Contents

[1. dia 1](#_Toc324720473)

[2. dia 3](#_Toc324720474)

[3. dia 5](#_Toc324720475)

[4. dia 14](#_Toc324720476)

[5. dia 23](#_Toc324720477)

[6. dia 28](#_Toc324720478)

[7. dia 36](#_Toc324720479)

# dia

**Definíció (Média):** Médiának nevezzük a kommunikáció bármely eszközét, amely biztosítja egy közlemény eljuttatását egy közlőtől a fogadóig, függetlenül attól, hogy a kommunikációban résztvevők egyének vagy csoportok. Ilyen értelemben bármilyen eszköz, amely egy információt téren és időn át juttat el egy forrástól a fogadóhoz, a kommunikáció médiumának tekinthető.

**Definíció (Médiatechnológia):** Azon mérnöki megoldások összessége, amely

* hang, kép, mozgókép, ezekhez kapcsoló kiegészítő információ
* előállítását, feldolgozását, továbbítását, tárolását, megjelenítését

teszi lehetővé.

**Médiatechnológia céljai:**

* Továbbítás

Média: hálózat (vezető, optikai szál, rádióhullám)

* + Átvitel (azonos időben)
  + Sokszorosítás (elosztás) (nem mindig)
* Tárolás

Média: permanens adathordozó (manapság jellemzően digitális)

* + Archiválás (későbbre megőrzés)
  + Továbbítás (nagy késleltetés) (adatsebesség?!)
  + Sokszorosítás
* Feldolgozás

Média: átmeneti, nem releváns; a tartalom változása a lényeges

* + Egy vagy több médiatartalomból egy vagy több másikat
* Hibrid
  + A fentiek közül egynél több alkalmazása

**Adattömörítés:**

1. Célja:

* A kódolt adat kisebb méretű, gazdaságosabban tárolható vagy vihető át

1. Jellemzője:

* dekódolással hibamentesen visszanyerhető belőle az eredeti adat.

**Jeltömörítés:**

A kódolt adatból visszaállított jel különbözhet az eredetitől, azaz veszteség lép fel. Az azonban cél, hogy a lehető legjobban hasonlítson egymásra az eredeti és a dekódolt jel.

1. Célja:

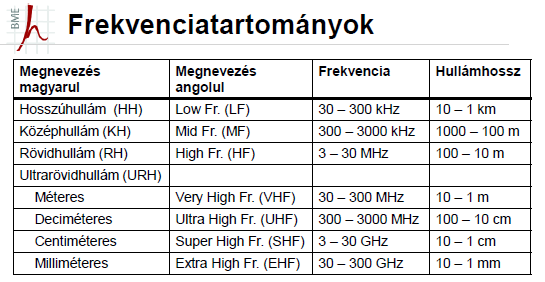
* valamilyen hűségkritérium alapján a legjobban megfelelő kódolt adatot rendeljük hozzá a jelhez

1. Jellemzője:

* nagyobb tömörítés → nagyobb információvesztés
* kisebb tömörítés → kisebb információvesztés

# dia

**Rádió feltalálója:** Tesla (1893)



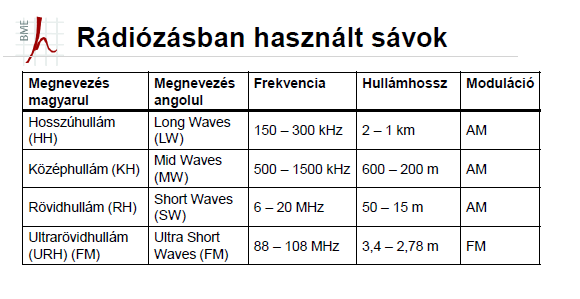
**Feladat:**

Mekkora antenna szükséges a 100 MHz-es FM adás vételéhez, ha λ/8-as antenna kell a jó minőségű vételhez?

**Megoldás:**

λ = = = 3m

**= m**



**Lakihegy adótorony:**

* Hely: Lakihegy (Szigetszentmiklós)
* Adóteljesítmény: 120 kW
* Magasság: 314 m
* Átadás éve: 1928
* Rekord:
  + Magyarország legmagasabb építménye
  + világ legnagyobb Blaw-Knox-antennája (szivar alakú acél)

**Solti adótorony:**

* Hely: Solt
* Adóteljesítmény: 2 MW
* Magasság: 304 m
* Építés éve: 1974-1977
* Adás: Kossuth Rádió (mr1) (540kHz)

**Yagi antenna működése:**



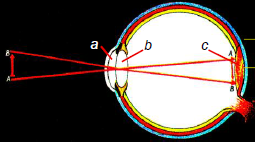
* Dipólus kerül csatlakoztatásra
* Hossz: λ/2, λ/4
* Reflektor kb. 5%-kal hosszabb
* Direktorok kb. 5%-kal rövidebbek

# dia

Az emberi látórendszer **(HVS: Human Visual System)** az elektromágneses spektrumnak csak egy egészen keskeny részét képes érzékelni: az *ultraibolya* (ibolyán túli) és az *infravörös* (vörösön „inneni) sugárzás közötti kb. **400-700 nm**-es hullámhossz tartományt.

A **látószerv** gyűjtőnév, beleértendő a szem, a látóidegek és az agy azon részei, melyek a fényingert ingerületté alakítják át.

A *pupillán* (a) áthaladó fénysugarakat a *szemlencse* (b) gyűjti össze és fókuszálja a *retinára* (ideghártyára) (c), ahol fordított állású, kicsinyített, valódi kép keletkezik.



**Látóideg-végződések:**

* ***Csapok***:
  + fényérzékeny elemek
  + biztosítják az erős fényhez hozzászokott szem fény és színérzékenységét -> a csapokkal látunk színeket, de csak világosban
  + a retina közepén helyezkednek el
* ***Pálcikák***:
  + biztosítják a gyenge fényhez hozzászokott szem fényérzékenységét
  + a retina széle felé haladva számuk fokozatosan nő

**Definíció (Láthatósági függvény):** az azonos intenzitású, de eltérő hullámhosszú fény hatására a szemben keletkezett fényérzet a láthatósági függvény szerint változik.

* a láthatósági függvény nagyon sok ember látásának átlagos érzékenységét tükrözi.
* a görbe maximuma kb. 555 nm hullámhossznál van.
* a láthatósági függvény eltér erős és megvilágítás esetén

*Nappal (jó megvilágításban)*

* A csapok működésének eredménye
* Látunk színeket
* Maximuma 555 nm-nél van (zöldessárga)

*Sötétben (gyenge megvilágításnál)*

* A pálcikák működésének eredménye
* Nem látunk színeket
* Maximuma a kék irányába tolódik el (kb. 505 nm – zöld)

**Definíció (Fényáram [lumen]):** a sugárzott teljesítmény (Φ(𝜆)) a láthatósági függvénnyel (𝑉(𝜆)) korrigálva, vagyis: amit a szem a sugárzott teljesítményből „lát”.

**Definíció (Fénysűrűség):** egységnyi felületre egységnyi térszögben eső fényáram

**Definíció (Kontraszt):** fénysűrűség- (világosság-) különbség.

**Definíció (Kontrasztérzékenység-küszöb):** az a két fénysűrűség érték, amelyek már éppen nem megkülönböztethetők egymástól.

**A kontraszt időbeli változása [Hz]**

* Függ az átlagos fényerősségtől, világosnál érzékenyebb (20-30-szoros a sötéthez képest)
* Sötét képnél max. 15 Hz, világosnál max. 50..60Hz

**A kontraszt térbeli változása [c/d = cycles/degree]:**

* 30 c/d felett lényegében nincs érzékelés
* Normál szemmozgásnál 2..4-szer érzékenyebb
* mérőszáma az 1 fokon belüli ciklusok száma (**c/d** = cycles / degree, vagy c/°)
  + **1 ciklus** (vagy periódus): a fénysűrűség érték változása egyik fénysűrűség értékről a másikra, azaz 2 db képelem
  + Ha 30 c/d felett lényegében nincs érzékelés bármilyen körülmény mellett, az azt jelenti, hogy a látás felbontása legjobb esetben is csak 60 képelemet képes megkülönböztetni 1 fokon belül -> a látás szögfelbontásának a hat ára 1/60 fok, azaz 1’
* Maximum: normál mozgásra 100 @2c/d, fixre 10 @4..5 c/d

**Raszteres letapogatás**

*Célja*: 2D (térbeli) képből 1D (időbeli) mintasorozatot készíteni

* Soronkénti letapogatás: lassan jobbra, majd gyors visszafutás (vízszintes eltérítés)
* Függőleges eltérítés: egyenletesen halad lefelé, majd gyors visszafutás felfelé
* Váltottsoros vagy progresszív

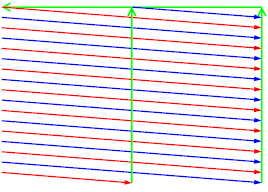


**Definíció (fúziós frekvencia):** a képelem váltása nem lehet 50-60 Hz-nél kisebb, mert a képelem villogni fog.

Ahhoz, hogy egy képsort **folyamatos**an mozgónak lássunk, elég **20-30 Hz** gyakorisággal felrajzolni a mozgás fázisait. Ekkor ugyanazt a képet 2-3-szor felrajzolva nem látható villogás és a mozgás is folytonosnak hat (pl. mozi).

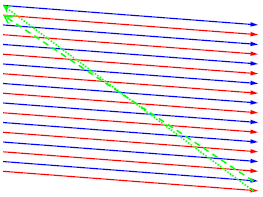
**DE:** Ugyanazt kétszer elküldeni **nem sávszélesség hatékony** módszer, és a teljes kép analóg tárolása nem megoldható, **ezért** a **2 félképes felbontás**t választották:

* A teljes képek 25 Hz (30 Hz) gyakorisággal követik egymást.
* Egy kép két félképből áll:
  + páros sorok (páros félkép)
  + páratlan sorok (páratlan félkép)
* Az azonos képhez tartozó félképeket egymás után küldjük el, először a páratlant, aztán a párosat. A félképváltás frekvenciája így 50 Hz (60 Hz) lesz, ami által a képernyő tartalma már nem fog villogni.



**Páratlan sorszámú váltott soros kép**

* Csak egyfajta függőleges és vízszintes visszafutás kell.
* A teljes képen páratlan sorszám van, ezért:
  + a legalsó sor csak félig van meg, a sor felénél a függőleges visszafutás lép életbe,
  + emiatt az ezt követő legfelső sor is csak fél sor.

**Páros sorszámú váltott soros kép**

* A módszer nagy hátránya a páratlan sorszámú képhez képest az, hogy csak egyfajta vízszintes visszafutás van, de a függőleges visszafutásból sajnos két különböző kell.
* Ez elég nehéz technológiai követelmény, emiatt a páratlan soros megoldást választották.
* Előnye viszont, hogy minden sor teljes sor.

