

OSVRT

-

BOJA I ZVUK U VIDEO KOMPRESIJI

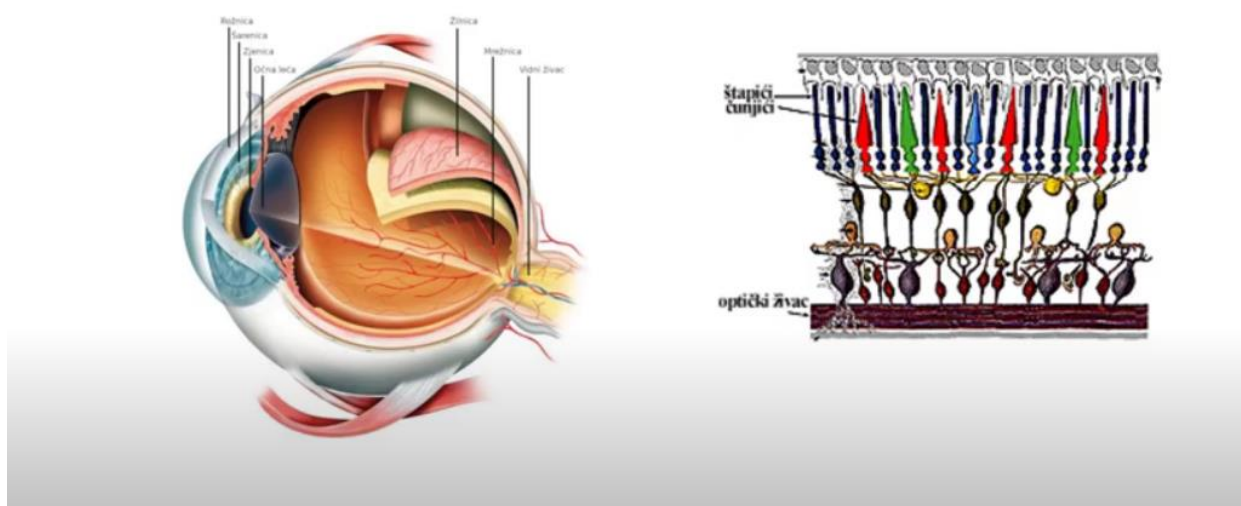
Kada imamo preveliku količinu podataka koja nije pogodna za efikasnu transmisiju podataka potrebno ih je komprimirati pomoću perceptualnog kodiranja. Perceptualno kodiranje uzima u obzir limite ljudskih organa koji primaju signal (uši i oči), te na taj način reducira podatke koje mi ne možemo primjetiti. Taj proces se odvija tokom transmisije video podataka ili prilikom samog snimanja fotoaparatom, s obzirom da svi imaju mali kapacitet pohrane potrebno je da izvedu nekakvu vrstu kompresije kako bi mogli spremiti podatke.

Kada radimo transmisiju video i zvučnog signala, moramo voditi brigu i o količini podataka koje šaljemo kako bi primatelj signala mogao nesmetano pratiti signal bez prevelikih gubitaka. Source coding tj. sažimanje podataka direktno tokom snimanja je sažimanje analognog signala koji dolazi na senzore uređaja za snimanje, za razliku od kodiranja podataka u postprocesiranju koji se prilagođava za različite medije. To znači da razlikujemo dvije vrste kodiranja podataka, prvo odmah na izvoru i drugo kasnije u postprocesiranju.

Nakon kodiranja izvora, količina podataka svest će se na 1 do 15 megabita, ovisno o algoritmu kompresije koju koristimo. Veličina kompresije ovisi o video CODEC-u i željenoj video rezoluciji slike, odnosno zvuka. Postoji reduciranje suvišnih (redundantnih) podataka - oni koji se ponavljaju više puta, i nevažnih (irelevantnih) podataka – oni koje ljudsko oko neće ni primjetiti da postoji. Oni mogu smanjiti veličinu podataka čak i preko sto puta.

Redundantni podaci mogu se lako i bez gubitaka izračunati nekim matematičkim algoritmima prilikom dekodiranja – to se naziva kompresija bez gubitaka ili losless kompresija. Prilikom prijenosa u toj kompresiji se originalni kodovi samo zamjenjuju kraćim kodovima. Kompresija irelevantnih podataka naziva se kompresija s gubitcima ili lossy kompresija.

Lossy kompresija se bazira na anatomiji oka, naše oči su više osjetljive na svjetlo/sjena promjenu nego na boje. Npr. U JPG formatu grublje strukture se bolje prijenose dok se finije strukture gube, jer nisu toliko primjetne ljudskom oku.



Ljudsko oko najviše je osjetljivo na zelenu boju, onda na crvenu, te plavu na zadnjem mjestu. To je rezultat evolucijskih faktora.

