

Médiatech órán feltett kérdések

[tavalyi kidolgozás](#)

Hogy jön ki a telefon hangminőségénél a 8 kbit/s bitsebesség?

mintavételezés:

Telefon, beszédhang esetén 300-3400 Hz között van a hallható és érthető beszéd frekvenciája. Ez a védősávokkal együtt kb 4000 Hz.

Mintavételi tétel(Shannon-Nyquist kritérium) : (2-szeres mintavétel kell)-> $4000\text{Hz} \cdot 2 = 8000\text{Hz}$ (8000 minta)

kvantálás:

8000 mintád van és ezeket lekódolni 2^8 sávban, minden egyes mintát 1 bájtól tárolunk digitalizálási vagy kvantálási mélység (az adatszó digitális számjegyeinek száma) a digitális telefonhálózaton: 1 bájt=8 bit ($2^8 = 256$ lehetséges érték)

kódolás:

minta(Hz)*csatornák száma (darab)*bitmélység (bit)

"Egy Mono 8Khz-es 8 bites kódolás esetén :"

$8000\text{Hz} \cdot 1\text{ch} \cdot 8\text{bit} = 64000\text{bit} = 64\text{kbit}$

64kbit méret 8biten szekundumonként: $64\text{kbit}/8/\text{s} = 8\text{kbit/s}$

(mebibyte-ban: $((8 \times 8000) / 1,048,576) / 8 \times 60 = 0.4577 \text{ MiB/min}$)

Miért viszi a levegő víz felett a hangot (erősíti)?

A víz hullámtanilag ritkább közeg -> segíti a hang terjedését

Mert a víz felszínén jobban terjed a hang (visszaverődik a vízről)?

??A Dunás kérdés nem emlékszem.

??Ha találkoznánk egy földönkívülivel akkor hogy magyaráznád el neki, az 1 nap fogalmát?(vagy 1 óra vagy 1 perc...)

Alapállapotú cézium 133 atom 2 hiperenergiaszintje közti átmenetének megfelelő sugárzás x periódusának időtartama. forrás:

https://hu.wikipedia.org/wiki/A_c%C3%A9zium_izot%C3%B3pjai

[Adnék neki egy ilyen órát.](#) Csodálatos. :D

Shannon féle csatornamodell helyes sorrendje

tömörítő -> titkosító -> csatornakódoló

Hány kérdésből találjuk ki hogy melyik hallgatóra gondol, igen nem kérdéssel?

50 hallgatónál 6 kérdés elég, mert $2^6 = 64$, minden hallgatónak van egy sorszáma. felírjuk a hallgatók (sor)számát 6 biten tárolva, majd bitenként kérdezzük, hogy 0 vagy 1

Hőtani és információelméleti entrópia közötti kapcsolat?

A rendszer rendezetlenségével kapcsolatos mindkettő.

Az entrópia és a rendezetlenség egyenértékűsége elvben még a termodinamikában felbukkan, de végleg [Erwin Schrödinger](#) tisztázta az életjelenségek kapcsán. Később – a formai hasonlóság alapján – [Neumann János](#) javasolta Shannonnak, hogy képletét nevezze entrópiának. -wikipédiáról

Milyen kódok egyértelműen dekódolhatóak?

Fix hosszúságúak, de rajtuk nem lehet tömöríteni.

A prefix kódokkal lefedjük az egyértelműen dekódolhatóság fogalmát.

Egy K kód prefix kód, ha egyik kódszó sem kezdődik egy másik kódszóval.

For example, a code with code words {9, 5} has the prefix property; a code consisting of {9, 5, 59, 55} does not, because "5" is a prefix of both "59" and "55".

McMillen egyenlőtlenség!!

Kraft egyenlőtlenség !!

Entrópia lehetséges minimuma és maximuma

n db elem esetén:

Min: 0 (ha az egyik elem valószínűsége 1, a többi meg 0) (ha pl egy vg-e adott, a többit úgy rendezzük, hogy az egyik megkapja a maradék vg-et, a többi pedig 0-t)

Max: $\log_2(n)$ (ha az összes elem valószínűsége azonos $P = (1/n)$) (ha pl egy vg-e adott, akkor a maradékot egyenletesen elosztjuk a többi vg közt)

Miért halljuk a mély hangokat távolabbról?

Hullámhaladás miatt, magas hang sugár irányba megy, mély elterül.

Audiókódolásban mik az alapok?

Nem egyenletes újrakvantálás, ...

Miért segít az egyenletes kvantálás?

Azáltal, hogy ..., nem kell dekódolni.

Mi az MP3?

MPEG-1 szabvány 3. rétegének a kódolója.

Miért nincs MPEG3 szabvány?

Hogy ne keverjék össze az MP3-mal.

Milyen szerepe van a Huffman kódolásnak a veszteséges tömörítésnél?

A belső adatstruktúrákat tárolja.

MPEG2, AAC eredmények, trükkök. (?)

AAC: fele akkora bitsebességgel felül lehet múlni az MPEG 1, 2. rétegét. és 70%-kal a 3. rétegét

MP3 vs AAC

AAC-vel 96kbit/s-el el lehet érni azt, amit MP3-mal 128kbit/s-el.

AAC hátránya: nem visszafele kompatibilis a régi verziókkal.

Ez is kell:

Adatsebesség [Mbit/s]	1 sec video	1 perc video	1 óra video	1000 óra video
Real video (56 kbit/s) Internet	7 Kbájt	420 Kbájt	25,2 Mbájt	25,2 Gbájt
MPEG-1 (1,5 Mbit/s) (LDTV: VHS)	187,5 Kbájt	11,25 Mbájt	675 Mbájt	675 Gbájt
MPEG-2 MP@ML (4Mbit/s) (SDTV minőség)	0,5 Mbájt	30 Mbájt	1,8 Gbájt	1,8 Tbájt
MPEG-2 MP@ML (8Mbit/s) (Szétosztási minőség)	1 Mbájt	60 Mbájt	3,6 Gbájt	3,6 Tbájt
MPEG 422P@ML (18Mbit/s) Hírminőség	2,25 Mbájt	135 Mbájt	8,1 Gbájt	8,1 Tbájt
DVCPRO25 (25Mbit/s) Hírminőség	3,125 Mbájt	187,5 Mbájt	11,25 Gbájt	11,25 Tbájt
DVCPRO50 (50Mbit/s) Újrafeldolgozási minőség	6,25 Mbájt	375 Mbájt	22,5 Gbájt	22,5 Tbájt
Digit Beta (126Mbit/s) (stúdió transzparens minőség)	15,75 Mbájt	945 Mbájt	56,7 Gbájt	56,7 Tbájt
D5 (218Mbit/s) (stúdió transzparens minőség)	27,25 Mbájt	1,635 Mbájt	98,1 Gbájt	98,1 Tbájt
D1, ITU-R 601 (270Mbit/s) (stúdió minőség)	33,75 Mbájt	2,025 Gbájt	121,5 Gbájt	121,5 Tbájt

Mi az oka, hogy ugyanúgy kompatibilis a fekete fehér TV, és a legmodernebb plazma?

