

ZVUK I RAČUNALA

SKRIPTA ZA 2. MI

REDUKCIJA PODATAKA

Npr. pri pohrani 1min.16 bitovnog stereo signala uz f.u. 44,1 kHz, dobiva se veliki tok podataka koji je potrebno smanjiti. Najjednostavnije bi bilo kompresorom i ekspanderom smanjiti dinamiku ulaznog analognog signala. No tada se pojavljuju izobličenja zbog nekomplementarnih karakteristika tih sklopova, te do izobličenja kvantiziranja zbog manjeg broja stupnjeva kvantiziranja. Općenito se može reći da na analognom signalu nije pogodno vršiti bilo kakvu redukciju. Bolje rezultate možemo postići sljedećim postupcima:

a) *nelinearno kvantiziranje* - nejednoliko stupnjevanje intervala kvantizacije. Za male amplitude koristimo male stupnjeve kvantizacije, a za velike velike stupnjeve kvantizacije. Nedostaci: nekomplementarne karakteristike daju izobličenja, trenutna vrijednost signala određuje napon smetnje i time uzrokuje čujna izobličenja.

b) *diferencijalno pulsno kodna modulacija (DPCM)*. Prenosi se razlika upravo uzorkovane vrijednosti od zadnje prethodne uzorkovane vrijednosti. Nedostatak je širenje pogreški. c) *delta modulacija* – koristi se naduzorkovanje s frek. većom 10 do 20 puta od frekvencije uzorkovanja. Razlika među uzorcima postaje 1 bit tj. vrlo male. Nedostatak je širenje pogreški. d) *konvertor s pomičnim zarezom*. Prema razini signala logika povećava ili smanjuje pojačanje. Tako je konvertor uvijek optimalno pobuđen. 12 bitova se smješta u 16 bitovnu riječ pomoću posmačnog registra. Ovaki konvertori povećavaju razinu šuma kvantiziranja i ne koriste se za profesionalne studijske uređaje.

Redukcija podataka se općenito može izvesti pomoću: a) *redundancije* – Osobina signala da sadrži dijelove koji su višetruki i koje ne treba prenositi. Dva podatka su ista pa se ne prenose oba. b) *irelevantnost* – nebitne komponente, odnosno nečujne jer se nalaze ispod praga maskiranja se ne trebaju prenositi. c) *entropija* – daje mogućnost da se one vrijednosti koje se često ponavljaju kodiraju s manjim brojem bitova. Huffmanov kod – vrijednosti koje se česti pojavljuju kvantiziraju se sa manjim brojem bitova.

Razlikujemo još : a) pojasno kodiranje (percepcijsko). Kodiraju se pojedini pojasevi (barkovi). Optimalno korištenje efekta maskiranja. b) transformacijsko kodiranje. Prvo analiza u frek. domeni pa onda kodiranje.

KODIRANJE SIGNALA U FREKVENCIJSKOJ DOMENI

(vezano je djelom i uz redukciju podataka)

1) PERCEPCIJSKO KODIRANJE

Koriste se kratkotrajne (short-term) karakteristike izvornog signala i psihoakustičke osobine maskiranja ljudskog uha da bi se smanjio tok podataka. Izvorni spektar dijeli se na frekvencijske pojaseve. Ljudski sluh ima 24 područja maskiranja (24 barka na bazilarnoj membrani) koja se dijele na kritične pojaseve širine 100 Hz do frekvencije od 500 Hz, a nakon toga širine 1/3 oktave. Odgovarajuća frekvencijska komponenta unutar nekog pojasa podignut će prag maskiranja u tom pojasu, te okolni signali, koji padnu ispod praga maskiranja, neće biti čujni. Takvo se maskiranje zove i simultano maskiranje.

Nagib funkcije praga maskiranja je strmiji prema nižim frekvencijama i obično iznosi oko 27 dB/barku. Nagib prema višim frekvencijama ovisi i o glasnoći maskera. Vrlo glasni signali pokazuju nagib od -6 dB/barku, a srednje glasni od -10 dB/barku.

Ukoliko se maskirajući signal sastoji od više maskirajućih tonova, potrebno je odrediti ukupni prag maskiranja kao funkciju frekvencije, a koji se dobiva iz pragova maskiranja svakog pojedinog pojasa. Konačno se dobiva SMR koji predstavlja omjer snage signala i ukupnog praga maskiranja.

Osim frekvencijskog, postoji i vremensko (temporalno) maskiranje koje nastaje prije i poslije pojave maskirajućeg signala. Vrijeme pred-maskiranja je znatno kraće od vremena post-maskiranja.

2) TRANSFORMACIJSKO KODIRANJE

Blok ulaznih uzoraka se linearno transformira diskretnom transformacijom u niz transformacijskih koeficijenata. Ovi koeficijenti se zatim kvantiziraju i prenose. U dekoderu inverzna transformacija vraća signal natrag u vremensku domenu.

Tipične transformacije su diskretna Fourierova transformacija (DFT), MDCT (modificirana diskretna kosinusna transformacija), itd. Diskretne transformacije se mogu predložiti kao niz uskih pojasnopropusnih filtara, a kod MDCT se koriste preklapajući blokovi sa 50% preklapanja. Na taj način stvara se dvostruko više koeficijenata nego što je potrebno, ali zbog TDAC (Time Domain Aliasing Cancellation, poništenja izobličenja preklapanja u vremenskoj domeni) iste komponente koje se pojavljuju u dva susjedna bloka bit će poništene tijekom rekonstrukcije.

Frekvencijsko kodiranje daje znatno bolji učinak ako je broj bitova koji se dodjeljuje za kvantizaciju svakog transformacijskog koeficijenta prilagođen kratkotrajnom spektru.

Rekvantizacija u koderu povećava kvantizacijski šum u cijelom vremenskom bloku, pa ako se tranzijent pojavi na kraju bloka, kvantizacijski šum će se povećati i na početku bloka, što će rezultirati u pojavi pred-jeke. Da bi se to izbjeglo, potrebno je koristiti blokove različite duljine.

Transformacijsko kodiranje koristi se npr. u MPEG 2 AAC koderu.

Kod koderu u frekvencijskoj domeni frekvencijske komponente se kvantiziraju s varirajućim brojem bitova, ovisno o čujnosti i pragu maskiranja pojedinih pojaseva.

Često se koristi **dinamička alokacija** bitova upravljana kratkotrajnom spektralnom ovojnicom izvornog signala, što se prenosi kao dodatna informacija.

Moguća je pojava pred-jeke, koja je prilično čujna, a izbjegava se smanjivanjem vremenske širine blokova, što dovodi do povećanja toka podataka jer se prenosi više dodatne informacije. Kompromisno se upotrebljavaju blokovi s dvije različite duljine trajanja, gdje se kraći blokovi koriste pri kratkotrajnim signalima. Vrlo kratka pred-jeka (nekoliko ms) bit će maskirana prirodnim vremenskim maskiranjem uha.

Kodiranje u *frekvencijskoj* domeni dijeli se na podpojasno kodiranje i adaptivno transformacijsko kodiranje, između kojih postoje razlike u:

- broju spektralnih komponenata;
- strategijama kvantizacije;
- maskiranju rezultirajućih kodnih grešaka.

Pri *redukciji* audio podataka koristi se:

- **redundancija** - osobina signala da sadrži dijelove koji su višestruki (redundantni) i koji se ne trebaju prenositi;
- **irelevantnost** - nebitne komponente, odnosno nečujne jer se nalaze ispod praga čujnosti ili maskiranja, ne treba prenositi;
- **entropija** - daje mogućnost da se one vrijednosti koje se često ponavljaju kodiraju s manjim brojem bitova.

