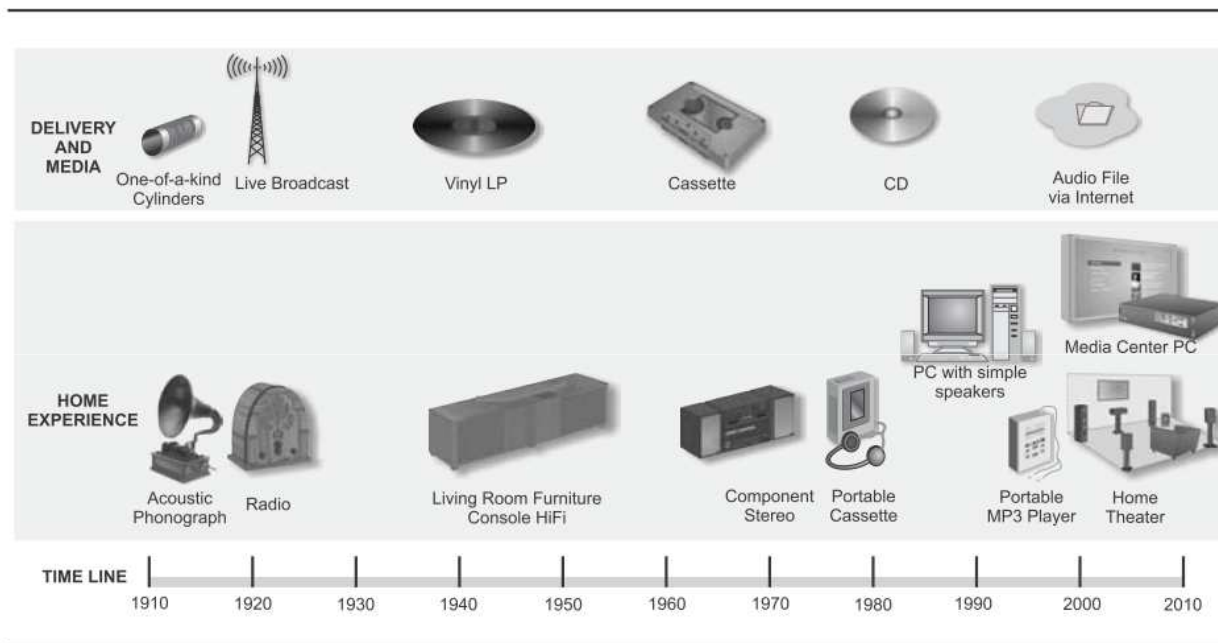




# Osnove obrade audio signala: Audio kodovanje

Algoritmi i Arhitekture DSP 2

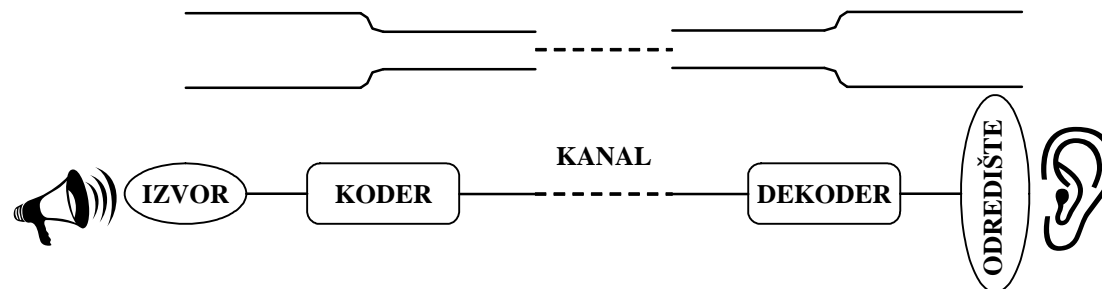
# Razvoj uređaja za reprodukciju zvuka



**Figure 1.1** Timeline of Home Audio Entertainment Devices

# Prostorna karakteristika zvuka

- Kanal je komunikacioni put između izvora i odredišta zvuka
- Do polovine 20. veka snima se i reprodukuje jedan kanal (centralni, mono)
- Stereo sistemi: snima se preko dva prostorno odvojena mikrofona koja istovremeno snimaju (slični kanali, ali imaju različito vreme dolaska zvuka i nivo zvučnog pritiska)
- Uvodi se kanal koji nosi niske frekvencije LFE
- Sredinom 90. godina 20 veka višekanalni sistemi

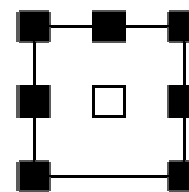
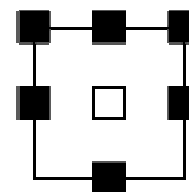
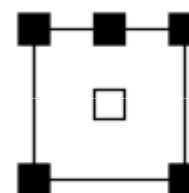
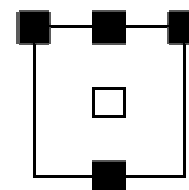


# Višekanalni audio sistemi

- Prednja slika:
  - Levi (L), Centar (C) i Desni (R)
- Kanali okruženja (eng. surround channels):
  - Levi okruženje (Ls) i Desni okruženje (Rs)
- Kanal niskih frekvencija (eng. Low Frequency Effect channel - LFE)

Notacija:

- 5.1 - 5 kanala punog opsega i 1 LFE
- 3/2.1 - 3 kanala ispred slušaoca (levi, centar, desni) + 2 kanala pored slušaoca (levi okruženje, desni okruženje) i jedan kanal niskih frekvencija.



# Ciljevi audio kodovanja

Predstaviti digitalni signal sa što manje bita uz što veće očuvanje kvaliteta

- Smeštanje zvuka u digitalnom formatu:
    - $44.1 \text{ kHz} * 16 \text{ bit} = 705.6 \text{ kbit/s}$  (stereo 1.41 Mbit/sec)
    - $48 \text{ kHz} * 16 \text{ bit} = 768 \text{ kbit/s}$  (stereo 1.54 Mbit/sec)
  - Prva generacija digitalnih audio aplikacija: kasete (DAT), kompakt diskovi (CD, 1985.)
  - Previše informacija za drugu generaciju uređaja: prenosivi multimedijalne uređaje, kao i prenos podataka (internet, bežični prenos..)
  - Dolazi do velikog istraživačkog rada u oblasti kompresije zvuka. Rezultat: definisanje standarda koji obezbeđuju kompatibilnost uređaja
- (Najveći doprinos standardizaciji dala je Internacionalna Organizacija za Standarde (ISO) i to ISO/JTC1/SC29/WG11 grupa, poznatija kao MPEG (eng. Moving Pictures Experts Group).



# Osnovni parametri audio koda

- **Kvalitet**: mera perceptualnog kvaliteta, najvažnija karakteristika sistema kodovanja (telefon-razumljivost, reprodukcija muzike-CD kvalitet..)
- **Bitska brzina** (eng. Data rate): zavisi od prenosa, načina smeštanja, propusnog opsega
- **Složenost**: preslika se na cenu fizičke arhitekture (MIPS, memorija)
- **Kašnjenje** (eng. delay)
- **Skalabilnost**
- **Otpornost na greške** (eng. error robustness)

# Izvorišno i odredišno kodovanje (1)

- Mogućnosti kompresije audio signala proizilaze iz poznavanja **izvora** i **odredišta** kao krajnjih tačaka komunikacionog lanca
- Iz modela **izvora** dobija se koncept *redundantnih informacija* i *izvorišno kodovanje* (eng. source coding.), a iz modela **odredišta** koncept *irelevantnih informacija* i *odredišnog kodovanja* (eng. sink coding)
- Ako su odbirci koje daje izvor statistički nezavisni (izvor bez pamćenja), redundantne informacije se dobijaju isključivo iz neravomerno raspoređene amplitude odbirka. Ako izvor ima pamćenje, odnosno ako su odbirci statistički zavisni, redundantne informacije se dobijaju iz međuzavisnosti odbiraka
- Za uklanjanje redundantnosti iz signala sa neuniformno raspoređenom amplitudom, koristi se entropijsko kodovanje dodeljivanjem kratkih kodnih reči simbolima koji se često ponavljaju, a dugačkih kodnih reči simbolima koji se retko ponavljaju. Redundantnost, koja proističe iz međuzavisnosti odbiraka, uklanja se prediktivnim kodovanjem ili transformacionim odnosno podopsežnim kodovanjem.

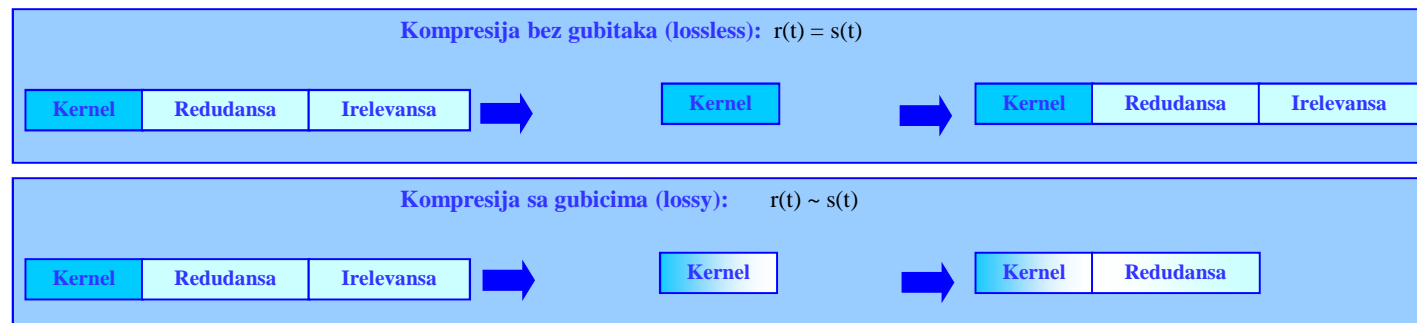
# Izvorišno i odredišno kodovanje (2)

- Visoko kvalitetni audio signal podrazumeva bilo koji prirodno ili sintetički prouzrokovan zvuk
- Model izvora komplikovan, jako nestatičan i bez Gausove raspodele
- Ovo uvodi novi koncept kodovanja: perceptualno kodovanje ili perceptualna entropija
- Perceptualno kodovanje oslanja se na **odredište** u komunikacionom lancu, pa se često zove i **odredišno kodovanje**.