Nem RGB-ben működik, hanem Y és színekülönbség. (R-Y)(B-Y)

Színekülönbség (B-Y, R-Y). miért csak ez a 2 van?

A zöldnek a legkisebb a jele, ezt elhagyjuk. (R: red, B: blue).

Emberi hallás miért szelektív?

A hallócsigák különböző részén csillapodik le a hanghullám.

Miért tudjuk megkülönböztetni egymás hangját?

Fogak, nyelv, koponyacsont egyedisége miatt.

Egyszerre 1khz és 5khz et szólaltatunk meg. melyik hangosabb?

5 khz, mert kevesebb Phon tartozik hozzá.

A -> B tagadása.

A és nem B, mert $A \rightarrow B = (\text{nem } A) \text{ vagy } B$;

nem B $\rightarrow$ nem A = B vagy (nem A)

Mi a láthatósági fgv?

Megmutatja, hogy egy nappali látáshoz alkalmazkodott emberi szem mennyire érzékeny a különböző hullámhosszú spektrumszínek energiája iránt.

Fej vagy írás játék, 50-50 az esély. mennyi az entrópia?

Entrópia maximális, ha a valószínűségek egyformák.

Max entrópia $\text{LOG } n$ (n ahány eset van, jelen példánál 2), így a megoldás 1.

Mi a szín?

Bizonyos hullámhosszú fény.

Miért nem tudjuk a mai technológiával a világ összes színét megjeleníteni?

A CMYK-ból nem lehet előállítani az összes szürkeárnyalatot (csak 50-e)t, valamint nem lehet a ... minden pontján tökéletes színt előállítani. Hiányzik az interferencia szimulálása.

Miért nem használja ki a 2 megapixeles kamera a True Color lehetőséget?

Azt hiszem azt mondta, hogy azért mert 16,777,216 color variations a True color, az sokkal több mint 2 mp. Tehát ha minden pixel különböző színű akkor se használtam fel az összes pixelt.

Forrás: http://en.wikipedia.org/wiki/True_Color#True_color_.2824-bit.29

1. előadásból ami tuti kell:

A bitsebesség-torzítás diagram!

Kódolók összehasonlítása rajz végiggondolni

Bitsebesség-minőség diagram

8 db ugyanolyan méretű súly közül 7 ugyanolyan tömegű. 1 nehezebb. Hány méréssel lehet megállapítani, hogy melyik az az egyetlen eltérő súly?

2 méréssel:

Menet:: veszel 3-3 súlyt

ha egyenlőek, akkor veszed a maradék kettőt, leméred és amelyik a nehezebb

ha nem egyenlőek akkor veszed a nehezebb 3-ast és abból választasz ki 2-őt

ha egyenlőek, akkor a 3. súly a nehezebb

ha nem egyenlőek, akkor a nehezebb

MPEG1 rétegei?

1. Szekvencia réteg: Magát a kódolt szekvenciát jelöli, fejléce tartalmazza a rendszeradatokat (képméret, bitsebesség, képfrekvencia, kvantálási mátrixok, stb.).

2. Képcsoport réteg (GOP=Group Of Pictures): Legalább 1 önmagában kódolt (intra) képet tartalmazó, bizonyos számú kép együttese, amely a véletlen hozzáférés egysége.

3. Képréteg: Egy kép kódolt adatait tartalmazza. A kódolatlan kép formátuma MPEG-1-ben 4:2:0, MPEG-2-ben 4:2:0, 4:2:2 vagy 4:4:4, struktúrájú YUV. Az MPEG-1 csak progresszív, míg az MPEG-2 alkalmas váltott-soros képek kódolására is.

4. Szelet (slice) réteg: MB-ok sorfolytonos csoportja, az újraszinkronizáció egysége. Ez a legalsó szint, amelyen a dekóder még képes feléledni bithiba esetén.

5. Makroblokk réteg: Az Y 16x16-os, az U,V 8x8 (4:2:0), 8x16 (4:2:2), vagy 16x16-os (4:4:4) blokkjaiból áll, a mozgáskompensáció egysége.

6. Blokk réteg: A MB 8x8-as blokkjai, a DCT kódolás egysége.

Miért 44100 Hz a mintavételezési frekvencia?

frekvencia osztók -> a prímszámok miatt jó + a hálózati áramforrás és TV, CD frekvenciájával is össze kellett vetni ($44100 = 7^2 \cdot 5^2 \cdot 3^2 \cdot 2^2$)

Hogy lássunk mennyi mennyiség kell fényből?

frekvenciától függ, hogy 5-6 fotont meglássunk ??

Látható fény

400-700 nm hullámhossz között

Mennyi a hullámhossza a fehér, fekete, szürke színeknek?

nincs mert kevert színek (fekete a fény hiánya) -> a törésmutató hullámhossz függő

Miért nehéz az ember hallását reprodukálni?

Mert túl nagy a tartomány, a szemnél is nagy tartományban érzékelünk.

Példa: egy nagyon halk hangok észlelésére épített szenzor tönkremegy, ha fellősz mellette egy rakétát. Egy nagyon hangos hangok észlelésére épített szenzor pedig nem észleli a halk neszeket.

Mi határozza meg az objektumok színét?

saját maga által kisugárzott szín, és a felületet megvilágító fény spektrális sűrűsége

Szem biológiája (robosztusságának oka):

Csapok: erős fény esetén, fény- és színérzékelés,

Pálcikák: gyenge fény esetén csak fényérzékelés

Miért nézi az amcsi gyerek messzebből a TV-t?

látószög

Európai SD (4:3, 575 hasznos sor):

A nézőtávolság a kép magasságának kb. 6-szorosa

Látószög = $2 \cdot \arctg(\frac{1}{2} / 6) = 2 \cdot \arctg(1 / 12) \approx 9,53^\circ$

Függőleges felbontás = $9,5^\circ \cdot 30c/^\circ \approx 286$ ciklus = 572 sor

Amerikai SD (4:3, 480 sor):

A nézőtávolság a képmagasság kb. 7,15-szorosa

Látószög = $2 \cdot \arctg(\frac{1}{2} / 7,15) \approx 8^\circ$

Függőleges felbontás = $8^\circ \cdot 30c/^\circ \approx 240$ ciklus = 480 sor

Miért nagyobb a fullHD TV?

mert ha kisebb akkor közelebbről kellene nézni

Jellemző SD képfelbontások

Európa, Ázsia: 625 sor, 50 Hz félkép, 25 Hz kép,

$625 \cdot 25 = 15625$ Hz sorfrekvencia

Amerika, Japán: 525 sor, 60 Hz félkép, 30 Hz kép,

$525 \cdot 30 = 15750$ Hz sorfrekvencia

Kezdetben:

- *Anglia:* 405 sor
- *Franciaország:* 819 sor

HD képfelbontások

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

Diszkó mellett a mély hangokat miért halljuk jobban?

minél nagyobb a frekvencia annál jobban érzékelhető. A mélyhangoknak nagyobb a hullámhosszuk -> hullámelhajlás - > jobban halljuk

Hullámelhajlás:

- Keskeny résen áthaladó hullám behatol az árnyéktérbe is.