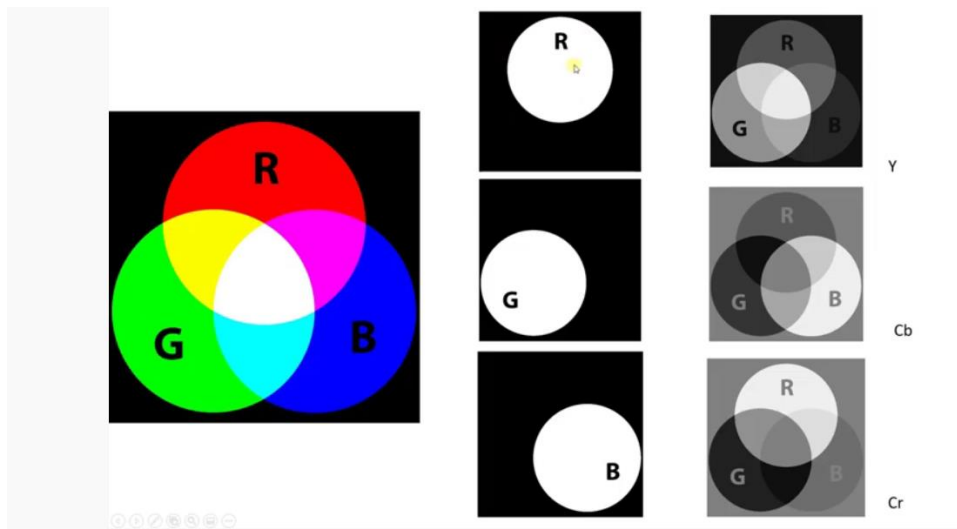
- Redukcija podataka za boju
- RGB -> luminantna i krominantna komponenta
- YCbCr (YUV)

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

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

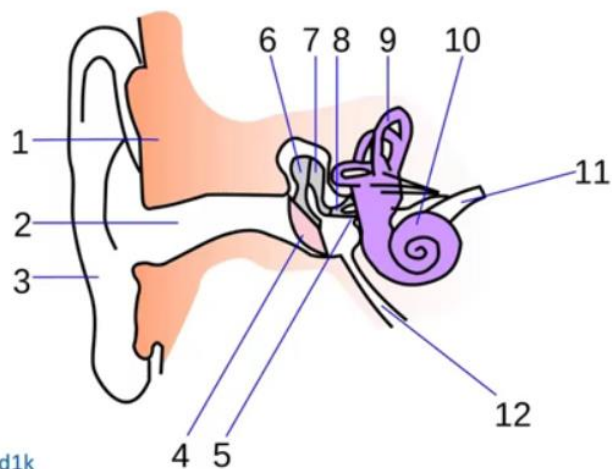
$$Cr = 0.71 \cdot (R - Y)$$

Jedan video signal se razlučuje na Luma (svjetlinu - doživljava se kao crno – bijela slika) i Chromu (podatci o boji – tj. Tonu). Omjer za sliku (odnosno za luma i chromu) može biti 4:4:4 što je originalna boja tj. Full sample, 4:2:2, 4:2:0 – dok u zadnje dvije skale tonova nije potpuna, te je malo grublja.



Ljudsko uho ima dinamički raspon od oko 140 decibela i raspon frekvencija od 20 herca do 20 000 herca. Valovi veće frekvencije od 20 000 herca 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.

- Digitalni audio signal



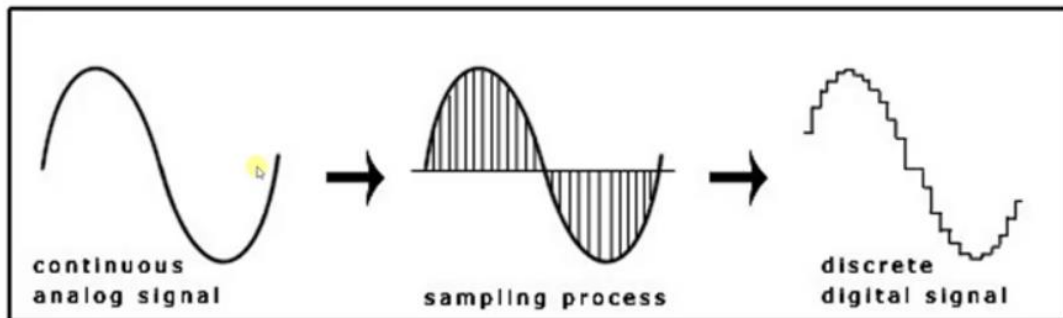
<https://www.youtube.com/watch?v=qNf9nzvnd1k>

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

Sampling rate tj. brzina uzorkovanja mjeri se kao broj snimljenih audio uzoraka unutar jedne sekunde. Ona nam govori kolika je finoća uzoraka u jedinici vremena.

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

Sampling rate – brzina uzorkovanja



$$\text{Bit rate} = \text{Sample rate} * \text{Sample depth}$$

$$= \text{uzorak / sek} * \text{bit / uzorku} = \text{bit / sek}$$

Kompresija podataka => redundantni i irelevantni podaci

- Najčešće korišteni sampling rate:
- 8 kHz – telefonija
- 44.1/48 kHz – TV / CD
- 96/192 kHz – blu-ray
- > 300 kHz