományokra

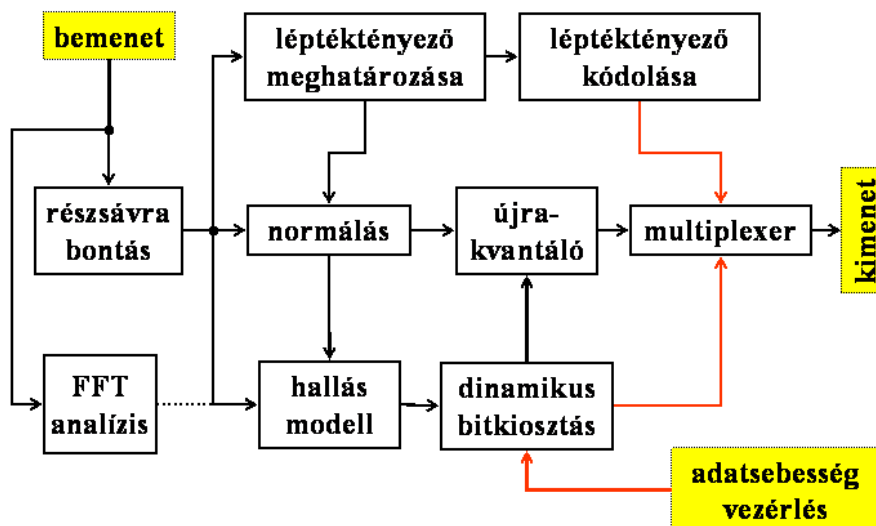
MPEG-1 Audió I. réteg

- A MUSICAM egyszerűsített változata
- Főbb jellemzői:
 - kis komplexitás, a kóder és a dekóder oldalon egyaránt
 - rövid kódolási/dekódolási időkésltetés
 - nagy időfelbontás
- Mára nincs jelentősége

MPEG-1 II. réteg

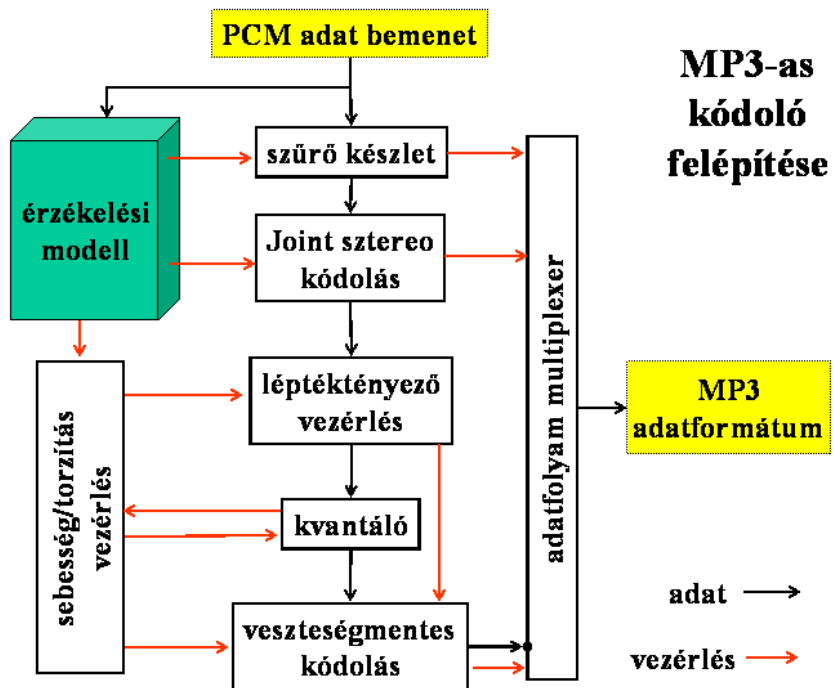
- Az I. réteghez képest több ponton javulás
- Főbb jellemzői:
 - bonyolultabb kóder, a tömörítési hatékonyság növelése érdekében
 - kis komplexitású dekóder (aszimmetrikus)
 - rövid időkésltetés (~45...50 ms)
 - editálhatóság, konstans kerethossz
 - erős bithiba védettség

Az MPEG-1 Audió II. réteg kódolója



MPEG-1 Audió III. réteg

- Jelentősen eltér az első kettőtől
- Részsávként MDCT transzformációs kódolás
- Huffman kódolás
- Adaptív bitsebesség vezérlés
- A kódolás/dekódolás késleltetési ideje megnövekszik
- Legfontosabb jellemzői:
 - bonyolult kóder és dekóder felépítés
 - hatékony tömörítés



MPEG-1 Audió alkalmazások

I. réteg

- DCC, Solid State Storage Audio
- ajánlott bitsebesség: 192...448 kbit/s/csatorna

II. Réteg

- DAB és általában a DVB műsorszórás
- műsorszétoosztás: 256...384 kbit/s
- műsorsugárzás: 80...256 kbit/s
- diszkbázisú editáló, tároló rendszerek
- archiválás, multimédia, CD-I
- ISDN, stb.