- Ha a rés kisebb mint a hullámhossz, a hullámelhajlás nagy mértékű.

A szokásos ajtó és ablak méret tartományába a mély hangok esnek.

Adattömörítés, jeltömörítés legfőbb különbségek?

adattömörítés: veszteségmentes tömörítés (tárolásra)

jeltömörítés: veszteséges tömörítés (szállításra)

Mi olvasható le a bitsebesség-torzítás diagramról?

A torzítás mértéke 'b' kimeneti bitsebességen

Minimális és maximális torzítás/bitsebesség

Mi a hangnyomásszint?

Nyugalmi légnyomáson szuperponálódó légnyomás változás

Minek a mértékegysége a Phon, mit jelent 1 Phon?

Egy hangnak annyi a Phon-ja, amennyi dB a vele azonos hangosságérzetet keltő 1 kHz-es hang hangnyomása.

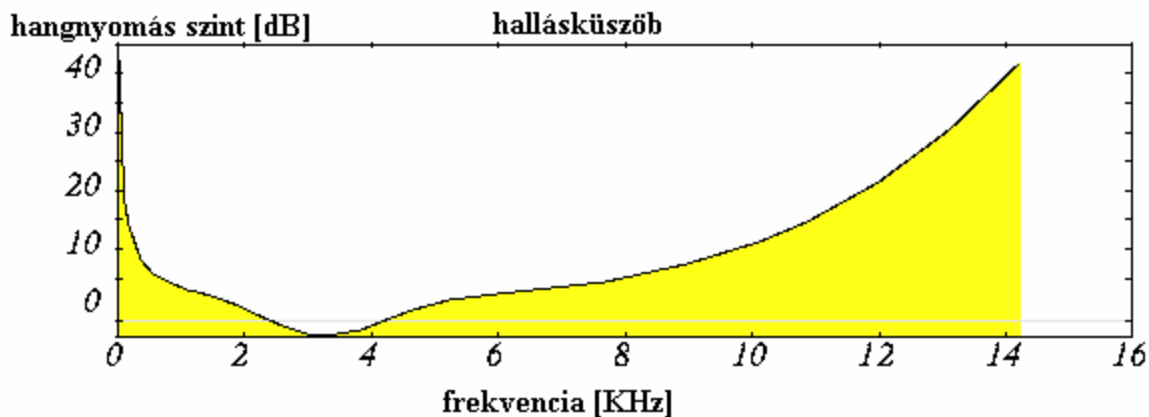
Frekvenciatartománybeli elfedés lényege?

A spektrumból kiemelkedő keskeny sávú zaj jellegű komponensek megemelik a frekvenciatartományban környezetükben lévő hallásküszöböt.

Pszichoakusztikus kódolás főbb fogalmai?

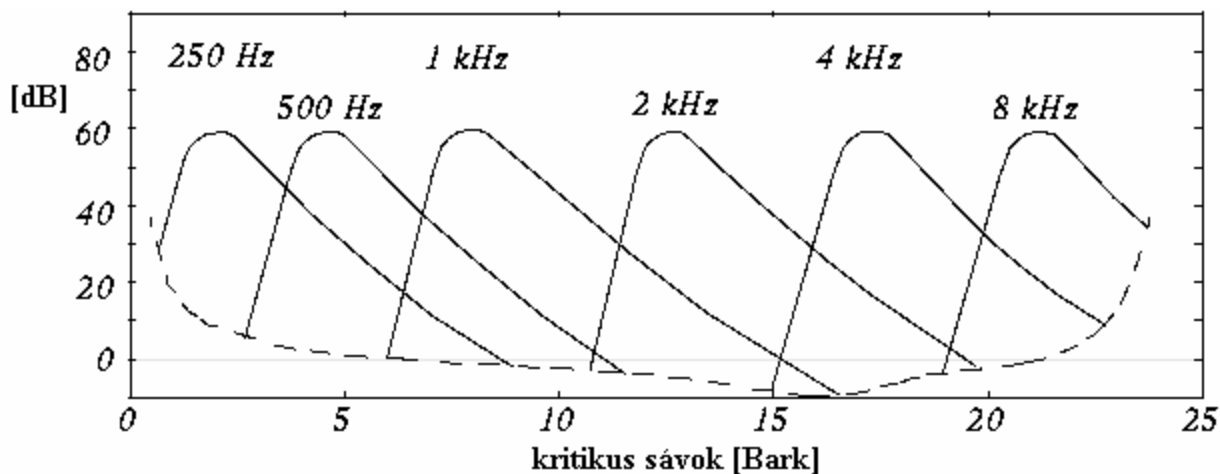
Abszolút hallásküszöb: a természetes alapzörejek elfedési görbéje





Kritikus sávok fogalma.

Hallásunk tulajdonságai alapján az észlelhető frekvenciatartományt fel lehet osztani jól meghatározott sávokra, az úgynevezett kritikus sávokra.



Frekvencia tartománybeli elfedés.

Dinamikus hallásküszöb.

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.

Idő tartománybeli elfedés.

A tonalitás.

Hallásunk a harmonikus hangok tisztaságára érzékenyebb, mint a zaj jellegű hangokéra.

Ezért a tiszta hangok spektrumvonalait pontosabban kell átvinni.

Ehhez azonban meg kell határozni, melyek a tonális illetve a zaj jellegű komponensek.

Frekvenciatisztaság igénye.

Milyen mikrofon elrendezést használunk műfejes sztereotechnikánál?

kardioid mikrofont (fülek halláskarakterisztikáját utánozó mikrofont)

Prediktív kódolás

CBR puffer?

Konstans bitráta

Mi alapján választották az audio CD mintavételi frekvenciáját?

44,1 kHz = 44100 Hz = $2^2 * 3^2 * 5^2 * 7^2 \Leftarrow$ első négy prím négyzetének szorzata.

technológiai okok miatt kis prímeket érdemes alkalmazni + az analóg TV sorfrekvenciájával kellett összehozni.

Van egy kódkészlet amiben a kódszók hossza a következő: 2,3,4,4,5. Lehet-e egyértelműen dekódolni?

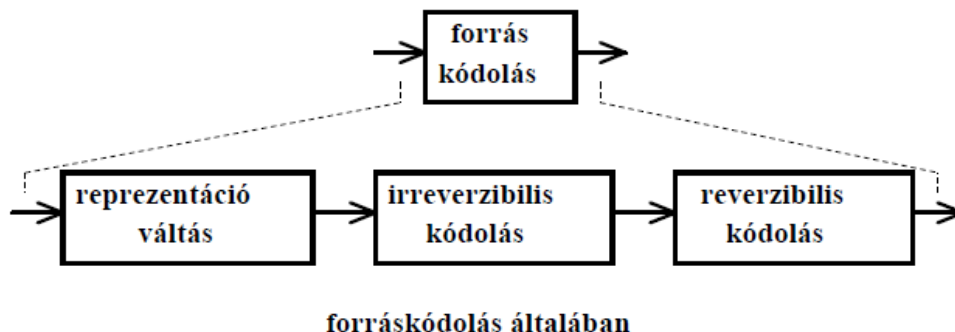
(McMillan) MC Milán YEAAH

$$2^{-2} * 2^{-3} * 2 * 2^{-4} * 2^{-5} = \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{32} = \frac{17}{32}$$

$$\frac{17}{32} < 1$$

Ha nem teljesül az egyenlőtlenség, akkor biztosan nem dekódolható. Ha teljesül akkor nem lehet eldönteni, de nincs kizárva.