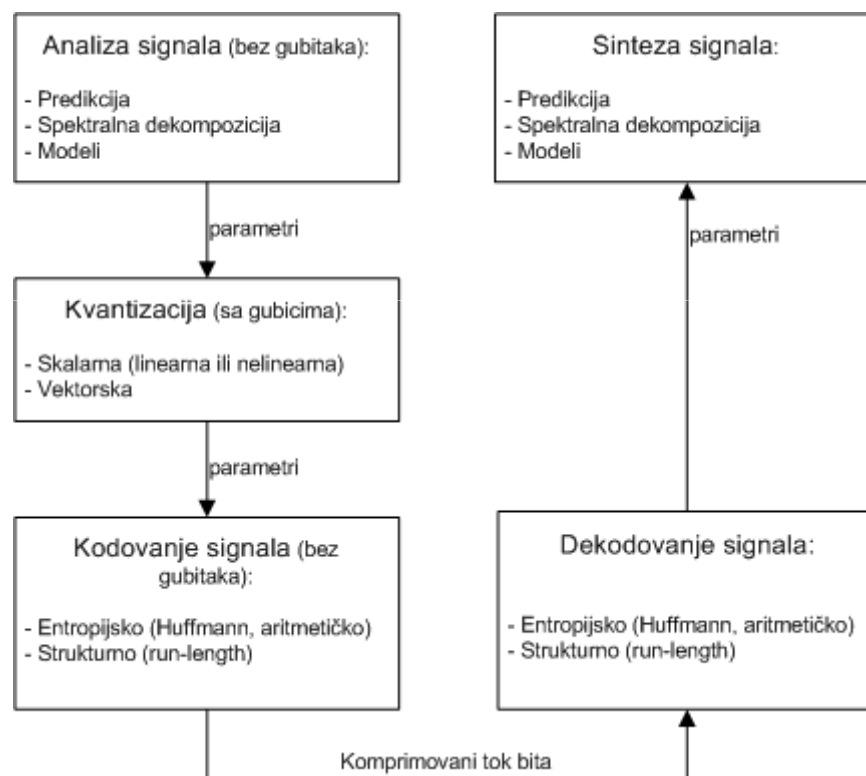


# Tipovi audio kompresije

- Generičke metode kompresije (npr. ZIP)
  - postižu mali stepen kompresije audio signala (oko 85% originala)
  - nisu pogodne za primenu u realnom vremenu
- Metode audio kompresije:
  - **sa gubicima** (eng. Lossy)
    - Dovodi do nepovratnog gubitka dela informacije
    - Stepen kompresije reda 10:1 i više – ali uz gubitak kvaliteta
  - **bez gubitaka** (eng. Lossless)
    - Audio informacija je u potpunosti sačuvana (bit-exact).
    - Stepen kompresije je  $\sim 2:1$

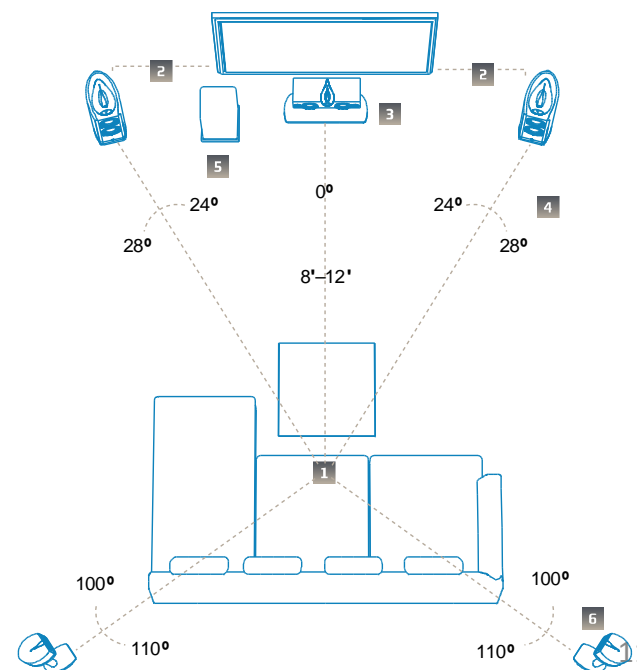
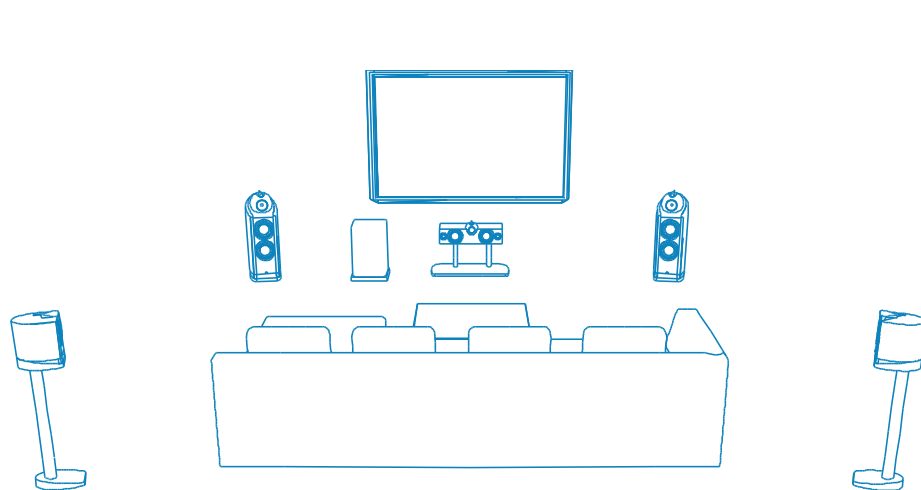


# Osnovne metode kompresije audio signala



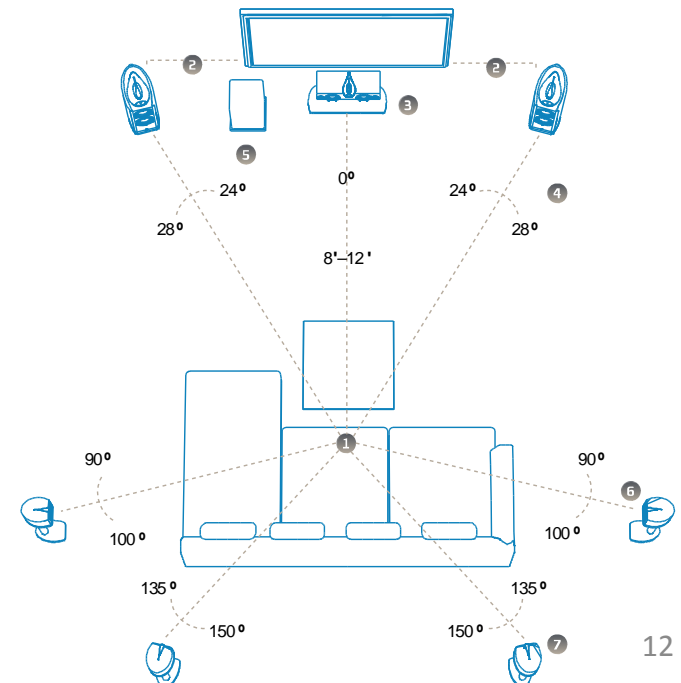
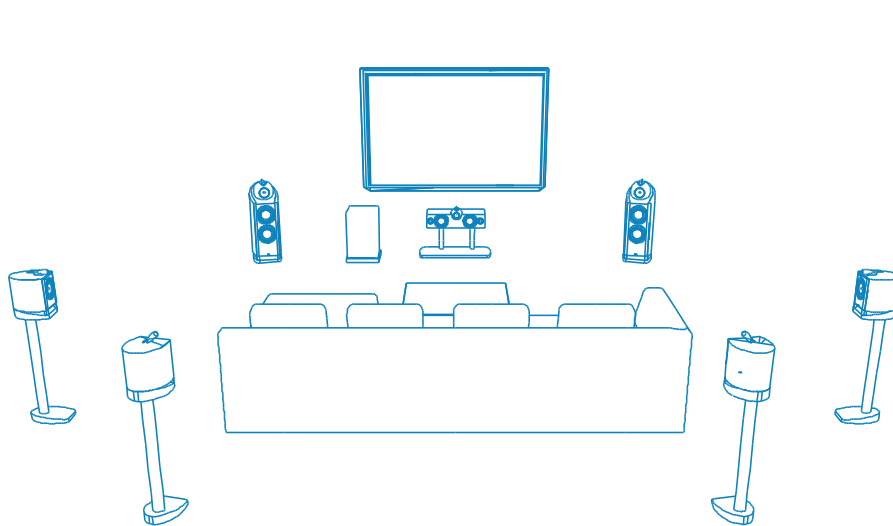
# Komercijalni standardi – audio kompresija sa gubicima

- Kapacitet prethodne generacije DVD diskova (4.7 GB sa jednim slojem, 8.7 GB sa dva sloja) je nametao korišćenje audio kodnih standarda utemeljenih na kompresiji sa gubicima. Primeri ovakvih standarda su:
  - Dolby Digital (AC3)
  - DTS Audio
- U osnovnoj verziji namenjeni su za bitske brzine do 3Mb/s i prenos pomoću S/PDIF veze, podržavaju maksimalno 5.1 konfiguracije kanala, 16 bita po odbirku, učestanosti odabiranja retko preko 48kHz.

