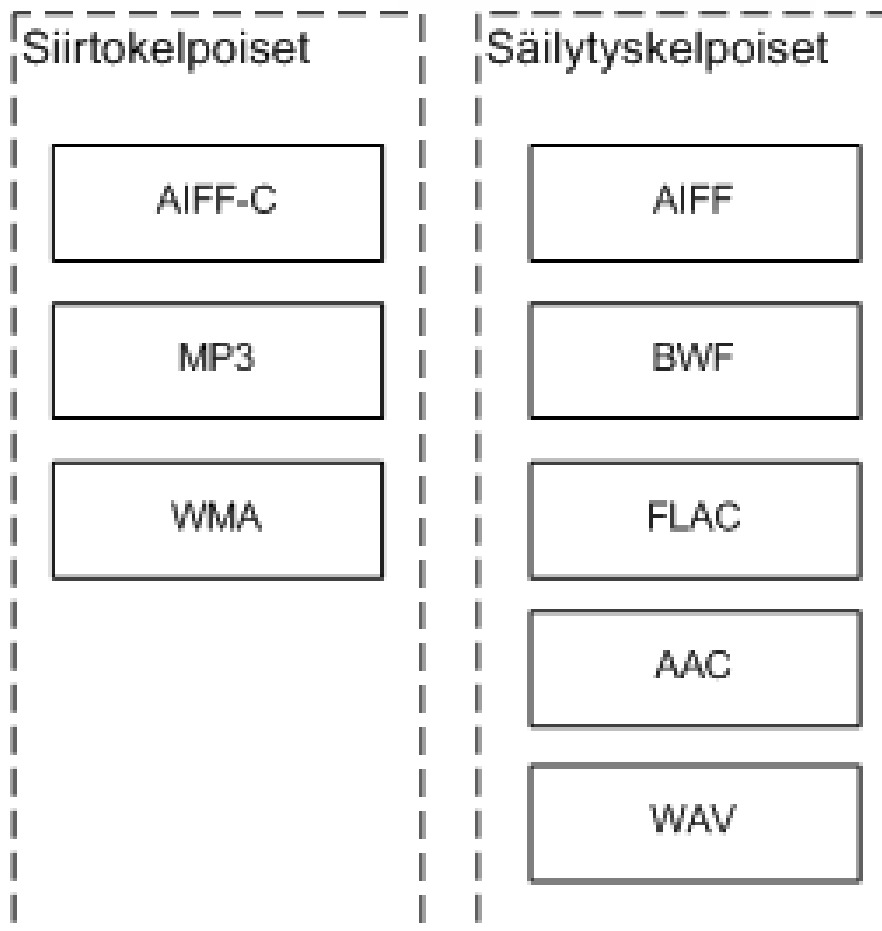




AV-muotojen migraatiotyöpaja - ääni

KDK-pitkäaikaissäilytys 2013 -seminaari
6.5.2013 / Juha Lehtonen

Äänimuodot



Ääneen vaikuttavia asioita

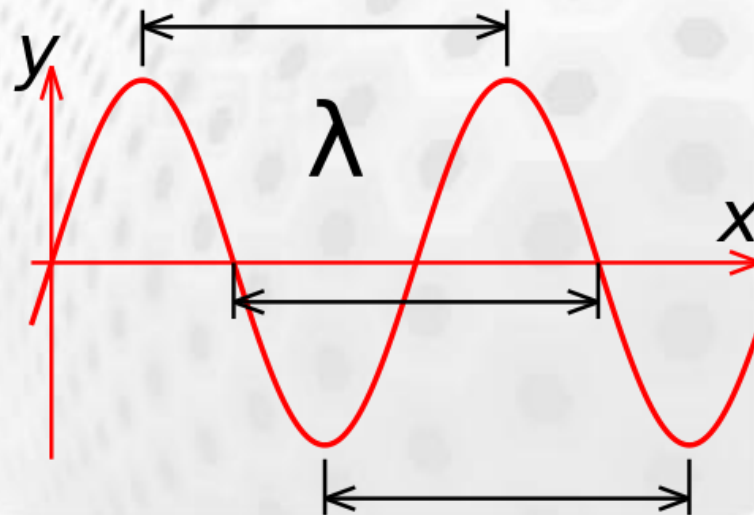
- Taajuudet
- Äänen voimakkuus
- Kanavien määrä
- Näytteistys
- Bittisyvyys
- Bittinopeus
- Datan koodausmenetelmä
- Metadata

Ääni

(...ja taas vähän fysiikkaa...)