Dvije su bitne veličine:

1. omjer amplitude signala prema minimalnom pragu čujnosti (SMR, signal to mask ratio), čime je određeno minimalno potrebno kvantiziranje. To je odnos između signala i praga maskiranja. Pozitivna vrijednost pokazuje da je signal čujan, a negativna da je maskiran.
2. Omjer razine šuma prema maskiranju (MNR, mask to noise ratio), što je razmak šuma i praga čujnosti. Taj omjer je usporediv s uobičajenim SNR, koji kod komprimiranog (reduciranog) signala jedva da ima značenje jer je šum prikriven signalom. $SNR = 6n + k$ [dB]

$$SNR = SMR + MNR$$

PERCEPCIJSKO KODIRANJE

Koriste se psihoakustički modeli (npr. efekti maskiranja) uz vremensku rezoluciju određenu svojstvima sluha. Ljudski sluh ima 24 područja maskiranja (barka) koji su široki od 100 Hz do 500Hz. Frekvencijska komponenta unutar nekog pojasa će podići prag maskiranja u tom pojasu, a okolni signali će u tom području biti maskirani. Uskopojasni šum više maskira ton nego obratno. Za pojedini pojas broj bitova se dodjeljuje prema: broj bitova = $((27 \text{ dB} (l_g - l_d))/6.2 \text{ (dB)})$ gdje je l_g gornji rub barka, a l_d donji rub barka. Princip rada: izvorni signal se privodi nizu pojasnopropusnih filtara najčešće 32 koji aproksimiraju kritične pojaseve, susjedne po frekvenciji. Oni dijele signal tako da niz podpojsanih uzoraka može ponovo dati približno originalan signal. Na izlazu iz svakog filtra dobiva se signal nekoliko puta manje f.u. koji se šalje u dekodek koji povećava f.u. svakog podpojsa, interpolira ga i na izlazu sintetizirajućeg filtarskog niza daje signal sličan originalu. Ovakvo kodiranje koristi se u MPEG-1 layer 1 i layer 2 koderima. Kod MPEG-layer 3 primjenjeno je hibridno kodiranje (podpojsano/transformacijsko).

TRANSFORMACIJSKO KODIRANJE

Diskretni vremenski blokovi audio-uzoraka konvertiraju se u frekvencijsku domenu pomoću brze FT (FFT), diskretne FT (DFT) ili modificirane DCT (MDCT).

Najčešće se koristi 50% preklapanje uzastopnih blokova zbog poboljšanja vremenske rezolucije.

Transformacija uzoraka dati će određeni broj (npr. 256, 512, 1024 itd.) spektralnih koeficijenata. Oni će se grupirati u određeni broj (npr. 32) pojasa, čime se aproksimira analiza kritičnim pojasi (barkovima). Spektralni koeficijenti se kvantiziraju u skladu s psihoakustičkim modelom: maskirane komponente se eliminiraju a kvantizacija se radi ovisno o veličini SMR.

Dovoljno velika vremenska rezolucija će se postići ako se koristi dovoljan broj podpojaseva (procesiranje dovoljnog broja uzoraka bliskih po frekvenciji).

Takav način kodiranja može se smatrati perceptualnim kodiranjem u frekvencijskoj domeni.

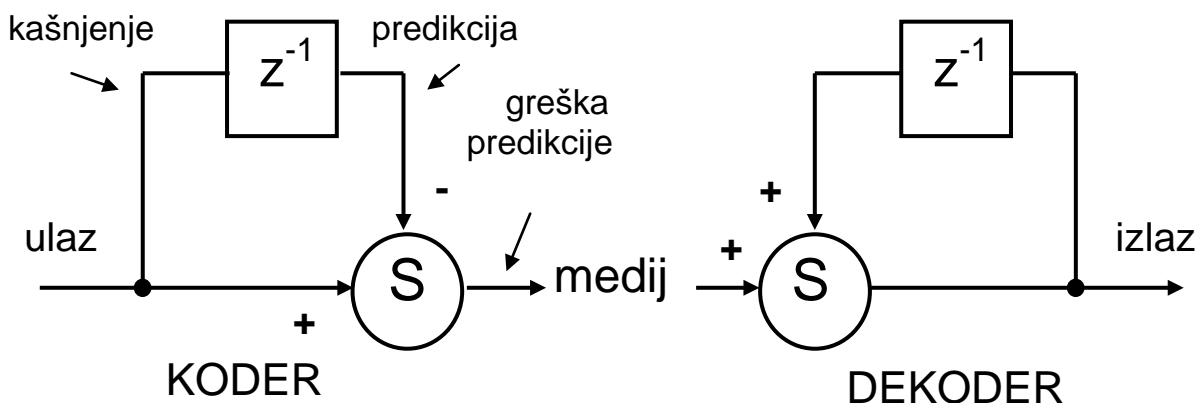
Zbog složenosti uobičajenih valnih oblika (npr. glazba) prozor transformacije ne smije biti predug, pa se zato režu u kraće valne oblike, odnosno odgovarajuće prozoriraju.

Svaki se ulazni uzorak transformira dva puta, ali s vrednovanjem ovisnim o njegovom položaju na vremenskoj osi.

Time se generira dvostruki broj podataka, što može dovesti do aliasinga u frekvencijskoj domeni. Ako je aliasing produkt u drugoj polovici prozora jednak po amplitudi, ali različit u polaritetu aliasing produktu u prvoj polovici sljedećeg prozora dolazi do poništavanja aliasinga pri rekonstrukciji. To je tzv. "time domain aliasing cancellation" (TDAC), koja se postiže lakom promjenom transformacije.

Transformacija je obično takve rezolucije da može direktno koristiti psihoakustički model.

METODA PREDIKCIJE



Medijem se prenose razlike između prethodne i sadašnje vrijednosti. Budući da su razlike manje od originalnih vrijednosti, prenose se s manje bitova.

Dekoder zbrajanje prethodne vrijednosti i greške predikcije koristi za rekonstrukciju tekuće vrijednosti uzorka. *Prediktor* je digitalni diferencijator s prijenosnom funkcijom $(1-z^{-1})$.

Prediktor n-tog reda imao bi prijenosnu funkciju $(1-z^{-1})^n$.

Vrijednosti greške predikcije prenose se u blokovima i nastoji se koristiti minimalan mogući broj bitova za svaki blok. Budući da u audio-signalu mogu postojati velike vršne razlike šteta je zbog možda samo nekoliko većih vrijednosti sve vrijednosti u bloku kodirati većim brojem bitova. To se rješava *Huffman*-ovim kodiranjem, kojim se prenose varijabilne dužine riječi. Česte vrijednosti reprezentirane su kraćim kodnim riječima, a dužima one koje se pojavljuju rjeđe.

MPEG-1

Kodiranje pokretnih slika i tona za digitalne medije do oko 1.5Mbit/s.

Ovaj standard omogućava jednokanalno (mono) ili dvokanalno (stereo) kodiranje uz frekvencije uzorkovanja 32, 44.1 i 48 kHz.

Slojevi (layers) prikazuju složenost kodnih algoritama (layer 3 ima najveću kompleksnost).

Slojevi su unatrag kompatibilni.

MPEG-1 standard sastoji se iz sljedećih dijelova:

1. dio se odnosi na problem kombiniranja jednog ili više audio i video tokova u jedinstveni tok podataka.

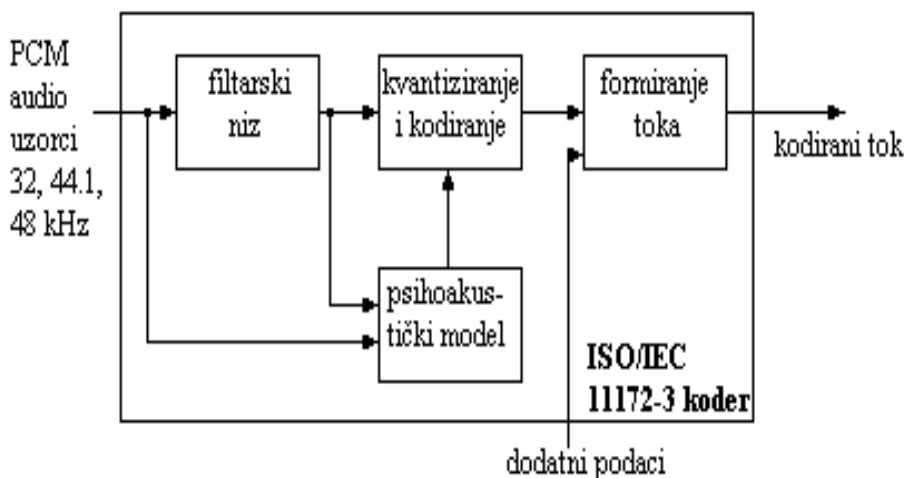
2. dio određuje kodiranje video sekvenci.