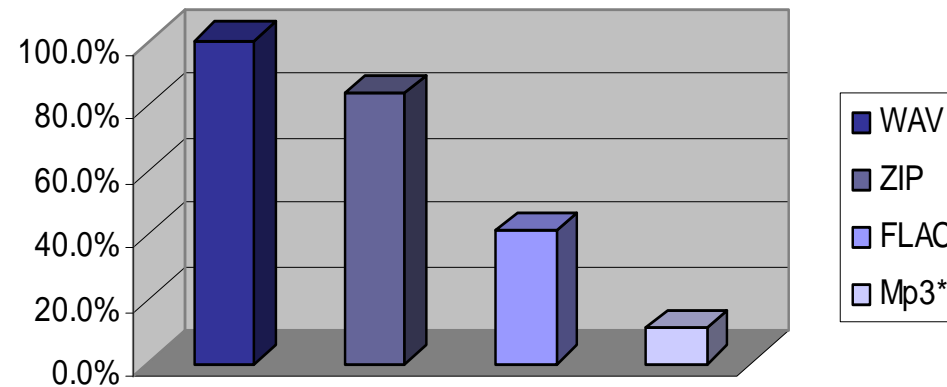


# Komercijalni standardi – audio kodovanje bez gubitaka

- Nova generacija digitalnih video diskova visoke definicije (Blu-Ray Disc i HD DVD) donosi znatno veći kapacitet (25GB sa jednim slojem, 50 GB sa dva sloja), što je omogućava korišćenje audio kodnih standarda bez gubitaka za kodovanje višekanalnog zvučnog zapisa. Primeri su:
  - Dolby TrueHD
  - DTS HD Master Audio
- U osnovnoj verziji namenjeni su za bitske brzine do 24Mb/s i prenos pomoću HDMI veze, podržavaju 7.1 konfiguracije kanala, 24 bita po odbirku, učestanosti odabiranja do 192kHz.



## Poređenje veličine bitskog toka kod različitih tipova kompresije:



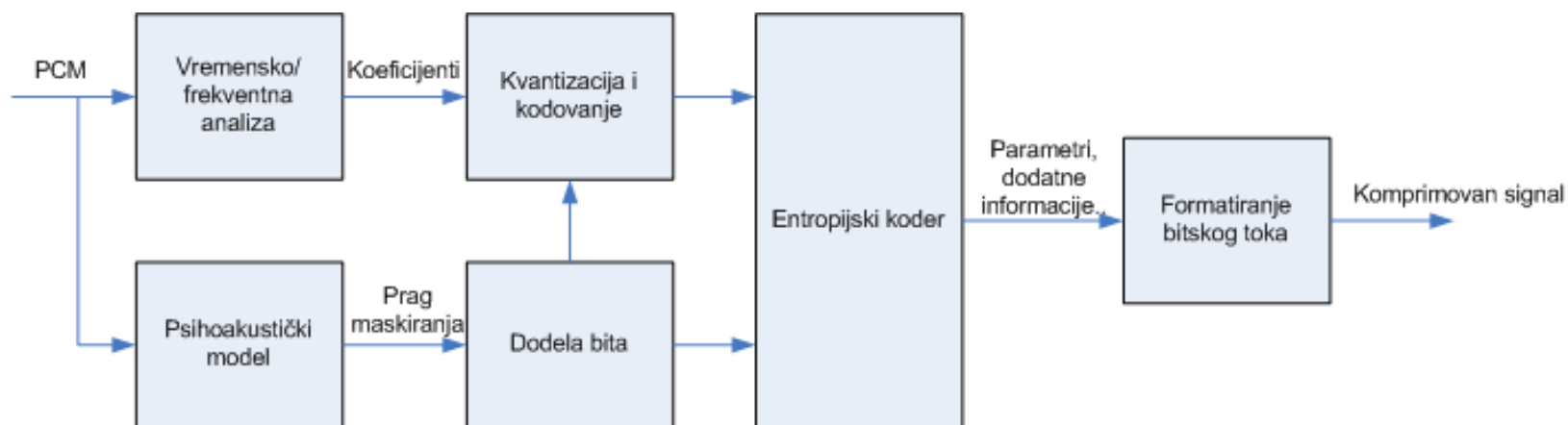
WAV	152,027,180	100.0%
ZIP	127,604,448	83.9%
FLAC	62,624,372	41.2%
Mp3	13,790,145	10.8%

Fs = 44.1kHz, 16bit, Stereo; Mp3 @ 128kb/s, CBR

Primer: L.W. Beethoven - Symphonie Nr. 9 - 1. Allego ma non troppo.wav

# Audio koderi sa gubicima – perceptualni koderi

- Oslanjaju se na nesavršenost ljudskog uha, odnosno odredište u komunikacionom lancu
- Koriste psihoakustički model da bi izvukli irelevantne informacije
- Osnovni blok dijagram perceptualnog audio koda:



# Perceptualni audio koder (1)

- Obrađuju blok ulaznih PCM odbiraka. Dužina ulaznog bloka zavisi od tehnologije: (najčešće od 256 do 2048). Dužina bloka može biti promenjiva u zavisnosti od prirode signala: stacionarni signali dozvoljavaju veće blokove obrade, tranzicioni male (npr: kod WMA koda se dužina okvira kreće od 128 do 2048 odbiraka).
- Filter Banka: preslikavanje ulaznog signala iz vremenskog u frekventni domen omogućava izbacivanje suvišnih informacija zasebnim kodovanjem svake frekventne komponente. Visoki tonovi, na primer, imaju vremenski sporo promenjive frekventne komponente. Na ovaj način se smanjuje količina podataka neophodnih za predstavljanje signala, u odnosu na predstavljanje signala u vremenskom domenu.
  - Prilagodljivi broj bita za svaku frekventnu komponentu.
  - Dodeljivanje različitog broja bita za različite frekventne komponente omogućava kontrolu nivoa kvantizacije svake komponente ponaosob i samim tim visoku preciznost kodovanja.
  - Primer: Pseudo Quadrature Mirror Filter Banks - PQMF ili Modifikovana Diskretna Kosinusna Transformacija – MDCT

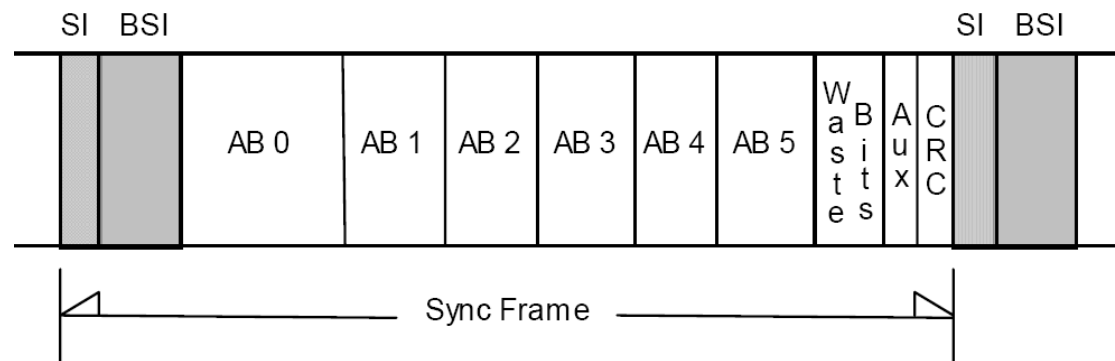
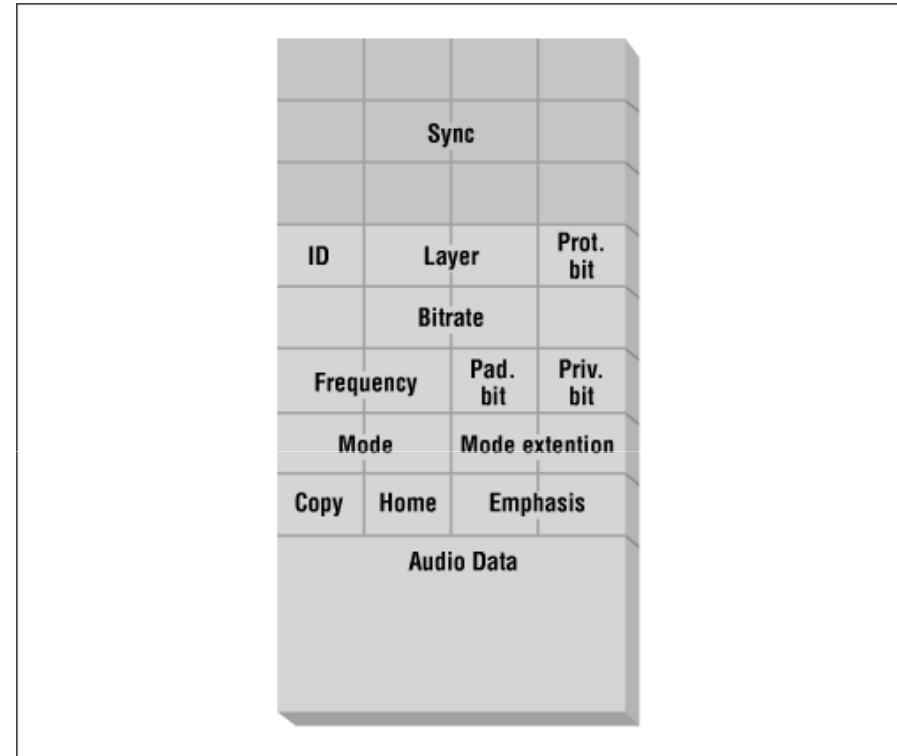
# Perceptualni audio koder (2)

- Nakon mapiranja, frekventne komponente se kvantizuju
- **Psihoakustički model** analizira ulazni signal, određuje nivoe maskiranja svake frekventne komponente i računa SMR vrednosti. U zavisnosti od jačine signala i SMR vrednosti, rutina za dodelu bita dodeljuje ograničeni broj bita predviđenih za mantise frekventim komponentama.
- **Kvantizacijom** upravlja **blok za dodelu bita**, tako da maksimalno uveća razliku odnosa signala i šuma (eng. Signal to Noise Ratio - SNR) i signala i maske (eng. Signal to Mask Ratio - SMR) na svakom bloku podataka.
- Entropijski koder (najčešće Hafman) izbacuje redundatne informacije iz signala
- Bitski tok nosi informacije o kodovanim audio podacima (npr: mantise, faktori skaliranja), a i podatke o dodeli bita. Pored ovoga, bitski tok često nosi kontrolne informacije kao što su: dužina bloka obrade, tip prozora i sl, koje su potrebne u procesu dekodovanja. Dodatni sadržaj bitskog toka mogu činiti zaštitni kodovi, sinhronizacione informacije i sl..



