

Univerza v Ljubljani  
Fakulteta za računalništvo  
in informatiko



# PRODUKCIJA MULTIMEDIJSKIH GRADIV (PMG)

## ZVOK

Borut Batagelj

V1.3  
2023



# Zvok

- Zvok ni vizualni medij (kot so ostali)
- Zaznavamo ga s čutilom za sluh
- Kompleksna mešanica fizičnih in psiholoških faktorjev
- Zvok v multimediji ni vedno potreben
  - Včasih je celo moteč ali popolnoma odveč
  - Včasih nam zelo popestri material
  - Včasih pa brez zvoka ne gre

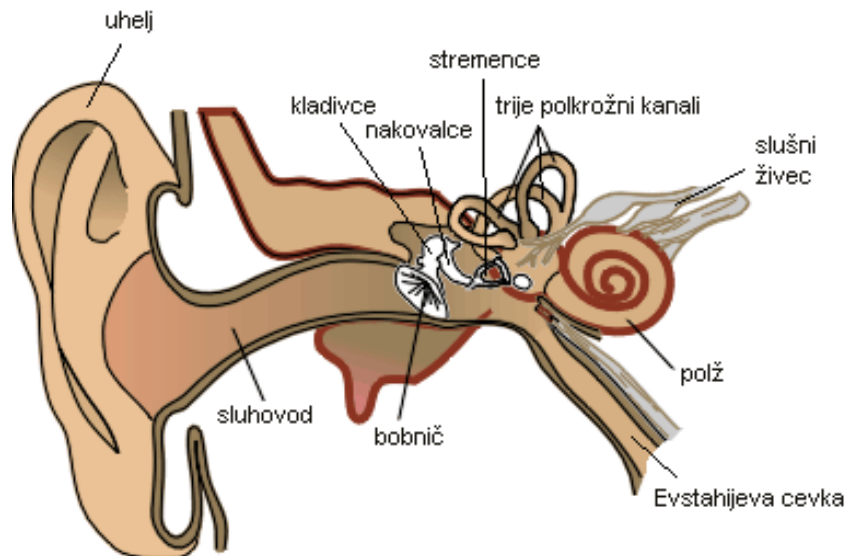


- **vedno moramo imeti možnost izključiti zvok**



# Narava zvoka

- Zvok se prenaša z valovanjem zraka (ali drugega medija)
- Vibracije se prenašajo na bobnič
- Prenos na notranje uho – sproži živčne impulze
- To zaznamo kot zvok – dejansko poslušamo z možgani
- Človek zaznava frekvence v razponu 20Hz-20Khz





# Valovanje zvoka

- Amplituda valovanja
- Statična predstavitev dinamičnega pojava
- Za analizo, urejanje, razpoznavanje, ...

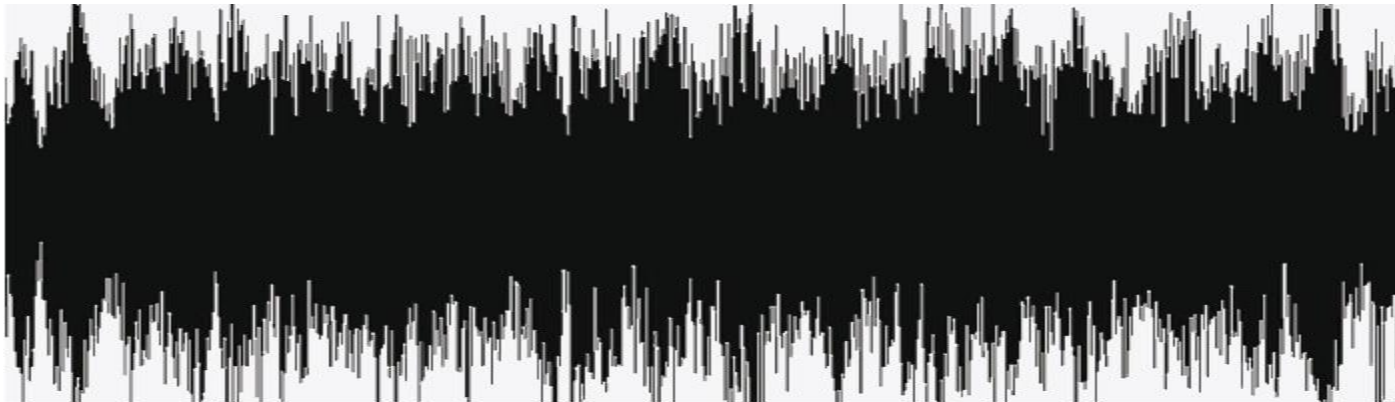
*„Feisty teenager“*





# Zvočni signal

*„Didgeridoo“*



*Boogie-woogie*





# Amplituda zvočnega signala

*Violina, čelo in klavir*



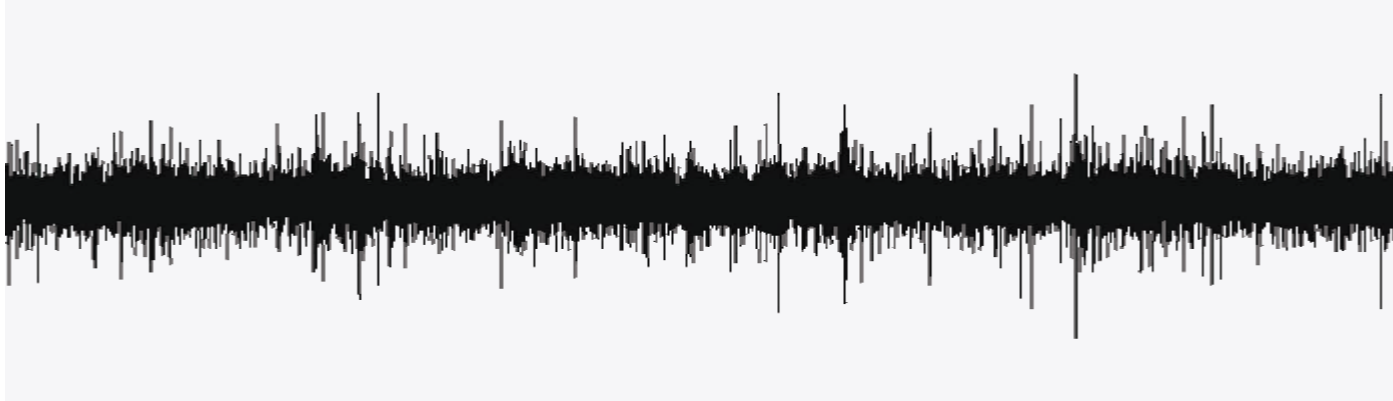
*„Men grow cold...“*



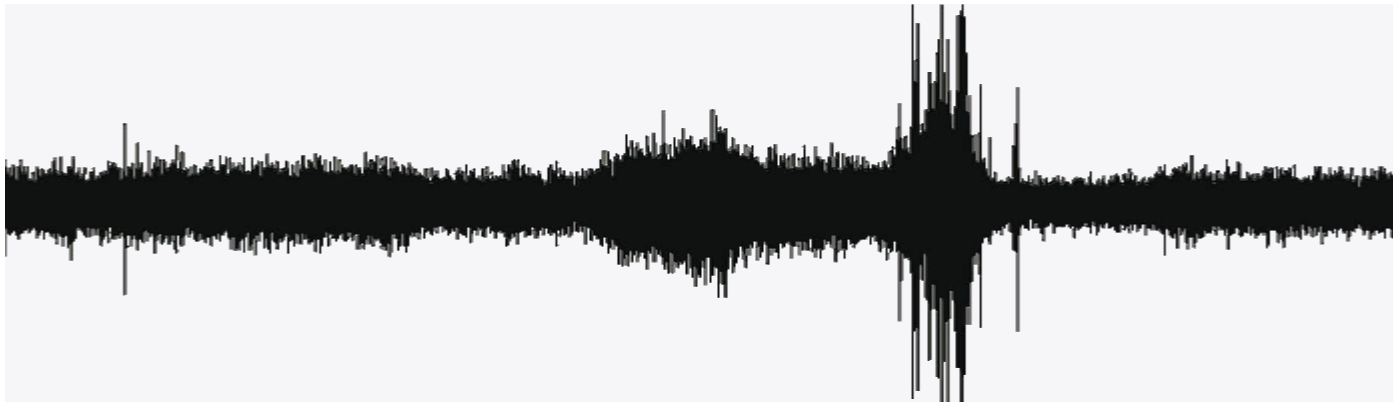


# Amplituda zvočnega signala

*Kapljanje*



*Morje*





# Različno dojetanje

- Glasnejši ton daje občutek nižje višine
- Kompleksnejši toni dajejo občutek nižje višine, kot čisti toni
- Sposobni smo slišati zvok svojega imena, čeprav drugega ni mogoče razbrati
- Zmožni smo razbrati govor, čeprav je šum glasnejši
- Stereografski zvok
  - Dojamemo izvor zvoka: glasnost, zamik (faza)



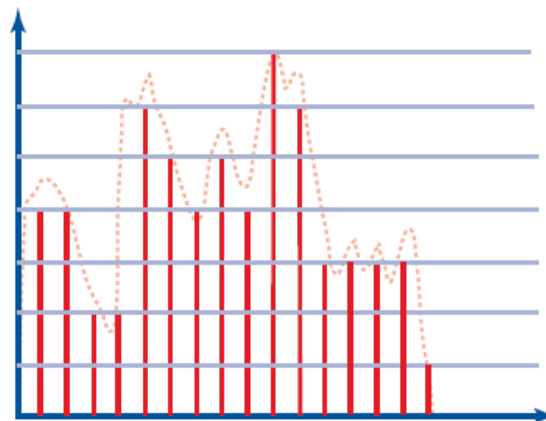
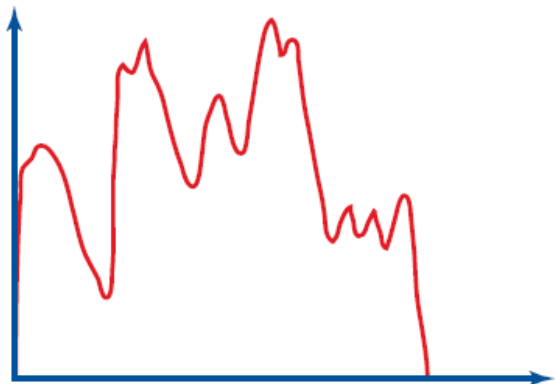


# Digitalizacija zvoka

- A/D pretvornik
  - Mikrofon, električna kitara

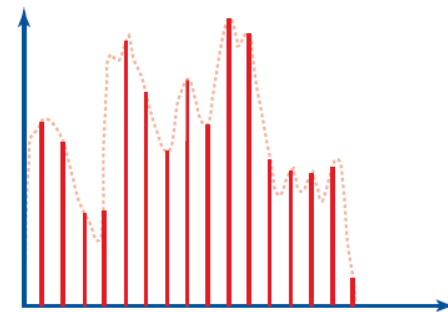


- Vzorčenje
- Kvantizacija



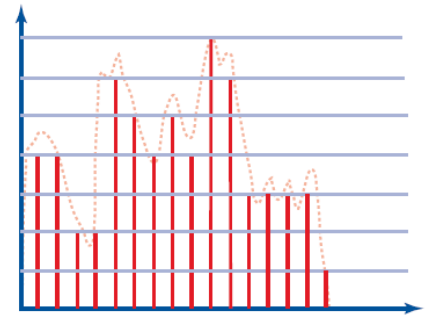


# Vzorčenje zvoka

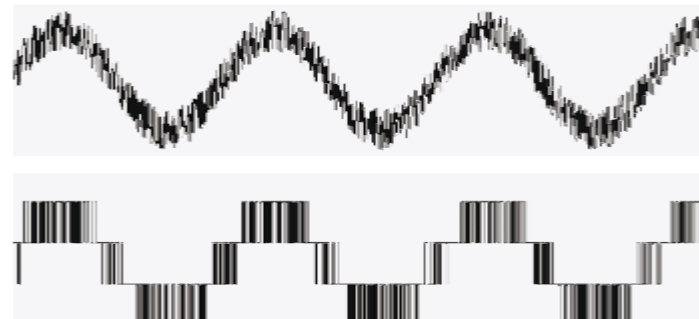
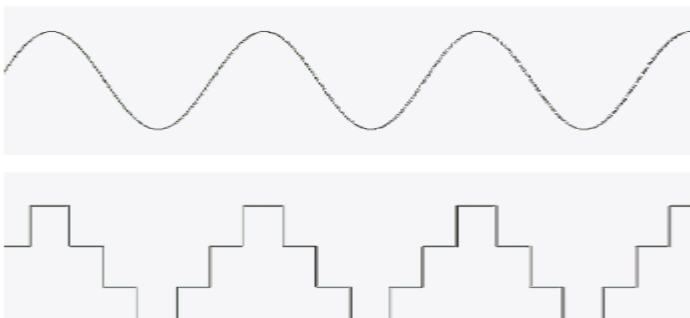


- Človeško uho zaznava frekvence do 20 kHz
- Teorem vzorčenja: potrebna frekvenca vzorčenja:  $\geq 40\text{kHz}$ 
  - Audio CD: 44.1 kHz
  - Manjše zahteve (internet): 22.05 kHz
  - Govor: 11.025 kHz
  - Bolj profesionalne naprave: 48kHz, tudi 96kHz in 192kHz
- Pred vzorčenjem se frekvence  $> 1/2$  frekvence vzorčenja porežejo

# Kvantizacija

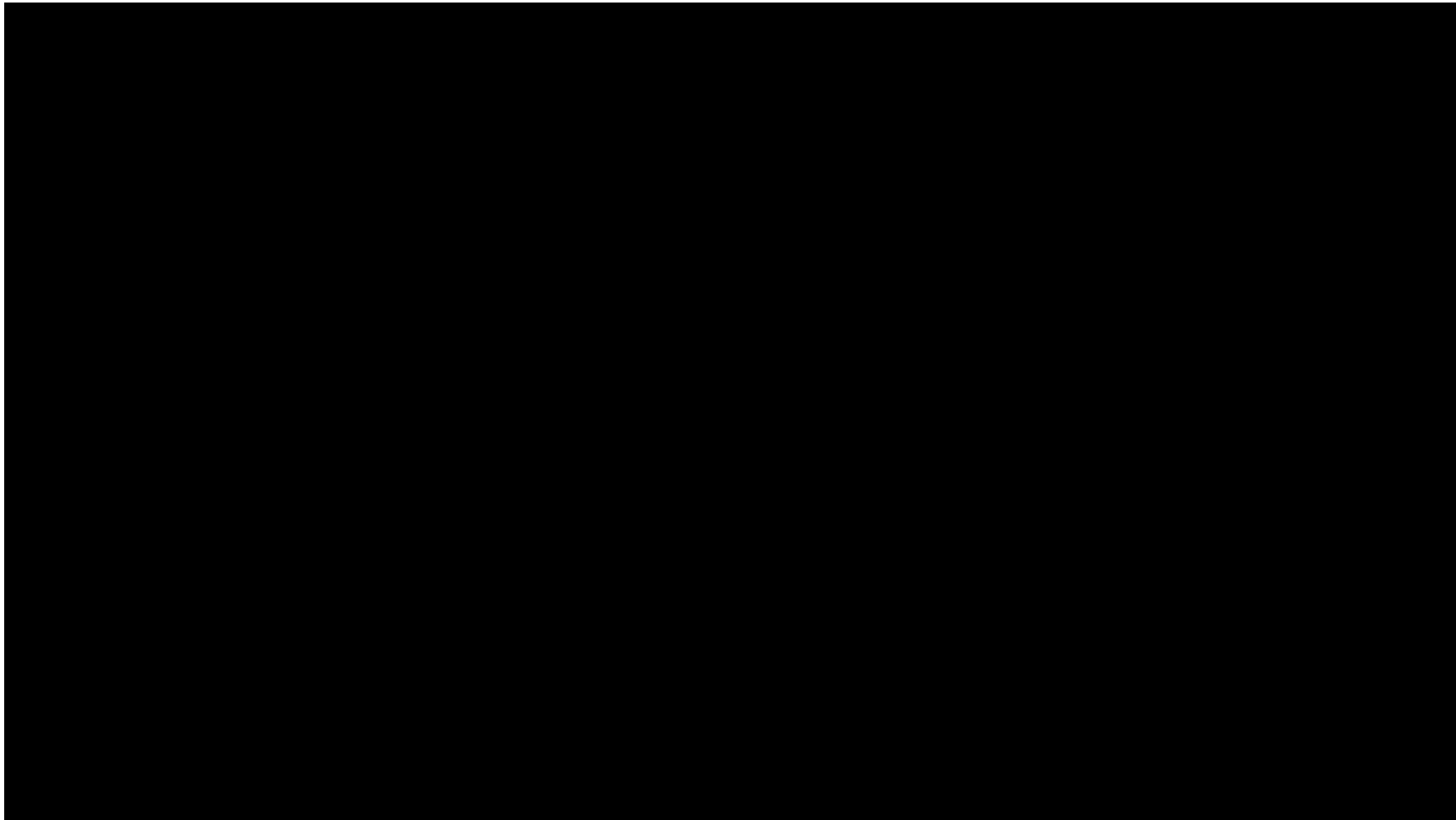


- Število kvantizacijskih nivojev:
  - Običajno: 16 bitov = 65.536 nivojev (CD audio)
  - Minimalno: 8 bit = 256 nivojev (komunikacija)
  - Nadstandardno: 24 bitov
- Kvantizacijski šum je večji pri signalih z nizko amplitudo
  - Povečamo število kvantizacijskih nivojev
  - Dodamo majhen naključni šum pred vzorčenjem ([youtube](#))
    - Zgladimo ozke prehode





# Kvantizacija





# Formati



- Različni (nekompresirani) formati na različnih platformah (so povsod podprti)
  - AIFF: Mac OS X
  - AU: Unix
  - WAV: Windows
- Brezizgubna kompresija (Lossless)
  - FLAC – Lossless codec developed by Xiph.Org Foundation., free in opensource
  - ALAC=Apple Lossless
  - APE ni tako razširjen ima pa boljšo kompresijo
- Izgubni (Lossy Formats)
  - MP3: MPEG Audio Layer III
  - AAC: Advanced Audio Coding, malce boljše kot mp3, iTunes ga je naredil popularnega, uporablja Apple (iPod)
  - Ogg Vorbis, opensource, ni toliko razširjen format
  - WMA = Windows Media Audio (Microsoft) ni toliko razširjen





# Formati

- MP3: za glasbo, internet, podcaste
  - V prvi vrsti je *encoder* ne format

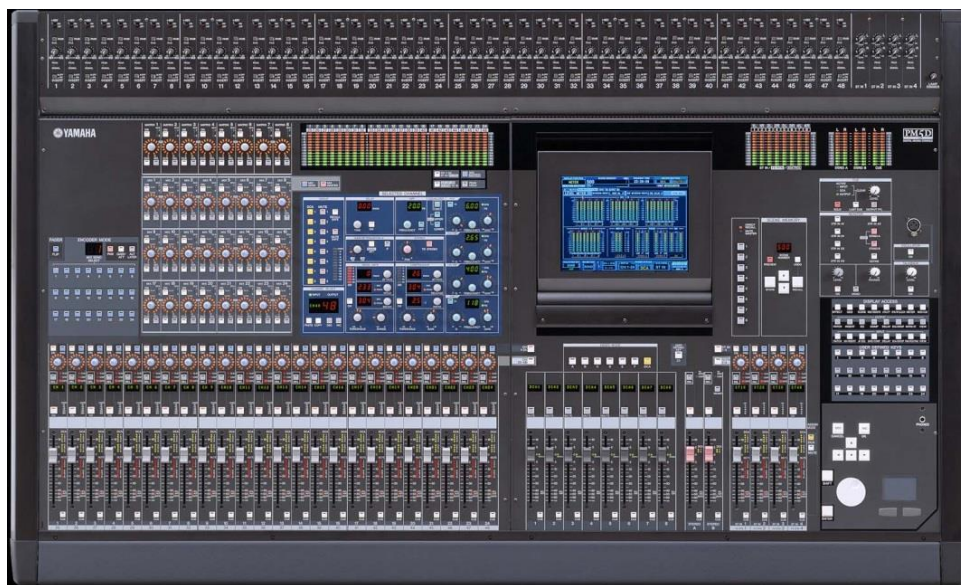


- Pretočni zvok
  - Ni potrebe po shranjevanju na disk
  - Predvajanje zvoka v živo
  - Bolj uspešno kot video (zahteva manjšo propustnost)
  - QuickTime (AAC kodek), Windows Media Audio (WMA), MP3



# Obdelava zvoka

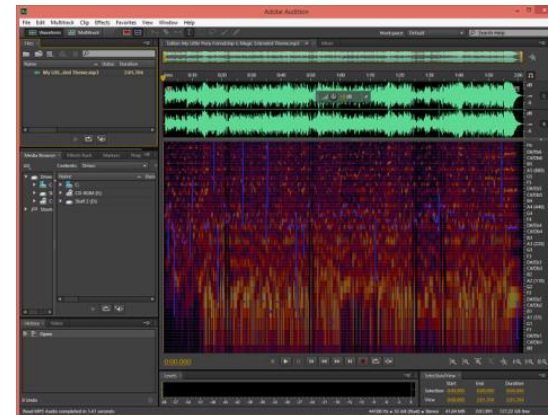
- Tradicionalno: profesionalna oprema
  - Kompleksne mešalne mize
  - Digitalne izvedbe





# Obdelava zvoka

- V multimediji
  - Manjše zahteve
- Programi za urejanje zvoka
  - Odprtokodni programi (Audacity)
  - Adobe Audition (prej Soundbooth)
  - Apple Logic Pro (GarageBand)
  - Programi za urejanje videa (Premiere, Final Cut)







# Zajemanje zvoka

- V računalnik vgrajen mikrofoni ne zadošča
  - Nekvaliteten mikrofoni
  - Šumi iz ventilatorjev, diskov, ...
- Zunanji mikrofoni
- Posebni snemalniki zvoka
  - Brez kompresije
- Zagotoviti moramo dobro akustiko prostora
- Snemalni studio



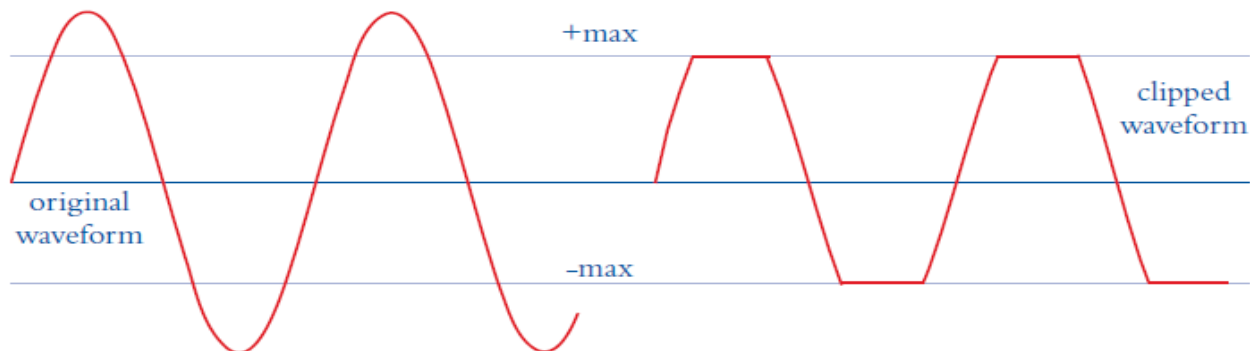
# Nastavitve pri zajemanju zvoka

- Frekvenca vzorčenja in število nivojev
  - CD kvaliteta: 44.1 KHz, 16 bit
    - $44.100 * 16 / 8 = 86$  kB/s oz. 5MB/min (mono)
  - Velikost datoteke prepolovimo:
    - Če prepolovimo frekvenco vzorčenja (bolje, manj opazno)
    - Če prepolovimo število nivojev
    - Če uporabimo mono namesto stereo tehnike



# Nastavitve pri zajemanju zvoka

- Nastavitve nivoja vhodnega signala
  - Na čim višji nivo, ki ne povzroči rezanja najvišjih amplitud
    - Zagotovimo maksimalni dinamični obseg
  - Problematično pri snemanju v živo
    - Težko predvideti najvišji nivo
  - Avtomatski nadzor ojačitve signala (gain control)
    - Se spreminja dinamično glede na amplitudo signala
    - Zmanjša dinamiko zvoka
  - Programska kasnejša normalizacija signala ne reši problema
    - Poveča dinamični obseg, ne pa doda podrobnosti





# Uvoz zvoka

- Uvažanje zvoka v digitalni obliki
  - Bolj enostavno
- Uvoz zvoka z avdio CD-ja
- Uvoz zvoka iz drugih zvočnih datotek
  - Z interneta
  - Običajno je glasba kompresirana
- Če se le da, uvozimo nekompresirane posnetke
- Kompresirani niso primerni za obdelavo

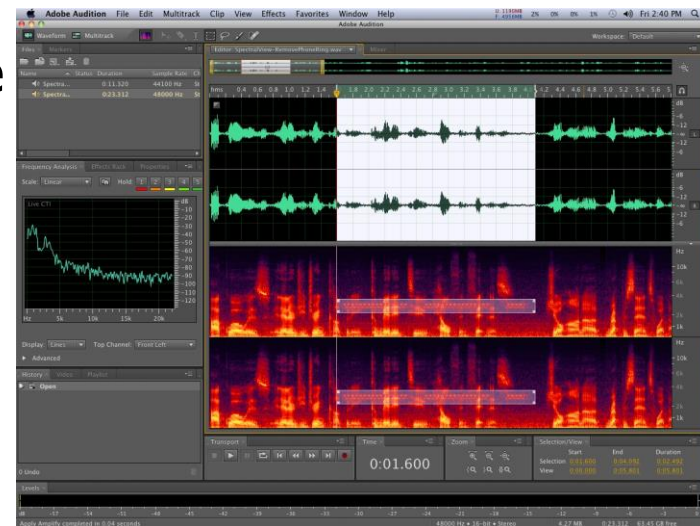


**Avtorske  
pravice !**



# Urejanje zvoka

- Podobno urejanju videa
  - Časovnica
  - Prikazana valovna oblika skozi čas (amplituda)
- Krajšanje, kombiniranje, preurejanje
- Več trakov (tracks) hkrati
  - Stereo zvok (ločena trakova)
- Kreiranje zank
  - Kratke za posamezne tone
  - Pomembni so prehodi med zankami (ničle)
  - Daljše zanke: ponavljajoči se delčki skladbe





# Postprodukcija

- Popravljanje napak, izboljšanje kvalitete, ipd
- Vrata (gates) in filtri
- Odstranjevanje šuma
  - **Protišumna vrata** (noise gate)
    - Odstranijo vse vzorce pod nekim pragom
      - Šum se bo vklapljal in izklapljal
  - **Filtri za odstranjevanje šuma**
    - Nizkopropustni filtri (Low-pass)
      - Odstranijo visoke frekvence
      - Izločimo praskanje/sikanje
    - Visokopropustni filtri (High pass)
      - Odstranijo nizkofrekvenčno „mrmranje“ (mehanske vibracije)
    - Notch filter
      - Odstrani ozek frekvenčni pas (npr. 50Hz)
  - Bolj napredni filtri
    - Analiza šumnega prostora in potem detekcija vzorca



# Postprodukcija

- Popravljanje napak nastalih med snemanjem
  - De-esser
    - Odstranitev šuma povzročenem z govorom ali petjem preblizu mikrofona
  - Click repairer
    - Odstranitev klikanja poškodovanih vinil plošč
  - Še vedno je najbolje zvok posneti brez napak



# Postprodukcija

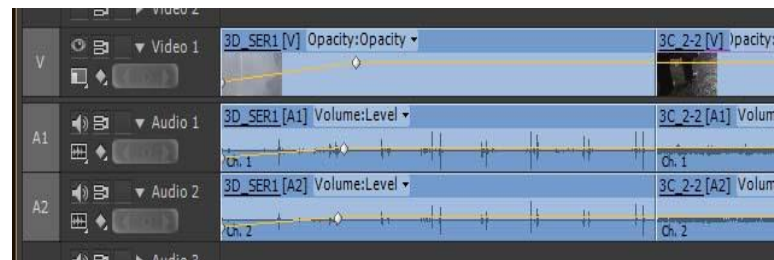
- Spreminjanje zvoka
  - Minorni popravki -> slaba izvedba ali snemanje
  - Redikalne spremembe -> nov zvok iz originala
- Reverb
  - Odmev: dodajanje zakasnelih kopij signala
  - Poustvarimo prostor
- Envelope shaping
  - Spreminjanje oblike amplitude zvoka skozi čas
    - Postopno višanje ali nižanje amplitude, poljubno spreminjanje
- Spreminjanje dolžine
  - Dodajanje/odvzemanje vzorcev
  - Sprememba dolžine brez spremembe višine glasu
    - Sinhronizacija
    - Spreminjanje hitrosti (tempo) glasbe





# Kombiniranje zvoka in slike

- Pri videu imamo ločeni sledi za sliko in zvok, ki sta sinhronizirani
  - Sledi sta zaklenjeni skupaj
- Sinhronizacija zvoka in slike
  - Najbolj opazno pri govoru
  - Opis mora sovpadati s sliko
  - Glasba s prikazanim
  - Naravni zvoki z dogajanjem
- Skupni časovni trak za video in zvok
  - Različno število vzorcev na sekundo (25:44100)
    - Osnova je okvir
  - Ugotavljanje točk ujemanja s poslušanjem (in vizualno)
- Prenos slike ločeno od zvoka
  - Spuščanje okvirjev videa, če le-ta zamuja





# Kompresija zvoka

- CD kvaliteta, stereo zvok: 3min = 30MiB
  - Če hočemo zmanjšati velikost -> kompresija
- Zvočni signal je kompleksen (šum)
  - Brezizgubna kompresija ni učinkovita
- Odstranjevanje tišine
  - Tišine ne vzorčimo
  - Detekcija tišine (protišumna vrata)
- Algoritmi za kompresijo
  - Se razlikujejo od tistih za sliko
  - Visoke frekvence so pri zvoku pomembne





# Kompresija govora

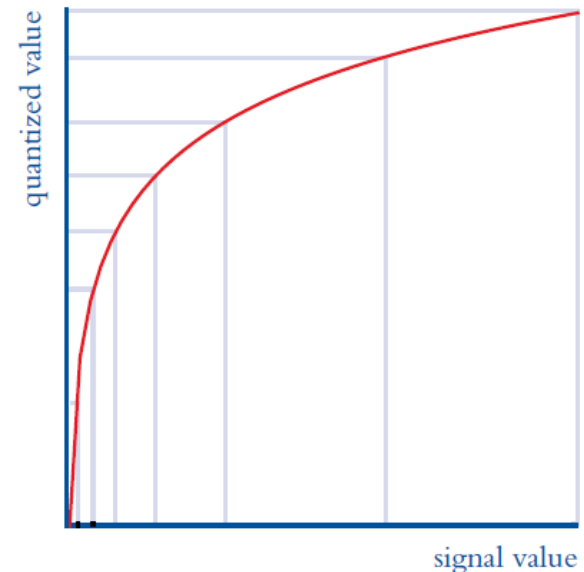
- Uporaba v prenosu govora preko telefona
  - 50-letna zgodovina
- Nelinearna kvantizacija (več nivojev za nizke jakosti) (*companding* = compressing/expanding)
  - Uporabljajo se različne log. funkcije:
    - Zakon  $\mu$ : S Amerika, Japonska, format AU
    - Zakon A

96 kbps -> 64 kbps

Telefonski signal:

frekvenca vzorčenja: 8kHz

dinamični obseg iz 12bit na 8bit





# Kompresija govora

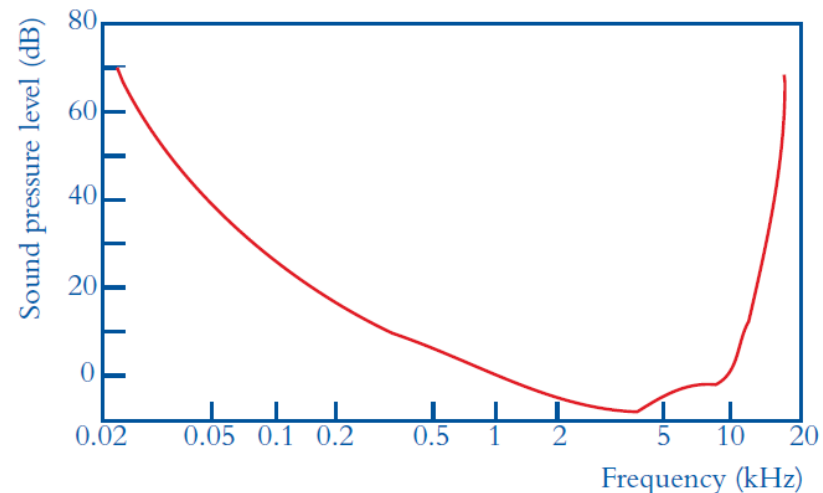
16 kbps in 32 kbps

- ***Adaptive Differential Pulse Code Modulation*** (ADPCM)
  - Shranijo se samo razlike med vzorci
    - Razlika z napovedano vrednostjo
  - Dinamično spreminjanje razpona nivojev
    - Majhne razlike – natančnejša kvantizacija
      - Ohranja se nivo podrobnosti
- ***Linear Predictive Coding*** 2.4 kbps
  - Uporaba matematičnega modela
  - Prenesejo se samo parametri
  - Ustvari „računalniški“ glas
  - Uporabno kjer je pomembnejše sporočilo in ne toliko zvok osebe



# Kompresija na osnovi zaznavanja

- Ni potrebno hraniti zvokov, ki ne morejo biti predvajani ali niso slišni
- Človeško uho tihih zvokov ne more zaznati
  - Slišni prag se spreminja glede na frekvenco
    - Psiho-akustični model:

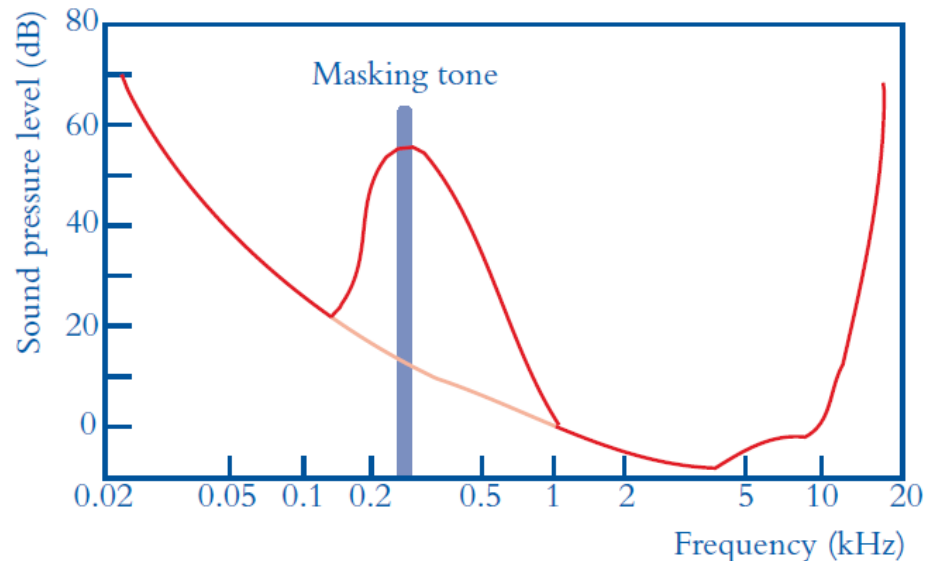


- Zvoke pod tem pragom (tihan zvoki) se ne shranjuje



# Kompresija na osnovi zaznavanja

- Glasnejši toni preglasijo tišje
  - Po tem in (zanimivo) tudi pred tem
  - Popravek praga v modelu:



- Maskiranje
  - Vsi toni pod krivuljo maske se ne slišijo = se ne hranijo



# MPEG standardi

- MPEG avdio uporablja kompresijo na osnovi zaznavanja
- MPEG-1 Part3 (3 del)
  - Layer I (Nivo 1): 192 kbps
  - Layer II (Nivo 2): 128 kbps
  - Layer III (Nivo 3): 64 kbps (MP3)
    - Razmerje kompresije 10:1
    - Dober kot fm-radio, približek CD kvaliteti
    - Lahko generiramo tudi druge kbps
- MPEG2 (Part 7), MPEG4 (Part 3)
  - Izboljšave na prostorskem zvoku
  - Uporablja AAC (Advanced Audio Coding)
    - Razvit v MPEG-2, uporabljen tudi v MPEG-4
    - Boljša kompresija kot pri MP3
      - AAC avdio pri 96kbps = MP3 pri 128 kbps



# MIDI

- Glasbo lahko prenesemo na dva načina:
  - Vzorčimo zvok in prenesemo vzorčen signal
  - Zakodiramo glasbo po neki konvenciji, prenesemo kode (parametre, note), ki jih nato rekonstruiramo (zaigramo)
- MIDI ()
  - Standardni protokol za komunikacijo med elektronskimi instrumenti
  - Instrumente se lahko nadzira s pošiljanjem nizov ukazov MIDI
  - Računalnik lahko generira, predvaja ukaze MIDI
  - Veliko večje kompresijske stopnje pri glasbi
  - Lahko zakodiramo samo glasbo
    - Lahko združimo z običajnim zvočnim zaporedjem

```
MFile 1 2 480
MTrk
0 Tempo 500000
0 PrCh ch=1 p=0
0 On ch=1 n=60 v=64
480 On ch=1 n=64 v=64
960 On ch=1 n=67 v=64
1440 Off ch=1 n=60 v=64
1920 Off ch=1 n=64 v=64
2400 Off ch=1 n=67 v=64
TrkEnd
MTrk
0 PrCh ch=2 p=0
0 On ch=2 n=40 v=64
480 On ch=2 n=45 v=64
960 Off ch=2 n=40 v=64
960 On ch=2 n=47 v=64
1440 Off ch=2 n=45 v=64
TrkEnd
```

[Program za ustvarjanje glasbe](#)