3. dio određuje kodiranje audio sekvenci, i mono i stereo.

4. dio MPEG-1 standarda određuje kako se mogu napraviti testovi za provjeru da li tokovi podataka i dekoderi ispunjavaju zahtjeve u dijelovima 1, 2 i 3 standarda.

5. dio MPEG-1 standarda predstavlja programski kod ostalih dijelova standarda.

MPEG audio koderi za generički audio, npr. sve vrste glazbe i govora, su percepcijski audio koderi, više nego koderi valnog oblika. U percepcijskim koderima, koder/dekoder ne nastoji zadržati valni oblik jednak ulaznom signalu, nego mu je zadaća osigurati da izlazni signal zvuči *ljudskom uhu identično*. Osnovni psihoakustički efekt koji koristi percepcijski audio koder je maskiranje.



Slika: Osnovni koder

Ulazni izvorni audio uzorci se dovode u koder. Na izlazu filteranskog niza je filtrirani i poduzorkovani prikaz ulaznog audio toka.

Psihoakustički model analizira ulazne signale u slijednim vremenskim blokovima i za svaki blok određuje spektralne komponente ulaznog signala, modelira karakteristiku maskiranja ljudskog auditornog sustava i određuje upravo zamjetljivu razinu šuma, odnosno prag maskiranja.

Psihoakustički model stvara niz podataka za kontrolu bloka kvantiziranja i kodiranja. U tom bloku koder pokušava alocirati pristigle bitove na način da se zadovolje potrebe maskiranja i dozvoljene brzine podataka.

Informacija o tome kako su bitovi distribuirani preko spektra se prenosi u toku podataka kao popratna informacija. Ovaj blok proizvodi niz kodiranih simbola.

Blok "formiranje toka" sklapa tok informacije iz izlaznih podataka drugih blokova, i dodaje ostalu informaciju ako je potrebno.

Dekoder je mnogo jednostavniji jer ne sadrži psihoakustički model i postupak alociranja bitova. Jedina zadaća mu je rekonstrukcija audio signala iz kodiranih spektralnih komponentata i popratne informacije.

REDUKCIJA PODATAKA RAZINE 1 (LAYER 1) PASC

MPEG-1 layer 1 poznat je i pod komercijalnim imenom PASC.

U uporabi je kod DCC. To je pojasno kodiranje prema psiho-akustičkom modelu dinamičkog maskiranja i uz adaptivnu alokaciju bitova. Dužina mantise koja može biti između 2 i 15 bitova se prilagođuje proporcionalno razmaku vršne energije pragu maskiranja. Tok podataka je 384 kbit/s podijeljen na 32 pojasa iste širine (u ovisnosti o frekvenciji uzorkovanja). U svakom podpojasu je smješteno 12 uzoraka u jednom bloku. Širina podpojasa je ovisna o frekvenciji uzorkovanja: 500Hz za 32kHz, 689 Hz za 44.1kHz, 750 Hz za 48kHz. Svaki blok traje 8 ms pri f.u. 48kHz. Dinamičko područje je $64 \times 2\text{dB} = 128\text{ dB}$

REDUKCIJA PODATAKA RAZINE 2 (LAYER 2) MUSICAM

MPEG-1 layer 2 poznat je pod komercijalnim nazivom MUSICAM.

Ulazni signal se dijeli na 32 podpojasa širine 750 Hz. Na niskim frek. je ta širina prevelika pa unosi pogrešku. Zato se to područje dodatno obrađuje pomoću FFT-a, koja radi u vremenskom prozoru sa 1024 uzoraka. Svaki se slijed od 12 uzoraka poveže u blok koji traje 8 ms. Maksimalna vrijednost bloka daje skalarni faktor (eksponent) i kvantizira se sa 6 bitova, što daje dinamičko područje od 120 dB. Tok podataka je $2 \times 96\text{ kbit/s}$. Koder može raditi sa: 64, 96, 128, 196 kbit/s za mono kanal. Frek. uzorkovanja je 32, 44.1, 48 kHz.

REDUKCIJA PODATAKA RAZINE 3 (LAYER 3)

Komercijalni naziv MP3. Pri ovoj redukciji podataka upotrebljava se hibridni polifazni/ MDCT filteranski niz zbog kompatibilost sa layer 2. 32 kanala iz polifaznog niza dovode se na 18 kanalni MDCT filteranski niz što daje $32 \times 18 = 576$ koeficijenata (spektralnih linija). Koriste se vremenski prozori različite duljine radi neutralizacije predjeke. Omogućuje brzine toka podataka od 8 kbit/s, za telefonski kanal, pa sve do 112-128 kbit/s za približno CD audio kvalitetu. Omogućuje redukciju u omjeru 1:12 i 1:14 uz brzinu od 112 do 128 kbit/s. Layer 3 (MP3) je najčešće upotrebljavan tip redukcije podataka iz standarda MPEG1.

Huffmanovo kodiranje smanjuje blok podataka za oko 25%. Dinamička dodjela bitova naročito je efikasna kod malih tokova podataka. MP3 jedini podržava MS – stereo kodiranje i intezitetno kodiranje. Za smanjenje toka podataka koristi stereofonsku redukciju. Za kvantiziranje i kodiranje

koristi sustav dvije povezane iteracijske petlje. Kvantiziranje se obavlja po eksponencijalnom zakonu (manje vrijednosti se automatski kodiraju s više točnosti). Kvantizirane vrijednosti se kodiraju Hoffmanovim kodiranjem koje je kao metoda uklanjanja redundancije, kodiranje bez gubitaka. Postupak određivanja optimalnog pojačanja i skalarnog faktora za dani blok, brzinu podataka i izlaz iz psihoakustičnog modela odvija se u dvije petlje :

1.) **unutarnja petlja** (petlja brzine toka podataka). Koriste se tablice Huffmanovog kodiranja. Ako broj bitova nakon kodiranja premašuje raspoloživi broj bitova onda se smanji ukupno pojačanje, tako da se dobije veća duljina kvantizacijskog koraka, dakle manjih kvantiziranih vrijednosti. Moguće je i obratno. Ova operacija se ponavlja dok se ne postigne zadani tok bitova. 2.) **vanjska petlja** (petlja kontrole izobličenja/šum). Treba odrediti šum kvantizacije prema pragu maskiranja za svaki frek. pojas. To se radi tako da se određuju skalarni faktori koji imaju početno stanje 1.0 u svakom pojasu. Ako je u nekom pojasu šum kvantizacije veći od praga maskiranja, skalarni faktor za taj pojas se namješta tako da se smanji kvantizacijski šum (manji šum kvantizacije zahtijeva veći broj kvantizacijskih koraka, pa se unutarnja petlja brzine mora napraviti ponovo i namjestiti zadanu brzinu). Prva petlja je smještena u drugoj petlji koja se ponavlja sve dok aktualni šum (izračunat iz razlike originalnih i kvantiziranih vrijednosti) ne dođe ispod praga maskiranja za svaki pojas skalarnih faktora.

MPEG – 2

Generičko kodiranje pokretnih slika i tona do 256 kbit/s.

Donosi *unatrag-kompatibilno višekanalno* proširenje MPEG-1. Može se kodirati do 5 glavnih kanala plus još jedan kanal za niskofrekventne efekte (LFE, Low-Frequency Enhancement).

Raspon ukupne brzine kodiranih podataka je proširen do 1 Mbit/s. Omogućene su niže frekvencije uzorkovanja: 16 kHz, 22.05 kHz i 24 kHz.