**Jellemző SD képfelbontások**

* Kezdetben a félképfrekvencia megválasztásánál az erősáramú hálózat frekvenciáját vették figyelembe,
* Az ok: a brumm a képen mozgó zajként jelenik meg, ha a két frekvencia jelentősen eltér (az álló zavarkép kevésbé zavaró, mint a mozgó).
* Európa, Ázsia: 625 sor, 50 Hz félkép, 25 Hz kép,
  + 625\*25=15625 Hz sorfrekvencia
* Amerika, Japán: 525 sor, 60 Hz félkép, 30 Hz kép,
  + 525\*30=15750 Hz sorfrekvencia

**HD képfelbontások**

* Cél: a még teljesebb vizuális élmény, a periférikus látómezőt is lefedni képtartalommal
* Ennek megfelelően a 4:3 képméretarány helyett szélesebb, jellemzően 16:9 képméretarány
* HD Ready:
  + képes 720 sor (progresszív)
  + 16:9 képméretarány
  + 50 és 60 Hz képváltási frekv.
  + DVI vagy HDMI digitális bemenet
* Full HD vagy HD Ready 1080p:
  + Képesség: 1080p sor (1920x1080p)
  + 16:9 képméretarány
  + 24, 25, 30, 50 és 60 Hz
  + DVI vagy HDMI digitális bemenet

**Alapvető 16:9 HD képfelbontások**

* A sorok száma és a képméretarány alapján kiszámítható a képpontok száma.
* Váltottsoros esetben a sorok száma a két félkép összes sorát tartalmazza.
* Csak négyzet alakú képpontra egyezik meg a képméretarány képpontból származó méretaránnyal
* Például:
  + 720p: 1280×720 progresszív
  + 1080p: 1920×1080 progresszív
  + 1080i: 1920×1080 váltottsoros, de itt a félkép frekvencia a képfrekvencia kétszerese
* De pl. a HDCAM/HDV 1440×1080 képméretaránya azért 16:9, mert a képpont nem négyzet (hanem 4:3)

**Színjelek mintavételezése a világosságjelhez képest**

* Világosságjelre a látás felbontó képessége 3-5-ször nagyobb, mint a színjelre.
* A színjel különböző mintavételi megoldásai a világosságjelhez képest:
  + 4:4:4 (ugyanaz)
  + 4:2:2
  + 4:2:0
  + 3:1:1
  + 4:1:1
* **Követelmények:**
  + 3 dimenziós térben ha van 1 darab (1D) világosságjel, akkor még egy 2 dimenziós jel kell a 3D információ leírásához.
  + A világosság jel a fénysűrűség (Y), amely nemnegatív.
  + Hordozza a színinformációt 2 darab színjel, míg a fénysűrűség tőlük független mennyiség legyen.
  + Az R = G = B = Y legyen valamelyik fehér (szürke) (megjegyzés: több pont is lehet fehér, de mi csak az egyiket jelöljük ki). (Továbbiakban fontos jelölés: 3 alapszín: R (red - vörös), G (green – zöld), B (blue - kék))
  + Ha az egyik fehér pontban vagyunk, akkor a két színjel legyen nulla, más pontokban a színjelek közül legalább egyik legyen nem nulla.
* Egy lehetséges megoldás a két színjelre:
  + Legyen a két színjel a három színkülönbségi jelből (R-Y, G-Y, B-Y) kettő.

**Definíció (világosság - tv-technika):**

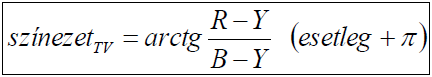
* megfelel a fénysűrűség (Y) színösszetevőnek.
* Értéktartománya 0-tól 1-ig terjed:
  + fekete szintnél Y=0
  + fehérnél (csúcs) Y=1 (vagy 100%).
* (Nemcsak az Y relatív skálája 0...1 közötti, hanem az RGB színösszetevőké is.)

**Az SD tv-technika színezet fogalma: YUV**

* Egy látható fényt 3 koordináta ír le:
  + Világosság: Y
  + A további 2 dimenziós adat két színkülönbségi jel: U = β·(B-Y), V=ρ·(R-Y)
* Legyen most „C fehér” és az R, G és B legyen a három FCC (*US Federal Communications Commission*) színösszetevő. Ekkor **Y = 0,3·RFCC + 0,59·GFCC + 0,11·BFCC**
* Az YUV és RGB közötti kapcsolat (β és ρ nélkül):
  + R = (R-Y) + Y
  + B = (B-Y) + Y
  + G = Y – (0,11·(B-Y) – 0,3·(R-Y))/0,59
* Megjegyzés: a G koordinátához szükség van az RGB és Y közötti kapcsolatra (ld. fent). Ezt felhasználva G számolható:
  + 0,59·G = Y - 0,11·B - 0,3·R = 0,59·Y - 0,11·(B-Y) - 0,3·(R-Y)

**A tv-technika színezet fogalma**

* A színkülönbségi jelek értéktartománya:
  + -0,7 <= (R-Y) <= 0,7
  + -0,89 <= (B-Y) <= 0,89
  + -0,41 <= (G-Y) <= 0,41
* A (G-Y) értéktartománya a legkisebb, ezért a jobb SNR (jel/zaj viszony) biztosítása miatt ezt hagyjuk el. Átvisszük így az (R-Y) és (B-Y) színkülönbségi jelet, valamint az Y világosság jelet.
* A világosságot tehát Y, a színinformációt pedig (R-Y) és (B-Y) hordozza:



**A tv-technika telítettség fogalma**

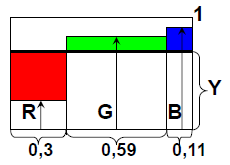
* A telítettség azt mutatja, hogy egy színinger mennyire van „hígítva” fehér színnel.
* A **telítettség értékét egységnek (100%) vesszük** akkor, ha a színpont rajta van az FCC (*US Federal Communications Commission*) színháromszög bármelyik élén vagy csúcsán, azaz **amikor fehér\_szint = 0**.
* A televízió technika referencia pontjában, a „C fehérben” a telítettség értéke nulla (0%).

**Definíció (spektrálszínek):** fényszórás vagy fénytörés által tovább fel nem bonthatók.

**Definíció (kvázi-spektrálszínek):**

* a spektrálszínek, illetve a bíborvonal megfelelői a televízió technikában
* a spektrálszínek és a bíborvonal analógiájára a kvázi spektrálszínek jellemzői, hogy egyszerre nem tartalmazzák mind a három alapszínt.
* ezért a kvázi-spektrál színek a televízió technika által visszaadható színeket megadó háromszög oldalaira vagy csúcsaira esnek

**A fehér pont és a tv-technika színkülönbségi jelei: R-Y, G-Y, B-Y**

* Legyen a fehér pont a „C” fehér.
* Az FCC alapszínek és a „C” fehér használata esetén az eredő szín fénysűrűség értéke:
  + Y = 0,3·RFCC + 0,59·GFCC + 0,11·BFCC
* Ezt átrendezve az alábbi egyenletet kapjuk:
  + 0 = 0,3·(RFCC-Y) + 0,59· (GFCC -Y) + 0,11· (BFCC -Y)
* Ez az egyenlet grafikusan egy területdiagram ábrázolásnak felel meg egy tetszőleges R, G és B értékhármas esetén.

**Az U és V jel digitális rendszerekben**

* Összefoglalás (ismétlés):
  + Az U és V jelre érzékenység 3-5-ször rosszabb, mint az Y-ra
  + Az U és V két színkülönbségi jel zsugorításából származik
  + Az U és V jelek értéktartománya a 0-ra szimmetrikus
* Következtetések:
  + Az U és V mérete lehet 3-5-ször kisebb minden irányban
  + 8 bit esetén – minden értéket kihasználva – az Y tartománya [0..255], de U és V tartománya [-127..127]
* Méretbeli eltérések az Y jelhez képest (U és V azonos méretű):
  + Mindkét irányban felezés: 4:2:0
  + Vízszintes irányban felezés: 4:2:2
  + Vízszintes irányban harmadolás: 3:1:1
  + Vízszintes irányban negyedelés: 4:1:1

**Planáris YUV formátumok**

* A lényege, hogy a memóriában az Y, U és V összetevők 3 különálló 2D mátrixban tárolódnak le.
* A sorok egymás utáni tárolása a 2D mátrixon belül:
  + Fentről lefelé vagy lentről felfelé is lehet
  + Két szomszédos sor nem feltétlenül W (kép oszlopainak száma) képpont távolságban van. A mátrix oszlopainak száma (stride) lehet több, mint a W, például azért, hogy a sor címe 4-gyel v. 8-cal v. 16-tal osztható legyen, és ezzel hatékonyabban végezhetők műveletek (pl. MMX, SSE)
* Példa: bájt alapú Y színösszetevő, az (x,y) pozícióban lévő érték WH képméretnél, ha egy sor Stride bájtból áll:
  + Fentről lefelé esetben: Kép\_Y[ y \* Stride + x ]
  + Lentről felfelé esetben: Kép\_Y[ (H-y) \* Stride + x ]

**Pakolt YUV (és RGB) formátumok**

* A lényege, hogy a memóriában egy 2D mátrixban tárolódik a kép, de a mátrix egy eleme egy több dimenziós makropixel.
* A mátrix mérete (sorok és oszlopok száma) nem feltétlenül egyezik meg a kép sor- és oszlopok számával, de a kettő összefügg az alábbiak szerint:
  + Sorszámmátrix = Makropixel magasság · Sorszámkép
  + Oszlopszámmátrix = Makropixel szélesség · Oszlopszámkép
* A gyakorlatban a makropixel magassága szinte mindig 1.
* A sorok távolsága itt is eltérhet a makropixel mérettől, és lehet pl. 4-gyel vagy 8-cal vagy 16-tal osztható.
* A planáris formátumoknál annyi mátrix van, ahány színösszetevő, ezzel szemben itt 1 mátrix van, és egy makropixelen belül van minden színösszetevőből.

**YUV és YCrCb különbségek**

* Az YUV egyes analóg TV rendszerek színrendszere (*PAL, NTSC*), míg az YCrCb egyes analóg rendszerek digitális ábrázolásának színrendszere.

**Képtömörítés: redundancia**

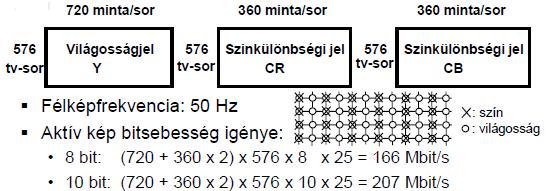
* Statisztikai redundancia: a képpontok (pixelek) adott környezetén belül (képen belül és időben is) hasonlók.
* Az emberi látás tulajdonságai: a videojel az emberi látórendszer (*HVS: Human Visual System*) számára lényegtelen részleteket is tartalmaz, amely eltávolítható. Tömörítéskor kihasználjuk a HVS tulajdonságait (részletek észrevétele, mozgás követése).
* A mozgóképben rejlő redundanciák (konkrétabban):
  + Térbeli: Intra-frame (kép) és Intra-field (félkép) kódolás,
  + Időbeli: Inter-frame kódolás (képek közötti).