# Perceptualni audio koder (3)

- Rezultat kodovanja jednog bloka PCM odbiraka je sinhronizacioni okvir
- Veličina sinhronizacionog okvira zavisi od stepena kompresije i od dodatnih informacija smeštenih u okvir.
- Na slikama su prikazani primeri sinhronizacionih okvira kod AC3 i MP3 koda:



# Perceptualni audio dekodler (1)

- Kodovani bitski tok se najpre raspakuje na sestavne delove, odnosno izdvoje se audio podatki, kontrolne informacije i dodatni sadržaj.
- Informacije o dodeli bita se koriste pri dekvantizaciji, da bi se što bolje rekonstruisali koeficijenti u frekventnom domenu.
- Korišćenjem filter banki sinteze (npr: PQMF ili Inverzna MDCT) koeficijenti se preslikaju iz frekventnog u vremenski domen i konačno u audio PCM signal



# Perceptualni audio dekodler (2)

- Rekonstruisani koeficijenti sadrže kvantizacioni šum koji bi trebalo da je nečujan (ili blizu nečujnog), ako je psihoakustički model dobro projektovan.
- **Psihoakustički model** i dodela bita se najčešće vrše samo na strani koda.
- Izuzetak je npr. Dolbijeve AC-3 standard kod koga se dodela bita u pojednostavljenoj formi vrši i sa strane dekodera pa je potrebno, u okviru bitskog toka, preneti i podatke koji će ovaj proces sa strane dekodera pojednostaviti.

# Vremensko frekventna transformacija

## *Izobličenja usled prednjeg odjeka*

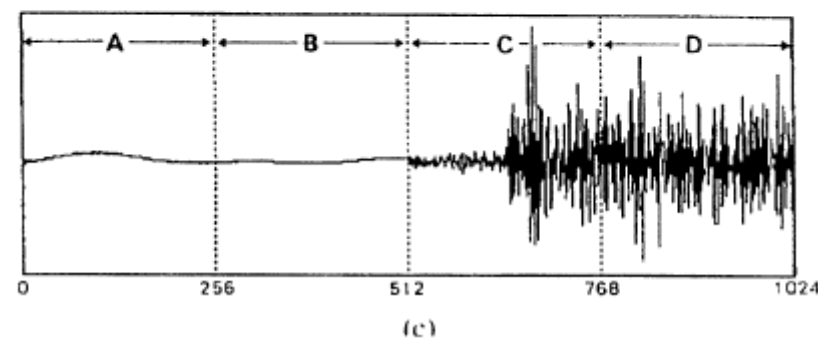
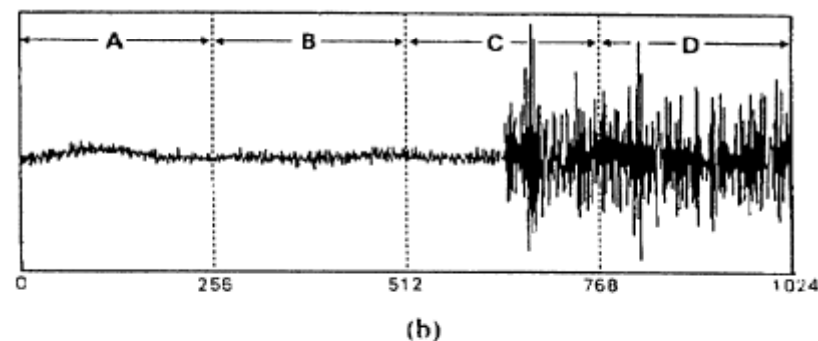
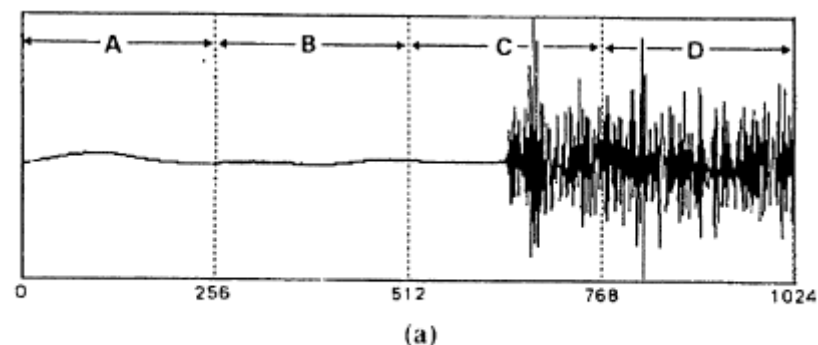
Problem: jak iznenadni zvuk se pojavi blizu kraja bloka obrade i prethode mu zvukovi niske energije

- Inverzna transformacija ravnomerno 'raširi' kvantizacionu grešku po bloku obrade, oslanjajući se na srednju procenu energije bloka, što dovodi do nemaskirane greške u delovima spektra sa malom energijom.
- Energija signala koji se koduje, odnosno ulaznog PCM bloka obrade, prenese se na vremenski period pre nastanka signala pa dolazi do **prednjeg odjeka**

# Vremensko frekventna transformacija

## *Adaptivni blokovi obrade*

- Primer kastanjete
- Rešenje: adaptivne filter banke
  - veliki blokovi kod stacionarnih signala
  - mali blokovi kod promenjivih signala
- Većina standardnih audio koda koristi ovaj pristup kodovanju (mp3, aac, wma, ac3..)



# Metode za dodelu bita (1)

- Dodela bita je proces raspodele ograničenog broja bita  $B$  na  $M$  frekventih opsega, tako da se zadovolji kriterijum minimalnog izobličenja. Rezultat predstavljaju pridruženi biti  $b(k)$ ,  $k=0,1,\dots,M-1$ , koji određuju dužinu reči frekventog koeficijenta u  $k$ -tom opsegu.

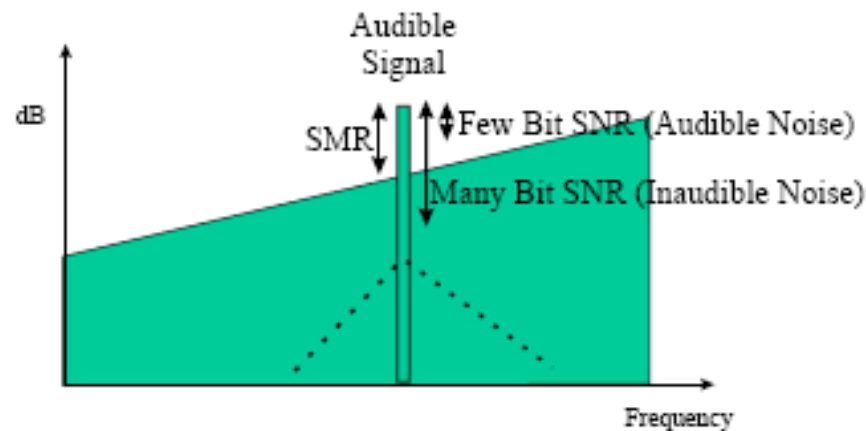
- Dodela bita mora da zadovoljava sledeći uslov:

$$\sum_{k=0}^{M-1} b(k) = B$$

- $B$  zavisi od zadate veličine sinhronizacionog bloka, kapaciteta kanala i sl.. Najčešće je konstan za ceo bitski tok
- U psihoakustičkom modelu se računa željena raspodela kvantizacionog šuma za svaki frekventi opseg, a biti se raspodeljuju tako da kvantizacioni šum aproksimira željenu raspodelu

# Metode za dodelu bita (2)

- Osnovna prednost obrade signala u frekventom domenu je mogućnost kodovanja pojedinačnih komponenti spektra sa različitom tačnošću.
- Dodela bita sledi visinu energije frekventne komponente: pojasi sa visokom energijom zahtevaju više bita od pojasa sa niskom energijom (koji se možda ne moraju ni prenositi).
- Na ovaj način se kontroliše kvantizaciona greška na nivou opsega, i uobličava se šum dobijen rekonstrukcijom signala:



# Metode za dodelu bita (3)

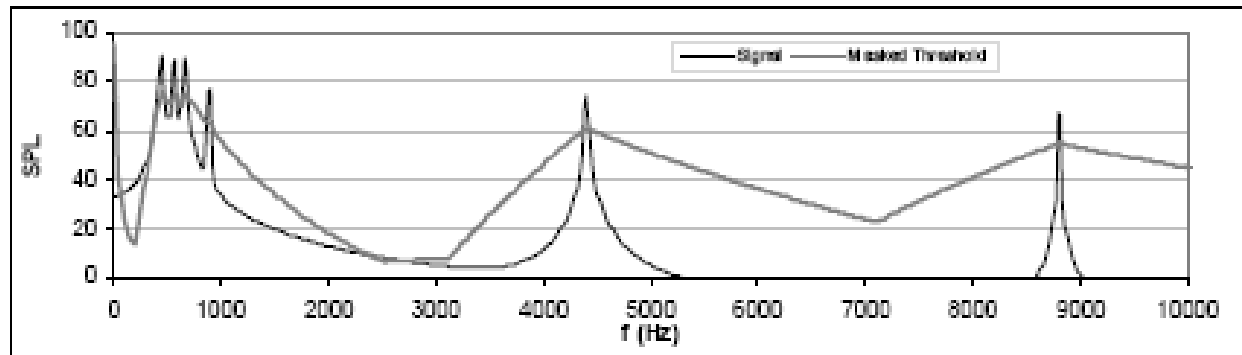
- Dodela bita treba da je takva da zadrži SNR iznad SMR kroz ceo spektar
- Ova razlika, odnos šuma i maske (NMR), daje meru izobličenja u odnosu na izračunati prag maskiranja.
- Kada nema dovoljno bita na raspolaganju, strategija raspodele bita treba da obezbedi minimalnu pozitivnu vrednost NMR-a.