Brzine podataka za jednokanalni signal su: 32 kbit/s do 256 kbit/s za Layer I, te 8 kbit/s do 160 kbit/s za Layer II i Layer III.

Backward compatibility se postiže na osnovi matrice

$$L0=LN+xC+yLS, D0=DN+xC+yDS$$

gdje je $x=y\sqrt{2}$ (tipično), c =centralni kanal, $L0$ i $D0$ su MPEG -1 lijevi i desni kanal, N =naprijed, S =straga.

Sistemske dio se odnosi na kombiniranje jednog ili više tokova tonskih, video ili drugih podataka u jedan tok ili višestruke tokove za prijenos ili pohranu. Definirana su dva oblika toka: programski tok i transportni tok.

Programski tok je jednak MPEG-1 sistemskom multipleksu. Dobiva se kombinacijom u jedan tok jednog ili više osnovnih tokova (PES, packetized elementary stream), koji imaju *zajedničku vremensku bazu*. Programski tok se može koristiti samo u okolini koja ne unosi greške, pa je dobar za programske aplikacije. Paketi programskog toka mogu biti različite i relativno velike duljine.

Transportni tok kombinira jedan ili više PES tokova koji imaju *različite vremenske baze* u jedan tok. Elementarni tokovi mogu imati i zajedničke vremenske baze i tada tvore program. Transportni tok se koristi u okolini koja unosi pogreške, kao što je transport preko medija koji ima utjecaj šuma. Paketi transportnog toka imaju fiksnu duljinu od 188 riječi.

Dio MPEG-2 standarda je **DSM-CC** (digital storage media command and control), upravljanje i kontrola digitalnim medijem pohrane. To je pobliži opis grupe protokola koji osiguravaju kontrolne funkcije i operacije specifične za upravljanje MPEG-1 i MPEG-2 tokovima podataka.

MPEG – 2 AAC (ADVANCED AUDIO CODING)

Nije unatrag kompatibilan s prethodnim algoritmima (Layer I, II i III), ali osigurava bolju i efikasniju redukciju audio signala. Frekvencije uzorkovanja su 8 kHz do 96 kHz. AAC na 96 kbit/s daje bolju ili jednaku kvalitetu nego MPEG Layer III na 128 kbit/s.

AAC koristi sljedeće alate:

1) Filtarski niz u vremenskoj domeni uzima blokove različite duljine uz mogućnost odabira oblika prozora. Koristi se MDCT.

2) TNS je vremenska obrada šuma (Temporal Noise Shaping) koja omogućava poboljšanu kontrolu vremenskog položaja kvantizacijskog šuma unutar bloka.

TNS oblikuje raspodjelu kvantizacijskog šuma po vremenu obradom u frekvencijskoj domeni. Najveće je poboljšanje za govorne signale.

3) Intenzitetni stereo osniva se na percepcijskoj analizi tonskog dijela viših frekvencija, osobito na obliku vremenske ovojnice visokofrekvencijskog dijela spektra. Intenzitetno stereo kodiranje omogućuje da stereokanalni par iskoristi redundancije između kanala za visokofrekvencijske dijelove, uz vrlo male gubitke kvalitete. Za svaki kanal zadržava se unikatna ovojnica pomoću koje se dekodiranjem dobiva originalna razina.

4) M/S (Mid/Side) stereo kodiranje je još jedno kodiranje koje se temelji na stereokanalnom paru. U ovom slučaju elementi kanalnog para se blok po blok analiziraju kao lijevi/desni i dobiva se suma/razlika kanala, iz formule $M=(L+D)/2$, $S=(L-D)/2$.

M je monofonska komponenta, a S stereo komponenta stereo signala.

5) Predviđanje (Prediction) je tehnika koja opisuje stacionarne ili polustacionarne signale u vremenskoj domeni da bi se tako smanjila redundancija. Uzima se ponavljajući uzorak i koristi se naredba ponavljanja. Upotrebljava se na dugim blokovima jer su kratki blokovi za nestacionarne signale koji se brzo mijenjaju. Proces predviđanja temelji se na unatrag adaptivnom modelu drugog reda u kojem se spektralne vrijednosti 2 prethodna bloka koriste za predviđanje idućeg. Predikcijski parametri se mijenjaju od bloka do bloka i ne prenose se nego se izračunavaju.

6) Kvantizacija i kodiranje podudara se s iteracijskim petljama u MPEG I Layer III. Spektralni podaci se kodiraju prema *psihoakustičkom* modelu, koji određuje frekvencijsku raspodjelu kvantizacijskog šuma. Koristi se *Huffmanovo* kodiranje.

Koristi se tehnika alokacije bitova, kao kod MPEG Layer III, što omogućava kombiniranje bitova po vremenu.

“Bešumno” kodiranje se koristi zajedno s Huffmanovim kodiranjem u cilju dodatnog smanjenja redundancije optimizacijom Huffmanovih kodnih tablica.

7) Formiranje toka podataka je vrlo fleksibilno, tok podataka se sastoji iz dva dijela: transporta i bloka. Blok se sastoji od 5 dijelova: programske konfiguracije, audio elementa, elemenata stereokanalnog para, elemenata popunjavanja i terminatora (kraja bloka).

PROFILI U MPEG – 2 AAC

1) GLAVNI PROFIL – osigurava najveću kvalitetu, koristi sve alate, može dekodirati tok podataka LC profila.

2) PROFIL MALE SLOŽENOSTI (LC) – koristi se kada su procesorska brzina i memorija ograničene.

3) PROFIL SKALABILNE FREKVENCIJE UZORKOVANJA (SSR) – koristi posebne alate za predprocesiranje.

PRIMJENA MPEG – 2 AAC

Radio difuzija i distribucija audio materijala preko Interneta, te posebno kao algoritam unutar MPEG-4 standarda.

MPEG – 4 KODIRANJE

MPEG-4 uvodi univerzalno, efikasno kodiranje različitih oblika audiovizualne informacije odnosno audiovizualnih objekata. Ti objekti mogu biti prirodnog ili sintetičkog porijekla. MPEG-4 standard omogućuje:

- novi način interaktivnosti s dinamičkim objektima (prije nego sa statičkim);
- integracija prirodnog i sintetičkog tona i videa;
- kodni opis koji može uzeti u obzir niže slojeve, dok programer aplikacije ne mora voditi brigu o njima;
- simultano korištenje materijala koji dolaze iz različitih izvora, i mogućnost slanja materijala na različite lokacije;
- integracija žive informacije i pohranjene informacije u istoj prezentaciji.

MPEG-4 podržava niz audio objekata u kojima su sadržani različiti audio signali i definira im pripadajuće algoritme kodiranja: **Govorni signali** – CELP ili HVXC imaju brzine toka podataka od 1,2 kbit/s do 24kbit/s po kanalu. Frekvencije uzorkovanja su 8kHz za uskopojasni i 16 kHz za širokopojasni govor. **Glazba** – za kodiranje glazbe pri vrlo niskim brzinama (veće od 4 kbit/s) koristi se HILN parametarski koder. **Opći audio signali** – uključuju govor i glazbu. Brzine toka podataka su iznad 6 kbit/s, a frek uzorkovanja je od 7.35 do 96 kHz. Koriste se generalni audio koderi (Time/Frequency koderi). **Sintetički audio signali** – podrška za sintetički zvuk je u obliku «strukturiranog audio dekodera» koji daje informaciju o načinu izvođenja glazbe zapisane u SAOL jeziku. **Sintezirani govor** – standardizirano je sučelje TTSI za text to speech (TTS) kodiranje govora koje omogućuje prijenos teksta i prozoričkih parametara pri brzinama od 0.2 do 1.2 kbit/s.

Za kodiranje prirodnog zvuka koriste se algoritmi: **paramteraski koderi** (govor i audio), **CELP**(coded excited linear predictive) koder govora, **T/F** generalni audio koder.

Za kodiranje sintetičkog zvuka: strukturirani audio (glazba), sučelje TTSI (govor).