Ääni

- Ääni on mekaanista aaltoliikettä, joka etenee väliaineessa (esim. ilmassa)
- Taajuus: Kuinka monta kertaa sama aaltoliikkeen jakso toistuu tietyssä ajassa?
 - Esim. 4000 Hz tarkoittaa, että aalto toistuu 4000 kertaa sekunnissa
- Mitä suurempi taajuus, sitä korkeammalta ääni kuulostaa.
- Mitä korkeampi aalto, sitä voimakkaampi ääni



Äänen taajuus, näytteistys ja bittisyvyys

- Ääntä voidaan tallentaa ottamalla siitä näytteitä riittävän nopein jaksoin.
 - Tällöin puhutaan näytteenottotaajuudesta, joka on siis eri asia kuin äänen taajuus/taajuudet
 - Näytteenottotaajuuden (jatkossa ”näytteistykseen”) on oltava vähintään kaksinkertainen verrattuna äänen korkeimpaan taajuuteen (Nyquistin teoreema)
 - Muussa tapauksessa korkeimpia taajuuksia ei saada talteen oikein
- Bittisyvyys: Tallennetun näytteen tarkkuus

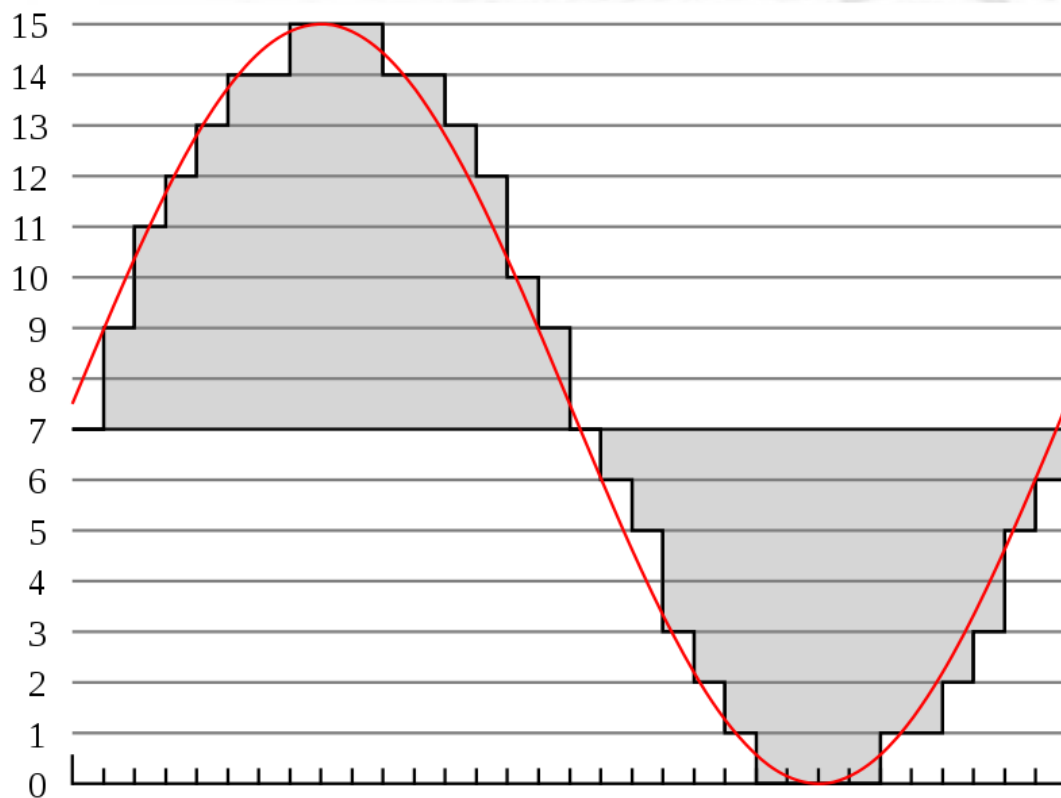
PCM-koodaus

- Pulse Code Modulation
- Kehittäjä: Alec Reeves (1937)
- Menetelmä äänitaajuussignaalin koodaamiseksi digitaaliseen muotoon
- Analogisesta signaalista otetaan tasaisin väliajoin näytteitä, jotka ilmaistaan numeerisesti
- CD-levyjen PCM-äänisignaali on laadultaan 44 100 näytettä sekunnissa (näytteistys siis 44,1 kHz) ja 16 bitin tarkkuudella (bittisyvyys siis 16-bit)

PCM-koodaus

● Kuvassa ääniaallon PCM-koodaus

- 4-bittinen: $2^4 = 16$ ”porrasta”
- Tallennetaan: 7, 9, 11, 12, 13, 14, 14, 15, 15,...



Äänen bittinopeus

- Kuinka monta bittiä kulutetaan sekuntia kohti
- Käytetään, kun pakataan häviöllisesti esim. PCM-koodattua digitaalista ääntä