**Tömörítetlen videó bitsebessége**

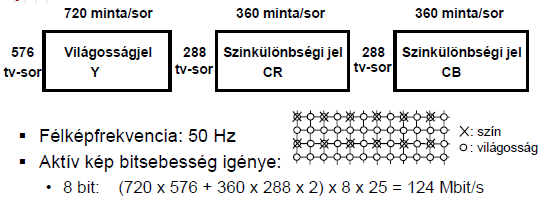
A számítás menete:

1. A képernyő sorok számát meghatározzuk a nézőtávolság és látószög alapján
2. A képernyő oszlopainak számát meghatározzuk a képpont méretarány és a képméretarány alapján
3. A képváltási frekvencia alapján meghatározzuk a másodpercenkénti képpont számot
4. Meghatározzuk, hogy egy képpontra átlagosan hány bit jut (pl. RGB esetén 3·8 bit)
5. A másodpercenkénti képpont számból és a képpont bitszámából meghatározzuk a bitsebességet

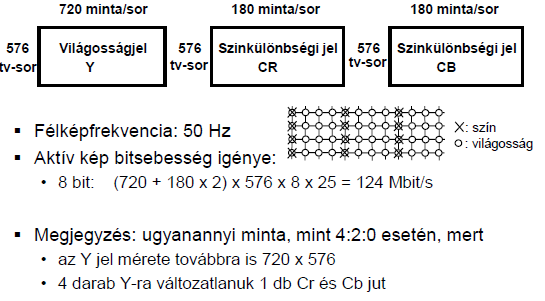
**4:2:2 stúdióformátum**



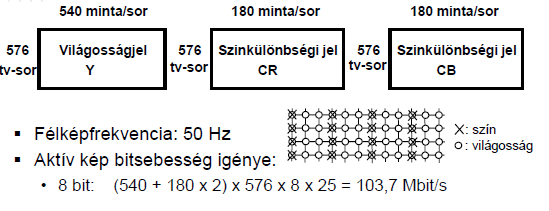
**4:2:0 formátum (MPEG, DVC, JPEG)**



**4:1:1 Komponensformátum (PAL, DVC, JPEG)**



**3:1:1 formátum (Non-linear Editing System)**



# dia

**Definíció (Hangnyomás):** a nyugalmi légnyomásra szuperponálódó légnyomás változás.

P(t) = Pnyugalmi + p(t) [1 Pa = 1 N/m2]

**Definíció (Hangnyomásszint - SPL):** egy adott p0 vonatkoztatási szinthez mért hangnyomás, ahol p0 a még éppen hallható 1 kHz-es hang hangnyomás értéke.

p0 = 20 · 10-6 Pa = 20 μPa

SPL = 20 lg( p / p0) [dB]

**Hangintenzitásszint**

* Hanghullámok esetén: v · c = f · λ, ahol c a hang terjedési sebessége (levegőben 340 m/s), v a részecskesebesség, f a frekvencia és λ a hullámhossz.
* A levegő sűrűsége ρ0 = 1,29 kg/m3

p = v · ρ0 · c

* **Definíció (Hang intenzitás):** időegység alatt a felület egységen áthaladó energia (Sound Intensity):

I = p · v = p2 / (ρ0 · c)

* **Definíció (Referenciaszint):** a még hallató 1 kHz-es hang intenzitása: I0 = 1 pW/m2 (= 10-12 W/m2)
* Hangintenzitásszint: SIL = 10 lg( I / I0) [dB]

**Hangosság és hangosságérzet (son és phon)**

* Szubjektív hangosságérzet a hangerősség.
* Tetszőleges hang hangerőssége annyi Phon, ahány dB a vele azonos hangosságérzetet keltő 1 kHz-es hang hangnyomás szintje.
* Ha a frekvencia függvényében ábrázoljuk az azonos hangerősségű pontokat, a Fletcher-Munson görbéket kapjuk.
* Ezekből meghatározható a frekvencia és a hangnyomásszint függvényében a hangerősség.
* Hangosság (son): =
* Hangosságérzet (phon): 10 20
* 40dB felett:

phon = , son =

**Az emberi hallás fiziológiai tulajdonságai**

* **Definíció (Hallásküszöb):** természetes alapzörejek elfedési görbéje
* Szinuszos jelek elfedési görbéi nem terjednek ki a teljes frekvencia tartományra. Nagy frekvenciákon nagyobb az elfedés mint kis frekvenciákon.
* Egy tiszta szinusz keskeny sávú zörejekre vonatkozó elfedési görbéje a komponens szintje alatt 2 dB-lel lévő csúccsal rendelkezik.
* Hasonló elfedési görbével rendelkezik egy kritikus sávnál lényegesen kisebb sávszélességű zaj is.
* **Definíció (Kritikus sávok)**: egy adott frekvencia sávon belül hallásunk egyidejű gerjesztés esetén intenzitás (energia) alapon összegzett hangosságot érzékel. Ezek a sávok a kritikus sávok.
  + (24 db) Bark-ban számozzuk, Barkhausen tiszteletére
  + Ezek sávszélessége a frekvencia növekedésével nő. (800 Hz alatt 100Hz, 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.
* **Hallás késleltetése**
* Dobhártya: „végtelenül” gyors
* Hallócsontok: 0,08 ms késés
* Csiga:
  + 20 Hz: 3 ms késés
  + 100 Hz: 1,5 ms késés
  + 1000 Hz: 0,3 ms késés
  + >3000 Hz: késés nélkül
* Ideg-impulzus időtartam: 1 ms
* Idegsejt feléledési idő: 1 ms
* **Összesen:**
  + Dobhártyától az agyközpontig: 3 ms – 6 ms

**Elfedés a frekvencia tartományban**

* A spektrumból kiemelkedő tonális vagy keskenysávú zaj jellegű komponensek megemelik a frekvenciatartománybeli környezetükben a hallásküszöböt.
* Kialakul a dinamikus hallásküszöb, ún. maszk.
* Ami a maszk alatt van, az nem hallható.
* Ez az elfedési jelenség 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.

**Audiókódolás jellemző paraméterei**

* A kódolt frekvencia tartomány:
  + Érthető beszéd, nem felismerhető beszélő (1-2 kHz).
  + Érthető beszéd, felismerhető beszélő (300 – 3400 Hz, telefon)
  + Rádió KH minőség (100 Hz – 4,5 kHz)
  + Közepes minőség (100 Hz – 7 kHz)
  + Rádió, FM minőség (40 Hz – 15 kHz)
  + Hifi minőség (20 Hz – 20 kHz)
* Jel-zajviszony
* Dinamika
* Torzítás (1%-os határ).

**Többcsatornás hangrendszerek**

* Az emberre a kétfülű hallás jellemző → térhallás.
* A két fül távolsága 18-22 cm, a késleltetési idők különbsége általában (!) 500-800 μs
* A hangvisszaadás szempontjából az egyik legfontosabb a hang irányának meghatározása, az irányérzékelés.
* Az irányérzékelés alapja a két fül által érzékelt hanghullámok közötti fázis- és intenzitáskülönbség.
* Az irányérzékelés frekvenciafüggő, más jellegű a kis- és nagyfrekvenciás hangoknál. A kétféle irányérzékelés határfrekvenciája ≈400-800 Hz

**Kisfrekvenciás irányérzékelés**

* Kis frekvenciákon a fej lényegében nem árnyékolja le a hangot a hang terjedési tulajdonságai miatt.
* Ekkor a két fül közötti intenzitáskülönbség kicsi, és nem hordoz lényeges információt a hangforrás irányáról.
* De: a hullámhossz összemérhető a fej méretével.
* Azonos frekvencián a két fülbe jutó jel fáziskülönbsége (Δφ + k·π) 1 perióduson belül van, ezért az fontos információt hordoz (k=0).

**Nagyfrekvenciás irányérzékelés**

* Nagy frekvenciákon a fej árnyékolása már jelentős, így az ellentétes oldalon lévő fülbe jelentősen csillapított hang jut.
* A hullámhossz kisebb a fej méreténél, azonos frek-vencián a két fülbe jutó jel fáziskülönbsége (Δφ + k·π) nem hordoz értékelhető információt (k=0,1,2,…?).
* A nagyfrekvenciás iránymeghatározásnál így az intenzitáskülönbséget használjuk ki.
* Azt, hogy adott hang elölről vagy hátulról szól, a fej önkéntelen, apró mozgatásával “kísérletezzük” ki, mivel a két esetben azonos fejmozgatáshoz eltérő észlelésváltozások tartoznak.

**Sztereó technikák**

* A hallás kétfülű modellje:
  + A megfelelő oldalon lévő fülbe a direkt jel jut, az ellentétes oldali fülbe késleltetett és csillapított jel:
  + m – csillapítás, τ – késleltetés a két jelút között.
  + f1(t) és f2(t) – bal, illetve jobb oldalon lévő jelforrás esetén a direkt jel (bal, illetve jobb fülbe akadálymentesen eljutó jel)
* A két fülbe jutó jel:
  + Fbal(t) = f1(t) + mbal·f1(t- τbal)
  + Fjobb(t) = f2(t) + mjobb·f2(t- τjobb)
* **Intenzitásos sztereótechnika**
  + Egy közös kapszulába épített állítható iránykarakterisztikájú mikrofon párt helyeznek el a hangtér közepén.
  + Mivel a mikrofonok egy helyen vannak, egy adott pontból érkező hang két vett jele között semmilyen időkésés sincs.
  + Az eltérő iránykarakterisztikák miatt azonban a két jel között jelentős intenzitáskülönbség lesz, ebből pedig bizonyos korlátokkal az irány meghatározható.
* **Időkéséses sztereótechnika**
  + Két külön mikrofont használunk, de egymástól egy pontosan megadott távolságban helyezzük el őket. A távolság akár több méter is lehet, a helyiség méretétől függően.
  + A két mikrofon iránykarakterisztikái azonosak.
  + A két jel között jelentős időkülönbség lesz, valamint részben intenzitás különbség is, ezekből pedig az irány meghatározható.
* **Műfejes sztereótechnika**
  + A műfejes technika alapötlete az, hogy leutánozza az emberi hallás környezetének külső geometriáját.
  + A felvételnél használt két mikrofonmembránt egy emberi felsőtestet és fejutánzatot tartalmazó bábun helyezik el a műfej füleinek dobhártyái helyén.
  + Ezzel azt modellezik le, ahogyan a hallott hang a két fülbe eljut, beleértve a test, az arc, orr és egyéb testrészek árnyékoló hatását is.
  + A műfejes technikával készített anyagokat elvileg csak fejhallgatón lehetne lehallgatni, éppen azért, mert a felvétel során a mikrofonmembránok az ember dobhártyái helyén fellépő hanghullámokat érzékelték.