Parametarski koderi. Ovi koderi ulazni signal analiziraju na osnovi poboljšanog izvornog modela i prenose ga u obliku niza parametara koji ga određuju, koji se zatim na prijamoj strani sastavljaju. U MPEG-4 koriste se sljedeći algoritmi za parametarsko kodiranje: HVXC (harmonic vector excitation coding) za govor, HILN (harmonic and individual lines plus noise) za glazbu.

HVXC, HILN, CELP KODER, T/F KODER, MPEG-4 AAC LD, TTS sinteza, sinteza glazbe.

PARAMETARSKI KODERI

HVXC podržava promjenjivu brzinu prijenosa (variable bit rate) u smislu da omogućava detektiranje intervala pozadinskog šuma za vrijeme kojeg nema signala, a prenose se samo parametri za generiranje tog šuma, što smanjuje tok podataka. Osnovne karakteristike HVXC kodiranja su:

- frekvencija uzorkovanja: 8 kHz;
- frekvencijski opseg govora: 100 – 3800 Hz;
- fiksna brzina toka podataka: 2 kbit/s i 4 kbit/s;
- srednja brzina varijabilnog toka podataka: 1.5 kbit/s i 3 kbit/s;
- algoritamsko kašnjenje: < 40 ms;
- duljina okvira: 20 ms;
- skalabilnost: 2 + 2 kbit/s.

HILN je parametarski koder glazbenog signala koji se koristi pri brzinama prijenosa od 4 kbit/s naviše. Tipične brzine su od 6 do 16 kbit/s. Koder se sastoji iz dijela *izdvajanja parametara* (analize) i dijela *kodiranja parametara* (prema psihoakustičkom modelu), koji se izvode za svaki okvir ulaznog signala.

Određuju se 3 grupe parametara, po jedna za svaki dio signala:

- harmonijski ton – određen svojom osnovnom frekvencijom, amplitudom i spektralnom ovojnicom pripadnih harmoničkih komponenata (LPC spektar);

- individualne linije – određene svojom frekvencijom i amplitudom;
- šum – određen svojom energijom i spektralnom ovojnicom.

Određivanje parametarskog koda za neki audio signal ovisno o vrsti signala (govor ili glazba) može biti automatsko, pri čemu se koristi alat za klasifikaciju koji na osnovi ponašanja parametara "jakosti i visine tona" i "energije u okviru" određuje vrstu audio signala. Kod govora su veće vremenske promjene tih parametara. Može se raditi samo u HVXC modu, samo u HILN modu, ili u kombiniranom (miješanom) modu.

Skalabilnost toka podataka moguća je korištenjem jednog osnovnog toka i dodatnih tokova. U osnovnom toku nalaze se percepcijski najvažnije komponente, a to su najvažnije harmonijske linije, individualne linije i šum. U dodatnim tokovima mogu biti dodatne individualne ili harmonijske linije ili dodatni parametri šuma.

Osnovne karakteristike HILN kodiranja su:

- frekvencija uzorkovanja: 8 kHz i 16 kHz;
- brzina toka podataka: > 4 kbit/s;
- duljina okvira: najčešće oko 32 ms.

CELP KODER

CELP (Coded Excited Linear Predictive) je hibridni koder govora koji u MPEG-4 inačici donosi neke proširene funkcionalnosti, kao što su višestruke brzine toka podataka, skalabilnost toka, skalabilnost frekvencijskog pojasa i skalabilnost složenosti.

Skalabilnost MPEG-4 CELP koda važna je za uređaje sa ograničenim procesorskim mogućnostima, kao što su mobilni telefoni. CELP koristi alat za kodiranje tišine, koji smanjuje tok podataka za vrijeme tišine, i kada se CELP koristi u modu promjenjive brzine prijenosa. VAD (Voice Activity Detector) algoritam omogućava razlikovanje glasa i tišine. Za vrijeme tišine prenosi se samo SID opisnik (Silence Insertion Descriptor), koji upravlja generatorom šuma u dekodiranju.

T/F KODER

T/F (Time/Frequency transformation) je koder za generalni audio signal, koji obuhvaća govor i glazbu.. Koristi se pri brzinama podataka višim od 6 kbit/s po kanalu, a kvaliteta je najbolja iznad 16 kbit/s po kanalu. Podržane su frekvencije uzorkovanja od 7.35 do 96 kHz.

T/F je percepcijski transformacijski koder. Zasniva se na MPEG-2 AAC algoritmu kodiranja, koji se proširuje s novim tehnologijama i alatima.

MPEG – 4 AAC LD (LOW DELAY)

Je verzija MPEG-4 AAC koda koja ima maksimalno algoritamsko kašnjenje od samo 20 ms, što omogućava dvosmjernu komunikaciju u realnom vremenu. Podržava razne tokove podataka (24–96 kbit/s).

Zbog moguće pred-jeke i kašnjenja AAC LD ima 2 puta manju duljinu bloka od AAC, i ne koristi promjenu duljine blokova, nego umjesto toga koristi promjenu oblika prozora. Kašnjenja još unose ulazni filtri i oblikovanje u okvire, što su standardni uzroci kašnjenja percepcijskih koda.

Skalabilno audio kodiranje u MPEG-4 omogućava stvaranje skalabilnog toka podataka, koji se može prilagoditi različitim primjenama ili potrebama prijenosa (Internet, mobilni uređaji, pretraživanje, itd.). Tok se sastoji od osnovnog toka ili prvog sloja, koji daje osnovni koder, te dodatnih tokova ili slojeva proširenja, od idućih kodnih modula. Na ovaj način, moguć je prijenos ili dekodiranje samo dijela signala uz određeno smanjenje kvalitete. Skalabilnim kodiranjem postižu se uštede u odnosu na "simulcast" kodiranje, gdje se prenose cijeli tokovi za svako novo proširenje.

TTS SINTEZA

TTS (Text-to-Speech) dekoderi generiraju govor prema tekstu, uz kojeg mogu biti neki prozodički parametri (visina, duljina i amplituda glasova) i neke kontrolne informacije za kontrolu TTS dekodera. Glasovi (fonemi) su najmanje jedinice ljudskog jezika. Koriste se međunarodni simboli za foneme, i tako je omogućeno TTS dekodiranje različitih jezika. Prozodički parametri mogu se eventualno generirati i u dekoderu razumijevanjem teksta. Brzine podataka su 200 bit/s (samo tekst) do 1.2 kbit/s (tekst sa prozodičkim parametrima i kontrolnim podacima).

MPEG-4 određuje standardizirano sučelje TTSI (Text-to-Speech Interface) za rad TTS sintezatora, ali ne i samu TTS sintezu. Na taj način omogućena je upotreba raznih već postojećih sintezatora. Izgovor teksta može se sinkronizirati s animacijom lica i usana.

SINTEZA GLAZBE

Za sintezu glazbe koristi se sintezacijski jezik **SAOL (Structured Audio Orchestra Language)**, standardiziran u MPEG-4. Pomoću ovog jezika moguće je definirati "orkestar" sastavljen od "instrumenata" (koji se također mogu prenijeti). Instrumenti su definirani nizom naredbi za procesiranje signala koje obuhvaćaju stvaranje novih, kao i upotrebu pohranjenih zvukova.

MPEG-4 ne definira jedinstvenu metodu sinteze, nego opisuje korištenje raznih metoda, npr. iz tablice valnih oblika (wavetable), FM, aditivnu, itd.