- Käytännössä korvaa alkuperäisen digitaalisen audion näytteistykseen erilaisilla menetelmillä
- Bittinopeus voi olla vakio tai vaihteleva
 - Vaihteleva bittinopeus pyrkii tasaiseen laatuun

Äänikanavien määrä

- Jokaisella kanavalla on oma äänisignaalinensa
- 1.0 Mono
 - Yksi äänikanava, esim. puhe
- 2.0 Stereo
 - Kaksi äänikanavaa, esim. CD-musiikki
- 5.1 Kaksi etukaiutinta, kaksi takakaiutinta, keskikaiutin ja subwoofer
 - Viisi äänikanavaa ja matalien äänien kanava, esim. DVD-elokuvat
- 7.1, 9.1, ...

Psykoakustinen malli

- Ihmisen kuulossa on ominaisuuksia , joita hyödynnetään häviöllisessä pakkauksessa
 - Ihminen kuulee noin 20 – 20 000 Hz alueen
 - Raja-arvoilla kuuleminen on huonompaa, jolloin siltä alueilta voi hävittää tietoa enemmän
 - Herkimmillään ihmiskorva on 500 – 4 000 Hz välillä
 - Jos äänessä on useita toisiaan lähellä olevia taajuuksia, niistä voimakkain peittää muut
 - Peitettyjä taajuuksia voi hävittää
 - Matalilla taajuuksilla äänen suunnan havaitseminen huononee
 - Esim. stereoäänen voi tältä alueelta koodata monoksi

Tiedostomuodot



Tiedostomuoto WAV

- Waveform Audio File Format
- Kehittäjä: Microsoft ja IBM (1991)
- Näytteistys ja kanavien määrä käytännössä rajaton
- Kokorajoitus 4GB (max. 32-bittisyyden vuoksi)
 - Noin 6,8 tuntia ääntä CD-stereolaadulla
- Ääni jaetaan paketteihin, koodausmenetelmä on teoriassa vapaa
 - KDK suosittelee, että WAV-säiliöön voi tallentaa vain häviötöntä PCM-koodattua ääntä.