**A sztereó hangtér visszaadása**

* A sztereó hangtér visszaállításához használt hangszóró elrendezés a rögzítésekor használt mikrofontechnikától független.
* Kétcsatornás esetben a két hangszóró általában egyvonalban, 2-5 m távolságra helyezkedik el.
* A két hangszóró összekötő egyenese adja a bázisvonalat, távolságuk a bázisszélesség.
* A visszaadható hangforrások iránya általában a bázisszélességen belülre korlátozódik.
* De lehetőség van bázisszélességen kívüli virtuális hangforrások realizálására is.
* Szubjektív lehallgatások eredményei azt mutatták, hogy a kétcsatornás sztereónál a sztereóhatás optimálisan csak egy szűk területen érvényesül.
* A hallgatónak középen, a két hangszóró szimmetriatengelyében kell ülnie, mindkét hangszórótól körülbelül bázisszélesség távolságban.
* Nagyobb létszámú lehallgatáskor egyáltalán nem biztosítható mindenkinek az optimális pozíció.
* A lehallgatási terület fent említett korlátozottsága volt az egyik mozgatórugója a többcsatornás rendszerek kifejlesztésének.

**Dolby Stereo**

* A 70-es években a Dolby Laboratories új hangformátumot dolgoz ki a 35 mm-es filmekhez, és Dolby Stereo-nak nevezi el.
* Optikai hangrögzítést alkalmaznak a filmen.
* Hogy az új formátum kompatibilis legyen a korábbival – ami mágneses volt – a plusz információt a régi csatornák helyén kellett elhelyezni.
* A rendelkezésre álló hely remek hangminőséget biztosított, de kettőnél több csatorna (sáv) felvitele – a zaj elfogadhatatlan mértékűvé növekedése miatt – nem volt lehetséges.
* Két csatorna a filmiparban nem elég, a bal és a jobb csatornákon kívül szükséges egy közép- és egy háttérhangokat közvetítő csatornát is beiktatni a rendszerbe.
* A megoldás a mátrixolás: a Dolby Stereo a 4 csatornát 2 sávra mátrixszolja.

**Dolby Surround**

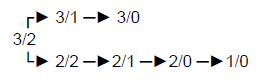
* A Dolby Surround kódolás elve megegyezik a filmiparban használt Dolby Stereo kódolás elvével, így a mozifilmek hangja egy az egyben átkerülhetett az új médiumokra.
* Elvei:
  + Mono/sztereo kompatibilis.
  + Oldalsó (surround) hangsugárzók: mono információt hordoznak.
  + L, R, C és S csatornák.

**Iránykiemelés: áthallások csökkentése**

* Ha kevesebb csatornán tudjuk átvinni, vagy kevesebb hangszórón tudjuk visszaállítani a műsort, mint az eredetileg készült, akkor a dekódolás után biztosan lesznek áthallások a csatornák között.
* A csatornák közötti áthallások zavaró hatása ellen az úgynevezett “iránykiemeléssel” lehet védekezni.
* A vissza-mátrixolás után a dekóder megvizsgálja, hogy van-e domináns irány a jelekben.
* Ha van domináns csatorna, akkor célszerű elvégezni az iránykiemelést, amely a jel domináns irányától függő erősítésszabályzást jelent: a domináns csatorna jelét erősítik, a többi csatorna jelét pedig csillapítják, de úgy, hogy az eredő lesugárzott hangteljesítményt ez a beavatkozás ne változtassa meg.
* Ezen és egyéb itt nem ismertetendő technikák alkalmazásával egy jól beállított Dolby Surround Pro Logic dekóder bármely két csatornája közötti áthallási csillapítás kb. 30 dB.

**Visszafelé kompatibilitás**

* **Kompatibilitás**: mind a korábbi kódolási eljárásokkal, mind a kisebb csatornaszámok fele. A korábbi kódolási eljárásokkal való kompatibilitás csökkentheti egy kódolás hatékonyságát.
* Mind az átlagos, mind a különlegesen jó lehallgatási körülmények között is lehetséges legyen a jó hangvisszaadás.
* **Megoldás**: a hanganyagot maximális dinamikával kódolják, és dinamika-szabályzó paramétereket szúrnak a kódolt audió programba, hogy a hangteljesítményt korlátok között tartsák.
* Többcsatornás rendszerek esetén a visszafele kompatibilitás:



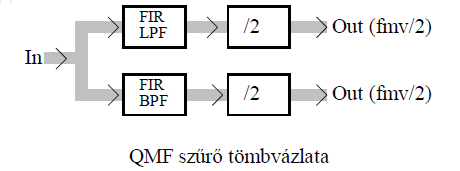
(**Jelölés**: Elöl lévő csatornák száma/Hátul lévő csatornák száma)

**5.1-es rendszer**

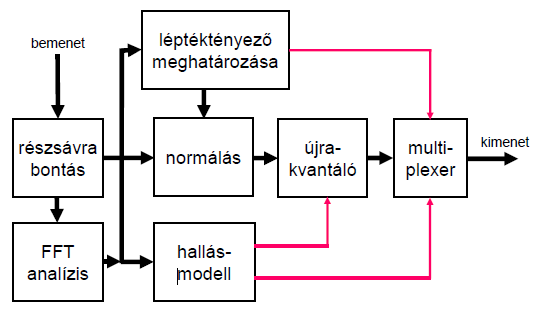
* A 3/2 rendszer kiegészíthető egy további kisfrekvenciás hangcsatornával. Ez az úgynevezett 5.1-es hangrendszer.
* Visszaállításkor a hangcsatornák kisfrekvenciás tartalmát egy közös, erre a célra készített mélyhangsugárzó adja le.
* A kisfrekvenciás hangsugárzó (subwoofer) a 20 -120 Hz közötti hangokat képes lesugározni. Erre a hangsugárzóra két kisfrekvenciás csatorna jelének összegét vezetjük rá:
  + **LFM** (*Low Frequency Main*) csatorna, ami nem külön átvitt csatorna, hanem csak egy dekódolt belső csatorna.
    - A 3/2 rendszerben használt hangszórók optimalizálása érdekében vezették be.
    - Mivel az emberi hallás irányérzékelése kisfrekvencián jelentősen leromlik, lehetőség van arra, hogy az összes csatorna kisfrekvenciás tartalmát egy közös, és ráadásul szinte tetszőlegesen elhelyezhető mélyhangsugárzón sugározzuk le.
    - Ez viszont azt jelenti, hogy az összes többi hangszórót “mentesítjük” a kisfrekvenciás komponensek lesugárzásától.
    - Ezáltal a hangszórók mérete, térfogata jelentős mértékben csökkenthető, és a fontos elektroakusztikai paramétereket kevesebb megkötés mellett lehet optimalizálni.
    - **Példa:** 
      * Ha ugyanazt a lesugárzott spektrumtartományt mélyhangsugárzó nélkül, öt teljesen egyenértékű (kisfrekvencián is sugárzó) hangszóróval oldanánk meg, akkor ~2,5-szeres nettó térfogatra lenne szükség.
      * Ehelyett egy közös mélyhangsugárzót és öt kisebb hangszórót használnak, amely esztétikailag is elfogadhatóbb megoldás, ráadásul segíti az 5.1 rendszer elfogadását lakószobákban és stúdiókban egyaránt.
      * A dekóder oldali LFM csatorna minden szükséges információt tartalmaz, ami a kisfrekvenciás tartalom visszaadásához kell, de opcionálisan kiegészíthető a járulékos, kóder oldalon beültetett LFE csatorna jelével.
  + **LFE** (*Low Frequency Enhancement*) csatorna, opcionális, kóder oldalon beültetett, ténylegesen átvitelre kerül, és a kisfrekvenciás tartományt egészíti ki.
    - Célja a kisfrekvenciás visszaadás teljesítmény szintbeli és frekvenciatartománybeli kiterjesztése.
    - Nagyszintű és kisfrekvenciájú jeleket tartalmaz.
    - Az LFE csatorna átvitele a meglehetősen kis sávszélessége (~20...120 Hz) miatt a 3/2 csatornák mellett nem igényel jelentős sávszélesség többletet, viszont a nyújtott hangélményt sokkal teljesebbé, telítettebbé teszi a megnövelt mélyhang-tartalom segítségével.

**Legfontosabb kódolási elvek**

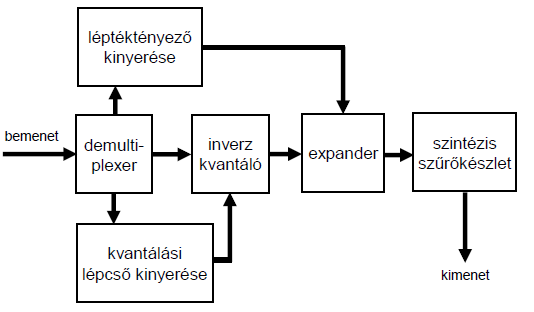
* **Prediktív kódolás**
  + az időbeni redundanciát csökkenti
* **Nem-egyenletes újrakvantálás**
  + Az egyenletesen kvantált mintákhoz egy keresési táblázat segítségével rendelünk új értéket.
  + Leggyakoribb változata a lebegőpontos kódolás.
  + Mintánkénti, mintacsoportonkénti megoldás is létezik.
  + Az újrakvantálási blokkidő: 1, 2 ms.
    - NICAM és a MAC: 14 bitről 10 bitre
    - DSR: 16 bitről 14 bitre
    - bitsebesség csökkentési faktor: 0,7-0,8.
    - hardware implementálás egyszerű.
    - a legmodernebb eljárások: 0,2 és 0,25 között.
* **Részsávos kódolás (Sub-band coding)**
  + Alapfeltételezések:
    - a hang spektruma nem egyenletes szerkezetű,
    - a legnagyobb komponens foglalja csak teljesen el a
    - dinamikatartományt,
    - a többi (lévén kisebb) nagy redundanciával kódolható.
  + Kihasználás:
    - a hangspektrum sávokra bontása,
    - Sőt! Hallás-elfedési jelenségek: kritikus sávok,
    - részsávonkénti maszkolási szint meghatározás,
    - ami a maszk alatt van, az nem hallható,
    - cél: újrakvantálási zaj ezen maszk alá kerüljön,
    - Kimeneti bitsebesség (fix?): bitújrakiosztás,
    - A bitújrakiosztás iteratív folyamat
  + **Két részsávra bontás (QMF)**
    - a hang spektrumának nem egyenletes voltát használja ki

****

* + - Jelmagyarázat:
      * + FIR: Finite Impulse Response (Filter)
        + LPF: Low Pass Filter (aluláteresztő szűrő)
        + BPF: Band Pass Filter (sáváteresztő szűrő)
        + QMF: Quadrature Mirror Filter
        + **A részsávos kódoló vázlata**



* + - * + **A részsávos dekódoló vázlata**

****

* **Transzformációs kódolás** 
  + a transzformációs "síkon" megjelenő redundanciát csökkenti
  + **Transzformáció**: áttérés egy másik síkra, az ott megjelenő redundanciát el lehet távolítani.
  + **DCT**-t (*Discrete Cosine Transformation*), vagy annak módosított változatát (*MDCT*) alkalmazzák.
  + Ha szükség van a frekvencia-tartománybeli analízisre, akkor az a leggyakrabban az **FFT** (*Fast Fourier Transformation*).
  + Dinamikus ablakméret váltás: az újrakvantálási zaj hallhatóvá válásának megakadályozására tranziens hangok esetén.
  + Ez utóbbit csak a legbonyolultabb eljárások alkalmazzák.
* **Pszichoakusztikus kódolások**
  + Alkalmazzuk az emberi hallás elfedési, vagy maszkolási modelljét.
  + Általában a hallás következő sajátosságait vesszük figyelembe:
    - hallás-elfedési tulajdonságok a frekvencia- és időtartományban,
    - a hang tonális és nem-tonális szerkezete,
    - a frekvenciától függő hallási pontosság.