Az alábbi képen miért nem lehet máshogy a dobozok sorrendje?



Itt egy hosszú levezetés áll ami $y=f(x)$ $g(x)=z$ ból $f^{-1}(f(x))=x$ lesz az eredmény, aki érti írja le mi is a lényege.

10 bitet kódolunk és true color színeket használunk. Mekkora képet kell használni, hogy meg térüljön?

39 oszlop 1000 sor 14 bit * x=39000 40/20 (???) hogy volt ez tovább?

Mi a nagy előnye a DCT-nek?

Minél tovább fut az algoritmus, annál pontosabb lesz a valószínűség.

Órai jegyzet

Bevezetés, alapfogalmak

Média fogalma: (def) a kommunikáció bármely eszköze, amely biztosítja a közleményt a forrástól a befogadóig

Definíció:

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.

Etimológia:

A média szó eredete a latin medium, és a közvetítő eszközre utal, amellyel egy üzenetet vagy hírt továbbítani lehet.

Média 3 jelentése:

1. A mondanivaló kifejezésére használatos **közvetítő közegek összessége**.
2. Nem maga a változó technikai közeg, hanem a mondanivaló kifejezésére használt fórumok; a sajtó, a rádió, a televízió és újabban a világháló kifejezési formáiban létrejött **nyilvános fórumok összessége**.
3. A média a **rádiózás és televíziózás műsorszolgáltatásainak összessége**, beleértve az Internet "tartalomszolgáltatását" is, és esetleg a távközlés rendszerint nem nyilvános mondanivalót továbbító legkülönbözőbb új szolgáltatásait. (pl. üzenetküldés interneten vagy mobiltelefonon keresztül) (szűk jelentés)

Médiatechnológia: mérnöki megoldások összessége, amely hang, kép, mozgókép + ezekhez kötődő infók előállítását, feldolgozását, továbbítását, tárolását, megjelenítését teszi lehetővé

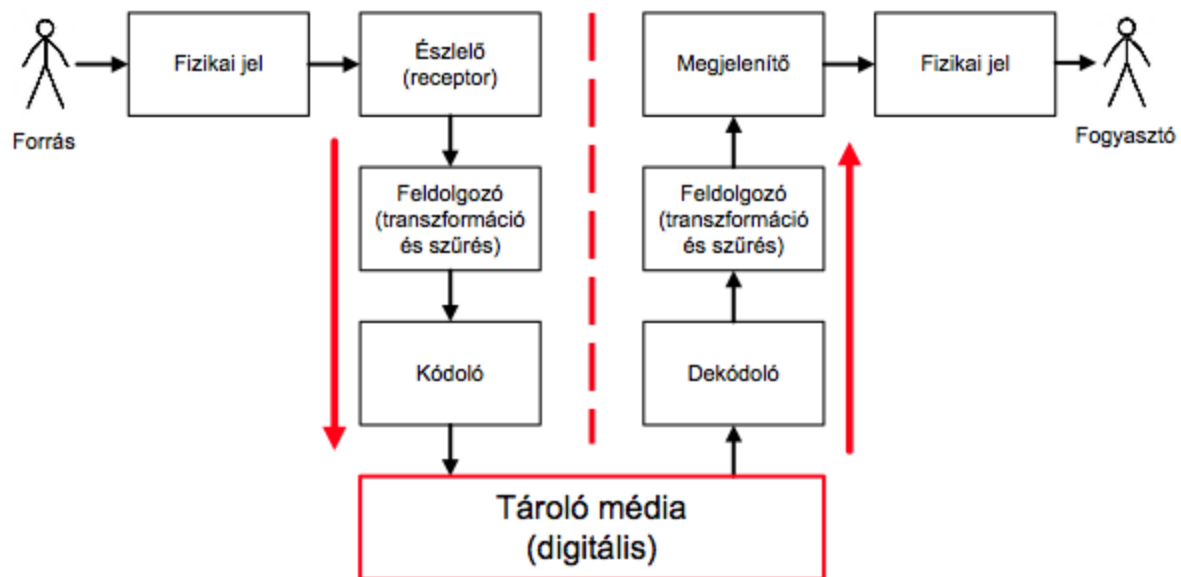
Céljai:

- továbbítás (média: hálózat (vezető, optikai szál, rádióhullám), átvitel: azonos időben)
- tárolás (média: permanens adathordozó) : archiválás, továbbítás, sokszorosítás
- feldolgozás (média: átmeneti, nem releváns): egy vagy több médiatartalomról több másik
- hibrid: a fentiek közül egynél több alkalmazása

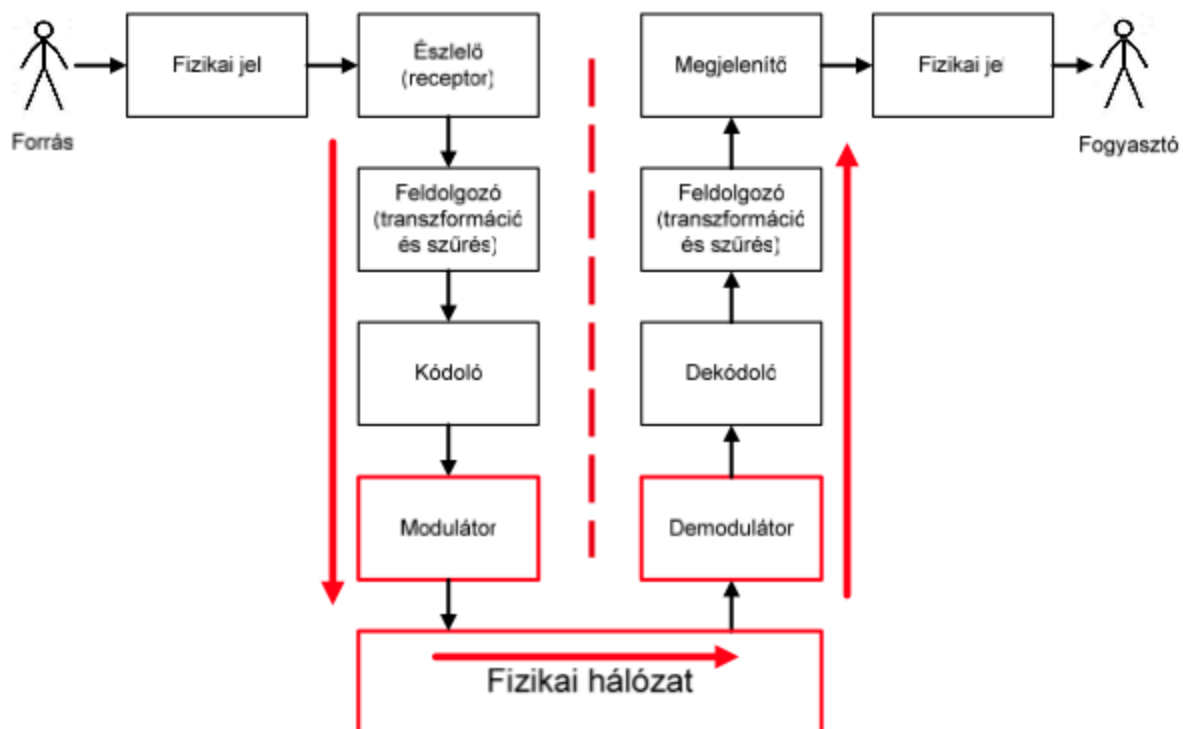
Rendszer elvi felépítése:

Forrás -> fizikai jel -> érzelő -> feldolgozó (transzformáció és szűrés) -> kódoló -> tároló média (digitális) -> dekódoló -> feldolgozó -> megjelenítő -> fizikai jel -> Fogasztó

Tárolás:



Továbbítás:



Adattömörítés: a bejövő adatot kisebb "mértékre" "kódolt adattá" kódolom át. dekódolva visszakapom az eredetit. veszteségmentes tömörítés -> kevesebb helyet foglal

Jeltömörítés: a kódolt adatból visszaállított jel különbözik -> van veszteség. minél nagyobb tömörítés, annál nagyobb veszteség. cél: lehető legjobban hasonlítson az eredetire a dekódolt jel

Információvesztés forrása:

Analog -> digitális

1. Időben folytonos jel időben diszkrét jel (mintavételezés)
2. Folytonos amplitúdó diszkrét érték (kvantálás)
3. • Az analógból digitálisba áttérés (mintavételezés + kvantálás) okozhat veszteséget: a folytonos jel mintavételével diszkrét idejű mintákat kapunk, majd a mintavett jel (még analóg) amplitúdóját kvantáljuk

Egyéb jellemző esetek:

- Digitális bemenetnél a veszteséget főleg az okozza
- Mintaszám-csökkentés:
- Minden n . mintát elhagyjuk (újra-mintavételezés)
- Transzformáció $\rightarrow$ minta elhagyás (vagy kinullázás) $\rightarrow$ inverz transzformáció: az eredeti mintaszám áll vissza, de csökkentett információ-tartalommal és a transzformált tartományban a mintaszám kisebb
- Speciális szűrési megoldások is okozhatnak veszteséget

Kódolás minőségét jellemző mutatók:

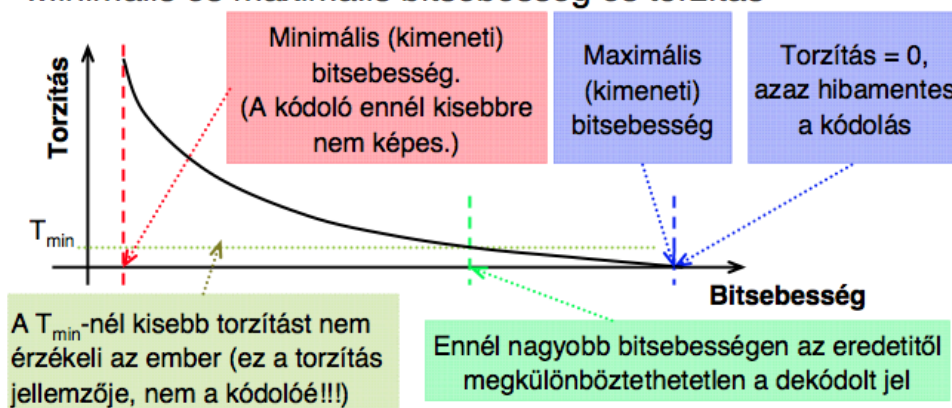
-**torzítás:** az eredeti és dekódolt jel különbsége valamilyen mértékkel kifejezve

-**tömörítési arány**

-**bitsebesség-torzítás diagram:** még mekkora tömörítést engedünk meg, hogy a felhasználói élmény ne sérüljön.

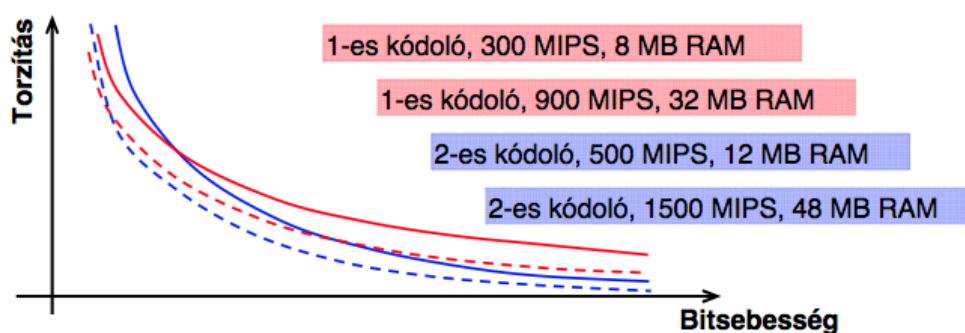
Leolvasható erről a diagramról:

- A torzítás mérték b kimeneti (kódolt) bitsebességen
- Az a minimális bitsebesség, amelyen a kódolás képes adott torzításnál jobb minőséget produkálni
- Minimális és maximális bitsebesség és torzítás



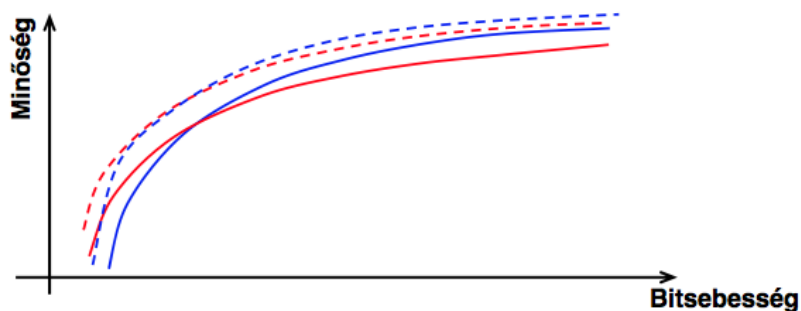
Kódolók összehasonlítása:

- A bitsebesség-torzítás diagramról:
 - Azonos torzítás mellett mekkora kimeneti (kódolt) bitsebességet produkálnak (kisebb a jobb)
 - Azonos bitsebesség mellett mekkora torzítást produkálnak (kisebb torzítás a jobb)
 - További szempontok is számítanak: pl. komplexitás, üzemeltetési költség, vételár



Bitsebesség-minőség diagram:

- A bitsebesség-torzítás diagram helyett olyan diagram, ahol minőségi mérték szerepel a torzítás helyén.
- A minőség esetén a nagyobb érték a jobb, a torzítás esetén a kisebb a jobb.
- Jól tervezett rendszernél egy monoton növekvő minőséget kell kapni a bitsebesség függvényében.