DOLBY AC 2

To je neovisan jedno kanalni sustav. Radi s brzinom podataka od 128 do 192 kbit/s. Koder može raditi sa 128 uzoraka (64 za tranzijentne signale) uz kašnjenje 7 ms, ali i sa 512 uzoraka (za sporo promjenjive signale) uz kašnjenje od 40 do 60 ms. Transformacija se bazira na TDAC, pa je preklapanje prozora 50%. Frekvencija koef. su u floating point prezentaciji. Grupiraju se u pojase usporedive po širini s kritičnim pojasima, a u svakom pojasu se određuje najveća amplituda. Koeficijent s maks. Amplitudom se pomiče u lijevo, a maks. Broj posmaka prije preljeva postaje eksponent (skalarni faktor) za taj pojas. Kombinacija eksponenata kreira spektralnu ovojnica sa 6 db rezolucijom. Zatim se uspoređuju bitovi tako da se izbjegne pojava nemaskiranog kvantizacijskog šuma. Algoritam za upravljanje raspodjelom bitova upravljan je procjenom spektralne ovojnice signala. Za korekciju pogreške upotrebljava se RSC kod.

U dekoderu se nakon demultipleksiranja i korekcije pogreške rekonstruira spektralna ovojnica. U dekoderu je ugrađen identičan psihoakustički model kao i u koderu.

Na transformacijske koeficijente se primijenjuje IMDCT.

DOLBY AC 3

Procesira višekanalni zvuk kao cjelinu. Za istu brzinu toka podataka daje bolju kvalitetu od AC-2 uz nešto veću složenost. Tok podataka je između 32 kbit/s i 640 kbit/s konfiguriran u različitim kanalnim konfiguracijama. Predviđen je za 5.1, 3.2, 3.1, 3.0, 2.2, 2.1 stereo, mono.

Najpoznatiji je "Dolby Digital Surround" (5.1) s izvrsnom kvalitetom uz 64 kbit/s po kanalu.

Serijski kodirani tok sastoji se od okvira koji sadrže: sinkronizacijsku riječ (SI), 6 audio-blokova (AB), zaglavlje (BSI) koje nosi informacije o kodiranoj audio-informaciji i toku podataka, te pomoćne informacije (AUX) i CRC riječ. Svaki blok ima 256 uzoraka.

Ulazni uzorci mogu biti dugi do 24 bita. Svaki AC-3 okvir ima 1536 PCM uzoraka.

Detektor tranzijenata odlučuje o preklapanju duljine dugog (512 uzoraka) transformacijskog bloka na kratki (256 uzoraka). Vrijeme kašnjenja je oko 100 ms.

KODIRANJE BEZ GUBITAKA INFORMACIJE

Uobičajeni načini računalnih kodiranja (npr. ZIP, ARJ itd.) nisu prikladni, jer je s njima postiziv komprimirani tok podataka između 400 kbit/s i 1100 kbit/s, u prosjeku oko 900 kbit/s. Ne vode računa o psihoakustičkim svojstvima tonskog signala, pa je kodiranje po uzorku u prosjeku iznad 10 bit/s (uz 16 bitovnu izvornu riječ), dok noviji "lossless" koderi postižu srednje vrijednosti ispod 6 bit/s (EBU SQAM).

Dva su osnovna pristupa:

- kodiranje bez gubitaka metodom predikcije (lossles prediction coding)
- kodiranje bez gubitaka metodom transformacije (lossless transform coding)

Što se tiče samog kodiranja s gubicima, imamo omjer kompresije veći od 10, odstranjuje se veliki dio zvučnih informacija (dekodirani signal nije jednak originalu), standardno se koriste MPEG, AC-3, ATRAC. Dok kod kodiranja bez gubitaka ne odbacujemo nečujne dijelove signala, mali je faktor kompresije (oko 2), ali je i dekodirani signal identičan originalu.

1. Metoda predikcije

Prvi koder CSX (Stautner, 1986.) koristio je set inverznih predikcijskih filtara i birao je filtar s najmanjom predviđenom greškom za svaki blok. Broj predikcijskih filtara odabrane za blokove i izvorni signal kodiran je linearnim kodom. Signal je kodiran Huffmanovim kodom. Korištena je samo jedna kodna tablica i fiksni kod.

Prosječni tok podataka je ipak veći nego kod perceptualnog kodiranja, ali manji nego kod klasičnih metoda.

2. Transformacijsko kodiranje

Koriste se strukture slične onima kod perceptualnog kodiranja. Audio-signal se transformira na frekvencijske komponente analizom filtarskih nizova. Ti frekvencijski koeficijenti su kvantizirani i kodirani entropijskim kodiranjem.

Stopa kvantizacije je kontrolirana da se zadrži pogreška u vremenskoj domeni ispod stupnja kvantizacije PCM signala.

Optimum je kombinacija perceptualnog audio-kodiranja i kodiranje bez gubitaka. Tada se od perceptualnog kodiranja koristi vjerojatnosna statistika, analiza u frekvencijskoj domeni i iterativna analiza. Uz naprednu statističku obradu kodnih shema moguća je vjerna rekonstrukcija signala.

GREŠKE KOD MAGNETSKOG ZAPISA

Osnovna karakteristika magnetskog medija i magnetske glave je *karakteristika pojasnog propusta*. Istosmjerna i NF komponenta se ne prenosi.

Osnovne smetnje kod magnetskog zapisa:

1) GREŠKA VREMENSKE BAZE:

Pojavljuje se zbog faznog kolebanja brzine vrpce, elastičnog rastezanja vrpce, proklizavanja između tonske osovine i pritisknog kotačića i ostalih mehaničkih utjecaja transportnog mehanizma.

2) GREŠKA U RAZINI SIGNALA:

Nastaje zbog prekida očitavanja podataka (Drop-Out), smetajućih impulsa, lošeg kontakta između glave i vrpce i kosog rasporeda glave.

3) ISI (posebno pitanje).

Magnetska vrpca je izložena različitim smetnjama i to zahtijeva razvoj specijalnih kanalnih kodova za direktno snimanje digitalnog signala.

ZAPIS DIGITALNOG AUDIO SIGNALA NA MAGNETSKU VRPCU POMOĆU ROTIRAJUĆIH GLAVA

Zbog velike pojasne širine upotrebljavaju se video rekorderi koji zapisuju *indirektno*, dakle video signal koji nosi digitaliziranu audio-informaciju. Bubanj i traka se vrte u različitim smjerovima. Zbog relativno velike brzine vrpce u odnosu na glavu postignut je širok frekvencijski pojas. Zbog male apsolutne brzine vrpce postiže se relativno velika gustoća zapisa. Druga mogućnost je *direktan zapis* rotirajućim glavama R-DAT sa azimutim zapisom.

Zahvaljujući različitim azimutima glave na bubnju, prigušenje preslušavanja između tragova je vrlo veliko pa nije potreban zaštitni pojas između tragova.

R-DAT. (ROTARY HEAD DAT)

Vrpca je visoke koercitivne sile, standardne debljine $13\mu\text{m}$ širine 3.81 mm, smještena u kasetu.

Vrpca je vođena oko bubnja s dvije glave. Bubanj je promjera 30 mm, vrti se 2000 okr/min. Na vrpcu se zapisuju kosi tragovi.

Stroj je dvokanalni, s linearnom 16 bitovnom kvantizacijom. Kod vlastitog snimanja radi s $f_u=48\text{ kHz}$, kod predsnimljenog materijala se koristi 44,1 kHz, a može raditi i s 32 kHz.

Kanalni kod je "8 u 10".

Ukupni tok podataka je 2,8 Mbit/s, od toga je

- pod-kod 273,1 kbit/s,

- identifikacijski kod 68,3 kbit/s.

Redundancija je 37,5%.

Na standardnu kasetu stane uz standardnu brzinu 120 min stereo-zapisa, a u posebnom modu (12 bitovna kvantizacija i $f_u=32\text{ kHz}$) može se zapisati do 4 sata audio-materijala.

Prijenos signala na bubanj je pomoću rotacijskog transformatora.

FORMAT ZAPISA.

Na svakom kosom tragu su podaci zapisani u blokovima različite dužine. U jednom tragu je zapisano ukupno 196 blokova.

U 128 blokova korisnih podataka naizmjenice su npr. u A tragu zapisane npr. lijeve parne (64 blokova) i desne neparne (64 blokova) riječi, a u B tragu desne parne i lijeve neparne riječi. Daljnji tragovi su tada nadalje jednako raspodijeljeni (slika). Time je postignut interleaving prvog stupnja.