# dia

**Az N szintű skalárkvantáló**

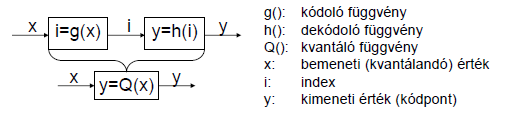
* Az N szintű skalár kvantáló a bemenetén egy folytonos értéket vár, a kimenetén viszont csak egy véges, N elemszámú halmazból vesz valós értékeket. Formálisan egy Q függvény, amelyre:
  + - Q: R {y1, y2, y3,..., yN}
  + ahol yi: i-dik kvantálási szint, (reprezentációs pont, kódpont).
* Az egy kódponthoz tartozó bemeneti értékek halmazát kvantálási cellának nevezzük. Formálisan
  + Ri = {x | Q(x) = yi }
* Az {Ri} kvantálási cellák az R (valós számok) egy partícióját adják:
  + Ri Rj = ésRj = R
* így minden bemenő értékhez egy és csakis egy kódpont tartozik.

**A granuláris és a nem korlátos tartomány**

* A kvantálási cellák két fajtáját különböztetjük meg:
  + nem korlátos cellák (overload cellák):
    - a kvantálási cella nem foglalható bele egy korlátos intervallumba
  + granuláris cellák:
    - a kvantálási cella befoglalható egy korlátos intervallumba (zárt vagy nyílt a vége a szomszédos celláktól függően)
* *A granuláris tartomány a granuláris cellák*, míg *az overload tartomány a nem korlátos cellák összessége.*

**Reguláris kvantálók**

* **Definíció (Reguláris kvantáló):** Egy skalár kvantáló reguláris, ha minden kvantálási cella egyetlen folytonos intervallum, melyet leír az alsó (egy cellánál ez lehet a –∞) és felső (egy cellánál ez lehet a +∞) határpontja, és a kódpont is benne van ebben az intervallumban.
* Reguláris skalár kvantáló megadásához az alábbi két adatok szükségesek és elégségesek:
  + a kvantálási cellák határpontjai (és hogy melyik cellához tartoznak)
  + kódpontok
* Így egy N pontú reguláris skalár kvantáló megadásához N darab kódpont és N-1 darab határpont megadása szükséges.
* **Optimális reguláris kvantáló**
  + A veszteséges tömörítés tervezésekor a legfontosabb cél, hogy adott tömörítési arányhoz a lehető legkisebb kvantálási torzítást érjünk el.
  + Egy optimális reguláris kvantáló megadásához meg kell határozni a kvantálási tartományok határát és a kódpontokat is.
  + Ezek együttes meghatározása általánosan nehéz feladat.
  + Ebben a részben a kódoló-dekódoló struktúrával vizsgáljuk meg, hogy milyen feltételek mellett lehet optimális kvantálót tervezni.
* **A kódoló-dekódoló struktúra**
  + A kvantálást általában az ún. kódoló-dekódoló struktúrával valósítjuk meg:
  + a kvantálási cellákat 1-től N-ig megszámozzuk
  + az y=Q(x) leképezést (x kvantálását) úgy hajtjuk végre, hogy számunkra nem az y kódpont, hanem a kvantálási cella indexe az érdekes
  + ezt az indexet küldjük el a dekódernek a csatornán
  + a dekóder aztán az index alapján vissza tudja állítani a kódpontot
  + A kvantálás így egy kódoló és egy dekódoló lépésből áll.



* **A legközelebbi szomszéd feltétel**
  + Más néven: optimális kódoló az adott dekódolóra tetszőleges d(,) torzítás-kritérium mellett
  + Adott a h() dekódoló függvény, vagyis az {y1, y2, y3,..., yN} kódpontok halmaza (kódkönyv, codebook).
  + A kódoló függvény feladata a partíciók meghatározása. Az optimális kódoló függvény:
  + g(x)=i d(x, yi) =minj[d(x, yj)]
  + Megjegyzés: ebből az is következik, hogy a döntés (kvantálás) nem függ az eloszlástól!

**A Lloyd-algoritmus kódkönyv tervezésére**

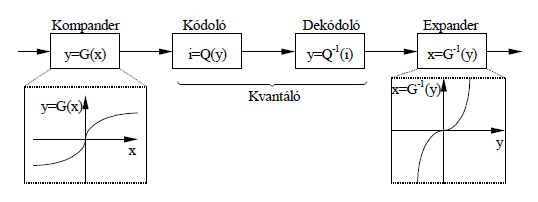
* A kvantálót {Ri} partíció és C={y1,y2,...,yN} kódkönyv adja meg. Cél a partíció és kódkönyv javítása lépésről lépésre.
* **A Lloyd-algoritmus alapötlete:** a kvantáló optimalizálása (a kódoló és dekódoló együttes optimalizálása) algoritmikusan nehéz feladat. Szuboptimum viszont elérhető úgy, hogy lépésről lépésre javítjuk a teljes rendszer torzítását azáltal, hogy felváltva optimalizáljuk a kódolót (a legközelebbi szomszéd feltétellel) és a dekódolót (a súlypont feltétellel). Megállapíthatjuk, hogy mindkét lépésben csökken a torzítás a megfelelő feltétel kielégítésével.
* **A Lloyd-algoritmus**

1. Vegyünk fel egy kezdeti kódkönyvet:
   * + C(0) = egy jó közelítés
     + m = 0
     + Számoljuk ki a torzítást!
2. Optimalizáljuk a partíciót a C(m) kódkönyvhöz
   * + a legközelebbi szomszéd feltétel kielégítése
3. Optimalizáljuk a kódkönyvet a kapott partícióhoz, így kapjuk C(m+1)-et.
4. Számoljuk ki, hogy mennyivel csökkent a torzítás.
5. Ha a torzítás már csak jelentéktelenül csökken, akkor vége, különben pedig folytassuk a 2. lépéstől m=m+1 értékkel.

* **Empirikus kvantálótervezés**
  + A Lloyd algoritmusban csak utaltunk a torzítás számítására, illetve a két optimalitási feltétel alkalmazására. A súlypont és a legközelebbi szomszéd feltételekben is feltételeztük, hogy ismerjük a sűrűségfüggvényt, de ez empirikus esetben nem feltétlenül ismert.
  + **Empirikus esetben:**
    - csak mintáink vannak a jelből, jelölje a tanító mintákat x1,x2,...,xM   
      (**tanítóminta:** ezekkel a mintákkal tanítjuk be a kvantálót).
    - súlypont számítás: cellán belüli átlagolás
    - legközelebbi szomszéd feltétel: a kvanálandó mintát összehasonlítom az összes kódponttal, és azt a kódpontot választom, ahol a torzítás a legkisebb

**Kompanderes kvantálótervezés**

* A kompanderes kvantálótervezés alapötlete: a kvantlást próbáljuk visszavezetni skalár kvantálásra úgy, hogy a bemeneti jel amplitúdóját egy limiter jellegű függvénnyel transzformáljuk olyan módon, hogy a kapott jel eloszlása közel egyenletes legyen, és erre már alkalmazhatjuk a skalár kvantálást, ami általában egyszerűen elvégezhető a mai processzorokkal (osztás vagy shiftelés).
* **Kompander:** olyan függvény, amely a skalár kvantálás előtt áttranszformálja a jelet
* **Expander:** a kompander függvény inverze a dekóder oldalon



**Az aszimptotikus kvantáláselmélet alapötlete:**

* nagyon nagy kódkönyvet használunk,
* ekkor a cellák már olyan kicsik, hogy azon belül az eloszlás már egyenletes eloszlással közelíthető, így a bonyolult sűrűségfüggvényt is lehet egyszerű módon közelíteni és kezelni.

**Vektorkvantálás**

* Egy k dimenziós vektorkvantáló a bemenetén egy valós értékekből álló többdimenziós vektort vár, a kimenetén viszont csak egy véges, N elemszámú halmazból vesz ki valós értékekből álló vektorokat. Formálisan egy olyan Q függvénnyel lehet megadni, amelyre:
  + Q: Rk {y1, y2, y3,..., yN}
  + ahol yi jelöli a kódvektorokat, ezeket nevezhetjük még reprezentációs pontnak vagy kódpontnak is.
* Az egy kódvektorhoz tartozó bemeneti értékek halmazát vektorkvantálás esetén kvantálási tartománynak vagy kvantálási cellának nevezzük. Egy kvantálási cella formálisan a következő módon adható meg:
  + Ri={x | Q(x) = yi }, ahol yi jelöli az i-dik kódvektort
* **Reguláris vektorkvantáló**
  + A reguláris vektorkvantálóra az igaz, hogy:
    - a kvantálási cellák konvexek
    - i-re yi Ri
  + Kódoló-dekódoló struktúra k dimenziós, N kódpontú kvantálóra az alábbi módon értelmezhető:
    - Rk { 1, 2, 3, ..., N} { y1, y2, y3,..., yN}
* **Definíció (Voronoi partíció):** Vektorkvantálás esetén az Rk partíciói közül kiemelten foglalkozunk azokkal, ahol minden partíción belüli kódvektor olyan, hogy a partíció összes pontjához közelebb van, mint az összes többi kódvektor. Ezeket a partíciókat Voronoi partícióknak (Voronoi cellák) nevezzük.

**A transzformációs kódolás**

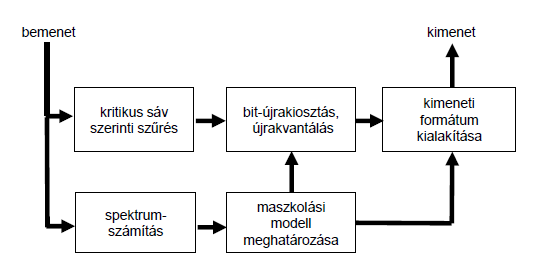
* A transzformációs kódolással az a célunk, hogy a bemenő jelet egy olyan jellé transzformáljuk, amelyiket adott bitszámmal kisebb kódolási hibával lehet tömöríteni.
* A transzformációt vektorokra végezzük, amely lehet egy skalár jelből kialakított blokk, vagy pedig eleve egy többdimenziós jel.
* A prediktív kódoláshoz hasonlóan a Benett-integrált vesszük alapul, amely szerint az optimális skalár kvantáló torzítása a jel prototípus sűrűségfüggvényétől és a szórástól függ (a Benett-integrált itt koordinátánként vizsgáljuk, hiszen azt skalár kvantálóra mondtuk ki).
* Akkor érdemes a transzformált jelet kvantálni az eredeti helyett, ha a transzformált jelek szórása jobban csökken az eredeti jelekéhez képest, mint amennyivel a prototípus sűrűségfüggvények tényezői nőnek
* **Menete:** a bejövő *K* dimenziós jelet egy *A* transzformáció után kvantáljuk, és a kvantált jelet az *A-1* inverz transzformációval visszaállítjuk, és azt reméljük, hogy az *A* transzformáció jó megválasztásával kisebb lesz az optimális kvantáló torzítása.