MPEG-1 Audió III. réteg alkalmazásai

- kezdetben: kommentár összeköttetések
- Jelenleg szinte mindenhol
- műholdas műsorszórás (WorldSpace)
- Internet on-line rádió RH minőség: 16 kbit/s, rádióadók műsorainak Interneten történő "sugárzása"
- Internet MP3 fájl letöltés

MPEG-1 III. réteg alkalmazásai

hangminőség	sávszélesség	mód	bitsebesség	tömörítés
kb. telefon	2,5 KHz	mono	8 Kbit/s	96:1
RH-nál jobb	4,5 KHz	mono	16 Kbit/s	48:1
AM-nél jobb	7,5 KHz	mono	32 Kbit/s	24:1
FM-nél jobb	11 KHz	sztereo	56-64 Kbit/s	26 – 24:1
közel CD	15 KHz	sztereo	96 Kbit/s	16:1
CD	20 KHz	sztereo	112-128 Kbit/s	14 – 12:1

MPEG-1 Audió hiányosságai

- Két csatornánál több csatorna nem kódolható
- Kis adatsebességek esetén a hangminőség rossz, mert pl. a mintavételi frekvencia 32 KHz-nél nem lehet kisebb

MPEG-2 Audió

- 1992: új kiterjesztések gondolata
- LSF (Low Sampling Frequency): kisebb mintavételi frekvenciák felé
- BC (Backward Compatibility): visszafelé kompatibilis többcsatornás rendszer
- Később AAC (Advanced Audio Coding): visszafelé nem kompatibilis többcsatornás rendszer

MPEG-2 Audió LSF

- A hang érzékelési entrópiája nem függ a mintavételi frekvenciától
- Ezért új mintavételi frekvenciák: 16, 22,05 , 24 kHz
- A lehallgatási tesztek szerint sokkal kevésbé zavaró a sávkorlátozás hatása, mint a kódolásból adódó melléktermékek megjelenése
- A hang minőség 64 kbit/s alatt lényegesen jobb lesz
- A kódolás közel azonos az MPEG-1 audióval
- A kóder és a dekóder könnyen implementálható

MPEG-2 Audió BC

- MPEG-1 kompatibilis többcsatornás kiterjesztése
- Az un. surround (3/2 és 5+1) zenei műsorok, és a több nyelvű (7) átvitel biztosítására
 - Visszafele (backward) kompatibilitás: a többcsatornás bitfolyamból egy szabványos MPEG-1-es dekóder képes legyen a 2/0 sztereo előállítására
 - Lefele (downward) kompatibilitás: a 3/2 hanganyagot kevesebb számú hangcsatornán (5+1, 3/2, 3/1, 3/0, 2/2, 2/1, 2/0, 1/0) is le lehessen hallgatni
 - Előre (forward) kompatibilitás: az MPEG-2 dekóder képes legyen dekódolni az MPEG-1-es bitfolyamot
- A kompatibilitás követelménye komoly hátrány

MPEG-2 AAC

- 1995-ben kezdték el fejleszteni, szabvánnyá 1997-ben vált
- Főbb célkitűzései:
 - minden mintavételi frekvencia támogatása (8...96 KHz)
 - csatornakiosztásra az 1/0-tól kezdve az 5+1-ig minden
 - az eredeti kevesebb csatornán is reprodukálható legyen
 - nagy bitsebesség-megtakarítás, a kompatibilitás árán is
 - az „eredetitől megkülönböztethetetlen” minőség 384 kbit/s /5 csatorna esetében
- Végül 320 kbit/s/5 csatorna bitsebesség mellett „az eredetitől megkülönböztethetetlen” minősítés az EBU szerint

Az MPEG-4 Audio célkitűzése

- Interaktív multimédia célokra
- Minden értelemben léptékelhető kódolási módok alkalmazása és definiálása
- A beszéd és a zene és minden egyéb audio objektum-orientált kezelése
- Az audio objektum fogalmának bevezetése
- Javított minőségű hangátvitel és kódolás
- Az addig élesen elkülönülő minőségi hangkódolásokat, a beszédkódolást és a számítógépes zenét együttesen, integráltan kezeli

MPEG-4 hanganyag típusk

- Minőségi hanganyagok mono, sztereo, illetve többcsatornás módban (AAC)
- A 64 kbit/s és afeletti bitsebességű hangkódoláshoz nem vezettek be alapvetően új eszközöket
- Közepes minőségű hanganyagok (MPEG-2 LSF)
- Szélessávú beszédjelek (7 kHz sávszélesség)
- Keskenysávú beszédjelek
- Érthető beszédjelek (információtartalom a fontos)
- Szintetikus beszéd (pl. fonéma- vagy más token-alapú szövegírással)
- Szintetikus hang (pl. zeneleíró nyelvvel)

- Természetes forrásanyagok: beszéd és zene, illetve zenei anyagok
- Különböző bitsebesség-tartományokban eltérő kódolási technikákat alkalmaznak az optimális kódolás érdekében
 - Parametrikus kódolás 2 - 4 kbit/s közötti beszédkódolásra és 4...16 kbit/s hangkódolásra
 - CELP (Code Excited Linear Predictive) kódolás, 6 - 24 kbit/s bitsebességű beszédkódolásra
 - MPEG-2 AAC 16 kbit/s és nagyobb bitsebességű hangkódolásra
 - Skálázható szélessávú kóder, minden bitsebességen használható, ám valamivel gyengébb hatásfokkal