Tiedostomuoto BWF

- Broadcast Wave Format, versio 2 (2011)
- Kehittäjä: European Broadcasting Union
- WAV:in laajennos, tiedostopääte .wav
 - Sama 4GB-rajoitus
- Mahdollista PCM- tai MPEG-koodauksen
- Laaja tuki metadatalle

Tiedostomuoto AIFF

- Audio Interchange File Format
- Kehittäjä: Apple (1989)
- AIFF: Häviötöntä PCM-koodattua ääntä
 - Säilytyskelpoinen
- AIFF-C: Eri kompressiomenetelmiä
 - Siirtokelpoinen, ei muita eroja kuin kompressio
- Käsittelee tietoja paketteina
- Voidaan tallentaa myös MIDI-musiikkia
- Muutamia kenttiä kuvailevalle metadatalle

Tiedostomuoto MP3

- MPEG-1/2 Audio Layer 3
- Kehittäjä: MPEG-komitea
- ISO/IEC-standardoitu (1993)
- Siirtokelpoinen muoto
- Patentoitu
- Näytteistys: 16-48 kHz
- Kanavia: 1-2 (MPEG-1) tai 1-5.1 (MPEG-2)
- Häviöllinen pakkausmenetelmä
 - Bittinopeus määräävänä tekijänä, mahdollisista bittinopeuksista on luettelo

Tiedostomuoto MP3

- Pakkauksessa psykoakustinen malli
 - Eri ohjelmien MP3-koodauksessa voi olla eroavaisuuksia, riippuen siitä, mitä äänen kuulemisen ominaisuuksia on pakkauksessa korostettu
- Myös vaihteleva bittinopeus (VBR, Variable Bit Rate) mahdollinen
 - Tavoitteena tasainen laatu
- Kuvaileva metadata esitetään yleensä ID3-tageina
 - Kokonaisuus ei sisälly viralliseen MP3-määrittelykseen

Tiedostomuoto MPEG-4 AAC

- MPEG-4 Advanced Audio Coding
- ISO/IEC-standardoitu (2005)
- Patentoitu
- Häviöllinen, mutta parempi kuin MP3
 - Näytteistys: 8-96 kHz
 - 1-48 kanavaa
 - Bittinopeuden määrittäminen vapaa
 - Tehokkaampi ja tarkempi koodaus pakkauksessa

Tiedostomuoto MPEG-4 AAC

- ➊ Pakkauksessa perustana
 - Psykoakustinen malli
 - Toistuvien osien eliminointi
- ➋ Modulaarisuus
 - Muotoon voidaan luoda profiileja
 - Eri profiilit toteuttavat koodauksen eri tavoin
 - Vaikuttavat koodauksen kompleksisuuteen, tehokkuuteen ja laatuun
 - Standardissa profiileja on määritelty toistakymmentä

Tiedostomuoto FLAC

- Free Lossless Audio Codec
- Kehittäjä: Xiph.org Foundation (2007)
- Avoin muoto
- Voidaan häviöttömästi pakata PCM-ääntä, jossa
 - Näytteistys: 1 Hz – 655 350 Hz
 - Bittisyvyys: 4 – 32
 - Kanavia: 1 – 8
- Metatiedot tiedoston alussa (oma muoto)

Tiedostomuoto WMA

- Windows Media Audio
- Kehittäjä: Microsoft
- Siirtokelpoinen muoto
- Koodaa audiota, jossa
 - Näytteistys max. 48 kHz
 - Kanavia max. 2
- Häviöllinen pakkaus
 - Psykoakustinen malli
 - Vaihteleva bittinopeus (VBA)
 - Samantapaisia piirteitä kuin AAC-muodossa

Ääni migraatiossa

- Mikä formaatti?
 - Onko edes tarkoitus migroida?
- Mitkä taajuudet?
 - Väärä taajuusalue voi kuulua jopa kokonaisien soittimien puuttumisena (esim. basso)
- Mikä bittisyvyys?
 - Liian pieni bittisyvyys heikentää äänen tarkkuutta
- Mikä bittinopeus?
 - Liian pieni bittinopeus heikentää äänen laatua

Ääni migraatiossa

- Montako kanavaa?
 - Uuden formaatin pitää tukea kanavamäärää
- Mikä koodausmenetelmä?
 - Pakkaus voi heikentää äänen laatua
- Häviöllisesti pakattua ääntä ei välttämättä kannata migroida pakkaamattomaan muotoon