**Decimálás és interpolálás**

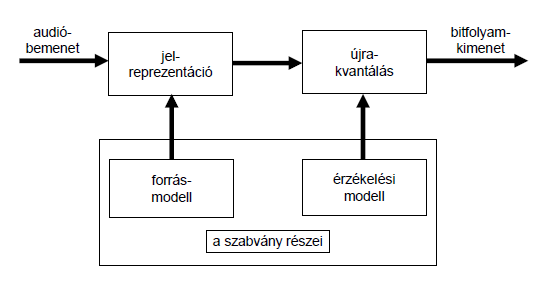
* Egy *N* mintaszámú jelet *K* sávra bontva *NK* mintát kapunk.
* A teljes frekvenciasávot *K* sávra bontva a sávszélesség (a frekvenciabeli felbontás finomsága) egy részsávnyi lesz, vagyis a teljes jelkészletre nézve a frekvenciabeli felbontás finomsága az eredetei *K*-szorosa.
* Viszont a határozatlansági reláció alapján így az időbeli felbontás finomsága *K*-szorosával kisebb lesz, tehát minden *K* minta hordoz csak információt.
* **Decimálás:** Így *K* mintából elegendő csak 1-et megtartani (ezt a műveletet decimálásnak hívjuk, és a **jelölése** *K*). Így végül is az *N* mintából szintén *N* minta lesz a részsávra bontás után.
  + Láttuk, hogy a kódoló minden részsávban *K* mintából csak 1-et tartott meg.
* A dekódoló a kódolási műveletek inverzét végzi el. Mivel a decimálás a kódolóban a részsávra bontást (sávszűrés) követi, ezért annak inverze, az **interpoláció** (jelölése: *K*) a részsávokból való visszaállítás előtt hajtódik végre.
* **Kérdés**, hogy mi kerüljön a decimálás során elhagyott minta helyére a dekódoláskor? A **válasz**: nulla interpolációt kell végezni.
* Így a decimálás és interpolálás hatása olyan, mintha egy K mintánként periodikus   
  …10K-1 10K-1… 10K-1… jellel szoroztuk volna meg a részsáv tartalmát.
* **A decimálás és interpolálás hatása** olyan, mintha a részsáv K mintáját megszoroztuk volna a 10K-1 jellel.
* A szűrés→decimálás→interpolálás→szűrés hatása ugyanaz, mint a szűrés→szorzás→szűrés hatása.

# dia

**Audiókódoló blokkvázlata**

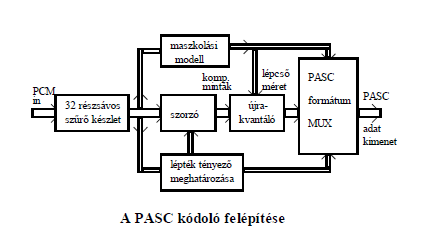


**A hagyományos audiókódoló blokkvázlata**



**PASC (Precision Adaptive Sub-band Coding)**

* Az egyik első bitsebesség csökkentési mód a DCC-kben (Digital Compact Casette).
* A rögzítési adatsebesség: 384 Kbit/s
  + 48 KHz-es mintavételi sebesség, 18 bites minták
  + De megengedett a 44,1 és a 32 KHz-es mintavételi frekvencia is.
* A szükséges tömörítési tényező 0,22.
* Tisztán részsávos kódolás
* Továbbfejlesztési lehetőség: pl. MPEG-1 Layer I

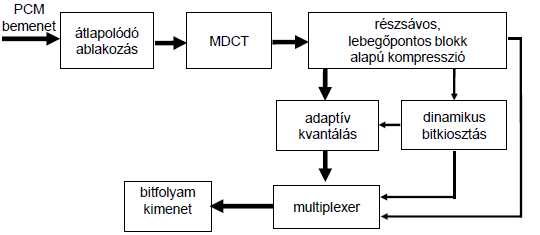


**ATRAC (Adaptive Transform Acoustic Coding)**

* A MiniDisc céljaira fejlesztették ki.
* Hibrid kódolás:
  + részsávos kódolás
  + lebegőpontos normálás és újrakvantálás
  + módosított diszkrét koszinusz transzformáció
  + blokkhossz-logika: a tranziensek függvényében, az újrakvantálás elő- és utóvisszhangja maszkolható
* Transzformáció utáni koefficienseket továbbítja
* Tömörítési tényező: 0,2

**AC-2 és AC-3 kódoló**

* Az amerikai ATSC hangátviteli szabványa. Az AC-2 sztereo, míg az AC-3 sokcsatornás hangátviteli eljárás
* Átlapolódó ablakozás és MDCT
* TDAC (Time Domain Aliasing Concellation)
* A részsáv együtthatóinak blokk alapú lebegőpontos normalizálása:
  + A blokk energiának az exponens részét továbbítják (vagyis log2() függvényét)
  + Ebből adódik ki a jel logaritmikus spektrumburkolója
  + A logaritmikus spektrumburkolóból számítják a maszkot
  + A bit újra-kiosztás bemenő adata így a blokk exponensek (ebből számítható a maszk)
* A dinamikus bitkiosztás a dekóderben is lefut, ellentétben az eddig ismertetett kódolókkal. Így sokkal kevesebb adatot kell átküldeni (mint pl.: maszk prototípus alakja (4 db egyenes)).

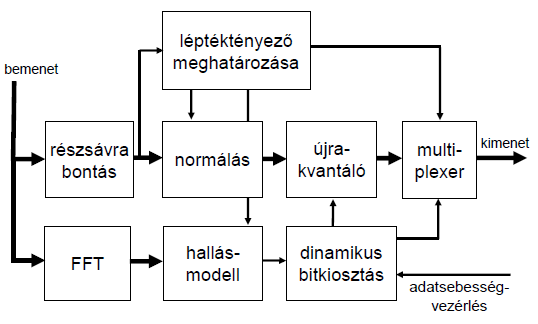


**MPEG kódolás**

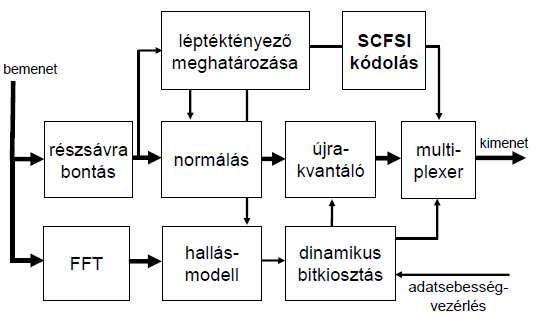
* **Cél**: egységes digitális kép és hang redundancia csökkentő rendszer-család kidolgozása
* MPEG-1: “ISO/IEC 11172 Information Technology Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mb/s”.
* A következő **fő szempontok**at vették figyelembe:
  + kiindulási hangminőség a CD
  + különböző tömörítési fokok és hangminőség
  + az átviteli hibákkal szembeni védettség
  + kódolás/dekódolás bonyolultsága, késleltetési idő
  + kaszkádosíthatóság
* Munkájuk eredménye az MPEG szabványok 3. része, az audiókódolás
* 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
    - elsősorban DAB (Digital Audio Broadcasting) 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

**Az MPEG-1 Audio**

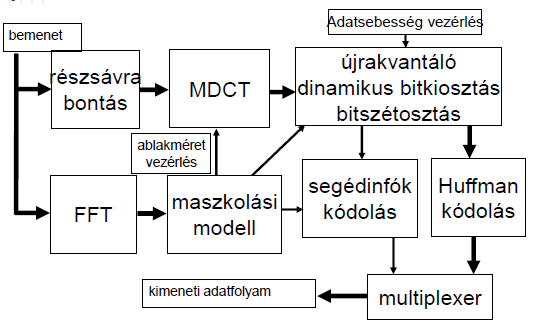
* **Általános**
  + Mintavételi frekvenciák:
    - 48; 44,1; 32 KHz
  + Kimeneti adatsebesség csatornánként:
    - 32 - 384 Kbit/s között diszkrét lépésekben állítható
  + Csatornák:
    - mono, sztereo, kettős, kapcsolt-sztereo (joint stereo)
  + Csak a kimeneti adatfolyam szabványos
  + 3 réteg definiált:
    - különböző alkalmazások és adatsebesség tartományokra
* **Layer I**
  + A MUSICAM egyszerűsített változata
  + PASC-kal összevetve két különbség:
    - a kimeneti bitsebesség vezérlés lehetőségében
    - jobb felbontását biztosító FFT-ben
  + Főbb jellemzői:
    - kis komplexitás (a kóder és a dekóder oldalon egyaránt)
    - rövid kódolási/dekódolási időkésleltetés
    - nagy időfelbontás



* **Layer II**
  + Az I. réteghez képest két ponton van eltérés:
    - lépték-tényező választási információ (SCFSI),
    - 1024 pontos FFT.
  + Főbb jellemzői:
    - bonyolultabb kódoló a tömörítési hatékonyság növelése érdekében
    - kis komplexitású dekóder (aszimmetrikus)
    - rövid időkésleltetés (~45…50 ms)
    - editálhatóság, konstans kerethossz
    - erős bithiba védettség

