

Audiokompressioon

Kui küsida, kas olete kunagi kuulnud midagi, mis on edastatud või salvestatud digitaalkujul, siis enamus inimesi vastab "ei". Tihti ei teata, et suur hulk raadiojaamu kasutab oma muusika ja reklaamide salvestamiseks digitaaltehnoloogiat ja kompressiooni, kuigi lõplik heli edastamine jääb ikkagi vana tuntud FM-i hooleks.

Audiokompressioon on teinud võimalikuks raadiosaadete reaalsajas edastamise internetis, väikesemõõduliste meediapleierite loomise jms.

Järgmises tabelis on toodud mõningate standardsete kodeerimata helisignaali ja mõningad kodeeritud helisignaali bitikiirused:

signaal	bitikiirus
mono, 11.025 kHz, 8 bit	88.2 kbps
mono, 22.05 kHz, 16 bit (tavaline kõne salvestus)	352.8 kbps
CD-DA, stereo, 44.1 kHz, 16 bit	1411.2 kbps
MP3 (stereo, <i>Internet standard</i>)	128 kbps
MPEG-2 AAC (CD kvaliteet)	96 kbps
MPEG-2 AAC 5.1 kanaliga eetrikvaliteediga muusika	320 kbps

Audiokompressioon kui idee on tegelikult vägagi vana, juba 80 aastat. 1930-ndatel aastatel katsetas Bell Labs kõnehälte kahandamisega nende komponent-sagedusteks, mida nimetati *formant*'ideks. Kõne asemel edastati vaid info, milliseid *formant*'e see sisaldab. Tulemus oli aga liiga tehnilik ja praktikas seda ei rakendatud.

Sellel ajal väljatöötatud tehnikat kasutatakse aga ka tänapäeval kõnetuvastuses, pop-muusikas ja heliefektides.

Praegusel ajal räägitakse kompressioonist (*compression*) ja kodeerimisest (*coding*), sageli vaadeldakse neid kahte terminit sünonüümidenä kuid neid on võimalik ka põhimõtteliselt eristada. Kodeerimisena (*coding*) vaadeldakse sageli meetodit, kuidas informatsioon üles märgitakse. Kompressioonina (*compression*) vaadeldakse meetodeid ja protsesse, mille eesmärgiks on talletatud andmete hulga vähendamine. Sageli toimuvad mõlemad protseduurid korraga ühe koodeki töö käigus.

Kuidas see töötab?

Suurim efektiivsus kodeerimisel (*coding*) saavutatakse algoritmidega, mis kasutavad inimese kuulmissüsteemi mudelit helitugevuse ja sageduse muutuste tajumisel (*perceptual coding*).

Kõik kodeerijad kasutavad sama põhistruktuuri. Kodeerimisskeemi või kirjeldada kui "tajutava müra vormimist" (*perceptual noise shaping*). Kodeerija analüüsib helisignaali spektraalkomponente ja rakendab psühhoakustilist mudelit määramaks märgatavat mürataset. Kvantimis ja kodeerimisfaasis üritab kodeerija kasutada vaid vajalikke helifragmente.

Dekooder on palju lihtsam. Tema ainus ülesanne on heli sünteesimine kodeeritud spektraalkomponentidest.

Kompressiooni läbipaistvus

Audikoodeki läbipaistvus on, nagu helikvaliteetki, subjektiivne ja sõltub sellest kui tuttav on kuulaja võimalike moonustustega, kuulamistingimustest, bitikiirusest jne. Sellegipoolest on püütud kokku leppida, millisel juhul peaks kompressioon olema läbipaistev suuremale osale inimestest suurema osa tehnika puhul. Näiteks loetakse läbipaistvaks:

- MP3 – 192 kbps (lähtefaili parameetrid 44,1 kHz, 16 bit, stereo);
- OGG – kvaliteeditase 5 (vastab umbes 160 kbps bitikiirusele).

Audiokompressiooni koodekid ja kodeeringud

Küllaltki hea nimekiri erinevatest koodekitest leiab Wikepeediast aadressil:

http://en.wikipedia.org/wiki/List_of_codecs

Sarnane ülevaade asub ka aadressil: <http://www.answers.com/topic/list-of-codecs>

Järgnevalt vaatleme tuntumaid kasutusel olevaid kodeerimisalgoritme ja koodekeid.

Kodeeringud

PCM

PCM (*Pulse Code Modulation*) on formaat, mida kasutatakse toore (*raw*) tihendamata (*noncompressed*) audiofailide salvestamiseks. PCM formaati toetab enamus tarkvara, see on nagu *text only* formaat tekstitöötuse maailmas.

ADPCM

ADPCM (*adaptive differential pulse code modulation*) ehk adaptiivne differentsiaalne võnke koodi modulaator on PCM-le baseeruv kodeering, mis püüab ette aimata analoogsignaali väärtust järgmisel diskreetimishetkel kasutades selleks algoritmi, mis intelligentselt kohandub signaalidega, mida ta just töödeldud on.

ADPCM jälgib helisalvestuse käiku diskreetsete punktide (*sample*) kaupa ja ennustab, mis toimub järgmisena, seejärel arvutatakse välja ennustuse ja tegeliku olukorra erinevus ning salvestab selle informatsiooni vaid 4 bitise diskreetimissuurusega.

ADPCM on küll kadudega kompressioon kuid tänu mudelile, millel ta baseerub, on tulemuseks originaalilähedane heli.

μ-law kodeering

Audio kompressiooni kasutasid juba varased digitaaltelefonivõrgud. Kaugekõne telefonikompaniid lasid analoogheli läbi odava mittelineaarse (*nonlinear*) ahela, mis võimendas vaikseid helisid rohkem, kui valjusid, seejärel muundati analoogheli digitaalseks. Osaline, mittelineaarne diskreetimistehnoloogia, mis on tuntud nime all **μ-Law-255** (lugeda: *mu-law*) võimaldab telefonivõrkudel kasutada 7 bitist digitaalheli ja saavutades samas praktiliselt sama kvaliteedi, mis oleks 14 bitisel helil ilma mittelineaarse võimenduseta. Sellise kompressiooniga (1:2) heli mahuks ühele CD-le 22 tundi.

CELP ja LPC

CELP (*code excited linear prediction*) ja LPC (*linear predictive coding*) on kodeerimisalgoritmid, mida kasutatakse väikese bitikiirusega kõne kodeerimiseks. Kasutavad bitikiiruseid 2.4 ja 4.8 kbps.

Kadudeta kompressioon

Küllaltki hea kadudeta audiokompressiooni koodekrite võrdluse leiab aadressilt:
http://www.bobulous.org.uk/misc/lossless_audio_2006.html

Shorten

Shorten (SHN) on failiformaat (*.shn), mida kasutatakse CD kvaliteediga (44,1 kHz, 16 bit, stereo PCM) helifailide kadudeta kompressiooniks. Tegemist on põhimõtteliselt ZIP või RAR kompressioonile sarnase kuid spetsiaalselt helimaterjali kompressiooniks optimeeritud vorminguga.

SHN loodi Tony Robinson'i poolt Cambridge Ülikoolis 1992 – 1993. aastal ja hiljem omandas selle SoftSound Ltd. Praegu seda vormingut enam ei arendata, asemele on tulnud paremad vormingud nagu FLAC, MPEG-4 ALS, Monkey's Audio, TTA ja WavPack. Sellegipoolest on tegemist ikka veel tavalise vorminguga, kuna audiofiilide seas on kasutusel veel suur hulk kontsertide salvestusi SHN vormingus.

Shorten vormingus failid on originaaliga võrreldes 50 – 70 % suurusega, mis on tänapäevaste vormingutega võrreldes kehv tulemus.