Emberi hallás és sajátosságai

-hangnyomás: a nyugalmi légnyomásra szuperponálódó légnyomás változás

$$P(t) = P_{\text{nyugalmi}} + p(t) \text{ [Pa]}$$

a hangnyomás effektív értéke alkalmas a hang fizikai jellemzésére

Hangnyomásszint(SPL): adott p_0 vonatkoztatási szinthez mért hangnyomás, ahol p_0 a még éppen hallható 1kHz-es hang hangnyomás értéke. $p_0 = 20 \text{ uPA}$

$$\text{SPL} = 20 \log(P/P_0) \text{ [dB]}$$

Hangintenzitásszint(SIL= $10 \lg(I/I_0)$ [dB]):

-levegő sűrűsége: $\rho = v \cdot r_0 \cdot c$

-hanghullámok esetén: $v \cdot c = f \cdot \lambda$

-hangintenzitás: időegység alatt a felület egységen áthaladó energia

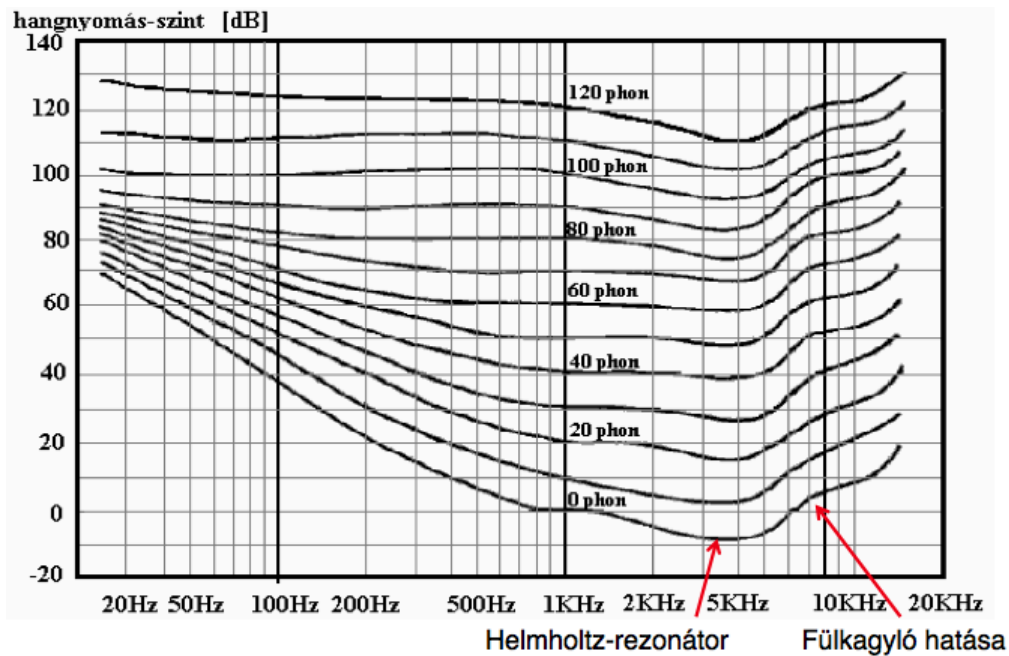
-referenciaszint: a még hallható 1kHz-es hang intenzitása: $I_0 = 1 \text{ pW/m}^2$

Phon

SPL (dB)	Jellemző hely	SPL (dB)	Jellemző hely
0	Hallásküszöb	80	Gyár
10	Grand Cayon North RIM	108	Mennydörgés
25	Jó stúdió	135	Fájdalom küszöb
42	Hallgatóság zajszintje	155	Sugárhajtású repülő
68	Forgalmas utca	180	Űrrepülő start

- 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.