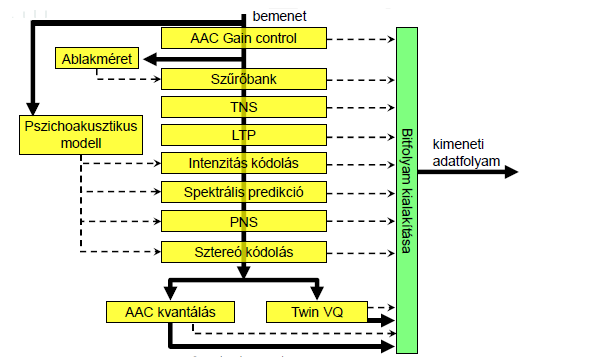


* 1. **Részsávos szűrő analízis**
     + 32 db azonos sávszélességű részsáv, 750Hz sávszélességgel (@48 kHz), sokfázisú szűrőkészlettel
  2. **Léptéktényező-számítás**
     + Alapegység a keret (keretidő = 24 ms @ 48 kHz), amely minden részsávból 36 mintát tartalmaz (összesen 36 x 32 = 1152 minta).
     + Egy részsáv mintáit 12-esével egyszerre normálják (8 ms).
     + A léptéktényező 6 bites, egy részsávban egy kerethez 3 db léptéktényező tartozik (12 + 12 + 12 = 36).
     + Definiálnak egy ún. léptéktényező kiválasztási információt **(Scale Factor Select Information; SCFSI)** amely az egymás utáni léptéktényezőkben meglévő redundanciát csökkenti a kódoláskor (3 egymás utáni léptéktényező értéke részsávon belül alig tér el egymástól).
  3. **Pszichoakusztikus modell**
     + Minden részsávban meghatározzuk az éppen észrevehető zajküszöböt (ezt maszknak nevezzük) és a jel nagyságát. Eredményként részsávonként a jel-maszk arányt kapják
     1. **FFT számítás** (adatok 48 kHz esetében):
        + 1024 pontos (21,3 ms hosszú), 16 komponens / részsáv
        + Hann-ablakfüggvénnyel, 100%-os átlapolódással,
        + A frekvencia felbontás: 48KHz / 1024 = 46,875 Hz
     2. **Szintszámítás minden FFT-komponensre**.
     3. **Csendszint-megfontolás**:
        + Minden részsávban minden komponensre egy táblázat, tartalmazza a hallhatósági küszöböt, pl. 46 Hz-re 42,1 dB, vagy pl. 3 kHz-re -4,45 dB, de pl. 15 kHz-re 51,04 dB.
        + 96Kbit/s felett minden értékből 12 dB-t el kell venni (!)
        + Elhagyjuk a maszk alatti komponenseket.
     4. **Tonalitás**: kritikus sávokon belül a maximumok megkeresése: a tonális komponensek megjelölése, a nem-tonális komponensek megkeresése
     5. **Tonalitás és nem-tonális maszkolók megkeresése**:
        + A maszkoló komponensek számának csökkentése
        + Több, kritikus sávon belüli maszkoló komponens által okozott együttes maszk meghatározása
     6. **Részsávonkénti egyedi maszkolási görbék meghatározása**
     7. **Teljes maszkolási görbe meghatározása**
     8. **Részsávonkénti minimális maszk meghatározása**
     9. **Jel-maszk arány számítása részsávonként**
* **Layer III**
  + Jelentősen eltér az első kettő rétegtől
  + ASPEC, MUSICAM leghatékonyabb elemeiből
  + Részsávonkénti 12 vagy 36 pontos MDCT: a transzformációs ablakméret vezérelhető
  + Huffman-kódolás
  + Adaptív bitsebesség-vezérlés
  + További átmeneti tárolási igény
  + A kódolás/dekódolás késleltetési ideje megnövekszik
  + **Legfontosabb jellemzői:**
    - bonyolult kóder- és dekóderfelépítés (off-line kódolás)
    - rendkívül hatékony tömörítés
    - ott használják, ahol a hatékony tömörítésen, és a nagyon kis bitsebességen van a hangsúly



**Az MPEG-2 Audio**

* **Általános**
  + **LSF (Low Sampling Frequency):** 
    - **kisebb mintavételi frekvenciák felé való kiterjesztés**
    - A hang érzékelési entrópiája nem függ a mintavételi frekvenciától: a lehallgatási tesztek szerint sokkal kevésbé zavaró a sávkorlátozás hatása, mint a kódolásból adódó melléktermékek megjelenése
    - Ezért új mintavételi frekvenciák: 16, 22.05 , 24 kHz
    - A kódolási nyereség megnő, ugyanakkora bitsebesség mellett a szubjektív minőség javul
    - A hang minőség 64 kbit/s alatt lényegesen jobb lesz
    - A kódolás közel azonos az MPEG-1 Audióval
    - Az LSF hátránya a durvább időfelbontás, így a tranzienseket nehezebb lekezelni
    - A kvantálási zaj elő-visszhangja hallhatóvá válhat
    - A kóder és a dekóder könnyen implementálható
  + **BC (Backward Compatibility):** 
    - **visszafelé kompatibilis többcsatornás rendszer**
    - MPEG-1 kompatibilis többcsatornás kiterjesztése.
    - A surround (3/2 és 5+1) zenei műsorok és a több nyelvű
    - 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 lehessen hallgatni
      * Előre (forward) kompatibilitás:
        + az MPEG-2 dekóder képes legyen dekódolni az MPEG-1-es bitfolyamot
    - A kompatibilitás követelmény: komoly hátrány
  + **AAC (Advanced Audio Coding):** 
    - **visszafelé nem kompatibilis többcsatornás rendszer**
    - **Főbb célkitűzései:**
      * Sok mintavételi frekvencia támogatása (8-96 KHz).
      * Csatornakiosztásra az 1/0-tól kezdve az 5+1-ig minden értelmes kiosztást megenged
      * Eredetinél kevesebb számú csatornán is reprodukálható legyen
      * Nagy bitsebesség-megtakarítás, a kompatibilitás árán is
      * Az „eredetitől megkülönböztethetetlen” minőség 384 kbit/s (5 csatorna esetében)
    - 1996. nov. 320 kbit/s (5 csatorna bitsebesség mellett) „az eredetitől megkülönböztethetetlen” minősítést kapta az EBU szerint.



* **Szűrőbank- és blokkméretválasztás**
* Időtartomány Frekvenciatartomány
* Transzformáció: MDCT
* Blokkméret:
  + stacionárius szakasz: 2048 minta
  + tranziens szakasz: 256 minta
* A blokkok 50%-ban átlapolódnak
* Különböző blokkméret átlapolása: az első és utolsó 8 rövid keretre külön-külön speciális ablakot definiáltak
* **TNS (Temporal Noise Shaping)**
* A transzformáció utófeldolgozási lépése (csak hosszú blokkra):
  + Az MDCT spektrumot szűrik a kvantálás előtt, azaz szűrt MDCT spektrumot kvantálnak, nem pedig az eredetit.
  + Elküldik a használt szűrőegyütthatókat
  + Jelenlegi implementációk (MPEG-2 AAC, MPEG-4 GA): a spektrumot különböző régiókra osztjuk, mindegyiknél 3 szűrő közül lehet választani
* A TNS egy kereten belüli dinamikus kódolás (ellentétben a spektrális predikcióval): Frekvenciatartománybeli szűrés, amely célja a kvantálási zaj alakjának időben változó átformálása a hallás számára kedvező módon
* **(Adaptív) spektrális predikció**
* Egy tonális tartalmú hang spektrumának megfelelő részei általában több kereten keresztül azonosak vagy nagyon hasonlóak.
* A frekvenciatartománybeli mintákat a két megelőző keretből becsüljük és a predikciós hibát visszük át (keretek közötti adaptív predikció).
* Skálafaktor tartományokként ki-be kapcsolható, egyébként adaptív a prediktor együtthatók meghatározása.
* Csak hosszú blokkokra használható, a rövidekre nem, ekkor automatikusan kikapcsolt (hiszen tranziens jellegű a szakasz, nem tonális).
* **PNS: Perceptual Noise Substitution**
* A zajok általában egyformák
* A zaj „színének”, paramétereinek visszaállítása kis bitsebességen alacsony prioritású
* Lépései:
  + Zajdetektálás minden skálafaktor sávban
  + Ahol csak zaj van: zaj paraméterek kódolása
  + Ahol nem csak zaj: az eredeti kódolás
* **Többcsatornás kódolás**
* Középcsatornás sztereó:
  + Előállítják a monó és a különbségi jelet
  + Minél inkább monó jellegű a jel, annál hatékonyabb.
  + Inkább magasabb bitsebességen használják
* Intenzitás/kapcsolt csatornás kódolás:
  + Közös spektrum részek, hang + irány elven kódolva
  + Inkább alacsonyabb bitsebességre
  + Függően összekapcsolt csatornák: azonos ablakméret váltás, a transzformáció előtt keverhetők az együtthatók
  + Függetlenül összekapcsolt csatornák: tetszőleges ablakméret minden csatornán, csak időtartományban keverhetők az együtthatók

**Az MPEG-4 Audio**

* **Célkitűzése**
  + Interaktív multimédia célokra
  + Minden értelemben léptékelhető kódolási módok alkalmazása és definiálása
  + A beszéd és a zene és minden egyéb audió objektum-orientált kezelése
  + Az audió objektum fogalmának bevezetése
  + Javított minőségű hangátvitel és kódolás
  + Az addig élesen elkülönülő minőségi hangkódolás, a beszéd forráskódolás, valamint a szintetikus zene együttes, integrált kezelése
* Az MPEG-4 által felkínált eszközkészlet: szabadabban konfigurálható adott célra, alkalmazásra, adott alkalmazás több komplexitási szintjéhez egyaránt
  + **Bitsebesség-skálázás:** Kis sávszélesség esetén az alacsony sebességű vevők csak a kis bitsebességet használják, míg a komplexebb vevők mindkét bitfolyamot
  + **Sávszélesség-skálázás:** A kis sebességű bitfolyamba a spektrum alkalmazás szempontjából “fontosabb” részét kódoljuk, hogy ezt az egyszerűbb vevők is dekódolni tudják
  + **Kódoló-komplexitás skálázás:** Különböző komplexitású kódolásokat tesz lehetővé a kódolóban
  + **Dekóder-komplexitás skálázás:** Különböző komplexitású dekódolásokat tesz lehetővé ugyanabból a bitfolyamból

# dia

**A képdigitalizálás előnyei**

* jel-zaj viszonyt nem befolyásolja a tároló és az átviteli közeg
* másoláskor nincs minőségromlás
* könnyen integrálható a digitális hang és kép tetszőleges adattal
* könnyű a titkosítás
* számítástechnikai alkalmazások, multimédia

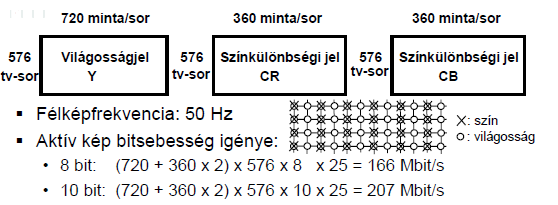
**Stúdiótechnikai színek, mintavételezés**

* A televíziós jel mintavételezése során a mintavételi frekvencia és a mintánkénti bitek száma a legfontosabb.
* A stúdiótechnika jelenleg szinte kizárólagosan az ún. komponens jeleket (Y,CR,CB) használja.
* A világosságjel mintavételi frekvenciája normál azaz 4:3-as képméret-arányú esetben: 13,5 MHz, míg a két színkülönbségi jel mintavételi frekvenciája: 6,75 MHz.
* 16:9-es képméretarány esetén 18 MHz.
* De különböző egyedi alkalmazástól függő frekvenciákat is használnak.

**Tömörítés szükséges**

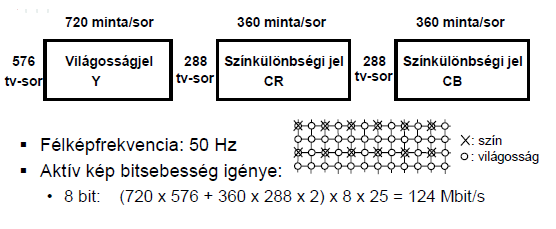
* Ezt tárolni nagyon nehéz (pl. D1, D5).
* De továbbítani szinte lehetetlen (sávszélesség).
* Az első digitális videó átemelések 140 Mbit/s-ot használtak (ETSI 140 tömörítés).
* Majd az ETSI 34 Mbit/s-os kódolás következett.
* Az adatsebességet csökkenteni kell!