# Metode za dodelu bita (4)

Na sledećoj slici je primer praga maskiranja i dodele bita koja je unela nečujni šum:

- Signal na slici ima frekventne komponente na 440,554, 660,880,4400,8800 Hz, a frekvenciju odabiranja 48 kHz.
- Za vremensko frekventnu analizu -FFT transformacija
- Prozorom veličine 2048 tačaka.
- SPL kriva pokazuje snagu originalnog signala,
- Aproksimaciju prenosne funkcije maskiranja korišćena trougaona funkcija



# Algoritam dodele bita

- Kako dekoderu preneti informaciju koji opseg je kodovan, a koji nije? Dekoder mora da zna na koji način su biti raspoređeni.
- Kako odlučiti koje opsege je potrebno kodovati? Za fiksni prag maskiranja različit broj bita je potreban za svaki blok! Ovo dovodi do promenljivog bitskog toka, odnosno do problema kod aplikacija u realnom vremenu. Ako se uzme konstanti broj bita po bloku, koder dodeljuje bite iz rezervoara bita dok ga ne isprazni.
- Kompleksan algoritam raspodele bita dovodi do dobre raspodele i puno kontrolnih informacija koje je potrebno preneti dekoderu to što opterećuje bitski protok.
- Dele se na unapred adaptivne (eng. *forward adaptive*) i unazad adaptivne (eng. *backwards adaptive*).

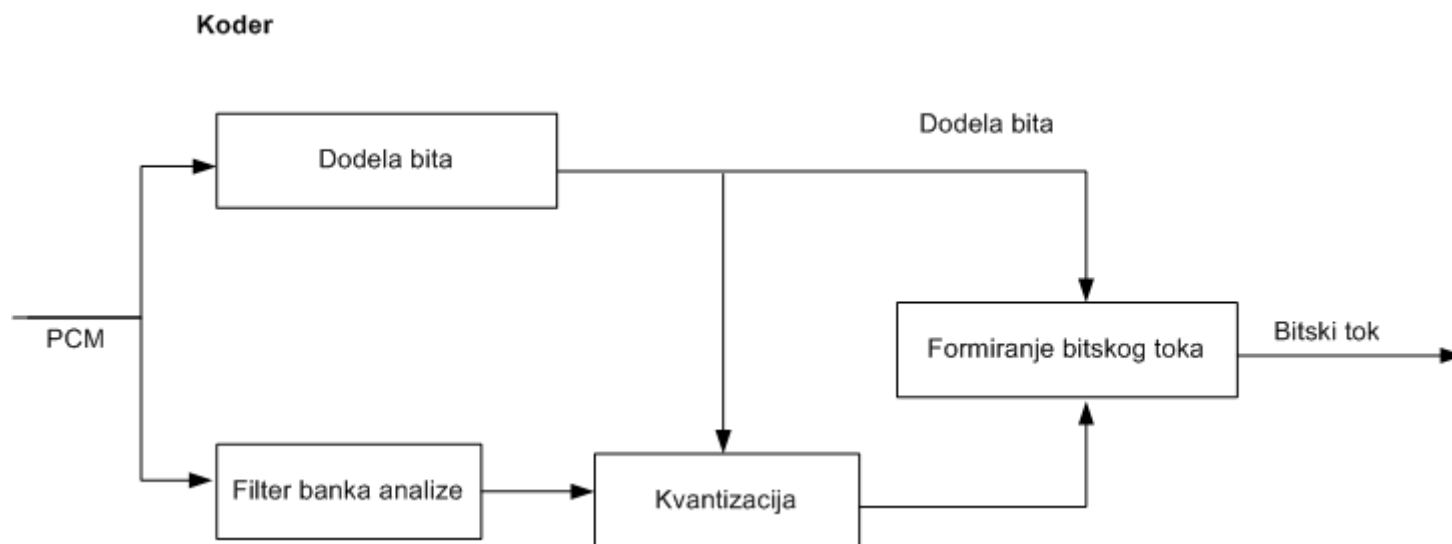
# Metode za dodelu bita

## *Unapred adaptivne metode (1)*

- Koriste u MPEG audio sistemima
- Koder izračunava dodelu bita  $b(k)$  i te informacije neposredno zapisuje u bitski tok.

### **Prednosti:**

- Veoma preciznu dodelu bita, jer koder ima sve informacije o ulaznom signalu.
- Složen psihoakustički model, kao i poboljšanja tog modela
- Jednostavan dekoder (raspakivanje okvira, dekvantizacija, filter banka sinteze).

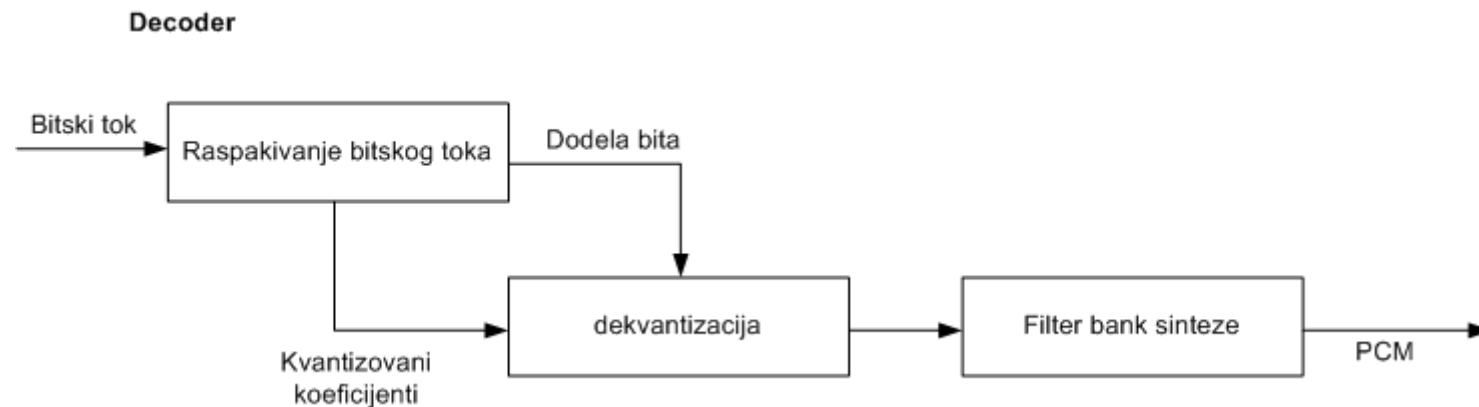


# Metode za dodelu bita

## *Unapred adaptivne metode (2)*

### Mane:

- Koder koristi značajan deo bitskog toka da bi poslao dekoderu informacije o dodeli bita, nije efikasan pristup na niskim bitskim brzinama.

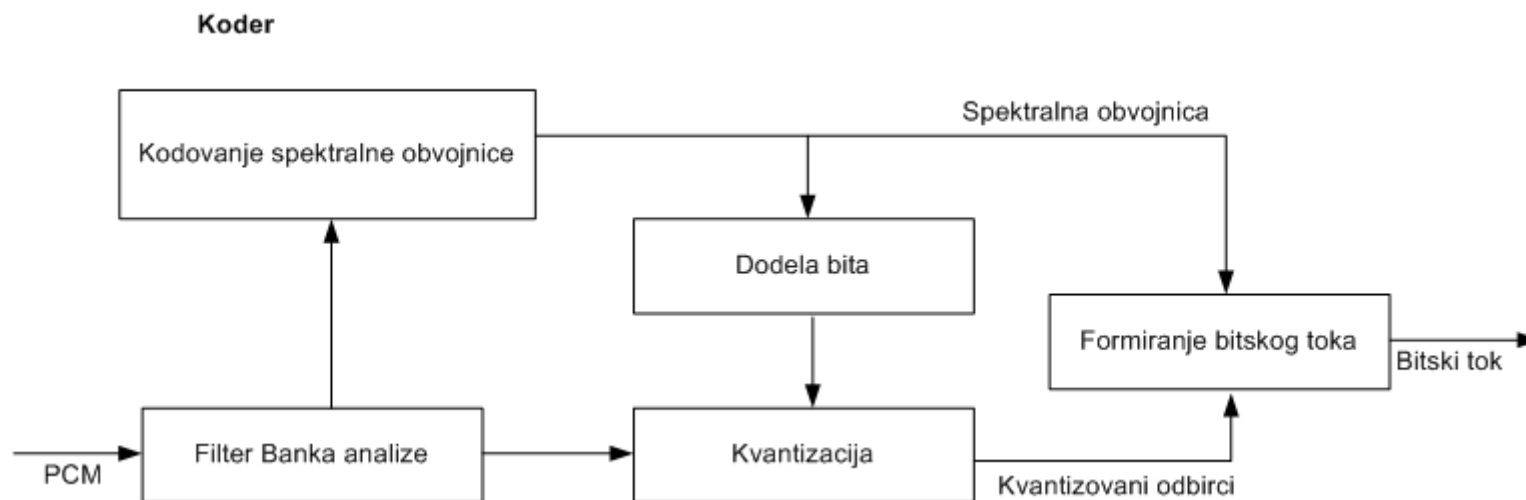


# Metode za dodelu bita

## *Unazad adaptivne metode (1)*

- Koristi se u AC2 audio sistemima, prethodnik AC3 tehnologije (Dolby)
- Dodela bita se računa i u koderu i u dekoderu

**Prednosti:** Nije potrebno prenositi dodelu bita u bitskom toku, što povećava broj bita za kodovanje audio podataka



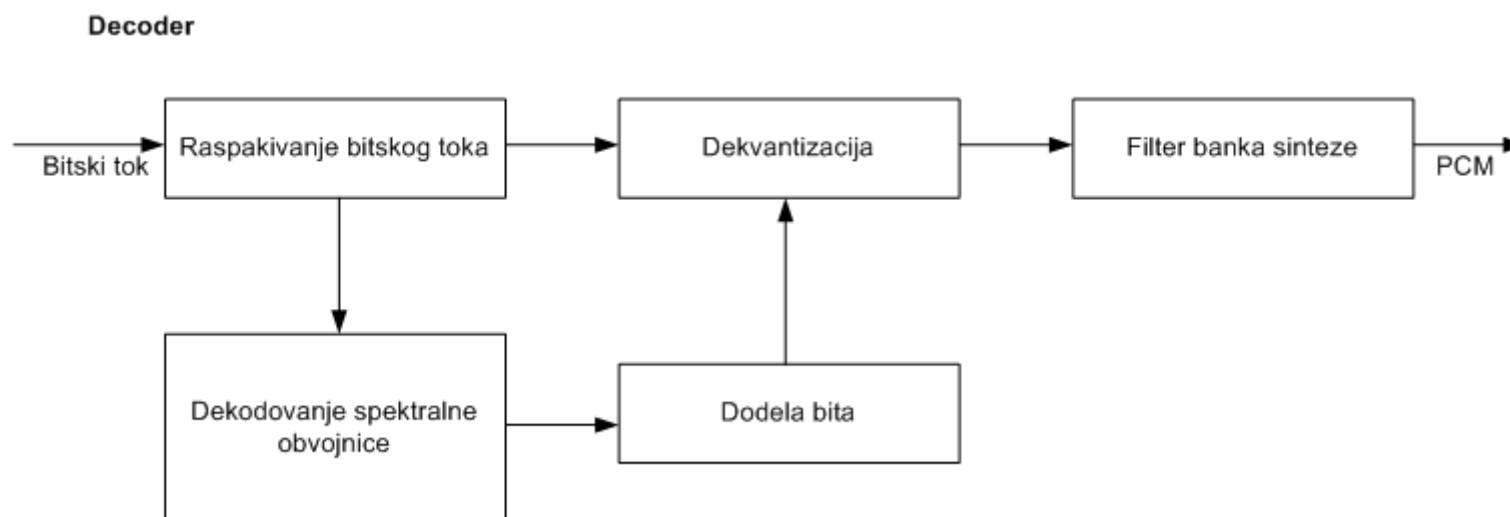
# Metode za dodelu bita

## *Unazad adaptivne metode (2)*

- Dodela bita se određuje iz kodovanog toka podata (spektralne obvojnice), bez neposrednih informacija od kodaera.