Ostali blokovi su podkodni, pomoćni, sinkronizacijski, ATF i odjelni blokovi. Početak i kraj traga su označeni sa 11 rubnih blokova koji čuvaju vrpcu od oštećenja. Blokovi su međusobno odvojeni razmakom u koji je zapisana f_4 . Svaki blok ima 288 izvornih bitova.

SLIJEĐENJE TRAGA R-DAT

Budući da glava A i B imaju različito zakošene raspore (azimut je $\pm 20^\circ$) potrebno je točno vođenje glave po "vlastitom" tragu. To se postiže očitavanjem odgovarajućih frekvencijskih uzoraka koji su individualno pridijeljeni glavama. Ti uzorci se nalaze smješteni na svakom tragu u 2×5 blokova pod nazivom "area-divided track following" (ATF, katkad i "automatic track finding").

Zapisani su sljedeći frekvencijski uzorci:

$f_1 \gg \gg \gg$ pilot: $130,67\text{ kHz} = f_{kan}/72$

$f_2 \gg \gg \gg$ glava A (+ azimut): $522,67\text{ kHz} = f_{kan}/18$

$f_3 \gg \gg \gg$ glava B (- azimut): $784,00\text{ kHz} = f_{kan}/12$

$f_4 \gg \gg \gg$ brisanje: $1,568\text{ MHz} = f_{kan}/6$

$f_{kan} = 9,408\text{ MHz}$

Zbog veće dužine raspore od širine traga glava čita i dio susjednog traga. Relativno niska f_1 će se moći pročitati bez obzira na azimut pa će svaka glava "znati" da se nalazi na ATF bloku. Glava će se tako namještati (promjenom brzine vrpce i broja okretaja bubnja) da čita "svoju" (f_2 ili f_3) frekvenciju maksimalne amplitude, a "tuđu" (f_3 ili f_2) minimalne amplitude. f_4 je opća međublokovska frekvencija i služi za brisanje zaostalih podataka.

DCC

Zbog kompatibilnosti analogne kompakt kasete i velikog broja uređaja razvijen je uređaj koji može snimati i reproducirati digitalni zapis, ali i klasičnu kompakt kasetu. Stacionarni zapis uz standardnu brzinu vrpce od 4,76 cm/s omogućen je redukcijom toka podataka pomoću PASC (MPEG 1 Layer 1). Upotrebljava se kombinirana glava koja može zapisati: 2 tonska kanala pomoću 9 digitalnih tragova, reproducirati 2 tonska kanala pomoću 9 digitalnih tragova i reproducirati 2 analogna kanala. Izvorno kvantiziranje 16 bit linearno. Kanalni kod je «8 u 10». F.u. je 32, 44.1, 48 kHz. Podatci smješteni u okvire duljine 104 448 bita. Jedan okvir obuhvaća 32 bloka na svakom od 8 tragova. Svaki blok ima slijedeće 8 bitovne riječi: 1 sinkronizacijsku, 2 adresne, 32 audio PASC, 16 paritetnih i sistemskih. Zaštita od pogreške se bazira na dvostrukom RSC-u koji djeluje na prvi interleaving od 3 okvira i drugi od 48 okvira. Redundacija iznosi oko 40 %.

NAČIN KODIRANJA ATRAC

U uporabi kod MINI DISCA. Sastoji se od tri glavna bloka u kojem se obavlja: 1. *vremensko frekvencijska analiza*; 2. *Spektralna kvantizacija*; 3. *Raspodjela bitova*. Analizirajući blok rastavlja ulazni audio signal na spektralne komponente grupirane u blok floating units (BFU). Stupanj za raspoređivanje dijeli raspoložive bitove između BFU blokova. U kvantizacijskom stupnju se svaki spektralni koeficijent odgovarajuće kvantizira na specificiranu duljinu riječi.

1) **Vremensko-frekvencijska analiza.** Ulazni digitalni signal se na početku rastavlja na tri pojasa pomoću QMF filtera i to: VF: 11025 -22050 Hz, SF:5502,5 Hz – 11025Hz, NF: 0-5502.5 Hz. Svaki od tri podpojasa transformira se u frek. domenu koristeći MDCT, čime je poboljšana vremenska rezolucija. Duljina transformacijskog bloka je adaptivna: dugi mod osigurava dobru frek. rezoluciju, ali se koristi kratki mod radi eliminacije predjeke kod tranzijentnih signala. ATRAC koristi 52 umjesto 24 kritična pojasa (nejednolika frek. podjela). 2) **Spektralna kvantizacija.** Spektralne vrijednosti se kvantiziraju koristeći dva parametara: skalarni faktor i duljina riječi. Skalarni faktor određen je amplitudom najvećeg uzorka u bloku i definira cijelo dinamičko područje kvantizacije za dotični blok. Duljina riječi određena je algoritmom za raspodjelu bitova i ona izravno utječe na preciznost područja kvantizacije. Za svaki okvir od 512 spektralnih vrijednosti zapisuju se: mod (dugi ili kratki), duljina riječi za svaki BFU, kod skalnog faktora za svaki BFU, kvantizirani spektralni koeficijenti. 3) **Raspodjela bitova.** Algoritam za raspodjelu bitova dijeli raspoloživi tok bitova između različitih blokova. Blokovi s velikim brojem bitova imati će nisku razinu šuma, a blokovi s malim broja imati će visoku razinu šuma. Algoritam mora osigurati optimalnu raspodjelu bitova s obzirom na šum i raspoloživost kapaciteta. ATRAC ne propisuje algoritam za raspodjelu bitova. Na disk se zapisuju svi potrebni parametri pa i o proizvođaču koda ovisi postignuta kvaliteta.

KOMPAKTNI DISK

Informacija u obliku izbočina («bump», sa strane čitanja), odnosno udubina («pit», s lakirane strane) na referentnoj površini («land») je smještena na spiralnom tragu koji se beskontaktno očitava. Konstantna brzina očitavanja je 1.2 m/s čime mikroprocesor dobiva konstantan tok podataka od 4.3 Mbit/s.

OPTIČKI SUSTAV

Zapis velikom gustoćom podataka zahtijeva malu valnu duljinu. Zato se za očitavanje podataka koristi laserski snop valne duljine 780 nm.

Zahtjev za što manjom osjetljivosti na oštećenja površine traži što manju dubinsku oštrinu.

Uzlazna ili silazna stijenka bumpa predstavlja promjenu stanja, dakle binarnu 1.

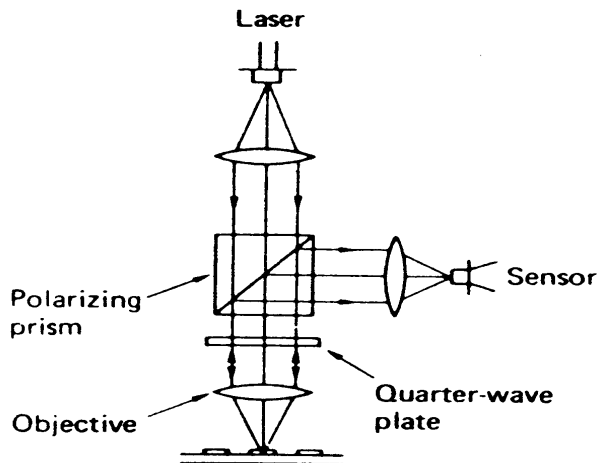
Međuprostor između 2 promjene stanja je binarna 0.

Propis kodiranja EFM koda traži najmanje 2 nule između 2 jedinice i dozvoljava najviše 10 nula između dviju promjena stanja.

Budući da je indeks loma transparentnog polikarbonata 1,46, valna duljina se smanjuje sa 780 nm na oko 500 nm. Tada je visina bumpa otprilike $\lambda/4$. Od bumpa reflektirana svjetlost laserskog snopa je time pomaknuta za $\lambda/2$, što odgovara 180° , te se zbog interferencije poništavaju upadna i reflektirana zraka. Budući da je promjer laserskog snopa veći od širine bumpa, jedan dio svjetlosti se uvijek reflektira od okolne informacijske površine pa dolazi samo do intenzitetne modulacije reflektiranog laserskog snopa.