**ITU-R- BT. 601 4:2:2 stúdióformátum**

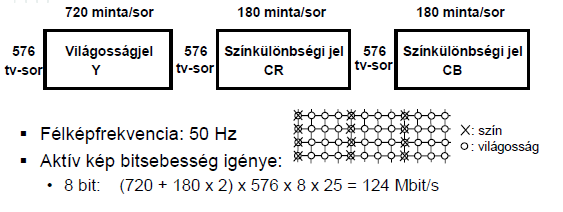
****

* Teljes bitsebesség igény: a teljes tartalmat mintavételezzük
* Y jelre: 13,5 MHz mintavételi frekvencia
* Cr és Cb jelre: 6,75 MHz mintavételi frekvencia egyenként
* Másodpercenként így 27M darab minta (kép+szinkron tartalom is)
* **Bitsebesség: 8 bit: 216 Mbit/s 10 bit: 270 Mbit/s**

**4:2:0 formátum (MPEG, DVC, JPEG)**

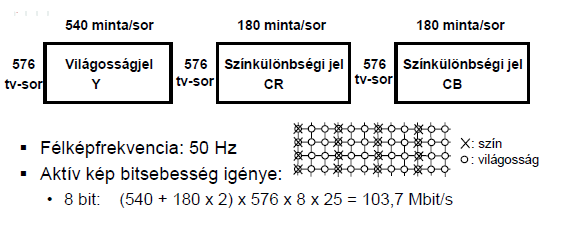
****

**4:1:1 Komponensformátum (PAL, DVC, JPEG)**

****

* Megjegyzés: ugyanannyi minta, mint 4:2:0 esetén, mert
  + az Y jel mérete továbbra is 720 x 576
  + 4 darab Y-ra változatlanuk 1 db Cr és Cb jut

**3:1:1 formátum (Non-linear Editing System)**



* Összehasonlítás az 576 soros 4:1:1 rendszerrel:
  + soronként ugyanúgy 180 Cr és Cb komponens van,
  + az Y komponensből lett kevesebb a 4:1:1-hez képest (540 / 720 vagyis 3 / 4)

**Képtömörítés: redundancia**

* Statisztikai redundancia: a képpontok (pixelek) adott környezetén belül (képen belül és időben is) hasonlók.
* Az emberi látás tulajdonságai: a videojel az emberi látórendszer (HVS: Human Visual System) számára lényegtelen részleteket is tartalmaz, amely eltávolítható. Tömörítéskor kihasználjuk a HVS tulajdonságait (részletek észrevétele, mozgás követése).
* A mozgóképben rejlő redundanciák (konkrétabban):
  + Térbeli: Intra-frame (kép) és Intra-field (félkép) kódolás,
  + Időbeli: Inter-frame kódolás (képek közötti).

**Veszteséges képkódolás minősége**

* Veszteséges kódolás minősítése:
  + Objektív úton:
    - MSE (Mean Square Error)
    - Peak SNR
  + Szubjektív úton (MOS = Mean Opinion Score)
  + Olyan egzakt pszichofizikai mérőszám, amely a HVS összes
* tulajdonságát figyelembe veszi jelenleg NINCS!

**Bitsebesség (bit rate) definíciók**

* + 1. Bitsebesség = képen belüli bitek (átlagos) száma

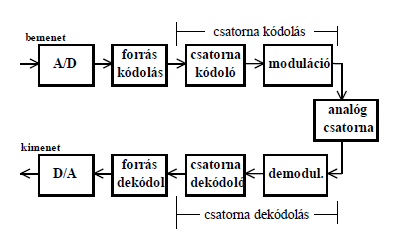
[bit/pixel, bpp], (álló és mozgó képre is),

* + 1. Bitsebesség = másodpercenként továbbított bitek száma

[bit/sec], (mozgó képre).

* **Kódolási technikák:**
  + Veszteségmentes:
    - változó szóhosszúságú kódolás
    - futamhossz kódolás
  + Veszteséges:
    - alul-mintavételezés
    - kvantálás
    - vektor-kvantálás
    - stb.

**Digitális csatornamodell**

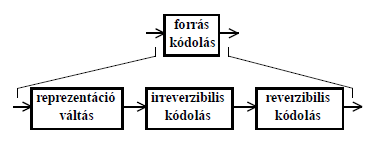


**Forráskódolás**

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

**Csatornakódolás**

* A csatorna átviteli tulajdonságait figyelembe véve választunk modulációs eljárást és hibavédelmi algoritmus.

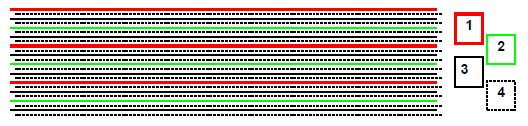


**A forráskódolás összetevői**

* Az új ábrázolási síkban kevesebb redundancia. Fajtái:
  + DPCM, predikció
  + transzformációs kódolás
  + részsávos kódolás
  + mozgáskompenzáció
* Irreverzibilis kódolás:
  + A pontosság csökkentése, a „lényegtelen” részek eltávolítása, megjelenik a kvantálási hiba, az elérhető kompressziós tényező nagy.
  + Megjelenési formái:
    - kvantálás
    - térbeli, időbeli alul-mintavételezés
* Reverzibilis kódolás:
  + Általában hatékony kódszó kijelölés, limitált kompressziós tényező, a jel reprezentáció redundanciáját csökkenti.
  + A leggyakrabban alkalmazott megoldások:
    - változó szóhosszúságú kódolás (VLC: Variable Length Coding),
    - futamhosszkódolás (RLC: Run-Length Coding).

**A GIF-KÓDOLÁS**

* **A GIF87a tömörítés**
  + A GIF87a egyetlen palettás képet tömörít veszteségmentesen.
  + **Paletta** méret: 1, 2, 4, 8 bit
  + Paletta színei: 24 bites RGB-ből
  + **Algoritmus**: Lempel-Ziv-Welch (LZ-78 Welch-féle kiterjesztése)
  + Max. 4096 (12 bit) kódszó, benne két kivételes szimbólummal:
    - **törlés szimbólum:** a szótárméret progresszíven növekszik, maximális értéke 4096, de ezután a módszer statikussá válik (a szótár nem nő tovább). Ha azonban a tömörítés hatékonysága romlani kezd, akkor a szótár kiüríthető (a kiürítést ez a kód jelzi a dekódernek) és így ismét adaptívvá válik a tömörítő.
    - **adatvég szimbólum:** ez a szimbólum jelzi, hogy elértük az utolsó képpontot. Elvileg felesleges, hiszen a képméretet tudjuk, és ez abból is kiderülne, mégis a hibadetektálás céljából szerepel.
* **A GIF89a tömörítés**
  + A GIF87a kiegészítése több új funkcióval, például:
    - a paletta egy színe 100%-osan transzparens lehet
    - mozgókép egymás utáni kockái lehetnek egy fájlban:
      * van egy globális képméret, a képek ezen belüli téglalapok (nem kötelező a teljes képet kitölteni, elég csak azt a téglalapot elküldeni, ahol változás történt)
      * megadható, hogy hány msec-ig kell egy képet kirajzolni
      * megadható, hogy új kép esetén mi legyen a régivel:
        + az új felülírja a régit (transzparens képpontok lehetnek, így az álló részek maradhatnak)
        + a régi képet egy háttérszínnel töröljük
        + a régi kép előtti képet állítjuk vissza
* **Alternatív letapogatás (GIF)**
  + **Lépései:**
    1. először minden 8-dik sort olvassuk ki (y 0 mod 8)
    2. ezután minden 4-dik sort olvassuk ki, ha már az előző körben kiolvastuk a sort, akkor azt nem kell még egyszer végigolvasni (y 4 mod 8)
    3. ezután minden 2-dik sort olvassuk ki, ha már az előző körben kiolvastuk a sort, akkor azt nem kell még egyszer végigolvasni (y 2 mod 4)
    4. ezután minden páratlan sorszámú sort (y 1 mod 2)

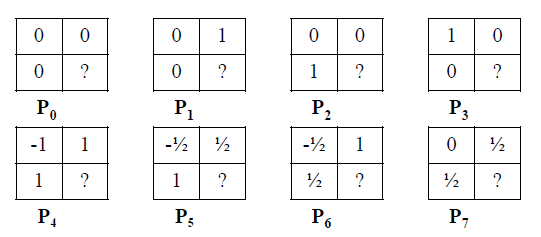


**A JPEG veszteségmentes tömörítő**

* Veszteségmentes prediktív differenciális kódolás a szomszédos minták alapján
* A predikciós hibát tömörítik Huffman- vagy aritmetikai kódolóval
* A becslés a képen belül már dekódolt mintákon alapul az alábbiak szerint:
  + A becslésben a bal (nyugati) szomszéd, az előző (felső) sorban lévő szomszéd (északi), illetve e kettő szomszéd közös szomszédja (észak-nyugati) használható.
  + 8 különböző becslőfüggvény van
  + A képet tartományokra lehet osztani, a becslőfüggvény a tartományon belül nem változik, kivéve a bal és felső széleket.

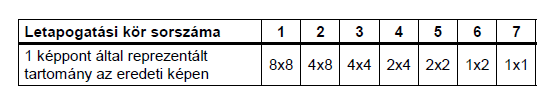
**A JPEG prediktorai**

* Összesen 8 prediktort használ
* A becslőfüggvény a tartományon belül nem változik, kivéve a bal oldali oszlopot, illetve a legfelső sort, itt mindig a P1, illetve a P2 prediktor használandó.

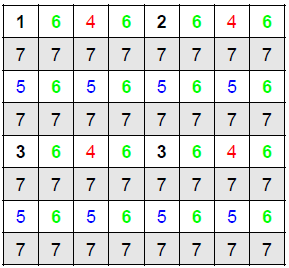


**A PNG-kódolás (Portable Network Graphics)**

* Palettás, szürkeskálás és RGB képekre
* A GIF és a veszteségmentes JPEG kódoló jellemzőit ötvözi
* Működése:
  + max. 8 bites komponensekre bontás
  + képpontok letapogatása
  + predikció a szomszédos mintákból
  + predikciós hibát az LZ77-tel tömöríti: szótárméret  32 kByte
  + Huffman-kódolás
* **Alternatív 2D letapogatás (PNG)**
  1. letapogatási kör: egy képpont egy 8x8-as területnek felel meg
  2. letapogatási kör: 4x8-as
  3. letapogatási kör: 4x4-es
  4. letapogatási kör: 2x4-es ....



* A letapogatás megadható egy 8x8-as területen belül, ez alapján a teljes képre kiterjeszthető a letapogatás.



* **PNG prediktorai**
  + Összesen 5 prediktort használ
  + 4 db szomszédságon alapuló lineáris prediktor