### **Mane:**

- Pošto se dodela bita računa sa strane dekodera na osnovu kodovanih audio podataka, tačnost informacija je ograničena i može doći do greške u proračunu.
- Psihoakustički model nije moguće poboljšati, jer bi to iziskivalo menjanje dekodera.
- Složeni dekodier. Teži se što jednostavnijim dekodierima zbog cene krajnjih uređaja (zahtevna obrada uslovljava kompleksniji namenski procesor, pa samim tim i veću cenu)



# Tehnike međukanalne dekorelacije

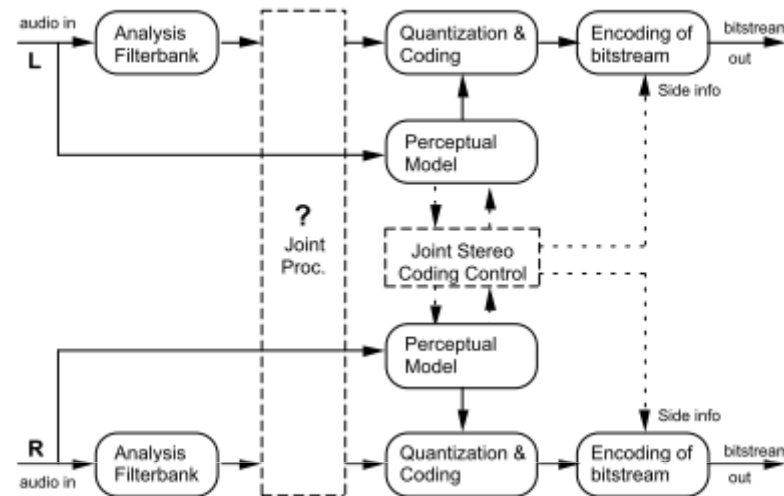


Figure 1: *Generic model of joint stereo encoding.*

- 1990/91 prva generacija audio kodera sa tehnikom međukanalne dekorelacije
- Tehnika nije smela da značajno poveća složenost kodera:
  - Obrada nad spektralnim koeficijentima (bez namenske FB)
  - Na gornjem crtežu: dva skupa funkcionalnih blokova obrade nad zasebnim kanalima.
  - Dodatak je združena obrada spektralnih koeficijenata kao i dodatna perceptualno zasnovana kontrola združenog kodovanja

# M/S Stereo kodovanje

*Razlika izmedju levog i desnog kanala je najčešće jako mala*

- Mid/Side tehnika:
  - Umesto čuvanja podataka levog (L) i desnog (R) kanala, čuvati sredinju vrednost (Mid) i razliku (Side)
$$\text{mid} = (L + R) / 2$$
$$\text{side} = (L - R) / 2$$

- srednja vrednost  
- razlika
  - Obratiti pažnju na znak. Konvencija: Pozitivni znak ukazuje daje levi kanal jači
  - Rekonstrukcija:
$$\text{left} = \text{mid} + \text{side}$$
$$\text{right} = \text{mid} - \text{side}$$
- Umesto dva slična kanala dobijamo jedan kanal (Mid) koji je mnogo veći od drugog (Side)!
- Za kodovanje Side kanala se koristi manje bita, pa se Mid kanal može kodovati preciznije. Ovo dovodi do tačnije rekonstrukcije kanala!



# M/S Stereo kodovanje - primer

**LEFT:** 20 000 21 000 22 000 23 000 22 000 22 000

**RIGHT:** 18 000 19 000 20 000 19 000 18 000 19 000

Pretpostavljamo da kompresija unosi grešku  $\sim \pm 2\%$

**LEFT:** 20 400 20 580 22 440 22 540 22 440 21 560

**RIGHT:** 18 360 18 620 20 400 18 620 18 360 18 620

Prebacivanje u Mid/Side format:

**MIDDLE:** 19 000 20 000 21 000 21 000 20 000 20 500

**SIDE:** +1 000 +1 000 +1 000 +2 000 +2 000 +1 500

Kanal razlike kodujemo sa manjom preciznošću, ali ukupna preciznost raste, pa je greška npr:  $\sim \pm 1\%$

**MIDDLE::** 19 190 19 800 21 210 20 790 20 200 20 295

**SIDE:** +1 040 +960 +1 040 +1 920 +2 080 +1 440

Prebacivanje u L/R format:

**LEFT:** 20 230 20 760 22 250 22 710 22 280 21 735

**RIGHT:1** 8 150 18 840 20 170 18 870 18 120 18 855

Rekonstruisani signal bolje odgovara originalnom signalu!

# M/S Stereo kodovanje

- **Dinamička aktivacija:** značajno povećanje kompresije kod sličnih signala (near monophonic), ali ne sme se koristiti kod dva mono signala. **Uklanja redundantne informacije!**
- **Perfektna rekonstrukcija**, ne unosi izobličenja
- **Zavisno od sadržaja signala** pojačanje kompresije varira (max ~50% kod približno istih kanala, negativan rezultat kod različitih signala )
- U potpunosti zadržava prostornu informaciju pa se može primeniti na **pun audio opseg**
- Koristi se kod MPEG 1/2 Layer i MPEG 2/4 AAC

# Intensity stereo kodovanje -IS

*U gornjim granicama čujnosti, ljudsko uho nije osetljivo na pravac zvuka*

- Za veoma niske i veoma visoke učestalosti, ljudsko uho nije u mogućnosti da locira izvor zvuka u prostiru sa punom tačnošću
  - Preko 2kHz percepcija zasnovana na spektralnoj obvojnici
- **Redukcija irelevantnih informacija**
- IS kodovanje unosi gubitak u signal. Koristi se za male bitske brzine (~96kb/s) jer unosi izobličenja koja manje smetaju od ostalih izobličenja!
- Velika ušteda u prenosu: prenosi se samo jedan kanal sa podopsežnim podacima i dodatne informacije (faktore skaliranja). Ako se primeni na pola frekventog opsega ušteda je ~20%. Maksimalna ušteda ~40%
- Korisna tehnika samo za visoke frekvencije. Ako se primeni na širem opsegu unosi velika izobličenja (neophodna pažljiva primena!)
- Primena: Svi nivoi MPEG 1/2 i MPEG 2/4 AAC
- Kod višekanalnih koda moguće je združiti frekvente koeficijente više kanala u jedan skup koeficijenata, uz dodatne informacije o skaliranju svakog kanala