Reflektirani snop se vraća na fotodetektor u kojem se informacija pretvara u napon.

Optička glava je dinamički element s dvije osi: za *slijeđenje traga* i *fokusiranje* laserskog snopa na informacijsku površinu.



Slika: Princip optičke glave

Laserska dioda (Al-Ga-As, $P \approx 5$ mW, $\lambda = 780$ nm) emitira slabo koherentni i astigmatični snop, koji prolazi kroz kolimatorsku leću, polureflektirajuću prizmu, polarizacijsku rešetku i fokusirajući objektiv. Nakon refleksije od diska i prolaska kroz polarizacijsku rešetku dio snopa se na prizmi reflektira prema prijemnoj fotodiodi. Prije dolaska na fotodiodu snop prolazi kroz cilindričnu leću, kojom se korigira astigmatizam, ali koja služi (kod velikog broja tipova reproduktora) i za dobivanje informacije o *fokusu* snopa.

OBLIK OKVIRA PODATAKA NA CDu

Okvir se sastoji od ukupno 588 bitova.

Okvir je sastavljen od:

- sinkronizacijskog uzorka (27 bit), oblika: 1 0000000000 1 0000000000 10XXX
- kontrolnog simbola (17 bit),
- 12 parnih i 12 neparnih korisnih, te
- 4 Q i 4 P simbola.

Svi simboli su izvorno 8-bitovni, pretvoreni u 17-bitovne kanalne.

Informacijski sadržaj pojedinih kontrolnih kanala označen je s P Q R S T U V W.

Iskorišteni su samo P i Q kanali (koji nemaju nikakve veze s paritetnim riječima P i Q, a ove oznake su nespretno odabrane); P najavljuje start nekog dijela, a Q-kanal sadrži sve informacije koje treba mikroprocesor za upravljanje reprodukcijom.

Za zaštitu od pogreške koristi se odgovarajuće premetanje simbola (interleaving) i Reed-Solomonov kod.

CD-R (Recordable)

Zagrijavanje organske polimerne boje uzrokuje širenje substratnog sloja u reflektivni pa se stvaraju udubine i miješaju s organskom bojom. Dolazi do modulacije reflektiranog svjetla. Snaga laserske diode je pri snimanju između 4 i 8 mW i tada se postiže temperatura oko 250°C u samom sloju boje, a pri reprodukciji je snaga oko 0.5 mW. Na supstratu je urezana predsnimljena spiralna brazda. Osim zapisa korisnih podataka (650MB) prije ulaznih spiralnih tragova nalazi se područje auto-kalibracije (PCA) te područje sadržaja (PMA).

DVD

DVD je specificiran u 4 verzije : 1. jednostrani jednoslojni 4,47GB 133 min ; 2. jednostrani dvoslojni 8,5GB 241 min; 3. dvostruki jednoslojni 9,4 GB 266 min; dvostuki dvoslojni 17 GB 482 min.

Valna duljina laserske diode je 650 nm, NA=0.6 (CD... λ =780 nm, NA=0.45). Kanalni kod DVD-a je EFMPlus, 8/16 bitova. Zahtijeva veći memorijski prostor od EFM koda ali ima veću efikasnost u sinkronizaciji, bolje potiskivanje DC i NF komponenti i ne treba bitove za povezivanje. Za zaštitu od pogreške koristi dvostruki RSC kod.

MINI DISK (MD)

MINI-DISK

Disk je promjera 64 mm i smješten je u kazetu. Snimanje i reprodukcija radi se optičkom i magneto-optičkom (MO) metodom. Prilikom čitanja magneto optičkog zapisa mijenja se polarizacijska ravnina reflektiranog laserskog snopa zbog magnetskog sloja diska. U ovisnosti o polarizaciji upravlja polarizacijska prizma raspodjelu svjetla na dva fotodetektora čiji izlazni naponi su ovisni o svjetlosnom intenzitetu, pa se time dobivaju 0 i 1. Ako se čitaju podaci optičkog diska radi se samo o optičkoj intenzitetnoj modulaciji.

Magneto-optički zapis.

Bazira se na Kerrovom efektu—ugrije li se magnetski sloj na određenu temperaturu (specifična za svaki materijal, a ovdje je oko 170-220 C) moguće mu je vanjskim magnetskim poljem promijeniti polarizaciju koja nakon hlađenja ostaje. Snimanje je moguće simultanim djelovanjem lasera i magnetske glave koja je smještena nasuprot laseru sa suprotne strane diska. Magnetsko polje glave za snimanje, koje je ovisno o signalu kanalnog koda koji treba upisati djeluje na mjesto fokusiranja laserskog snopa. Ovisno o digitalnom podatku (0 ili 1) upisati će se N ili S. Okretanjem diska zagrijano mjesto napušta područje djelovanja lasera, hladi se i temperatura mu pada ispod Curijeve temperature. Tako materijal zadrži nametnutu polarizaciju. Blok podataka je isti kao kod CD-a. Kanalni kod je EFM. Medij je podijeljen na sektore i predtiskanu brazdu. Tako imamo dobro adresiranje.

Podaci se, prije pretvorbe u kanalni kod, a nakon ATRAC redukcije, grupiraju u 212 lijevih i 212 desnih 8-bitovnih riječi. Jedna grupa ima 424 riječi, a 11 takvih grupa čine dva uzasopna sektora. Svaki sektor ima sinkronizaciju i adresu. 32 sektora s korisnim podacima, 1 podkodni sektor i 3 vezna sektora čine jedan okvir (36 sektora=1 cluster).

ZVUČNE KARTICE

Postoje dva glavna načina sinteze zvuka: FM i Wavetable.

Ugrađeni su 16 bitovni D/A i A/D konverteri i programabilni generatori takta koji određuju frekvenciju uzorkovanja. Standardna je 44,1 kHz, a niže se dobivaju dijeljenjem. Pojavljuje se i 48 kHz. Većina kartica koristi jedan ili više DMA (Direct Memory Access) kanala za čitanje i pisanje digitalnih audio podataka. Kvalitetnije kartice opremljene su direktnim digitalnim ulazom i izlazom preko optičkog ili koaksijalnog SPDIF konektora.

FM-frekvencijska modulacija je bila prva tehnika za generiranje zvuka. Zasniva se na mješanju sinusnog nosioca s drugim sinusnim signalom pri čemu nastaju novi valni oblici. Broj harmonika koji u spektru imaju značajniju amplitudu ovisi o modulacijskom indeksu: $I = A_m K_f / F_m$ gdje je A_m amplituda signala, K_f konst. frekvencija devijacija, a F_m modulirajući signal. Kada je I malen osnovna frek. može ostati u spektru nezamjetna. S povećanjem I (modulacijski indeks) povećava se i broj harmonika simetrično s obje strane nosioca. To je nedostatak FM sinteze jer ne može ostvariti asimetričan spektar.

Wavetable– ne upotrebljavaju se nosioci niti modulatori nego se iz memorije na kartici pozivaju već snimljeni uzorci pravih instrumenata. Kvaliteta ovisi o kvaliteti originalne snimke, frek. uzorkovanja kod snimanja instrumenata, broju uzorka upotrijebljenih za simuliranje instrumenata i metodi kompresije za pohranu uzoraka. Redukcija uzoraka većom brzinom od originalne rezultira višom frekvencijom tona. To omogućuje simulaciju instrumenata kroz više oktava samo jednim uzorkom. Zbog kvalitete se ova metoda koristi u većini profesionalnih kartica.

Komercijalne kartice

Prednosti: niska cijena, visok omjer kvaliteta/cijena

Nedostaci: loša podrška midi, usporuje procesor

Profesionalne kartice:

Prednosti: kvalitetna podrška midi, digitalni ulazi i izlazi, obrada zvuka (DSP procesor na kartici)

Nedostaci: cijena