

Kuulon malleihin perustuva audiokoodaus, osa II

Lähteet: Kahrs, Brandenburg, (Editors). (1998). "Applications of digital signal processing to audio and acoustics" Kluwer Academic.
Bernd Edler. (1997). "Low bit rate audio tools". MPEG meeting.
Spanias et al. (2007). "Audio signal processing and coding", Wiley & Sons
Yand et al (2006). "High-fidelity multichannel audio coding", Hindawi.



Sisältö:

- Todellisia koodausjärjestelmiä
- MPEG-1, MPEG-2 BC
- MPEG-2 AAC
- Vertailua
- MPEG-4
- MPEG:in ulkopuoliset koodausjärjestelmät

1 Todellisia koodausjärjestelmiä: Audiokoodaus 2 DA / Klapuri & Virtanen MPEG Audio 1 ja 2

- MPEG (Moving Picture Experts Group) standardisoi videon ja audion kompressiotekniikoita
- Kolme matalan bittinopeuden audiokoodausstandardia on valmistunut
- MPEG-1 Audio
 - suunniteltu soveltumaan useisiin tarpeisiin: tallentamiseen, digitaalisiin radiolähetysiin, ja lähetyksiin tietoverkossa
 - järjestelmä koostuu kolmesta eri tasosta (engl. *layers* 1, 2, 3)
 - bittinopeustavoitteet: 192 kbit/s (layer 1), 128 kbit/s (layer 2), 64 kbit/s (layer 3)
 - tukee näytteistystaajuuksia 32 kHz, 44.1 kHz, ja 48 kHz
 - moodeja: mono, stereo (bittiallokoinnissa molemmat kanavat huomioidaan, mutta koodataan toisistaan riippumatta), dual (täysin riippumattomat kanavat), joint stereo (stereokoodaus)

MPEG Audio 1 ja 2

Audiokoodaus 3
DA / Klapuri & Virtanen

- MPEG-2 Backwards Compatible coding (BC)
 - MPEG-1:n kanssa taaksepäin yhteensopiva *monikanavainen* koodaus
 - koodaus myös matalammilla näytteistystaajuuksilla (16 kHz asti)
 - ei uusia koodausalgoritmeja
 - MPEG-2 Advanced Audio Coding (AAC)
 - toisen sukupolven audiokoodekki geneeriseen stereo- ja monikanavasignaalien koodaukseen, mukaanlukien 5.1 ja 7.1 konfiguraatiot
 - ei taaksepäin yhtensopiva MPEG-1:n kanssa
- MPEG-2 tuo kaksi erilaista ja isoa lisäystä MPEG-1 koodekkien rinnalle

1.1 MPEG – ensimmäinen sukupolvi Audiokoodaus 4 DA / Klapuri & Virtanen

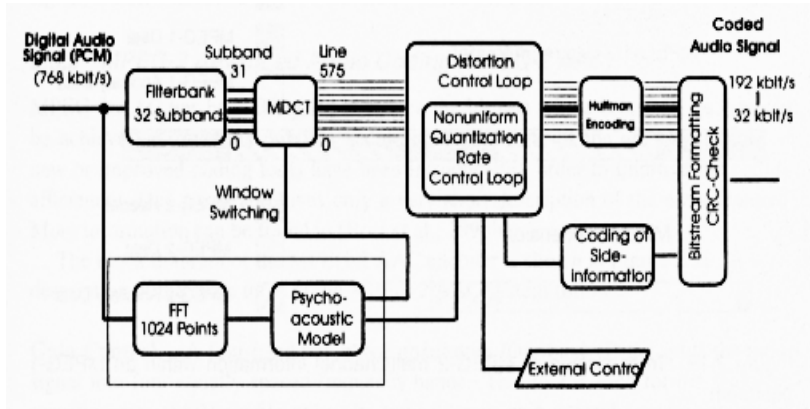
MPEG-1 tasot 1 ja 2:

- Pseudo-QMF-suodatinpankki
 - 32 alikaistaa, kiinteä aika/taajuusresoluutio (taso 1: koko ajan 8 ms aikakehyys)
 - 511-tappinen prototyyppisuodin
- Psykoakustinen malli määrää bittiallokoinnin
- Kvantisointi käyttäen blokki kompandointia
 - kvantisoidaan 12:n näytteen lohkoissa, maksimiarvo skaalataan 1:een
 - skaalaustekijät kvantisoidaan eksponentiaalisesti, 2 dB välein
 - 6:lla bitillä → 64 eri arvoa → 2 dB x 64 → ≈120 dB dynaamiseksi alue
 - skaalatut arvot kvantisoidaan tasavälisesti
- Tason 2 parannuksia verrattuna tasoon 1
 - yksittäisiä pieniä lisäkeinoja bittiallokoinnin yms. pakkaamiseen
 - erilainen aikakehyksiin jako (24 ms Layer 2:ssa, 8 ms Layer 1:ssä)
 - bittiallokointi optimoidaan iteratiiviseen tapaan siten, että signaali-maski suhde saadaan parhaaksi mahdolliseksi

MPEG – ensimmäinen sukupolvi

MPEG-1 taso 3 (.mp3)

- Kuva: lohkokkaavio MPEG-1 tason 3 enkooderista



MPEG – ensimmäinen sukupolvi

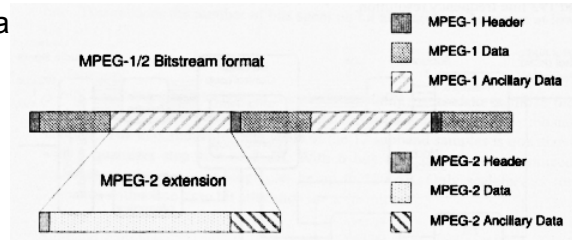
MPEG-1 taso 3 (.mp3)

- Piirteitä tasoista 1 ja 2
- Lisää koodaustehokkuutta saavutettu
 - paremman taajuuserottelun avulla
 - koodattujen arvojen häviöttömän Huffman-koodauksen avulla
- Hybridi-suodinpankki, vaihteleva aika-/taajuusresoluutio
 - kytkeytyy kolmen vaihtoehdon välillä: 576, 216, tai 192 taajuusviivaa, selostettiin adaptiivisten suodinpankkien yhteydessä
- Muita koodaustyökaluja
 - erilainen kvantisointi, edelleen epätasavälinen skaalaustekijöiden kvantisointi blokki kompandoinnissa
 - kvantisointikohinan kontrolli "analysis by synthesis"-periaatteella
 - stereokoodaustekniikoita
 - tukee M/S- ja intensiteetti-stereokoodauksen yhdistelmää

MPEG – ensimmäinen sukupolvi

MPEG-2 taaksepäin yhteensopiva monikanavakoodekki

- Ominaisuuksia
 - kehysformaatti on identtinen MPEG-1 bittivirran kanssa
 - lisäkanavat siirretään MPEG-1:n "ancillary data"-kentässä
 - MPEG-1 dekooderi osaa downmiksata 5-kanavaisen informaation
 - kolme versiota, tasot 1, 2 ja 3, jotka vastaavat MPEG-1:stä
- Kuva: MPEG-2 monikanavainformaation siirto MPEG-1 bittivirrassa



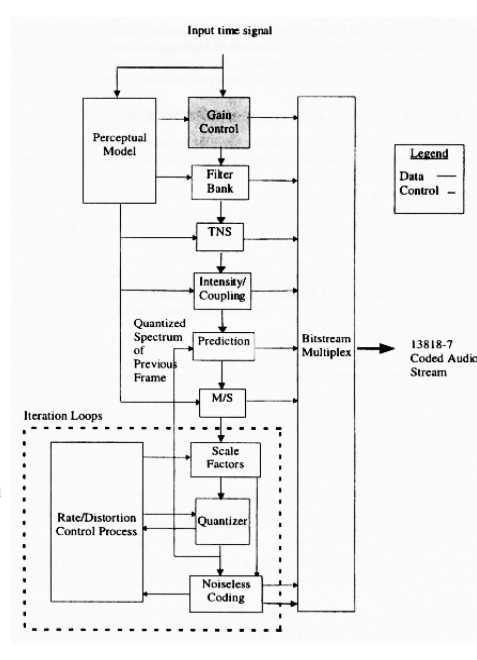
1.2 MPEG-2 AAC

- State-of-the-art geneerinen audiokoodekki
 - Suunnittelutavoite: laskea bittinopeutta mahdollisimman paljon kuitenkin säilyttäen "broadcast" laadun
 - saavutetaan matalahkolla bittinopeudella (64 kbit/s per kanava, tai vähempi)
 - Useita uusia koodaustyökaluja tehokkuuden parantamiseksi
 - joissain sovelluksissa AAC:llä voidaan pakata stereosignaali matalammillaan 16 kbit/s bittinopeudella
 - Kompleksisuusnäkökohtia
 - eri sovelluksissa tarvitaan erilaisia koodekin kompleksisuus vs. laatu kompromisseja
 - AAC määrittelee kolme eri kompleksisuusprofiilia
 - pääprofiili
 - matalan kompleksisuuden profiili (esim. ennustus ei ole käytössä, koska ei ole käytettävissä paljon muistia)
 - skaalautuvan näytteistystaajuuden profiili
- AAC yleensäkin: hyvä koodaushyöty kohtuullisella kompleksisuudella

MPEG-2 AAC

Lohkokaavio:

- Tason säätö
 - optionaalinen ominaisuus (käytössä skaalautuvan näytteistykseen profiilissa)
 - aikataso signaalin tasoa voidaan säätää esikaiku-efektien vähentämiseksi
- Suodinpankki
 - MDCT: 1024 tai 128
 - taajuusviivaa (aikaresoluu- tio parhaimmillaan 2.7 ms)
 - adaptiivinen: kaksi erilaista ikkunamuotoa



MPEG-2 AAC

... Lohkokaavio (ks. selostus tekniikoista materiaalin osassa Koodaus 1)

- Temporal noise shaping
 - tehdään säätämällä vahvistuksia eri taajuuskaistoille
- Intensiiteettikoodaus useille kanaville
- M/S-stereokoodaus
- Ennustus
 - toisen asteen taaksepäin adaptiivinen ennustaja
 - kullekin spektrikomponentille 16 kHz taajuuksiin asti
- Skaalaustekijät
 - spektri jaetaan useisiin spektrikomponenttien ryhmiin, jotka jakavat yhden skaalaustekijän
- Kvantisointi
 - epätasavälinen kvantisointi, askeleen koko 1.5 dB
- Huffman-koodausta käytetään koodaamaan kvantisoitu spektridata
 - häviötön koodaus

1.3 Vertailua

- Lähde: [MPEG-4 dokumentti]
- Tulokset ovat keskenään vertailukelpoisia vain kunkin paksiilla vaakaviivoilla erotetun ryhmän sisällä (kanavamäärä vaihtuu)
- AAC tuottaa broadcast-laatua viisikanavaisesti 320 kbit/s

Coding tool	Number of channels	Total bit rate	Typical subjective quality
AAC	5	320 kb/s	4.6
1995 Backward Compatible MPEG-2 Layer II	5	640 kb/s	4.6
AAC	2	128 kb/s	4.8
AAC	2	96 kb/s	4.4
MPEG-2 Layer II	2	192 kb/s	4.3
MPEG-2 Layer III	2	128 kb/s	4.1
AAC	1	24 kb/s	4.2
Scalable: CELP base and AAC enhancement	1	6 kb/s base, 18 kb/s enh.	3.7
Scalable: Twin VQ base and AAC enhancement	1	6 kb/s base, 18 kb/s enh.	3.6
AAC	1	18 kb/s	3.2
G.723	1	6.3 kb/s	2.8
Wideband CELP	1	18.2 kb/s	2.3

2 MPEG-4 Audio

» Luettavuuden helpottamiseksi MPEG-4 käsitellään omassa kappaleessaan, vaikka se loogisesti kuuluisikin edellisen alle