# Međukanalna dekorelacija – višekanalni koderi (1)

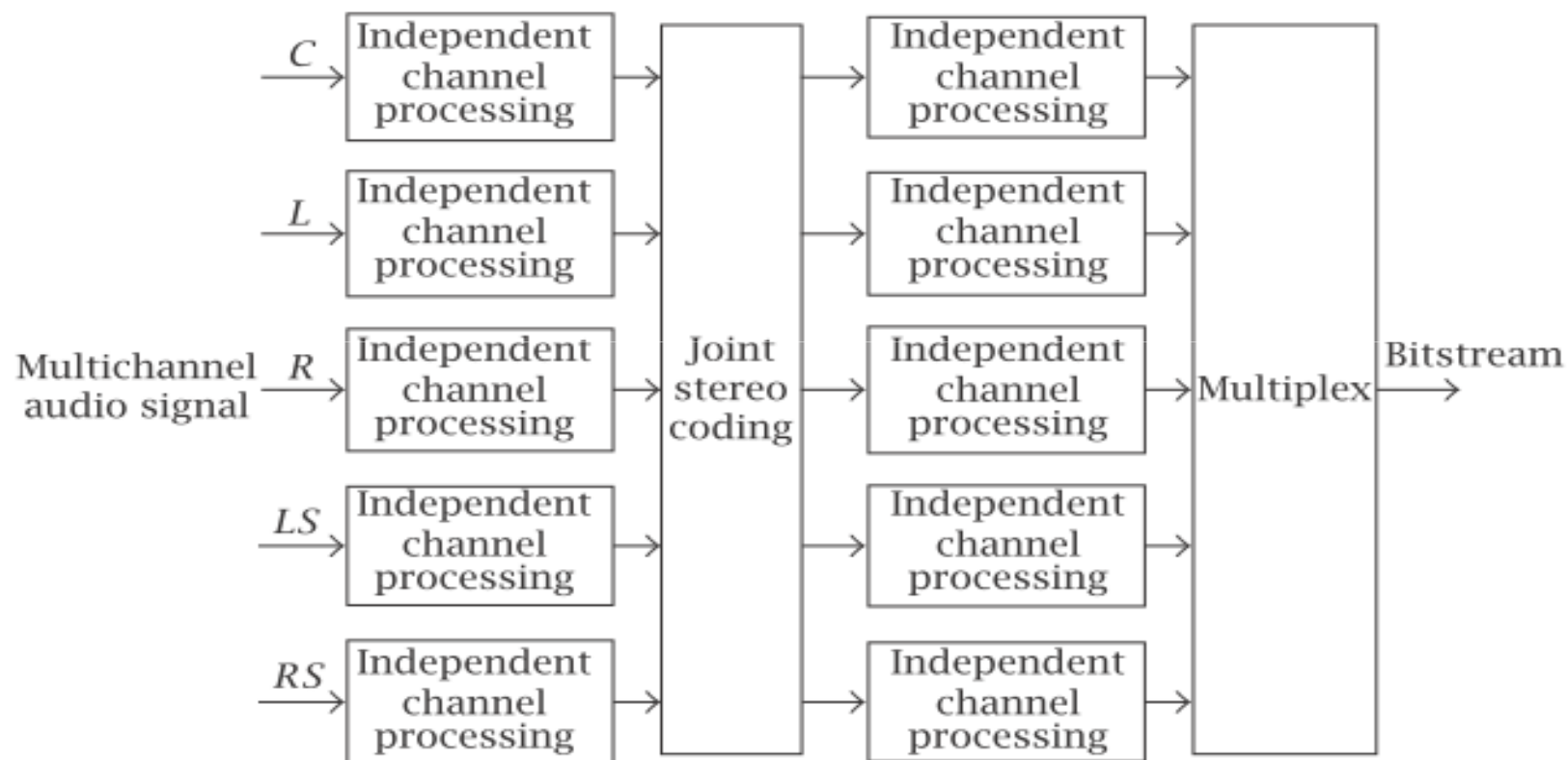


FIGURE 7.7. Illustration of a multichannel joint stereo coder.

# Međukanalna dekorelacija –višekanalni koderi (2)

- M/S i IS se mogu primeniti nad istim signalom
- M/S tehnika (donji deo spektra):
  - Združuju se kanali simetrični u odnosu na korisnika: levi i desni, levi okruženje i desni okruženje
- IS tehnika (gornji deo spektra)
  - $L(i)$  i  $R(i)$  se zamenjuju sa:

$$(L(i) + R(i)) * \sqrt{\frac{E_l}{E_r}}$$

( $E_l$  i  $E_r$  enegrija levog odnosno desnog podospega)

# Izobličenja audio signala (1)

- Predstavljaju sve promene rekonstruisanog signala u odnosu na originalni signal: odsecanje (eng. *clipping*), izobličenja harmonika i mešanja signala
- Izobličenja možemo podeliti i prema mestu na kom su unešena u okviru komunikacionog lanca
  - Izobličenja uneta procesom izvornog kodovanja
  - Izobličenja usled grešaka nastalih tokom prenosa signala
- Izobličenja koja nastaju usled izvornog kodovanja audio signala su, u slučaju perceptualnog koda, najčešće greške usled kvantizacije (kvantizacioni šum).
- Projektovanje audio koda zahteva pravilno raspoređivanje kvantizacionog šuma u odnosu na brzinu protoka podataka: više bita povećava brzinu protoka podataka, ali smanjuje kvantizacionu grešku. Osnovni zadatak bloka za dodelu bita je kontrola kvantizacione greške i predstavlja jezgro audio koda.
- Jedan od načina opisivanja kvantizacione greške je poređenje ulaznog signala sa izlaznim signalom :

$$q(t) = x_{out}(t) - x_{in}(t)$$

- Mogo češća mera kvantizacionog šuma data je odnosom signala i kvantizacione greške:

$$SNR_{audio} = 10 \log \left( \frac{\langle x_{in}^2 \rangle}{\langle q^2 \rangle} \right)$$

# Izobličenja audio signala (2)

Pored kvantizacione greške, čujni artefakti mogu nastati i zbog:

- **Prednji odjek** (eng. *pre-echo*): nastaje kada se energija signala koji se koduje prenese na vremenski period pre nastanka signala, odnosno kada se jak iznenadni zvuk pojavi blizu kraja bloka obrade, i kada mu prethode zvukovi niske energije. Jedan od klasičnih primera ovakvog signala su 'kastanjete'. Rešenje podrazumeva adaptivne filter banke.
- **Preklapanje** (eng. *aliasing*): u normalnim uslovima, ovaj artefakt je jedva čujan kod pod-opsežnih filter banki (kao što je PQMF). Kod MDCT pristupa, i pored perfektne rekonstrukcije, usled kvantizacije može dovesti do vremenskog preklapanja.
- **Efekat „ptičica“** (eng. *birdies*): nastaje kod malih bitskih brzina, gde se zbog velike kompresije dodela bita visokim frekvencijama vrši na nivou opsega, pa pojedine frekvencije često nastaju i nestaju. Ovaj neprijatan efekat može da se izbegne korišćenjem nisko propusnog filtra u slučaju velike kompresije.
- **Reverberacija govora**: dolazi usled korišćenja velikih blokova obrade pri obradi govora pri maloj bitskoj brzini. Proizvodi metalni, neprirodan zvuk, a može se izbeći korišćenjem filter banke koja se dinamički prilagođava vrsti signala (govor, muzika i sl..)

# Mere kvaliteta audio signala

## *Objektive mere (1)*

- Fundamentalni koncept audio kodovanja sa gubicima, ali ga je teško na objektivan način opisati
- Jednostavne objektivne mere kvaliteta:
  - odnos signala i šuma – *SNR*
  - vršni odnos signala i šuma – *PSNR*

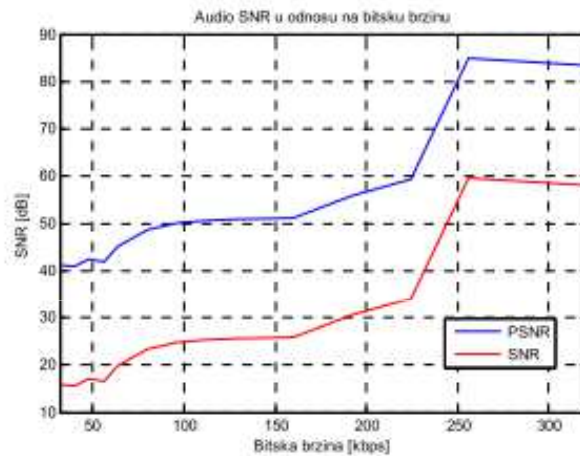
$$P = \frac{\sum_{n=0}^N (x(n) - r(n))^2}{N} \quad PSNR = 10 \log \left( \frac{2^{b-1}}{P} \right) [dB] \quad SNR = 10 \log \left( \frac{\sum_{n=1}^N r(n)^2}{P} \right) [dB]$$

gde je  $r$  referentni signal,  $x$  signal koji se ispituje,  $N$  veličina okvira u odbircima, a  $b$  broj bita po odbirku.

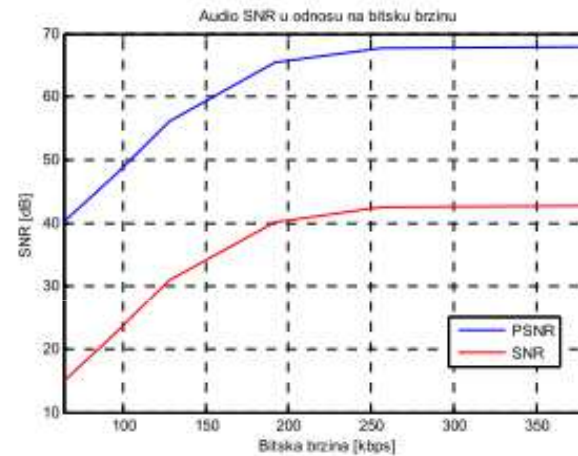
- Ove mere u potpunosti zanemaruju psihoakustičke efekte audio kodovanja



# Odnos kvaliteta rekonstruisanog audio signala u odnosu na stepen kompresije (kbps)



a) Lame MP3



b) AC3

# Mere kvaliteta audio signala

## *Objektive mere (2)*

- Neadekvatnost korišćenja ovih mera kod perceptualnih kodera objasnili su istraživači Johnston i Branderburg krajem osamdesetih godina prošlog veka. Oni su izveli sledeći eksperiment:
- U originalni signal uneli su dve vrste šuma:
  - šum koji je prilagođen, odnosno oblikovan, prema psihoakustičnom modelu maskiranja i nalazi se ispod praga maskiranja
  - beli šum koji je ravnomerno raspoređen u frekventom domenu, znači uglavnom iznad praga maskiranja.
- Oba signala su imala isti *SNR* od 13 dB, ali drugačiji perceptualni audio kvalitet: signal sa perceptualno prilagođenim šumom imao je jedva čujna izobličenja za razliku od signala sa belim šumom kod koga su čujna izobličenja bila velika. Ovaj efekat je nazvan 'čudo od 13 decibela' (eng. '*13 dB miracle*')
- I pored ovoga, *SNR* i *PSNR* daju dobru kvantitativnu meru pa se često koriste u oceni kvaliteta audio signala naročito u fazi razvoja audio tehnologije.