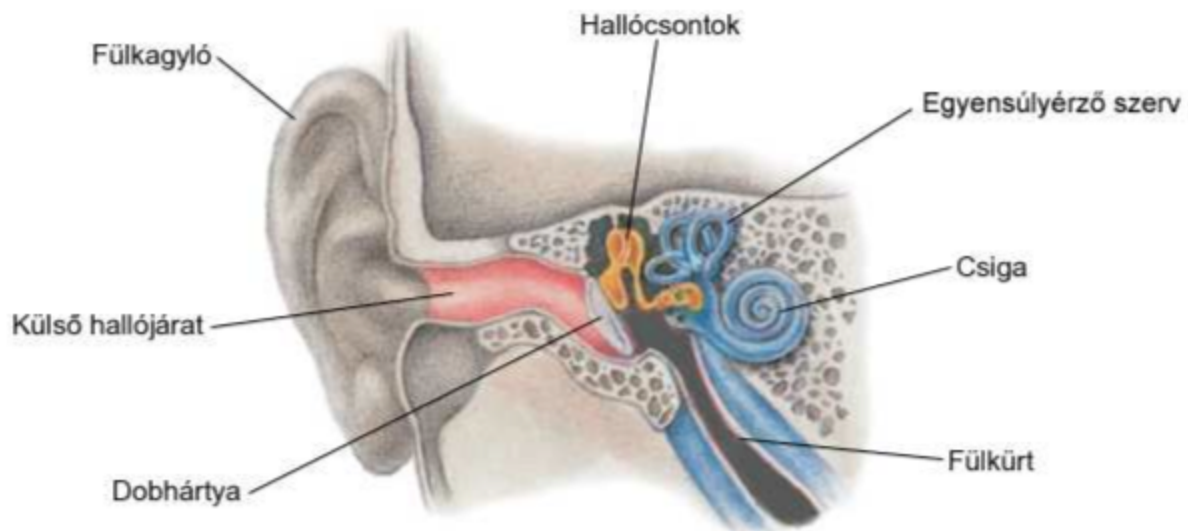
Fletcher-Munson görbék

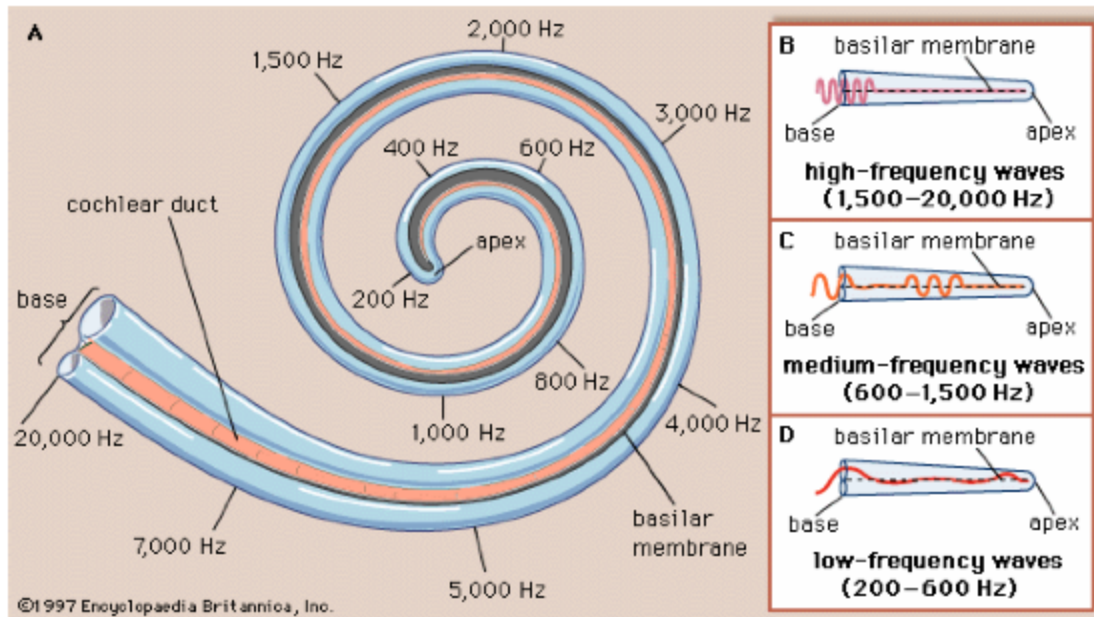


Hallásküszöb: természetes alapzörejek elfedési görbéje.

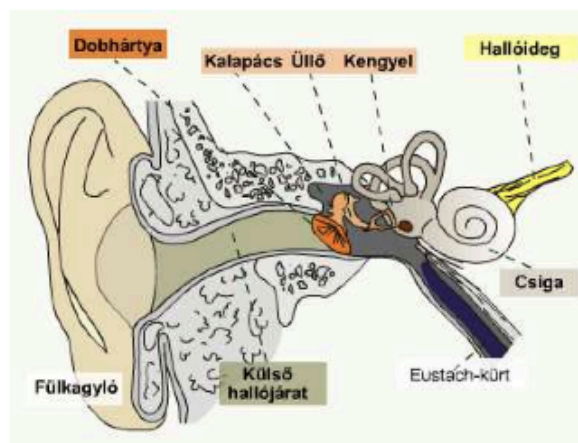
Kritikus sávok: egy adott frekvencia sávon belül hallásunk egyidejű gerjesztés esetén intenzitás (energia) alapon összegzett hangosságot érzékel. Ezek a sávok a kritikus sávok.

Fül felépítése:





- 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 feléléedési idő: 1 ms



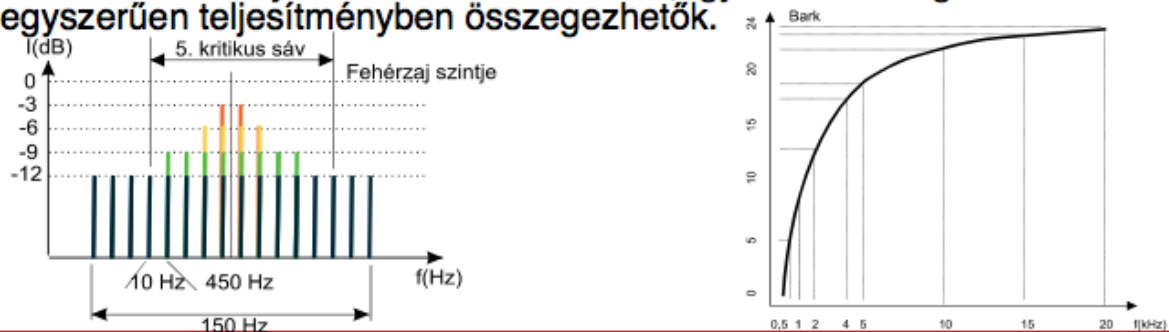
Összesen:

- Dobhártyától az agyközpontig: 3 ms – 6 ms

Kritikus sávok:

Elfedési jelenségek, elfedési görbe: van egy maszk. ha adott hang szól adott intenzitással

- Hallásunk tulajdonságai alapján az észlelhető frekvenciatartományt fel lehet osztani jól meghatározott sávokra, az úgynevezett kritikus sávokra.
- A kritikus sávokat (24 db) Bark-ban számozzuk, Barkhausen tiszteletére.
- 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.



és vele együtt szól egy nagyobb/kisebb frekvenciájú hang, de még a maszk alatti intenzitással, akkor, nem érzékeli fiziológiailag az ember a különbséget (mp3 kódoló, audio kódolók alapja)

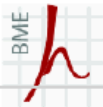
- A spektrumból kiemelkedő tonális vagy keskenysávú zaj jellegű komponensek megemelik a frekvenciatartománybeli környezetükben a hallásküszöböt.
- Kialakul a dinamikus hallásküszöb, ún. maszk.
- Ami a maszk alatt van, az nem hallható.
- Ez az elfedési jelenség frekvencia- és szintfüggő.
- Az elfedési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

Pszichoakusztikus kódolás: Miért építik ma is az erősítőket elektroncső alapon, miért nem félvezető tranzisztor alapon?

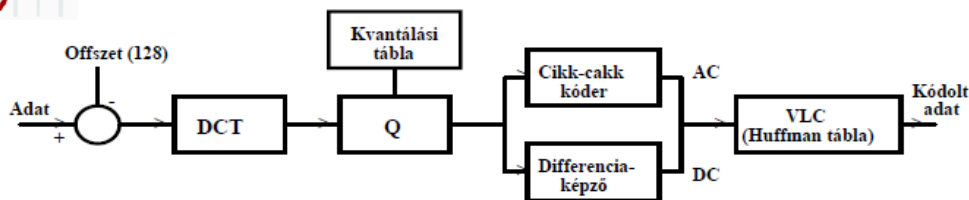
-mert az elektroncsőben vákuumban terjed a jel, míg a tranzisztorban a félvezető rétegek között. és ezért nagyobb a termikus zaj.

- A hallás következő sajátosságait vesszük figyelembe elfedési-maszkolás modellben:
 - Abszolút hallásküszöb.
 - Kritikus sávok.
 - Frekvenciatartománybeli elfedés.
 - Dinamikus hallásküszöb.
 - Időtartománybeli elfedés.
 - A hang tonális és nem tonális szerkezete.
 - A hallás frekvenciától függő pontossága.

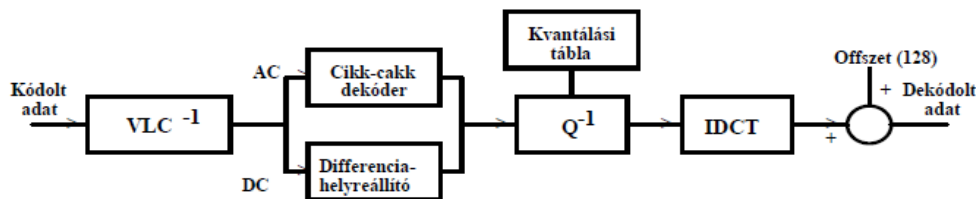
Sztereo: eredeti jelentése tér.



DCT-alapú szekvenciális kódolás



DCT alapú szekvenciális JPEG kódoló



DCT alapú szekvenciális JPEG dekódoló

DCT: Diszkrét Koszínusz Transzformáció

Q: Kvantáló

VLC: Változó szóhosszra kódoló egység

IDCT: Inverz Diszkrét Koszínusz Transzformáció

Q^{-1} : Inverz kvantáló

VLC^{-1} : Változó szóhosszt dekódoló egység

MPEG VIDEÓKÓDOLÁS

Célja:

- a videó és audio digitális, bitsebesség-csökkentett formában történő hatékony ábrázolása

Fontosabb részei:

- 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 **audió specifikáció**: a bitsebesség csökkentett audió bitfolyam szintaxisa és az audió modell dekóder specifikációja.
- 4. rész a **Conformance test** (tesztelési és megfelelési feladatok összefogására).
- 5. rész a **DSM-CC (Digital Storage Media Command Control) (csak MPEG-2) az interaktív csatorna protokoll specifikációja**
- 6. –11. rész **szoftver és egyéb pl. audio kérdések**.

- A szabvány nem specifikálja a kódert (kódolási algoritmusokat) sem a videó, sem az audio részben! Csak a kódolás kimeneti adatfolyamainak bitszintaxisa specifikált.

MPEG-1 (ISO 11172 /1993/): Alacsony bitsebességű multimédiás alkalmazások (Video-CD, CD-I), kb. 1,5 Mb/s VHS képminőség, SIF felbontású kódolt videó.

MPEG-2 (ISO 13818 /1994-95/):

- Műsorszórás (Digital Video Broadcasting, DVB), 2-8 Mb/s terjesztési minőség (distribution quality)
- Stúdiótechnika, 18-50 Mb/s újrafeldolgozási minőség (contribution quality)
- Digital Versatile Disc (DVD), 3-7 Mb/s (VBR!) jobb mint a PAL minőség
- Általában ITU-R BT-601 a felbontás mindegyikben
- Az MPEG-2 kibővített MPEG-1, a kódolási elv mindkét eljárásban azonos

MPEG-4 (ISO/IEC 14496): Alacsony bitsebességű kódolások (kiindulás x 64 kbit/s) szabványának indult (1994), de jelenleg az interaktív multimédia szolgáltatások objektum orientált szabványává vált.

MPEG-7: A tartalom leírás (metaadat) szabványa.

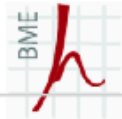


MPEG videókódolás

- Kifinomult, DCT-alapú mozgáskompensációt 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 bájtáron kezdődő egyedi startkóddal rendelkezik, így kezdetük dekódolás nélkül megtalálható

Az MPEG videókódolás rétegszerkezete

- **1. Szekvencia réteg:** Magát a kódolt szekvenciát jelöli, fejléce tartalmazza a rendszeradatokat (képméret, bitsebesség, képfrekvencia, kvantálási mátrixok, stb.).
- **2. Képcsoport réteg** (GOP=Group Of Pictures): Legalább 1 önmagában kódolt (intra) képet tartalmazó, bizonyos számú kép együttese, amely a véletlen hozzáférés egysége.
- **3. Képréteg:** Egy kép kódolt adatait tartalmazza. A kódolatlan kép formátuma MPEG-1-ben 4:2:0, MPEG-2-ben 4:2:0, 4:2:2 vagy 4:4:4, struktúrájú YUV. Az MPEG-1 csak progresszív, míg az MPEG-2 alkalmas váltott-soros képek kódolására is.
- **4. Szelet (slice) réteg:** MB-ok sorfolytonos csoportja, az újra-szinkronizáció egység. Ez a legalsó szint, amelyen a dekóder még képes feléledni bithiba esetén.
- **5. Makroblokk réteg:** Az Y 16x16-os, az U,V 8x8 (4:2:0), 8x16 (4:2:2), vagy 16x16-os (4:4:4) blokkjaiból áll, a mozgás-kompenzáció egysége.
- **6. Blokk réteg:** A MB 8x8-as blokkjai, a DCT kódolás egysége.



Képtípusok

- **I (Intra coded):** Önmagában kódolt kép. A dekódolásához szükséges minden adatot tartalmazza. A kóder oldali dekódolása kötelező.
- **P (Predictive coded):** Prediktíven kódolt kép, referenciája egy előző I vagy P kép (csak múltbeli referencia), a dekódolásához a referenciára szükség van, a kóder oldali dekódolás kötelező.
- **B (Bidirectionally coded):** Kétirányból kódolt kép, referenciája az előző I vagy P (múltbeli ref.) és a következő I vagy P kép (jövőbeli ref.). B kép nem lehet további referencia, így kóder oldali helyreállítása (referenciaként) nem kötelező.
- **D (DC coded):** olyan speciális I kép, amelyben csak a DC együtthatók kódoltak (gyorskeresés céljából – csak MPEG-1)

ZH előtti konzi

Média fogalma, 3 aspektusa

Shannon-i csatorna modell

Milyen rétegei vannak a kódolásnak, helyes sorrend, bitsebesség, mi változik minek a függvényében? -> aszimptoták, metszéspontok

Információ mérése (bitek, hány biten lehet kódolni vmit), fizikai mérés (különbség azonosság a fizikai mennyiségekkel)

Igen-nem kérdés

“minden azonosan valószínű” modell: mi a hibája, entrópia, az információnál a valószínűség is számít, entrópia fv (min, max, levezetés) -> fv. tulajdonság mi és mit jelent

Változó hosszúságú forráskódolás (elvi korlát) -> hány biten reprezentálható?

Milyen feltételek vannak az egyértelmű dekódolásra?

Veszteséges tömörítés elvi lehetőségei

Hullámterjedés, hullámjelenségek + példa (hullámtan)

Hallás biológiája (beduguló fül a liftben miért és hogy dugul ki)

Pszichoakusztikus kódolás, mik azok a tulajdonságok ami lehetővé teszi a nagy bitsebességű kódolást

Audio jelenségek, kvantálás! (matematikai precizitással) def + jelentés, skalárkvantálás, kvantáló struktúra, komponens, kódolás dekódolás

Vektor kvantálás (def + érteni)

Szint szám csökkentésének lehetősége

Transzformációs kódolás (részszávos is egyfajta transzf. kódoló)

Részszávos kódoló blokkvázlata

Audio kodek (hallásmodellel összekapcsolva),

MPEG szabvány (hangban)

MPEG 1-2 rétegei

MP3

látás, spektrum, hőmérséklet, sugárzás, színhőmérséklet, pálcikák, csapok, emberi látás működése

összeadó, kivonó színkeverés, megvilágosítás hatásai, reprodukálhatósága

Analóg TV technika

Színeltérítés

nem RGB-ben működik a TV

látás tulajdonságai

képfájl formátumok (veszteségmentes)

cikkcakk

jpeg

Videókódolás: szabvány részei mi a 6 szintű hierarchia, milyen részei vannak, I, P, B kép mozgásvektor