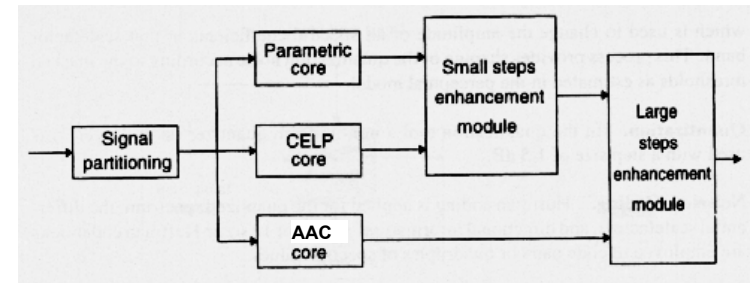
- Standardi saatiin valmiiksi 2000
- **Koostuu perheestä koodausalgoritmeja**, jotka tähtäävät eri bittinopeuksiin ja eri sovelluksiin
- Poistaa kuulun puheenkoodauksen, kuulon malleihin perustuva koodauksen, ja signaalien synteessin väliltä
 - bittinopeudet vaihtelevat välillä 2 kbit/s – yli 128 kbit/s per kanava
- Ominaisuuksia
 - skaalautuvuus erilaisille päätelaitteille
 - objektipohjaisuus: päätelaite voi valita äänivirrasta vain osia dekodattavaksi, ja voi uudelleen miksata tai manipuloida synkronisoituja ääniobjekteja
 - hyvä koodaustehokkuus (bittinopeus vs. laatu) saavutetaan valitsemalla kullekin signaalille sopiva koodausalgoritmi

MPEG-4 Audio

- Taustamotivointia – koodauksen dilemma:
- Yleiskäyttöiset *kuulon malleihin* perustuva audiokoodekit
 - saavutetaan hyvä laatu kaikentyyppisille audiosignaaleille bittinopeudella 64 kbit/s (MPEG-2 AAC)
 - perustuu kuulon kannalta irrelevantin informaation poistamiseen
 - puhetta tai yksittäisen soittimen ääntä *ei* voida koodata muita signaaleja ratkaisevasti pienemmällä bittinopeudella (AAC: puhekin vaatii ≥ 24 kbit/s)
- Puhekoodekit, esim. ITU-T
 - saavutetaan hyvä laatu puhesignaaleille yli 4 kbit/s nopeuksilla
 - perustuu pääasiassa signaalin *lähteen*, puheentuottoelimistön, *mallintamiseen*
 - ei oikein sovellu musiikkisignaalien koodaamiseen
- Mahdollisia ratkaisuja
 1. rajoitetaan tukemaan vain tiettyntyyppistä audiomateriaalia
 2. määritellään mielekäs yhdistelmä erilaisia koodekkityyppejä→ joku valinta täytyy tehdä

MPEG-4 Audio

- Kuva: MPEG-4:n skaalautuva konfiguraatio
 - *parametrinen ydin*: hyvin matalan bittinopeuden koodekki, joka perustuu äänten parametriseen esitystapaan
 - *CELP-ydin*: puhekoodekki (Code Excited Linear Predictive coding)
 - *AAC-ydin*: aika/taajuus –esitysmuotoon perustuva kuulomalliin perustuva koodekki, *MPEG-2 AAC*
 - *enhancement*: valinnaista lisäinformaatiota, joka lisätään bittivirtaan koodatun minimi-informaation lisäksi (skaalautuva laatu)



MPEG-4 Audio

- MPEG-4 tukee edellä mainittujen koodekkien ja niiden yhdistelmien käyttämistä
- Saavutetut bittinopeudet
 - puhesignaalit: 2 – 24 kbit/s
 - dekooderissa syntetisoitu puhe (siirretään tekstimuodossa mahdollisten kontrolliparametrien kera): 200 – 1200 bit/s
 - kompleksiset musiikkisignaalit: 6 – yli 128 kbit/s (monikanavatuki)
 - ei-kompleksiset musiikkisignaalit (synteettinen audio): ≥ 4 kbit/s
 - siirretään nuotti- ja kontrolli-informaatio, sekä tietyllä kielellä määritellyt instrumenttiäänät
 - kokonaiset äänimaisemat voidaan koota eri algoritmeilla koodatuista osista

MPEG-4 Audio

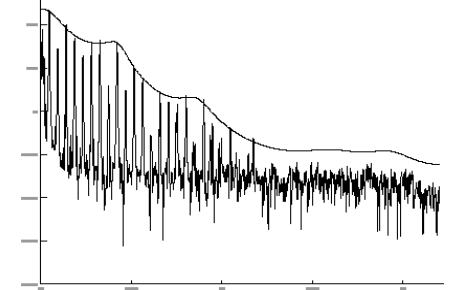
- Skaalautuva audiokoodaus
 - pelkästään osa bittivirrasta riittää, jotta voidaan dekodata ja generoida merkityksellinen audiosignaali
 - bittivirrassa on erikseen koodattuna minimi-informaatio, ja sen lisäksi "enhancement" osia, jotka yhdistämällä laatua voidaan parantaa
 - tavoite: multicast on tehokkaampaa kuin simulcast — on tehokkaampaa siirtää samaa bittivirtaa kaikille vastaanottajille, joista jokainen valitsee sen osan kuin kykenee dekodamaan/esittämään vastaanottimessaan
- Objektipohjaisuus
 - ääniä voidaan mikсата ja tuottaa päätelaitteessa, audio-objekti voidaan sijoittaa avaruudellisesti haluttuun paikkaan
 - binääriformaatti äänimaisemien määrittämiseen (engl. *binary format for scene description*): käytetään selostamaan miten ääniobjektit miksataan, kunkin efektit, yms. post-produktio
 - sisältöön perustuva interaktiivisuus (MPEG-4:n perusfilosofioita)
 - esim. virtuaalitodellisuudessa äänilähteestä voi kävellä kauemmaksi
 - valikoiva dekodaus: äänisignaalin eri osat voidaan koodata erikseen, ja silti synkronisoida yhteen. Tämä mahdollistaa signaalin eri osien koodaamisen eri algoritmeilla, ja valikoivan dekodamisen

2.1 CELP-puhekoodekki

- Puheinformaatio on moduloituneena spektrin karkeaan muotoon
 - foneemit erottovat toisistaan karkean spektrimuotonsa perusteella
 - korostuneita taajuusalueita nimitetään *formanteiksi*
 - spektrin hienorakenne ei muuta puheen foneemisisältöä
 - tieto äänteen soinnillisuudesta ja äänenkorkeus välitetään erikseen
- Lähdemalli
 - perustuu malliin ihmisen puheentuoelimestöstä. Tietyn tyyppiset signaalit ovat "mahdollisia", ei esim. äärettömän nopeat laajat muutokset
 - soinnillinen / soinniton heräte + ääniväylä, jota mallinnetaan suodattimella
 - muuttaen ääniväylän muotoa (esim. kielen ja leuan asentoa) foneemi vaihtuu
 - ks. seuraava sivu
- Kuulomalli
 - kvantisointikohinan spektrin muodon pitäisi olla samanlainen sisääntulevan signaalin kanssa
 - implementoitu spektriä painottavalla suodattimella
- Bittinopeudet
 - CELP: puhe 4 kbit/s (kelvollinen laatu) – 24 kbit/s (erittäin hyvä laatu)

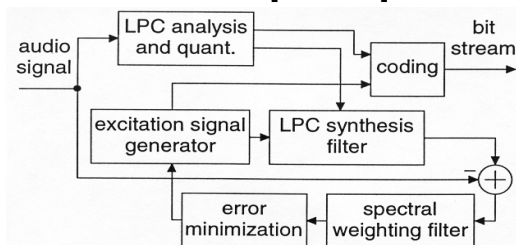
CELP-puhekoodekki

- Matala-asteinen suodatin sopii hyvin spektrin karkean muodon mallintamiseen, siis koodaamaan foneemitiedon
 - käytetään ns. ennustavia suotimia (LPC: linear predictive coding)
 - käytännössä käytetään lähes aina IIR-suodattimia, joissa on vain napoja (ei lainkaan nollia, osoittaja = 1)
 - suodatin on muotoa $H(z) = 1 / A(z)$, missä $A(z)$ on z:n polynomi
 - tämä johtuu siitä, että korostuneiden taajuusalueiden, formattien, ajatellaan välittävän olennaisen foneemi-informaation
 - estimoitavan suodattimen parametrimäärää saadaan pienemmäksi
- Kuva: signaalinpätkän magnitudispektri, sekä siitä LPC-analyysillä estimoitu spektrin karkea muoto



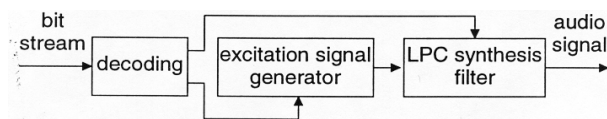
CELP-puhekoodekki

- Enkooderin lohkoakaavio [Edler97]



- Dekooderi

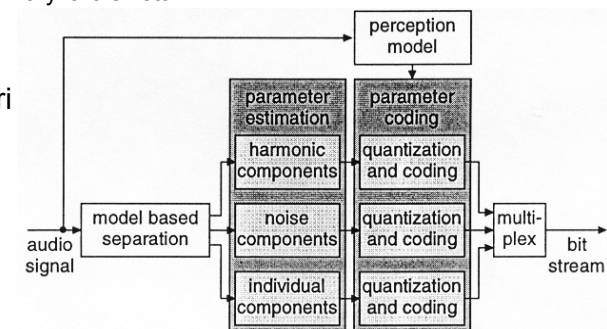
- herätesignaalin generointi perustuu koodikirjaan



2.2 Parametrinen koodekki

- Lähdemallit
 - harmonisia ääniä
 - kohinaa
 - yksittäisiä sinikomponentteja
 - yhdistelmä yllä olevista

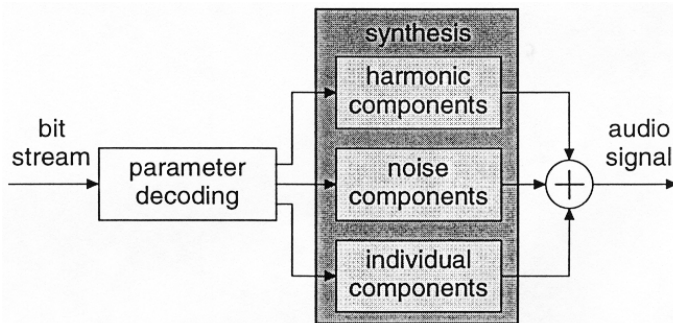
- Kuva: enkooderi [Edler97]



MPEG-4 Audio Parametrinen koodekki

Audiokoodaus 21
DA / Klapuri & Virtanen

- Kuva: dekooderi [Edler97]



MPEG-4 Audio Parametrinen koodekki

Audiokoodaus 22
DA / Klapuri & Virtanen

- Kaksi toimintamoodia
 - harmonisia ja kohinakomponentteja
 - soveltuu puheenkoodaukseen 2–4 kbit/s bittinopeuksilla
 - yksittäisiä sinikomponentteja
 - vähemmän kompleksisten musiikkisignaalien koodaukseen, (esim. yksittäinen instrumentti) nopeuksilla 4–16 kbit/s
 - yhdistelmä molemmista moodeista
 - syntaksi tukee, määrittellään transiitikohta
 - automaattinen moodin valitsin
- Parametrisen koodekin mahdollistamia lisätoiminnallisuuksia
 - toistonopeuden muuntelu vaihtamalla synteesin ajastusta
 - äänenkorkeuden muuntelu
 - skaalaamalla taajuusparametreja
 - puheelle spektrin karkea muoto pidetään paikallaan, jolloin pään (ääniväylän) dimensioiden aistimuksellinen kutistuminen voidaan välttää
- Bittinopeudet: puheelle 2–4 kbit/s, ei-kompleksinen musiikki 4–16 kb/s

MPEG-4 Audio 2.3 Koodaustekniikoiden yhdistely

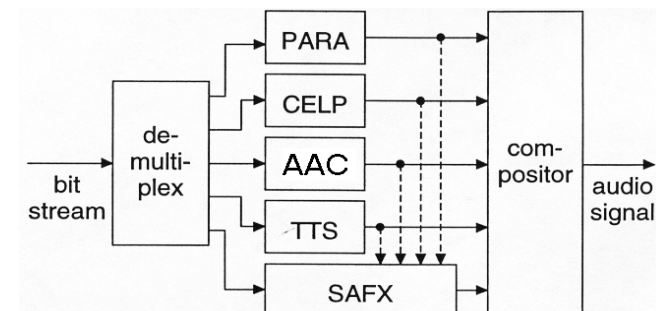
Audiokoodaus 23
DA / Klapuri & Virtanen

- Kohdesignaalina esim. puhe + taustamusiikki
- Signaalin osien erillinen koodaus voi parantaa suorituskykyä
- Eri osat oltava erillisinä signaaleina
 - äänilähteiden separointi: sisäänmenevästä polyfonisesta signaalista erotellaan kukin äänilähde omaksi signaalikseen
 - automaattinen analyysi ja separointi onnistuu nykytekniikalla ainoastaan hyvin yksinkertaisissa tapauksissa
 - signaalin eri osat on pidettävä erillään tuotantovaiheessa
 - koodekin valinta kullekin signaalille voi tapahtua käsin tai automaattisesti

MPEG-4 Audio 2.4 MPEG-4 dekooderi

Audiokoodaus 24
DA / Klapuri & Virtanen

- Aiempien MPEG-standardien tapaan vain bittivirta ja dekooderi on määritelty
 - enkooderien optimointi tulevaisuudessa käy mahdolliseksi
- Kuva: MPEG-4 dekooderi [Edler97]



MPEG-4 Audio Dekooderi

Audiokoodaus 25
DA / Klapuri & Virtanen

- Dekooderissa näkyy kaksi uutta lohkoa
 - TTS (text-to-speech synthesis): tekstistä puheeksi syntetisointi
 - SAFX: structured audio and effects
- Näiden abstraktiotaso on niin korkea, ettei tarvitse valita/määritellä mitään tiettyä koodekkityyppiä
 - tekstin syntetisointiin tarvittavat parametrit ovat yleisesti tiedossa
 - musiikkiparametrit tiedossa (MIDI)
- Huomaa: vaikka MPEG määrittelee vain dekooderin, määrää esim. puhekoodekin bittivirran määrittely kuitenkin *koodekkityypin*, samoin geneeriselle koodekille
 - näin ei ole kahdessa yllä mainitussa tapauksessa, kyse on korkeammasta, semanttisen tason esitysmuodosta

MPEG-4 Audio Dekooderi

Audiokoodaus 26
DA / Klapuri & Virtanen

- MPEG-4 tukee ns. synteettisen ja luonnollisen koodauksen hybridiä
 - dekooderissa voidaan yhdistää korkean tason semanttisesta tiedosta syntetisoitua ja koodattua ääntä
- Ääniä voidaan mikсата, tuottaa ja avaruudellisesti lokalisoida terminaalissa
 - binääriformaatti äänimaisemien määrittelyyn mahdollistaa selostuksen miksausesta, efekteistä, yms. post-produktiosta
 - omia efektejä voi määritellä SAFX:n orkestrointikielellä

MPEG-4 Audio 2.5 Tekstistä puheeksi synteesi

Audiokoodaus 27
DA / Klapuri & Virtanen

- Standardi sisältää määrittelyn prosodisista parametreista, joilla puheeseen saadaan ääntämyksellistä elävyyttä yms.
 - puheen nopeus ja äänenkorkeus
 - puhujan ikä ja sukupuoli voidaan määritellä
 - eri kielille ja aksenteille on omat koodinsa
 - kansainvälisesti kattava joukko foneemeja, joille on kullekin omat koodinsa
 - huulisynkronisointi kuvan kanssa on mahdollista
 - kasvojen animaatioparametrit voidaan välittää

MPEG-4 Audio 2.6 Strukturoitu audio (SA)

Audiokoodaus 28
DA / Klapuri & Virtanen

- Audion koodaus korkealla semanttisella tasolla
- SA-dekooderi: musiikin ja ääniefektien synteesi
- Äänidata
 - *structured audio orchestra language*
 - **instrumenttiäänten** synteesimenetelmien kuvauskieli (vrt. Csound, NetSound)
 - mitään yhtä synteesimenetelmää ei ole valittu, vain niiden kuvauskieli
 - käytettävissä peruselementtejä, kuten oskillaattoreita ja suotimia
 - *structured audio sample bank format*
 - ääniä voidaan esittää myös aaltomuotodatan avulla
 - perustana ns. wavetable sampling synteesi (syntetisaattoreissa käytetty)
 - suunnattu halvoille ja yksinkertaisille päätelaitteille
- Nuotti- ja kontrollidata: *structured audio score language*
 - formaatti, jolla **nuottidata ja äänten kontrolliparametrit** esitetään
 - käytetään ohjaamaan ääniä, jotka on määritelty yllä mainituilla tavoilla
 - vrt. MIDI, joka on myös otettu mukaan vaihtoehtoiseksi ja yksinkertaisemmaksi kuvausformaatiksi

2.7 Avoimia kysymyksiä

- Enkooderien optimointi jatkuu
 - valittujen koodekkityyppien puitteissa edelleen mahdollista, koska vain dekooderi on speksattu
- Optimaalisen koodaustekniikan valinta automaattisesti kulloisellekin signaalille
- Automaattinen signaalien separointi, jotta voidaan käyttää sopivaa koodekkia kullekin ääniobjektille

Kokonaiskuva MPEG:stä

- MPEG-1
 - VHS-laatu
 - video-CD
 - mp3 audio
- MPEG-2
 - DVD-laatu
 - DVD, digi-TV boxit
 - AAC audio
- MPEG-4
 - skaalautuva laatu
 - skaalautuva jakelu: matkapuhelin vs. satelliitti-TV
- MPEG-7: sisällön kuvaus, MPEG-21 jakelu: oikeuksien hallinta
 - ei koodausstandardeja

3 Onko MPEG:n ulkopuolella elämää?

- Kehittyneimmät kuulomalleihin perustuvat audiokoodekit on rakennettu MPEG Audion viitekehyksessä
- Eri yhtiöiden tuotteissa on omia koodekkivariaatioitaan:
- MiniDisc: adaptive transform acoustic coding (ATRAC)
 - Tavoite: CD-laatu 146 kb/s bittinopeudella
 - Hybridi: 3-kaistainen QMF + MDCT (32 / 128 / 256 taajuusviivaa)
 - Adaptiivinen suodatinpankki 1.45 / 2.9 / 11.6ms
 - Blokki kompandointi
- Digital Compact Cassette: precision adaptive subband coding (PASC)
 - Perustuu MPEG 1 layer 1:een

Onko MPEG:n ulkopuolella elämää?

- AT&T: perceptual audio coder (PAC)
 - MDCT-suodatinpankki, 128/1024 taajuusviivaa
- Dolby: AC-3 –koodaus
 - Laajasti käytössä monikanavakoodauksessa
 - Intensiteetti-monikanavakoodaus: yksi tehospektri, skaalaustekijät 18 kaistalle joka kanavalle
 - MDCT-suodatinpankki
- MP3Pro: perustuu MPEG 1 layer 3:een
 - lisäksi korkeat taajuudet koodataan karkeasti ja rekonstruoidaan käyttäen myös matalampia taajuuksia

Onko MPEG:n ulkopuolella elämää?

- Ogg Vorbis: kehitetty eteenpäin MPEG 1 layer 2:sta
 - avoimen lähdekoodin koodekki
 - aika/taajuus-esitys: MDCT
 - vektorikvantisointi
 - taustakohinamalli välttää useille koodekeille tyypilliset ”liverrys”-artefaktat matalilla bittinopeuksilla
- Windows Media Audio
 - aika/taajuus-esitys: MDCT, 5 eri ikkunan pituutta
- Noudattelevat suurinpiirtein samoja peruslinjoja kuin MPEG-koodekit