# Mere kvaliteta audio signala

## *Subjektivne mere (1)*

- Slušni testovi (eng. *listening tests*) , zahtevaju veliki broj obučenih slušalaca u kontrolisanom okruženju
- Čujna razlika između originalnog signala i signala koji se ispituje posmatra se kao oštećenje i ocenjuje se u skladu sa skalom **ITU-R BS.1284 standarda**
- Kao što je rečeno, slušni testovi zahtevaju veliki broj obučenih slušalaca i namensku opremu, pa su skupi i zahtevaju puno vremena

# Mere kvaliteta audio signala

## *Subjektivne mere (2)*

- Subjektivna ocena odstupanja (eng. *Subjective Difference Grade - SDG*)  
 **$SDG = \text{Ocena Ispitnog Signala} - \text{Ocena Referentog Signala}$**
- Negativna vrednost *SDG* znači da slušalac uspešno razlikuje originalni signal od ispitnog signala. *SDG* je nula kada je kvalitet transparentan i kada su izobličenja nečujna

<i>Subjektivni kvalitet signala</i>	<i>Ocena</i>	<i>SDG</i>	<i>Izobličenje</i>
Veoma dobar	5.0	0	Nečujno
Dobar	4.0	-1.0	Čujno ali ne smeta
Srednji	3.0	-2.0	Malo smeta
Slab	2.0	-3.0	Smeta
Loš	1.0	-4.0	Veoma smeta

**Subjektivna skala zasnovana na ITU-R BS.1284 standardu u opsegu od 1.0 do 5.0**

# Mere kvaliteta audio signala

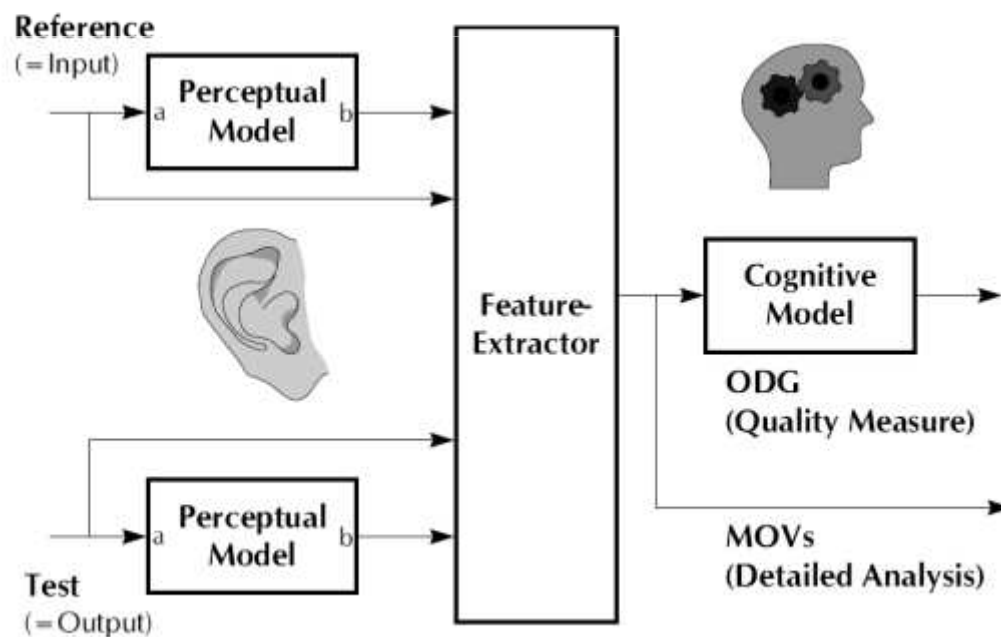
## *Objektivne perceptualne mere - PEAQ*

- Predviđa audio kvalitet na osnovu objektivnih mera zasnovanih na psihoakustičkim principima, odnosno odgovarajućim modelima ljudskog uha
- Perceptualno ocenjivanje audio kvaliteta (eng. *Perceptual Evaluation of Audio Quality*, u daljem tekstu PEAQ) prihvaćeno je od strane ITU-R u okviru BS.1387 [ITU-R BS.1387-1] preporuke 1998. godine.
- PEAQ je trenutno jedini prihvaćen standard za ocenu kvaliteta širokopojasnog audio signala i koristi se kako za ocenu audio tehnologija tako i uređaja.
- PEAQ algoritam zahteva dva ulaza: originalni neizmenjeni signal i ispitni signal

# Mere kvaliteta audio signala

## *Objektivna ocena odstupanja - ODG*

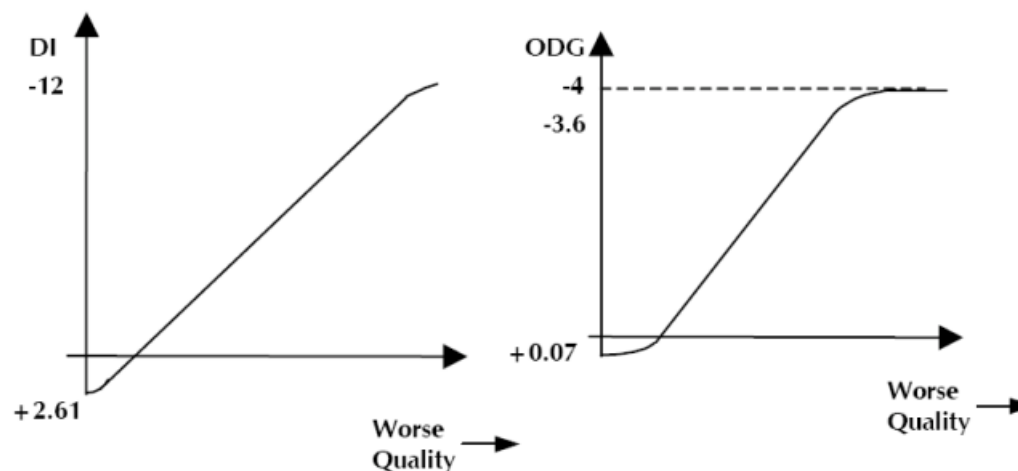
- Prvi korak je obrada signala u okviru perceptualnog modela
- Nakon toga, poređenjem izlaza iz perceptualnog modela, algoritam modeluje čujno izobličenje koje je prisutno u ispitnom signalu i generiše izlazne promenljive modela (eng. *Model Output Variables, -MOV*) koje daju detaljnu analizu signala.
- Krajnji cilj PEAQ algoritma izvođenje jedne mere kvaliteta, nazvane *objektivna ocena odstupanja* (eng. *Objective Difference Grade - ODG*), MOV podaci prolaze kroz kongitivni model ljudskog auditornog sistema, odnosno neuralnu mrežu



# Mere kvaliteta audio signala

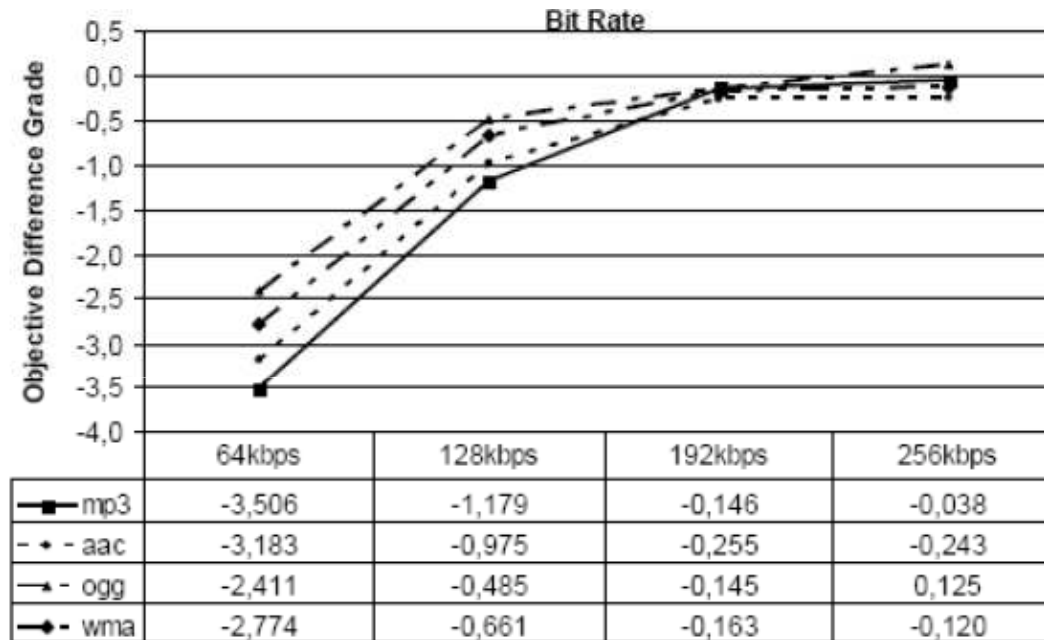
## *Indeks izobličenja- DI*

- Indeks izobličenja (eng. *Distortion Index - DI*) ima isto značenje kao i *ODG* mera, ali se opseg vrednosti ove dve promenjive razlikuje.
- Ako je vrednost *ODG* mere manja od -3.6 preporučuje se korišćenje *DI* mere. U suprotnom, *ODG* se koristi kao mera kvaliteta.
- Na slici je prikazan odnos audio kvaliteta i *DI* vrednosti, kao i odnos audio kvaliteta i *ODG* vrednosti:



# Mere kvaliteta audio signala

## *Prikaz ODG vrednosti različitih audio kodera*

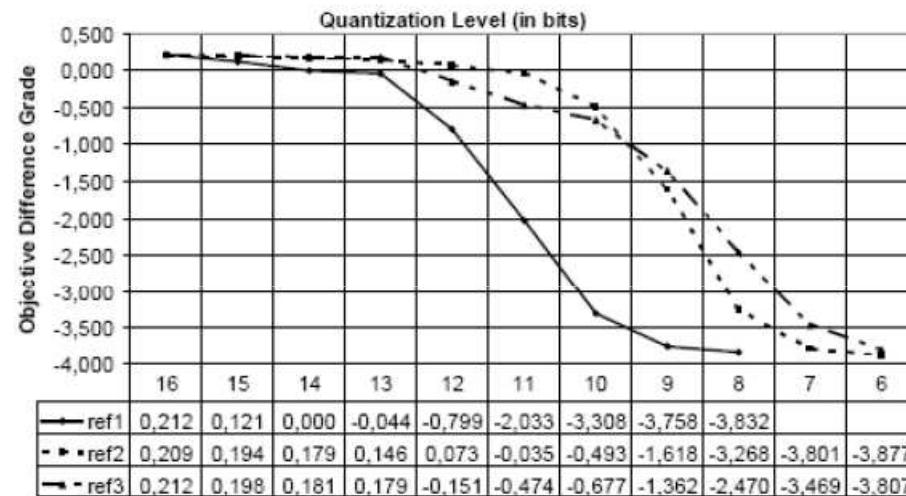


- Najveća razlika *ODG* mere se uočava na nižim bitskim brzinama
- Preko 192 kbps sve četiri tehnologije imaju sličan kvalitet



# Mere kvaliteta audio signala

## *Zavisnost ODG vrednosti u odnosu na stepen kvantizacije*



- Ref1 predstavlja signal govora, dok su ref2 i ref3 muzički signali
- Jasno se vidi da kvantizacija sa više od 12 bita gotovo da ne unosi izobličenje u signal: