

Osvrt na predavanje

Bola i zvuk u video kompresiji

Prijenos podataka (Bitrate) nekomprimiranog video signala je u rasponu od 270 megabita po sekundi za SDTV, 1.5 gigabita za HDTV ili 3 ili više gigabita po sekundi za Ultra High Definition TV. Takva količina podataka prevelika je za učinkovitu pohranu ili transmisiju. Zbog toga svi podaci moraju biti komprimirani odnosno sažeti koristeći perceptualno kodiranje kako bismo smanjili količinu podataka koju šaljemo.

Perceptualno kodiranje uzima u obzir karakteristike ljudskih organa za primanje signala, tj. oči i uši da bi se ustanovile granice naše percepcije i shodno tome reducirali podaci koje ne možemo primijetiti. Te operacije se izvode prilikom transmisije video signala ili prilikom samog snimanja kamerama, mobilnim uređajima itd. Svi oni imaju mali kapacitet pohrane i moraju izvoditi neku vrstu kompresije da bi mogli spremati podatke. Također, prilikom transmisije video i zvučnog signala, moramo voditi i o količini podataka koje šaljemo kako bi primatelj signala mogao neometano pratiti signal bez prevelikih gubitaka.

Kada govorimo o sažimanju podataka direktno prilikom snimanja govorimo o kodiranju izvora (Source Coding) jer sažimamo analogni signal koji dolazi na senzore uređaja za snimanje, za razliku od kodiranja podataka u post-procesiranju kako bi se on prilagodio za različite medije. Tu razlikujemo dvije vrste kodiranja odmah na izvoru i kasnije u post-procesiranju. Nakon kodiranja izvora, količina podataka svest će se na 1 do 15 megabita, ovisno o algoritmu kompresije koju koristimo.

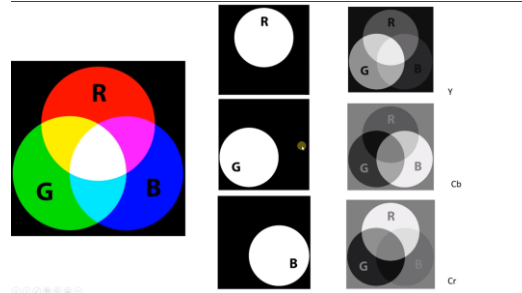
Postoji reduciranje suvišnih i nevažnih podataka. Suvišni (redundantni) podaci su oni koji se ponavljaju više puta, a nevažni (irrelevantni) su oni koje ljudsko oko ne će primijetiti ukoliko nedostaju. Oni mogu smanjiti količinu podataka i preko 100 puta. Redundantni podaci mogu se lako i bez gubitaka izračunati nekim matematičkim algoritmima prilikom dekodiranja (lossless kompresija). Kompresija irrelevantnih podataka naziva se kompresija s gubitcima (lossy kompresija).

- Redukcija podataka za boju
- RGB -> luminantna i krominantna komponenta
- YCbCr (YUV)

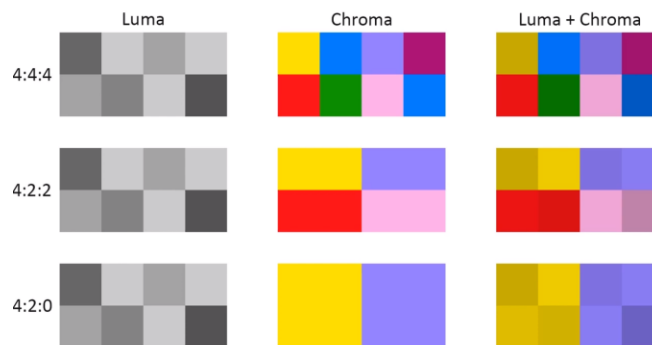
$$Y = (0.3 \cdot R) + (0.59 \cdot G) + (0.11 \cdot B)$$

$$Cb = 0.56 \cdot (B - Y)$$

$$Cr = 0.71 \cdot (R - Y)$$

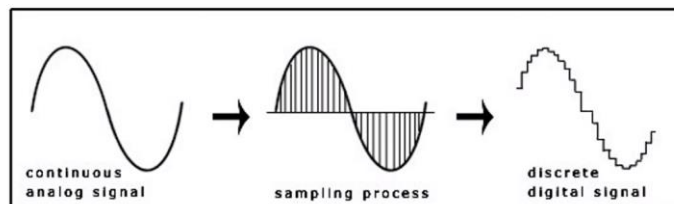


Ljudsko uho ima dinamički raspon od oko 140 decibela i raspon frekvencija od 20 herca (Hz) do 20 kiloherca (kHz). Valovi veće frekvencije od 20 kiloherca nazivaju se ultrazvuk, a frekvencije manje od 20 herca nazivaju se infrazvuk. Audio signal koji želimo transmitirati digitalnim putem mora imati karakteristike koje poštuju raspone čujnosti ljudskog uha.

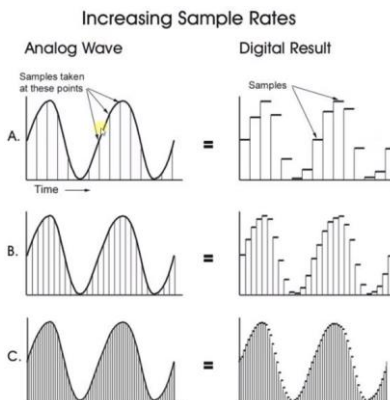


Analogni signal snima se različitim uređajima te se putem filtera limitira njegov raspon prije nego što se digitalizira. Taj proces zove se sempliranje ili uzorkovanje signala. Time analogni signal činimo mjerljivim.

Sampling rate – brzina uzorkovanja

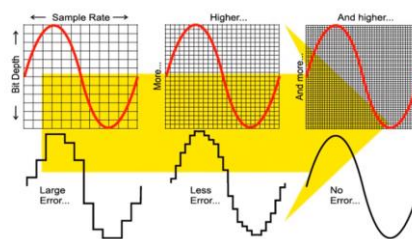


Sampling rate ili brzina uzorkovanja se mjeri kao broj snimljenih audio uzoraka unutar jedne sekunde.



Sampling depth ili preciznost uzorka se mjeri u bitovima po uzorku. Ta mjera određuje koliko mogućih razina amplitude može postići zvučni signal.

Sampling depth – preciznost uzorka



Bit rate = Sample rate * Sample depth = uzorak / sek * bit / uzorku = bit / sek

Kompresija podataka => redundantni i irelevantni podaci

Najčešće korišteni sampling rate:

- 9 kHz – telefonija
- 44.1/48 kHz – TV / CD
- 96/192 kHz – blu-ray
- > 300 kHz