TTA

TTA ehk *True Audio* (<http://www.true-audio.com>) on lihtne, reaalaajas töötav kadudeta audiokoodek. Koodek põhineb kohanduvatel ennustavatel (*adaptive prognostic*) filtritel.



TTA pakub:

- kuni 30% kompressiooni;
- reaalaajas kodeerimist ja dekodeerimist;
- kiirust ja vähest ressursivajadust;
- kasutatavust erinevatel arvutiplatvormidel;
- tasuta dokumentatsiooni, avatud lähtekoodi;
- riistvara tuge.

TTA pakub kadudeta kompressiooni 8, 16 ja 24-bitistele wav-failidele. Kompresseeritud faili maht moodustab 30% - 70% originaalfailist. TTA toetab ka ID3v1 ja ID3v2 info *tag*-e. Seda koodekit kasutades saab ühele DVD-R plaadile mahutada kuni 20 CD-DA jagu muusikat ning koos ID3 formaadis metaandmetega.

FLAC

FLAC (*Free Lossless Audio Codec*) pakub 30% - 50% kadudeta kompressiooni. FLAC kasutab lineaarset ennustamist (*linear prediction*) teisendades audio valimid (*audio samples*) korreleerumata numbriteks (*uncorrelated*) mis salvestatakse Golomb-Rice kodeerimisalgoritmi kasutades. Lisaks rakendatakse RLE (*run-length encoding*) kodeerimist.



FLAC koodeki eelis teiste kadudeta koodekrite ees seisneb võimaluses kasutada voogesitust (*streaming*).

FLAC toetab diskreetimissuuruseid 4 kuni 32 bitti ja diskreetimissagedusi 1 Hz kuni 65535 Hz. Töötab paljudel platvormidel: enamuse Unixitest (sealhulgas Linux, BSD, Solaris), Mac

OS X, Windows, BeOS, OS/2 ja teised. Mitmetel GNU/Linux'i distributsioonidel on FLAC eel-installeerituna olemas.

Monkey's Audio

Monkey's Audio on kadudeta formaat, helimaterjali jaoks kasutab *.ape faile ning metaandmete jaoks *.apl faile. Kuna pole tasuta ega avatud lähtekoodiga, siis ei ole ta saadaval näiteks Linux'i jaoks. Ametlikult ongi ta saadaval ainult Windowsi platvormile.



Nagu kadudeta kompressiooni puhul ikka, sõltub faili maht lähteandmetest kuid üldiselt saavutatakse pisut väiksem fail kui FLAC formaadi puhul ja palju väiksem fail võrreldes vanema Shorten formaadiga, samas toimub kodeerimine ja dekodeerimine aeglasemalt. Tüüpiline bitikiirus CD kvaliteediga heli jaoks on 600 – 700 kbps.

Aeglase otsingu (*seek*) ja vähese veakindluse tõttu ei sobi kasutamiseks kantavates seadmetes.

Tarkvara ja infot saab ametlikult kodulehelt: <http://www.monkeysaudio.com/>.

MPEG-4 ALS

MPEG-4 Audio Lossless Coding on MPEG-4 audiostandardi laiendus võimaldamaks kadudeta kompressiooni. Arendustööd alustati aastal 2002 kui *Moving Picture Experts Group* kuulutas välja konkursi kadudeta kodeerimisprotseduuride loomiseks. Tilman Liebchen, Marcus Purat ja Peter Noll Berliini Tehnikaülikooli (*Technical University Berlin*) telekommunikatsiooni instituudist (*Institute for Telecommunications*) pakkusid välja ennustava kadudeta audiokompressiooni algoritmi (*Lossless Predictive Audio Compression* ehk LPAC), mis 2003. aasta juulis kuulutati tulevase MPEG-4 ALS standardi aluseks. Tehnoloogia valmis detsembris 2005.

MPEG-4 ALS sarnaneb tööpõhimõttelt FLAC koodekile. Toetab PCM audiot diskreetimis-suurusega kuni 32 bitti, kasutada võib mistahes diskreetimissagedusi (*arbitrary sample rates*). Toetab mitme kanaliga (*multi-channel*) heli ja kuni 65536 helirida (*multi-track*). Tehnoloogia võimaldab kiiret pöördumist (*fast random access*) mistahes kodeeritud andmete osale, märgistamist (*tagging*), voogesitust (*streaming*) ja videoandmetega multipleksimist (*multiplexing*). Loomulikult sisaldab ka veaparandusmehhanisme.

Täpsemat informatsiooni leiab MPEG-4 ALS ametlikult kodulehelt: <http://www.nue.tu-berlin.de/forschung/projekte/lossless/mp4als.html>

MLP

MLP ehk *Meridian Lossless Packing* on kadudeta (*lossless*) kompressioon, mis loodud Meridian Audio Ltd poolt ning mõeldud PCM (*Pulse Code Modulation*) audio jaoks. MLP on standardne kadudeta



kompressioonimeetod DVD-Audio jaoks ning võimaldab enamuse helimaterjali jaoks kompressiooniastet 2:1. Kõik DVD mängijad on varustatud MLP dekodeerimisega kuid tema kasutamine on tootjatele valikuline. Järgnevas tabelis on MLP kasutus erinevates mediaformaadides.

HD DVD			Blu-ray			DVD-Audio			DVD-Video
Staatus	kanaleid	suurim bitikiirus	Staatus	kanaleid	suurim bitikiirus	Staatus	kanaleid	suurim bitikiirus	Staatus
Kohustuslik	2 to 8	18 Mbit/s	valikuline	2 to 8	18 Mbit/s	Kohustuslik	2 to 6	9.6 Mbit/s	pole kasutusel

MLP on tuntud ka nimetuse Packed PCM ehk PPCM all ning uue põlvkonna DVD puhul tuntakse teda nimega Dolby TrueHD, toetab kõrgemaid diskreetimissagedusi, suuremaid bitikiiruseid, 8 kanaliga heli, ulatuslikku metaandmestikku, muudetavat kõlarite paigutust (*custom speaker placements*) ja ajakoodi (*timecode*).

WavPack

WavPack on tasuta, avatud lähtekoodiga kadudeta (*lossless*) kompressiooniga formaat, mille väljatöötajaks on David Bryant. Faililaiendiks *.WV.



Ühildub praktiliselt kõigi PCM audioformaatidega, toetab diskreetimissuuruseid 8, 16, 24, 32 ja 32 *float*, diskreetimissagedusi 6 – 192 kHz (sealhulgas ka ebestandardseid sagedusi). Toetatakse ka mitme kanaliga heli ja kõrgeid diskreetimissagedusi. Nagu teistegi kadudeta koodekite puhul sõltub kompressiooniaste lähtematerjalist, vaikeseades kadudeta režiimis toimib WavPack täpselt nagu WinZip programm ning pakub tüüpilise poppmuusika puhul kompressiooni 30% – 70%.

WavPack kasutab unikaalset hübriidrežiimi (*hybrid mode*), mis pakub kõiki kadudeta kompressiooni hüvesid koos lisavõimalustega. Selles režiimis tekitatakse kaks suhteliselt väikesemahulist faili. Üks on kõrgekvaliteediline kadudega fail, mida võib ka eraldi kasutada ning teine on veaparandusfail (*correction*), mis kombinatsioonis esimesega võimaldab täielikku kadudeta taasesitust.

WavPack'i loomisel on kasutatud vaid vabalt kasutatavaid tehnoloogiaid. Kõikvõimalikke patenteeritud tehnoloogiaid on välditud, see tagab, et formaat on ka edaspidi avatud ning tasuta.

WavPack toetab erinevaid arvutiplatvorme (Windows, Linux, OS X) ning voogesitust (*streaming*). Metaandmete jaoks kasutatakse nii ID3v1 kui ka APEv2 märgiseid. Veelgi paremaks kompressiooniks kasutab asümmeetrilist režiimi.

WavPack koodeki ja muu tarkvara saab ametlikult kodulehelt: <http://www.wavpack.com/>.

Apple Lossless

Apple Lossless (tuntud ka kui *Apple Lossless Encoder* – ALE või *Apple Lossless Audio Codec* – ALAC) on Apple poolt loodud kadudeta helikompressioon. Esmakordselt tutvustati seda 28. aprillil 2004 QuickTime 6.5.1 komponendina.

Andmed salvestatakse MP4 konteinerisse faililaiendiga .m4a, mida kasutatakse Apple poolt ka AAC audio salvestamisel MP4 konteinerisse. Seejuures tuleb meeles pidada, et Apple Lossless kui kadudeta formaat ei ole üheks AAC variatsiooniks.

Erinevalt klassikalistest, selgetest kadudeta audiokompressioonidest, kasutab *Apple Lossless* sarnaselt FLAC ja Shorten kompressioonile lineaarset ennustust (*linear prediction*). Apple Lossless ei kasuta mingisugust DRM skeemi kuid konteineri arhitektuuri 'st tulenevalt võib DRM-i lisada sarnaselt teistele QuickTime konteineri failidele.

Apple sõnul saavutatakse andmemahu kokkuhoid 50%, testijate sõnul on muusikafailid mahult 40% - 60% originaalidest sõltuvalt muusika liigist.

Apple Lossless audiot mängivad loomulikult kõik iPod-id. Taasesitusega saavad hakkama ka kõik avatud lähtekoodiga teegil libavcodec põhinevad meediumipleierid, nende hulgas näiteks ka VLC media player. Mõneti üllatusena peaks Apple Lossless audio taasesitamisega

hakkama saama ka Creative Labs ZEN, dokumentatsioon väidab, et toetatakse iTunes-i loodud .m4a faile.

Kadudega kompressioon

MPEG audiokompressioon

Rahvusvaheline Standarditeorganisatsioon ISO nägi juba ammu ette video- ja audiokompressiooni meetodite tähtsust ning lõi seetõttu ekspertide grupi MPEG eesmärgiks oli luua ja standardiseerida need meetodid. MPEG on olnud väga edukas videokompressiooni vallas (standardid MPEG-1 ja MPEG-2). MPEG-1 Layer-3 on tuntud audiokompressioonistandardina (MP3).

MP3

1987. aastal alustas Fraunhofer Institut Integrierte Schaltungen EUREKA projekti "*Digital Audio Broadcasting*" raames uurimistööd audiofailide kodeerimise alal. Koostöös professor Seitzeriga Erlangeni ülikoolist loodi võimas algoritm, mis avalikustati 1991. aastal ja on nüüd tuntud standardina ISO-MPEG Audio Layer-3.

Tüüpiliselt koosneb CD kvaliteediga audiofail 16 bitistest diskreetsetest andmekogumitest, mis on salvestatud diskreetimissagedusega 44,1 kHz. See tähendab, et ühe sekundi stereoheli salvestamisel saame rohkem kui 1400 Mbit (179 kB) andmeid. Kasutades MPEG audiokompressiooni saame andmehulka vähendada kuni 12 korda, seejuures kvaliteeti kaotamata. Kompressiooniastme 24 juures on helikvaliteet ikka veel parem, kui vähendades diskreetimissagedust ja diskreetimisuurust. See on saavutatud kodeerimistehnoloogiaga, mis arvestab inimekõrva eripäradega helivõngete vastuvõtul.

Kasutades MPEG audiokompressiooni, võib saavutada kompressiooniastme:

1:4	Layer 1 (andmehulk 1 sekundi stereoheli kohta on 384 kbit)
1:6...1:8	Layer 2 (andmehulk 1 sekundi stereoheli kohta on 256...192 kbit)
1:10...1:12	Layer 3 (andmehulk 1 sekundi stereoheli kohta on 128...112 kbit)

säilitades siiski CD kvaliteedi.

Kasutades stereoeffekte ja piirates audio sagedusriba (*bandwidth*) võivad audiokompressiooni vahendid saavutada rahuldava kvaliteediga heli isegi veel väiksema andmehulgaga. MPEG Layer-3 on kõige võimsam MPEG audio kompressioonimehanismide pere liige. Etteantud kvaliteedi korral saavutab ta väikseima andmehulga.

Helikvaliteet

MPEG Layer-3 iseloomustavad järgmised suurused:

helikvaliteet	sagedusriba	resiim	andmehulk	kompressiooniaste
telefon	2.5 kHz	mono	8 kbps *	96:1
parem kui lühilaine raadio	4.5 kHz	mono	16 kbps	48:1
parem kui AM raadio	7.5 kHz	mono	32 kbps	24:1
lähedane FM raadiole	11 kHz	stereo	56...64 kbps	26...24:1
lähedane CD-le	15 kHz	stereo	96 kbps	16:1
CD	>15 kHz	stereo	112..128kbps	14..12:1

*) Fraunhofer kasutab parema tulemuse saamiseks MPEG Layer-3 ebastandardset laiendust "MPEG 2.5"

MPEG Layer-3 on oma võimet, säilitada helikvaliteet andmekompressiooni 1:12 (umbes 64 kbit/sek audiokanali kohta), suurepäraselt tõestanud kõigil rahvusvahelistel testidel. Kui tarkvara lubab kasutada piiratud sagedusriba (umbes 10 kHz), saavutatakse rahuldav kvaliteet isegi kompressiooni 1:24 puhul.

Mitmete testide tulemuste põhjal soovitatakse kvaliteetse tulemuse saamiseks kasutada:

- klassikalise muusika kodeerimisel vähemalt 192 kbps
- pop- ja rockmuusika jaoks vähemalt 160 kbps

Väiksemate bitikiiruste puhul on kvaliteedikadu juba märgatav.

mp3PRO

Senise mp3 formaadiga oli muusika kvaliteet suuremate andmemahtude juures (*bitrate*), näiteks 128 Kbps, väga hea. Kahjuks aga hakkasid madalamate mahtude (suurema kompressiooni) puhul kõrged sagedused kaduma. Bitikiiruse 64 Kbps juures ja alla selle, hakkab muusika tuhmilt kõlama. Põhjuseks on see, et nii madalate bitikiiruste juures ei suuda mp3 enam muusikat piisava detailsusega kodeerida.

Et parandada mp3 kvaliteeti madalame bitikiiruste juures loodi kodeerimiseks intensiivistatud tehnoloogia, mis säilitab ka kõrged sagedused. Seda tehnoloogiat nimetatakse "*Spectral Band Replication*" (SBR).

Kombinatsioon vanast mp3-st ja SBR tehnoloogiast genereerib tihendatud audiosignaali, millel on suur sagedusulatus ka madalate bitikiiruste korral. Uut tehnoloogiat nimetatakse mp3PRO ja see koosneb kahest osast, mp3 osa madalate sageduste jaoks ja SBR või PRO osa kõrgete sageduste jaoks.

Kuna PRO osa vajab vaid paari kilobitti sekundis, siis saadi formaat luua selline, et ta on vana mp3-ga täielikult ühilduv. See asjaolu võimaldab vanade mp3 pleieritega mängida ka mp3PRO faile, seejuures nad lihtsalt ignoreerivad uut PRO osa. Ainus nõue vanadele pleieritele on toetus lisaks 32; 44,1 ja 48 kHz diskreetimissagedustele ka 16; 22,5 ja 24 kHz diskreetimissagedustele. See nõue on täidetud kõigi tarkvaraliste pleierite puhul kuid kahjuks mitte kõigi riistvaraliste (kaasaskantavate) korral.

Esimestelt katselistelt kuulajatelt saadud tagasiside ütleb, et 64 kbps bitikiirusega mp3PRO fail tagab samasuguse helikvaliteedi nagu 96 – 128 kbps tavaline mp3. See tähendab, et sama kvaliteedi juures saab sama mahuga mäluseadmele salvestada ligi kask korda rohkem muusikat kui vana mp3 puhul, see on aga hea uudis kaasaskantavate pleierite tootjatele ning tarbijatele.

mp3PRO toetab tervet hulka bitikiiruseid. Madalamate mahtude korral on mp3PRO ideaalne formaat veebilehtedel kasutamiseks ja Interneti raadiotes. Kõrgemate bitikiiruste juures on mp3PRO ülimalt sobilik kõrge kvaliteediga helimaterjali edastamiseks ja arhiveerimiseks. 2001 aasta lõpuks ilmuvad mp3PRO tooted toetavad järgmisi bitikiiruseid (*bitrate*):

- mono: 24; 32; 40 ja 48 kbps
- stereo: 32; 40; 48; 56; 64; 80 ja 96 kbps

SBR selgitus

Seni on parimad pertseptiivsed (*perceptual*) audio koodekid saavutanud CD kvaliteedi või CD-lähedase kvaliteedi umbes 128 kbps bitikiiruse juures, mis tähendab ligikaudu 12:1 kompressiooni. Vähema andmemahu juures hakkab kvaliteet tuntavalt kehvema. Koodekid kas vähendavad helimaterjali sagedusulatust ja muudavad stereopilti (*stereo image*) või tekitavad häirivaid moonutusi (*coding artifacts*). Igal juhul muutub helikvaliteet teatud hetkel vastuvõetamatuks, 64 kbps bitikiiruse puhul võimaldab mp3 heli ribalaiust vaid umbes 10 kHz või toob sisse hulga moonutusi.

SBR (*Spectral Band Replication*) suurendab helimaterjali ribalaiust nii, et see on samal tasemel või isegi ületab tavalise FM radio heli ribalaiust (15 kHz).

SBR võib parandada ka kõne koodekite töötulemust pakkudes kompressiooniga kõne ribalaiusega isegi kuni 12 kHz. Selliseid koodekeid kasutatakse näiteks mitmekeelsete (*multilingual*)

ülekannete korral. Kuna enamuse kõne tarvis väljatöötatud koodekeid on väga piiratud ribalaiusega, siis SBR mitte ainult ei paranda helikvaliteedi vaid suurendab ka kõne arusaadavust ja kaasahaaravust (*comprehension*).

SBR on pea-asjalikult järelprotsess (*post-process*), see tähendab, et ta toimub peamiselt dekodeerimise käigus, kuid koodris (*encoder*) teostatakse pisut ka eelprotsesse (*pre-process*), et dekodeerimisprotsessi juhtida.

Tehnilisest vaatepunktist on SBR uus, suure efektiivsusega meetod kõrgete sageduste kompressiooniks audiokompressiooni algoritmide seas. Koos SBR-ga rakendades, on aluseks võetud kompressioonialgoritm vastutav ainult madalate sageduste kodeerimise eest. Kõrged sagedused genereeritakse SBR-i poolt. Selle asemel, et kõiki kõrgeid sagedusi edastada (*transmit*), rekonstrueerib SBR kõrged sagedused võttes aluseks põhikoodeki poolt edastatud madalate sageduste analüüsi. Tagamaks rekonstruktsiooni täpsust, edastatakse siiski väga väheses mahus juhendavat infot.

Rekonstrueerimine on tõhus nii harmooniliste helide (muusika) kui ka mürasarnaste (*noise-like*) helide jaoks ning tagab korraliku tulemuse nii ajalises (*time domain*) kui ka sageduslikus (*frequency domain*) mõttes. Tulemuseks on täis-ribalaiusega (*full bandwidth*) audiokompressioon väga väikeste andmemahtudega, kompressiooniaste on võrreldes originaalkoodekiga tunduvalt suurenenud.

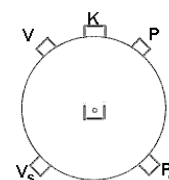
MPEG-2 AAC

Rahvusvahelise koostöö tulemusena Fraunhofer Instituudi ning firmade AT&T, SONY ja Dolby vahel on loodud siiani kõige efektiivsem MPEG audiokompressioonimeetod. Meetod kannab nime MPEG-2 Advance Audio Coding AAC ja kinnitati rahvusvaheliseks standardiks 1997. a. aprilli lõpus.



Milleks MPEG-2 AAC?

Liikumapanevaks jõuks AAC loomisel oli vajadus efektiivse *surround* helisignaali (na gu näiteks 5 kanaliga: V-vasak; P-parem;K-keskmine;V_S-vasak surround;P_S-parem surround) kodeerimismeetodi järele. Vajalikud algoritmid on MPEG-2 standardis olemas juba pikka aega, kuid tehnilistel ja ajaloolistel põhjustel pole nad saavutanud optimaalset efektiivsust. Sellepärast püstitati eesmärk saavutada vajaliku bitikiiruse (*bitrate*) tuntav vähenemine.



Mis on MPEG-2 AAC?

MPEG-2 AAC on tõeliselt eduka MPEG Audio Layer-3 (MP3) kodeerimismeetodi järeltulija. Temas on ühendatud suure efektiivsusega kodeerimine ja suur paindlikkus. Selles meetodis on arvestatud audio sektori tulevikuga (toetab diskreetimissagedusi (*sampling rate*) 8 kHz kuni 96 kHz ja kuni 48 audiokanalit). Võrreldes hästi tuntud kodeerimismeetoditega, nagu MPEG-2 Layer-2 ja Layer 3 (mp3) suudab AAC saavutada kuni poole väiksema bitikiirusega, seda ilma kadudeta subjektiivses kvaliteedis. AAC suudab kõrvaldada kuni 90% audiomaterjali mahust. Tagades mp3 failiga võrreldes parema kvaliteedi, jääb failimaht samas umbes 30% väiksemaks. ITU-R eetrikvaliteet 5.1 kanaliga heliülekandeks saavutatakse vaid 320 kbps bitikiirusega.

Kuidas MPEG-2 AAC töötab?

Nagu kõik tajutavusel põhinevad kodeerimisskeemid (*perceptual coding schemes*), kasutab MPEG-2 AAC andmete hulga vähendamisel ära inimese kuuldeaparaadi filtreerimise, maskimise (*masking*) omadusi. Kvantimismüra (*quantization noise*) jagatakse sagedusribadeks nii, et see maskeeritakse kogu signaali poolt välja, s.t. ta pole kuuldav. Kuigi

selle kodeerimismeetodi põhistruktuur erineb vaevu eelkäijate omast, esineb detailides uusi aspekte. Erinevused MPEG-2 AAC ja eelkäija MPEG Audio Layer-3 vahel on järgmised:

- Filtripank (*filter bank*): erinevalt MP3 hübriid filtripangast, mis on valitud ühilduvuse eesmärgil, kuid omab mitmeid struktuuri nõrkusi, kasutab MPEG-2 AAC tavalist modifitseeritud diskreetset koosinusteisendust MDCT (Modified Discrete Cosine Transform). Koos suurendatud ridade arvuga teisenduse kohta (2048 rida endise 1152 asemel) ületab MDCT eelmiste kodeerimismeetodite filtripankasid.
- Ajutine müravormimine TNS (*Temporal Noise Shaping*): tõeline uudis aeg/sagedus t/f (*time/frequency*) kodeerimisskeemides. TNS vormib kvantimismüra jagamise ajas vastavalt ennustusele sageduse valdkonnas. Helisignaali muutuvad tänu TNS-le detailides tunduvalt paremaks.
- Ennustamine (*prediction*): tehnika, millel põhineb kõne kodeerimine. See kasutab ära asjaolu, et teatud tüüpi helisignaali on kergesti etteaimatavad.
- Kvantimine (*quantization*): võimaldades paremat kontrolli kvantimisresolutsiooni (diskreetimissuuruse) üle, saab etteantud bitikiirust (*bitrate*) jõudsamalt ära kasutada.
- Bitivoo vorm (*bit-stream format*): edastatav informatsioon läbib entroopse kodeerimise (*entropy coding*), et hoida ebavajaliku info osakaalu võimalikult madala. Nende kodeerimismeetodite optimeerimine koos paindliku bitivoo struktuuriga on kodeerimiseefektiivsust väga palju suurendanud.

MPEG-2 AAC rakendused

Tänu oma suurele kodeerimiseefektiivsusele on AAC peamine kandidaat kõikide digitaalsete saatejaamade jaoks.

Esimesena on praktiliselt kõigis oma digitaalsaatejaamades AAC kasutusele otsustanud võtta Jaapan. AAC kodeeritud saadete edastus pidi seal plaani järgi algama juba aastal 2000, paljud elektroonikafirmad on juba alustanud spetsiaalsete MPEG-2 AAC dekodeerimise valmistamist. Oma efektiivsuse tõttu hakkab MPEG-2 AAC mängia tähtsat rolli ka muusika edastamisel Internetis.

Lisaks kõigele on MPEG-2 AAC (mõningate muudatustega) ainus kõrg-kvaliteetse audio kodeerimisskeem, mida kasutatakse ka MPEG-4 standardis.

MPEG-4 Audio

MPEG-4 Audio baseerub MPEG-2 AAC tehnoloogial ning tema peamiseks uuteks omadusteks, lisaks paranenud helikvaliteedile ja vähenenud failimahule, on:

- bitikiiruse skaleeritavus (*bitrate scalability*), võimaldab andmevoogu jaotada väiksema mahuga voogudeks, mida on siiski veel võimalik arusaadavaks signaaliks dekodeerida. Jaotamine võib toimuda nii dekodeerimise kui ka ülekande (*transmission*) ajal.
- Ribalaiuse (sagedusriba laius, *bandwidth*) skaleeritavus, bitikiiruse skaleeritavuse erijuht, mille korral ülekande või dekodeerimise ajal võib teatud sagedusi edasiandev andmevoog arvestamata jääda.
- Mono/stereo skaleeritavus, mille käigus stereosignaali jaoks vajalikud andmevood võivad arvestamata jääda ning tulemuseks on monofooniline heli.
- Kodeerimise keerukuse skaleeritavus (*encoder complexity scalability*) võimaldab erineva keerukusega enkoodritel genereerida arusaadavaid signaale.
- Dekodeerimise keerukuse skaleeritavus (*decoder complexity scalability*) võimaldab dekodeerimisel kasutada erineva keerukusega dekodeerimise.
- Veaparandusmehanismid (*error robustness tools*) võimaldavad parandatud ülekannet vea-aldide ülekandemehanismide korral. Nende vahendite hulka kuuluvad enkoodriga

vastavuses olevad "vea-elastsus tööriistad" (*error resilience*) ja üldine veakaitse vahend (*error protection tool*).

- Vähesel viivitusega audio kodeerimine (*Low Delay Audio Coding*) võimaldab kõrge kvaliteediga helisignaalide kodeerimist viivitusega vaid umbes 30 ms, mis on ülimalt oluline reaalajas ülekannete, eriti kahesuunaliste (*bi-directional*) ülekannete jaoks (MP3 ja MPEG-2 AAC korral on viivitus madala bitikiirusega signaali puhul umbes 100 ms). Tehnoloogia kannab nime AAC LD (*Advanced Audio Coding Low Delay*).

MPEG-4 Audio pakub optimaalseks kasutamiseks erinevaid profiile:

- The Speech Audio Profile kasutab parameetrilist kõne enkoodrit (*parametric speech coder*), CELP kõne enkoodrit ja *Text-To-Speech* liidest.
- The Synthesis Audio Profile pakub väikese bitikiirusega heli- ja kõnesünteesi.
- The Scalable Audio Profile on kõneprofiili supervariant, mis võimaldab skaleeritavat kõne ja muusika kodeerimist ülekandeks näiteks Internetis ja Digitaalses raadios.
- The Main Audio Profile on eelmiste kolme ühend. Sisaldab vahendeid nii naturaalsete helide kui ka sünteesitud helide kodeerimiseks.
- The High Quality Audio Profile sisaldab CELP kõne enkoodrit ja madala keerukusega (*low complexity*) AAC enkoodrit.
- The Low Delay Audio Profile sisaldab parameetrilist ja CELP kõne enkoodrit, AAC LD enkoodrit ja *Text-To-Speech* liidest.
- The Natural Audio Profile sisaldab kõiki MPEG-4 standardi all olemasolevaid enkoodreid naturaalsete helide jaoks.
- The Mobile Audio Networking Profile sisaldab skaleeritavaid ja lühikese viivitusega AAC enkoodreid. See profiil on mõeldud kasutamiseks rakendustes, kus tarvatakse kõrgkvaliteetseid non-MPEG audio kodeerimisalgoritme.

Musepack

Musepack ehk MPC on avatud lähtekoodiga kadudega koodek, mis on optimeeritud stereoheli läbipaistvaks (transparent) kompressiooniks bitikiirustega 160 – 180 kbps. Varem tunti seda koodekit kui MPEGPlus, MPEG+ või MP+.



MPC arendamist alustati 1997. aastal Andree Buschmann'i poolt, peagi jätkas Frank Klemm kes aitab praegust arendajat *Musepack Development Team* (MDT).

Koodek on saadaval MS Windows'is, Linux'is ning Mac OS X jaoks koos *plug-in*idega paljudele mängijatele. Nende hankimiseks tuleb külastada Musepack'i kodulehte <http://musepack.net/>.

Musepack loodi lähtudes MP2 koodekist, lisatud on mitmeid omadusi mis on sarnased näiteks AAC koodekile kuid neid on tunduvalt edasiarendatud. Bitikiirus on varieeritav vahemikus 3 – 1300 kbps.

Loomulikult toetatakse ka metaandmete kasutamist (*APEv2 tag metadata*).

WMA

Windows Media on Microsoft'i tehnoloogiakomplekt, mille koosseisu kuuluvad *Windows Media Player*, *Windows Media Services*, *Windows Media Tools*, *Windows Media Audio SDK* (*Software Developer Kit*) ning hulk audio ja video koodekeid. *Windows Media* komplekti kõik komponendid on tasuta saadaval Microsofti koduleheküljel.



WMA (*Windows Media Audio*) on Microsofti vastus mp3 formaadile. WMA pakub CD kvaliteediga heli vaid poole mp3 faili andmemahu juures. See annab võimaluse portatiivsete seadmete mälli salvestada 2 korda rohkem muusikat, Internetist faile 2 korda kiiremini allalaadida jne. *Windows Media* kvaliteeti on juba tunnustanud sellised juhtivad nimed muusikatööstuses nagu Sony, EMI, Warner Music Group ja BMG. Sellised tuntud firmad nagu TVT, Entertainment Boulevard, Launch, House of Blues, Liquid Audio, Intertainer ja ICast kodeerivad oma muusikat *Windows Media* formaati.

WMA ülevaade

Microsoft'i *Windows Media Audio* koodek on loodud igasuguse helimaterjali jaoks, alates kõnest, mis salvestatud 8 kHz diskreetimissagedusega kuni 48 kHz diskreetimissagedusega kõrgkvaliteetse stereomuusikani.

See formaat on väga vastupidav *packet loss*'i korral tekkivate kadude suhtes, sest ta ei kasuta *interframe* mälu. Viimane asjaolu teeb ta eriti sobivaks Internetis *streaming* audio edastamiseks.

WMA on skaleeritav ja pakub kõrgekvaliteedilist mono- ja stereoheli väga mitmesuguste ribalaiuste puhul. See võimaldab valida parima diskreetimissageduse ja ribalaiuse kombinatsiooni.

WMA on Microsofti poolt valitud alates 7. versioonist *Windows Media* vaikimisi kasutatavaks audiokoodekiks ning lisanud selle *Software Development Kit*'i.

Audio kodeerimine WMA formaati

Audiomaterjali kodeerimiseks võib hankida näiteks järgmised vahendid: *Windows Media Encoder* ja *Windows Media Recourse Kit*.

Soovitav on kasutada faili laiendit *.wma, sest see viitab järgmistele asjaoludele:

- fail on kodeeritud *Windows Media Audio* koodeki abil;
- fail sisaldab vaid audiomaterjali;
- kasutajad saavad seda faili allalaadida või mängida *stream*'ina.

Sisendfailil võib olla mistahes diskreetimissagedus vahemikus 8 kHz kuni 48 kHz, diskreetimissuurus 8 või 16 bitti ning üks või kaks kanalit (mono või stereo). Siinjuures soovitatakse parema kvaliteedi huvides 8 bitiseid faile siiski vältida ja mono sisendfailidest mitte stereofaile luua.

Windows Media Audio koodek võimaldab luua faile, mis vajavad ribalaiust 5 Kbps kuni 192 Kbps. Maksimaalse ribalaiuse korral on tulemus originaalist praktiliselt eristamatu. Sellist kõrgkvaliteetset materjali nimetatakse *CD transparency*.

Mitmete testide tulemuste põhjal soovitatakse kvaliteetse tulemuse saamiseks kasutada:

- klassikalise muusika kodeerimisel 160 kbps
- pop- ja rockmuusika jaoks 128 kbps

Väiksemate bitikiiruste puhul on kvaliteedikadu juba märgatav.

WMA koodekiga kodeeritud faile võib "pakkida" *Windows Media Packager*'i abil, mis kuulub *Microsoft Media Rights Manager*'i koosseisu, mis aitab rakendada *Digital Rights Management (DRM)* faile. "Pakitud" fail on krüpteeritud ja seda ei saa taasesitada ilma dekrüpteerimisvõtmeta, mis antakse kasutaja litsensiga kaasa. Kasutajad saavad litsensse registreerudes failide levitaja või looja juures.

ATRAC

ATRAC (*Adaptive TRansform Acoustic Coding*) on Sony loodud kõrge kompressiooniastmega (1:5) ja kõrge helikvaliteediga kadudega (*lossy*) audiokoodek, mida esmakordselt tutvustati seoses Sony MiniDisc seadmetega 1991. aastal. ATRAC3 on selle koodeki uuendus aastast 1999. Tähelepanu väärib, et RealAudio 8 on ATRAC3 rakendus.



Praeguseks on loodud juba ka ATRAC3plus koodek (aastal 2002), mida kasutatakse HiMD mängijates, Memory Stick mängijates, VAIO Pocket pihuarvutites ja ATRAC CD-mängijates.

Ta lubab luua väga väikesemõõtmelisi seadmeid ja on seetõttu väga sobiv just portatiivsete audioseadmete jaoks.

ATRAC (ehk ATRAC1) võimaldab läbipaistvat (*transparent*) kompressiooni bitikiirusel 292 kbps, see on originaalne MiniDisc standard. Sony väitel saavutatakse ATRAC3 koodekiga CD-lähedane (*near-CD*) kvaliteet bitikiirusel 64kbps ning see on ametlike väidete kohaselt võrdväärne mp3 heliga bitikiirusel 128 kbps.

ATRAC tehnoloogia kohta saab infot ametlikult kodulehelt:

<http://www.sony.net/Products/ATRAC3/>

Heli salvestamiseks ATRAC3plus formaati saab kasutada näiteks tasuta tarkvara SonicStage.

OGG Vorbis

OGG formaadi arendamine algas 1993. aastal Xiph.Org Foundation'i (<http://www.xiph.org/>) poolt, alguses kasutati koodnime "Squish" (lirts, lirtsatus). OGG on algusest peale olnud avatud lähtekoodiga projekt (*open source project*) ja seetõttu vaba igasugustest patentidest. OGG loodi mp3 ja WMA formaatide asemikuks (võistlejaks) ning on praeguseks pea-aegu sama populaarne kui mp3. Algoritmi arendatakse üha edasi, helikvaliteet muutub aina paremaks ning uuemate versioonide failid on tagasiühilduvad (*backwards compatible*), see tähendab, et uusi faile saab mängida vanade pleieritega.

Sarnaselt mp3-ga pakub OGG varieeruva bitikiirusega kompressiooni. Standardse OGG faili bitikiirus jääb CD-DA audio kompressiooniseadme puhul vastavalt valitud kvaliteediseadete vahemikku 45 – 500 kbps. Kvaliteediasetused on järgmised:

- | | | | | | |
|-----|----------|-----|----------|------|----------|
| • 0 | 64 kbps | • 4 | 128 kbps | • 8 | 256 kbps |
| • 1 | 80 kbps | • 5 | 160 kbps | • 9 | 320 kbps |
| • 2 | 96 kbps | • 6 | 192 kbps | • 10 | 500 kbps |
| • 3 | 112 kbps | • 7 | 224 kbps | | |

Paljude kasutajate arvates saavutab OGG läbipaistvuse (*transparency*) juba kvaliteediasetuse 5 puhul.

Sageli kasutatakse kvaliteediasetust 3, mis on parem kui mp3 128kbps. Kvaliteediasetuse 5 pakub CD-lähedast kvaliteeti ja on sageli kasutusel muusika jaoks. Kvaliteetsema muusika jaoks kasutatakse sageli kvaliteediasetust 6.

OGG võimaldab ka mitme kanaliga heli. Sarnaselt eelkäijatega on ka OGG kasutatav voogesitusega (*streaming*) heli edastamiseks.

OGG formaadist on olemas mitmeid erinevaid edasiarendusi (*tuned versions*) nagu näiteks: Garf, aoTuV, MegaMix ja Tremor.


OGG Vorbis formaadi ametlik kodulehekülj on: <http://www.vorbis.com/>

VQF

VQF formaat sündis aastal 2000. Formaat kannatab peamiselt kahe puuduse all: esiteks on ta praktiliselt tundmatu ja teiseks kulub PC arvutil helimaterjali kodeerimiseks VQF formaati 3 kuni 4 korda enam aega kui teiste formaatide puhul.

bitikiirusega 80 kbps pakub VQF sama kvaliteeti kui mp3 128 kbps bitikiirusega, mis tähendab, et VQF failid võtavad 30% vähem salvestusruumi. Kahjuks koormab VQF failide mängimine protsessorit 2 korda enam kui mp3 mängimine.

Dolby

 Dolby Laboratories on helimaailmas tuntud olnud juba 1960ndate lõpust, 1970-ndate algusest, kui loodi Dolby A-tüüpi müratasandi (*noise reduction*) professionaalseks kasutamiseks ja Dolby B-tüüpi müratasandi kodutarbijate jaoks. 1970ndate lõpus tekitas Dolby revolutsiooni filmimaailmas luues **Dolby Stereo** süsteemi, mis pakkus 4 kanaliga heli (kolm kanalit ees; vasak ja parem muusika ning efektide jaoks, keskmine dialoogi tarvis ja neljas *surround* kanal efektide ja atmosfääri jaoks).

1980ndate lõpus ja 1990ndate alguses tõi Dolby kodukasutajate jaoks turule Dolby Surround ja hiljem Dolby Pro Logic kodukino süsteemid, mis kasutasid Dolby Stereo tehnoloogiat.


Dolby helikodeerimisalgoritmid kuuluvad inimese kuulmistaju iseärasusi arvestatavate (*perceptual coding*) algoritmide hulka.

Praeguseks on välja töötatud suur hulk erinevaid digitaalset heli kodeerimismeetodeid, mille kohta saab infot näiteks Dolby kodulehelt:


http://www.dolby.com/consumer/technology/tech_overview.html





Lisaks on Dolby osalenud ka mitmete teiste tehnoloogiate väljatöötamisel, näiteks AAC (*Advanced Audio Coding*).

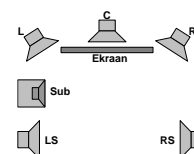
Dolby Digital

 Dolby Digital on audiokodeerimise formaat, mida esmakordselt kasutati 1992. aastal kinos, täpsemalt filmis "*Batman Returns*".

Dolby Digital on kasutusel väga erinevates konfiguratsioonides (erinevate kanalite arvuga). On olemas mono, stereo, surround (3 kanalit) ja 5.1 surround (kokku kuus kõlarit) variandid.

 5.1 surround konfiguratsioon sisaldab kolme valjuhääldid ees (ekraani ees ja kõrval), stereokõlarite paari taga ja madalasageduslikku efektikanalit (*low frequency effects channel*). Sageli kasutatakse selle kohta lühendit LCRSS (*Left, Center, Right, (left) Speaker, (right) Speaker*).

Icon	Sound mode
	Mono
	Stereo
	Dolby Surround
	5.1 surround




Dolby Digital on audiokompressioonialgoritm, mis kasutab ära vanu, juba 30 aastat vanu kogemusi, kõrvaldades helimaterjali nendest sagedusribadest, milledes esineb väga piiratud energiaga signaal. Lisaks rakendatakse veel *Noise Shaping* protseduuri, mille käigus eemaldatakse heliinformatsiooni nendest sageduspiirkondadest, kus inimene selle puudumist ei märka. Dolby Digital kasutab ka rist-korrelatsiooni (*cross-correlation*) meetodit, mis seisneb surround või stereo kanalite informatsiooni võrdlemises, nende erinevuste salvestamises ja sellega üleliigseks muutuva info kõrvaldamises. Tulemuseks on kuue


kanaliga (5.1 *surround*) audiovoog, mille maht (*bitrate*) ja nõuded salvestusruumile on väiksemad kui 44,1 kHz, 16-bit stereoheelil.


Dolby Digital kasutab sõltuvalt lähtematerjalist erinevaid bitikiiruseid ja kodeerimissüsteeme.


NB! NTSC telesüsteemiga maades on Dolby Digital, nii stereo kui ka 5.1 surround, peamiseks DVD audioformaadiks.

 Praeguseks on välja töötatud Dolby Digital Plus, mis võimaldab DVD helist paremat kvaliteeti (*Better-than-DVD*) kuna vajab vähem kompressiooni ja pakub rohkem kanaleid (kuni 7.1).


Teised Dolby kodeerimistehnoloogiad

 Dolby TrueHD on väidetavalt ainus audioformaad, mis pakub tõelist peeneraldukslikku (*true high-definition*) heli, kasutatakse kuni 7.1 kanalit, kadudeta kompressiooni. Dolby TrueHD on kasutusel HD DVD ja Blu-ray mängijates ning tuleviku Audio/Video vastuvõtjates (*receiver*) ning allalaetavas (*downloadable media*).

 Dolby Digital EX lisab tavalisele ruumilisele helile kuuenda kanali (*center back*) saavutamaks veel paremat lokaliseerimist ruumilistes heliefektides (*surround effects*). Kasutusel kodustes meelelahutuskeskustes.

 Dolby Digital Surround EX võttis 5.1 Dolby Digitali juures kasutusele tagumise keskmise (*center rear*) kanali, mis muudab ette-taha panoraamimisefektid (*pan*) selgemaks ning muudab kuulaja seljatagused helid realistlikumaks. Kasutusel filmitööstuses, kinodes.

 Dolby Digital Live on kasutusel arvutites ja mängukonsoolides, teisendab mistahes stereo või mitme kanaliga heli Dolby Digital formaati, et suurendada heli dramaatilisust, elamust.

 Dolby Surround kodeerib neli helikanalit (*left, center, right, mono surround*) vaid kaheks, et seda saaks kasutada näiteks televisioonis ja VHS lindile salvestatava filmiga. Süsteemid, mis toetavad Dolby Pro Logic tehnoloogiat, rekonstrueerivad esialgse nelja kanaliga heli. Kasutusel kodustes meelelahutusseadmetes ja kinos.

Koduste meelelahutusseadmete ja arvuti abil saavad kodukasutajad ise Dolby kodeeringuga helimaterjali luua järgmiste tehnoloogiate abil:


- Dolby Digital Recording – stereosalvestused koduste DVD – seadmetega.
- Dolby Digital Stereo Creator – Dolby Digital heli DVD loomisel. Tehnoloogiat toetab arvutitarkvara, DVD-le salvestavad kaamerad (*DVD camcorder*) jms.
- Dolby Digital 5.1 Creator – võimaldab salvestada ja luua 5.1 kanaliga Dolby Digital heliklippe. Kasutusel arvutitarkvaras, DVD-le salvestavates kaamerates, DVD-salestajates jne.



Dolby ruumilise heli avardamistehnoloogiad

 Dolby Pro Logic dekodeerib Dolby Surround heli nelja kanaliga (*left, center, right, mono surround*) heliks. Kasutusel kodustes meelelahutuskeskustes.

 Dolby Pro Logic II muudab mistahes stereosignaali viie kanaliga ruumiliseks heliks. Kasutusel näiteks Apple iPod-is, mängudes, kodustes meelelahutuskeskustes.

 Dolby Pro Logic IIx teisendab mistahes stereoheeli või 5.1 kanaliga heli kuue või

seitsme kanaliga ruumiliseks heliks. Kasutusel iPod-ides, kodustes meelelahutuskeskustes, mängudes, arvutites.

Dolby virtuaalse heli tehnoloogiad



Dolby Headphone pakub realistlikku heli mistahes kõrvaklappe kasutades. Kasutusel kodustes meelelahutuskeskustes, arvutites, mängudes.



Dolby Virtual Speaker pakub ruumilist 5.1 heli vaid kahe kõlari abil.

DTS

DTS (*Digital Theatre System*) on mitme kanaliga *surround* heliformaat, mida peamiselt kasutatakse filmidel ja DVD-del. Arendajafirma üheks asutajaks ning kaasomanikuks on filmirežissöör Steven Spielberg.



Töö selle formaadi kallal algas 1991. aastal. Esmakordselt leidis formaat kasutust 1993. aastal filmis *Jurassic Park*. See oli küll aasta pärast *Dolby Digital* formaadi esmakasutust kuid DTS-i kasutati ka sama filmi laserplaadil.

Kõige levinum versioon sellest formaadist on 5.1 ehk 5 peamist kõlarit ja *subwoofer*, mille nimetuseks on LFE (*Low Frequency Effects*) kanal. Kinokasutuses salvestatakse heli bitikiirusega 1,5 Mbit/s.

Lisaks standardsele 5.1 DTS Surround koodekile pakutakse veel järgmiseid tehnoloogiaid:

- DTS-ES, millele on lisatud tagumine keskmine kanal ehk tegemist on 6.1 süsteemiga.
- DTS NEO:6 on süsteem mis teisendab vanad Dolby Pro-Logic ja Dolby Stereo formaadis helid 5.1 või 6.1 formaati. Süsteem suudab ka 5.1 kanaliga DTS helile lisada kuuenda kanali (teisendada selle 6.1 DTS heliks).
- DTS 96/24 võimaldab 24 bitise 96 kHz DTS heli otsest taasesitust (*direct playback*) kasutades integreeritud digitaalanalooomuundureid (DAC). Süsteem on tagasiühilduv ja heli saab väiksema võimekusega helikaartidel kehvema kvaliteediga taasesitada.
- DTS Interactive on formaat, mis on loodud peamiselt mängude jaoks, kasutab reaajas taasesitust.
- DTS HD/Lossless

AMR

Adaptive Multi-Rate (AMR) on inimkõne (*speech*) kodeerimiseks optimeeritud kompressioon, mis valmis 1998. aastal ja on nüüd laialdaselt kasutusel GSM mobiilsidevõrkudes.

AMR kasutab bitikiiruseid 12,2 kbps; 10,2 kbps; 7,95 kbps; 7,4 kbps; 6,7 kbps; 5,9 kbps; 5,15 kbps ja 4,75 kbps mille vahel valitakse vastavalt ühenduse olukorrale. Info jagatakse 20 millisekundi pikkusteks osadeks (*frame*). AMR kasutab erinevaid tehnikaid nagu ACELP (*Algebraic Code Excited Linear Prediction*), DTX (*Discontinuous Transmission*), VAD (*voice activity detection*) ja CNG (*comfort noise generation*).

AMR puhul on kuuldava heli selgust (*clarity*) pisut kvaliteedi ja veakindluse tagamiseks ohverdatud.

Pärast AMR-i on loodud AMR-WB (*Adaptive Multi Rate – WideBand*), mis tänu suuremale sagedusulatussele (50 – 7000 Hz) pakub kõne kodeerimisel suurepäraselt kvaliteeti.

NB! AMR-WB on ITU-T standardiks G.722.2.

Sarnaselt AMR-ile, töötab ka AMR-WB erinevate bitikiirustega: 6,60 kbps; 8,85 kbps; 12,65 kbps; 14,25 kbps; 15,85 kbps; 18,25 kbps; 19,85 kbps; 23,05 kbps ja 23,85 kbps.

Müravaba keskkonna puhul on madalaim suurepärasest kvaliteeti võimaldav bitikiirus 12,65 kbps, suuremad bitikiirused on kasulikud taustamüra ja muusika puhul.

AMR-WB on UMTS (*Universal Mobile Telecommunications System*, üks 3G mobiilside tehnoloogia) standardiks.

AMR-WB laiendusena on välja töötatud AMR-WB+, mis lisab stereosignaali toe ja suuremad diskreetimissagedused. Ühe täiendusena on võetud kasutusele ka võimalus kasutada ACELP kodeerimist (*Algebraic Code Excited Linear Prediction*).

AMR-WB+ kasutab bitikiiruseid 5,2 kbps – 48 kbps (mono signaali puhul 5,2 – 36 kbps ja stereo puhul 6,2 – 48 kbps) ning reprodutseerib sagedusulatuse, mis läheneb CD-kvaliteedile.

AMR-WB+ loodi GSM ja 3G mobiilside süsteemide jaoks. Silmas peeti eelkõige PSS (*Packet-Switched Streaming service*), MMS (*Multimedia Messaging Service*) ja MBMS (*Multimedia Broadcast and Multicast Service*) teenuseid.

Speex

Speex on tasuta tarkvaraline kadudega (*lossy*) kõne (*speech*) koodek, mis on väidetavalt sõltumatu kõigist patentidega seotud piirangutest. Speex on litsenseeritud BSD litsensiga ja kasutatav koos OGG formaadiga või edastatud otse UDP/RTP (*User Datagram Protocol/User Datagram Protocol*) protokolle kasutades. Speex koodek kasutab OGG bitivoo (*bitstream*) formaati ning tema arendajad näevad oma projekti Vorbis üldise audiokompressiooni projekti täiendusena.



Erinevalt enamusest kõne kompressiooniks loodud koodekist pole Speex loodud mobiiltelefonidele vaid rohkem VoIP (*voice over IP*) rakendusi silmas pidades. Tema arendamisel oli eesmärgiks luua koodek, mis tagaks väga hea kõnekvaliteedi ja madala bitikiiruse ning see viis mitme bitikiirusega koodeki loomiseni. Hea kvaliteet tähendab ka laia sagedusulatuse (*wideband*, 16 kHz diskreetimissagedus) tuge lisaks tavapärasele telefonikvaliteedi kitsale sagedusulatusele (*narrowband*, 8 kHz diskreetimissagedus). VoIP rakenduste silmaspidamine tähendab ka seda, et Speex peab suutma hakkama saada pakettide kaotsimineku (*packet loss*) aga mitte vigaste pakettidega kuna UDP hoolitseb, et paketid jõuavad kohale tervetena või ei jõua üldse kohale. See on põhjuseks, miks Speex jaoks valiti CELP (*Code Excited Linear Prediction*) kodeerimisalgoritm, mis on end aja jooksul tõestanud ning saab hakkama nii madalate kui ka kõrgete bitikiirustega.

Speex on loodud peamiselt kolme diskreetimissagedust silmas pidades: 8 kHz telefonikõnede edastamiseks (*narrowband*), 16 KHz (*wideband*) ja 32 kHz (*ultra-wideband*). Seejuures on need erineva sagedusulatusega andmed integreeritud ühte bitivoogu.

Speex kasutab ka varieeruvat bitikiirust (VBR), mille puhul bitikiirus muutub dünaamiliselt vastavalt kodeeritava kõne keerukusele. VBR tagab küll üldiselt parema kvaliteedi kuid VoIP puhul esineb sellega seoses probleem, et on piiratud maksimaalne bitikiirus, mis peab kommunikatsioonikanali jaoks olema küllaltki madal. Selle probleemi lahendamiseks kasutatakse keskmist bitikiirust (AVR – *Average Bit-rate*), mis etteantud maksimaalse bitikiiruse (*target bit-rate*) hoidmiseks muudab dünaamiliselt VBR kodeerimise kvaliteeti.

Kasutusel on ka VAD (*Voice Activity Detection*), mis teeb kindlaks, kas parajasti kostab kõne või vaikus ning sellele vastavalt määratakse kodeerimisel bitikiirus. Varieeruva bitikiiruse (VBR) kasutamisel on VAD alati kasutusel, konstantse bitikiiruse (CBR) puhul saab seda ise valida. Vaikuse puhul tagab VAD selle, et kodeerimisel kasutatakse vaid nii palju bitte, kui on tarvis taustamüra reprodutseerimiseks. Seda nimetatakse CNG (*Comfort Noise Generation*) ja

see genereerib kunstliku taustamüra, kasutatakse raadios ja teistes juhtmevabades kommunikatsioonivahendites.

Kui taustamüra on muutumatu ja ja kõne ei kosta, siis rakendatakse VAD/VBR koostöö lisavõimalust DTX (*Discontinuous transmission*), mis katkestab ülekande. Kui toimub faili kodeerimine (*file-based operation*), siis kirjutatakse vaid 5-bitiseid andmeplokke (*frame*), mis vastab bitikiirusele 250 bps, sest faili kirjutamist ei saa lihtsalt seisma jätta.

Dekodeerimisel võidakse kasutada kvaliteedi parandamist (*perceptual enhancement*), mis püüab vähendada kodeerimise/dekodeerimise käigus tekkinud müra.

Igal kõne (*speech*) koodekil on ülekande viivitus (*delay*). Speex koodekil on see viivitus 8 kHz diskreetimissageduse (*narrowband*) korral 30 ms, 16 kHz (*wideband*) korral 34 ms. Need arvud ei arvesta ajaga, mis protsessoril kodeerimiseks kulub.

Speex tarkvara ja infot koodeki kohta saab ametlikult kodulehelt: <http://www.speex.org/>.